University Texts in the Mathematical Sciences

Raju K. George
Abhijith Ajayakumar

A Course in Linear Algebra

Springer

University Texts in the Mathematical Sciences

Textbooks in this series cover a wide variety of courses in mathematics, statistics and computational methods. Ranging across undergraduate and graduate levels, books may focus on theoretical or applied aspects. All texts include frequent examples and exercises of varying complexity. Illustrations, projects, historical remarks, program code and real-world examples may offer additional opportunities for engagement. Texts may be used as a primary or supplemental resource for coursework and are often suitable for independent study.

Raju K. George · Abhijith Ajayakumar

A Course in Linear Algebra

Springer

Raju K. George
Department of Mathematics
Indian Institute of Space Science
and Technology
Valiamala, Kerala, India

Abhijith Ajayakumar
Department of Mathematics
Indian Institute of Space Science
and Technology
Valiamala, Kerala, India

ISSN 2731-9318 ISSN 2731-9326 (electronic)
University Texts in the Mathematical Sciences
ISBN 978-981-99-8679-8 ISBN 978-981-99-8680-4 (eBook)
https://doi.org/10.1007/978-981-99-8680-4

Mathematics Subject Classification: 15Axx, 15Bxx, 46Bxx, 46Cxx, 47Axx, 47Bxx, 15A03, 15A04, 15A06, 15A09, 15A10, 15A18, 15A20, 15A63, 15B10, 46B20, 46C50, 47A05, 47A15, 47A30, 47B02, 47B15

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To all mathematics lovers, whose curiosity and passion inspire the pursuit of knowledge and discovery.

And to our dear friends and families, whose support and encouragement light the path of our journey.

Preface

Linear Algebra is one of the essential mathematical disciplines that undergraduate science and engineering students need for their studies. The subject stays at the fore of mathematics to date, be it in theory or applications as it provides an elegant and effective framework for comprehending and controlling the fundamental structures of space and transformation.

Our book is an exposition of basic linear algebra. We aim to present an introduction to linear algebra and basic functional analysis in a simple manner that will be helpful for readers regardless of their field of study. The book is self-contained, apt even for an upper undergraduate in their first brush with the subject. Abstract concepts are dealt with comparatively less rigor, keeping in mind a first-time reader.

The book aims mainly at graduate and engineering students and can be used as a primary text in a suitable course in Linear Algebra or as a supplementary reading. The subject matter is brought to life with numerous examples and solved problems, included as terminal chapters. This feature of the book would help students familiarize themselves with the subject matter faster and stimulate further interest in the subject. The book is written in such a way that students who are attempting competitive exams for higher studies can master the subject and attain problem-solving skills in the subject matter.

The book contains 13 chapters, of which the first seven chapters in Part 1 deal with the basic theory of linear algebra, basic functional analysis and a glimpse of applications of linear algebra. The last six chapters include solved problems based on the theory discussed in Part 1.

Chapter 1 sets the ground for the reader by dealing with preliminary topics. This chapter provides a basic understanding of elementary set theory, metric spaces and properties, and matrix theory, which is unavoidable in a linear algebra and functional analysis study. Solutions to the system of linear equations are also discussed. This chapter will motivate a beginner and help a proficient reader refresh the basics required to learn the upcoming topics.

Chapter 2 introduces the primary object in Linear Algebra, viz. Vector Spaces. Numerous examples follow the definition. Important notions of subspaces, linear dependence, basis, and dimension are given due respect and are elaborated. The chapter concludes with a section on sums and direct sums.

Having dealt with vector spaces, Chap. 3 focuses on mappings between vector spaces, particularly those which preserve the vector space structure, that is, linear transformations. Important terminologies, including range space, null space, rank, nullity, etc., are defined, followed by several important theorems. In order to bring a parallelism with matrix theory, matrix representations of linear transformations are discussed, and most of the abstract concepts related to linear transformations are dealt with in terms of matrices. The algebra of linear transformations is discussed, which would aid in constructing a vector space of linear transformations. This chapter includes a geometrical overview of varied linear transformations in $\mathbb{R}^2$, which would kindle the readers' geometric intuition. Topics like the change of coordinate matrix, linear functionals, dual space, etc. are also discussed in this chapter.

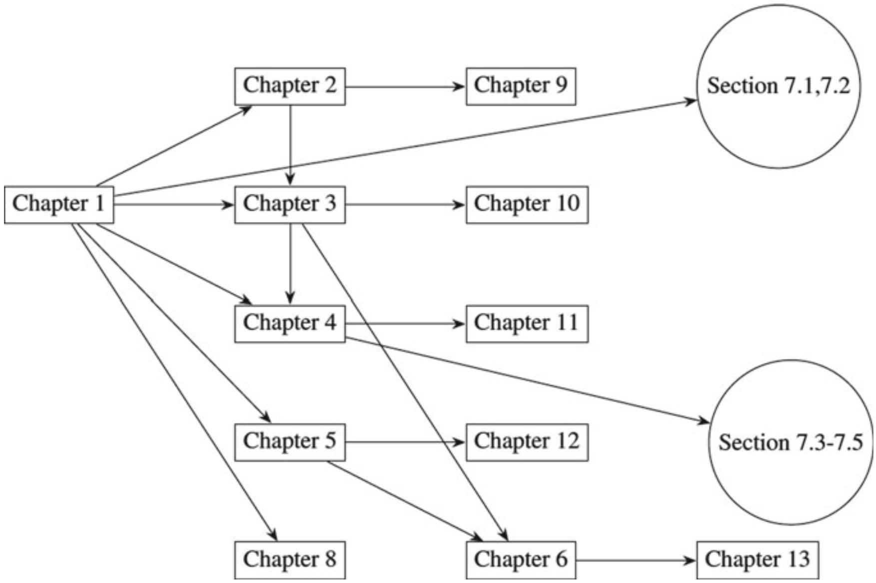
Chapter 4 is of paramount importance, as it discusses the spectral properties of matrices. Here, we study linear operators between finite-dimensional vector spaces in terms of matrices, employing matrix representations defined in Chap. 3. Important notions like eigenvalues, eigenvectors, and some useful classes of polynomials arising from matrices are defined. Important theorems like the Cayley–Hamiton theorem, Schur triangularization theorem, etc. are stated and proved in this chapter. An entire section is devoted to diagonalization. The idea of generalized eigenvectors and Jordan canonical form are also studied in detail, along with the discussion of algebraic and geometric multiplicities of eigenvalues of matrices.

We start Chap. 5 by introducing the distance notion by defining norms on arbitrary vector spaces and discussing the properties of normed linear spaces. Further, we discuss the idea of the usual dot product on $\mathbb{R}^2$ to generalize the concept to arbitrary vector spaces and obtain inner product spaces. Basic notions are introduced and theorems on inner product are proved, followed by a discussion on orthonormal sets, orthogonal projection, and the famous Gram–Schmidt Orthonormalization process. In this chapter, we revisit the notion of completeness of abstract spaces (introduced in Chap. 1), which helps to introduce Banach and Hilbert space notions. In short, this chapter briefly introduces fundamental ideas of functional analysis, a major mathematical discipline with roots in algebra and analysis.

Chapter 6 gives a flavor of operator theory by discussing bounded linear maps and their properties. Also, fundamental theorems on the adjoint of an operator, self-adjoint operators, normal operators, unitary operators, etc. are proved in this chapter. Singular value decomposition (SVD) and pseudo-inverse of matrices are discussed in detail. This chapter ends with a discussion on the least square solutions of system of linear equations.

In Chap. 7, we delve into the intriguing world of real-life linear algebra applications. Although there are many applications, we discuss only a few to give an idea on how the concepts in linear algebra are used in the real-world systems. As we progress through this chapter, we will see how the diverse tools of linear algebra open doors to innovation and new pathways for problem-solving across a wide range of disciplines.

Each chapter is provided with an ample number of examples and exercises. Solutions to selected exercises are given at the end of Part I. Chapters 8–13 comprise more than 500 solved problems of varying difficulty levels based on the topics discussed in Chaps. 1–6. Detailed solutions are given for each question to provide a better understanding of the ideas discussed. The following chapter-wise dependent chart demonstrates the sequential progression of topics throughout the book.



We are deeply indebted to all the authors whose works on linear algebra and functional analysis influenced our understanding of the subject. We take this opportunity to express our sincere gratitude toward them. We wish to acknowledge the support we received from our institution and the moral support from our colleagues and friends during each stage of manuscript preparation. We thank our academic fraternity, who have made valuable suggestions after reading through various parts of the manuscript. We would especially like to thank Manilal K. (Professor, University College, Trivandrum), Thomas V. O. (Professor, The Maharaja Sayajirao University of Baroda), Mahesh T. V. (Assistant Professor, MG College, Trivandrum), Mathew Thomas (Assistant Professor, St. Thomas College, Thrissur), Aleena Thomas (Research Scholar, Indian Institute of Space Space Science and Technology, Trivandrum), and Anikha S. Kumar (Research Scholar, Indian Institute of Space Space Science and Technology, Trivandrum) for their fruitful suggestions and constant support. Last but not the least, we wish to thank our family members for their patience and support during the preparation of this manuscript.

No work is ever complete until it has had its fair share of criticism. Readers are welcome to comment on our dispositions, which will help us improve the book.

Thiruvananthapuram, Kerala
October 2023

Raju K. George
Abhijith Ajayakumar

Contents

Part I
Theory of Linear Algebra

Chapter 1
Preliminaries

We introduce a wide range of fundamental mathematical concepts and structures in this chapter on foundation of mathematics. Understanding their fundamental operations and attributes, we start with sets and functions. We then delve into the metric space universe, which offers a framework for comprehending distance and convergence. Moving on to algebraic structures, we examine the distinctive qualities and illustrative instances of groups, rings, and fields. Polynomial rings and their essential properties are introduced, as are matrices and their rank, trace, and determinant, all of which are highlighted as they have vital roles in the coming chapters. The latter sections of the chapter provide an overview of Euclidean space and demonstrate how to solve systems of linear equations using techniques like Cramer's rule, LU decomposition, Gauss elimination, etc. These fundamental ideas in mathematics serve as the building blocks for more complex mathematical research and have numerous applications in science and engineering.

1.1 Sets and Functions

Set theory is the core of modern mathematics and serves as a language for mathematicians to discuss and organize their ideas. It is a crucial and elegant concept at its core; a set is simply a collection of objects, similar to a bag containing multiple objects. These objects can be anything from numbers, characters, shapes, or other sets. The way set theory lets us classify, compare, and evaluate these collections is what makes it so powerful. This section will discuss some of the essential concepts in set theory. Though the notion of set is not well-defined in wide generality as it leads to paradoxes like Russell's Paradox, published by *Bertrand Russell (1872–1970)* in 1901, we start with the following simple definition for a preliminary understanding of a set.

Definition 1.1 (*Set*) A set is a well-defined collection of objects. That is, to define a set X, we must know for sure whether an element x belongs to X or not. If x is

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_1

an element of X, then it is denoted by $x \in X$ and if x is not an element of X, then it is denoted by $x \notin X$. Two sets X and Y are said to be equal if they have the same elements.

Definition 1.2 (*Subset*) Let X and Y be any two sets, then X is a subset of Y, denoted by $X \subseteq Y$, if every element of X is also an element of Y. Two sets X and Y are equal if and only if $X \subseteq Y$ and $Y \subseteq X$.

A set can be defined in a number of ways. Commonly, a set is defined by either listing all the entries explicitly, called the *Roster form*, or by stating the properties that are meaningful and unambiguous for elements of the set, called the *Set builder form*.

Example 1.1 Here are some familiar collection/sets of numbers.

$\mathbb{N}$−the set of all natural numbers $-\{1, 2, 3, \ldots\}$
$\mathbb{W}$−the set of all whole numbers $-\{0, 1, 2, \ldots\}$
$\mathbb{Z}$−the set of all integers $-\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
$\mathbb{Q}$−the set of all rational numbers $-\{\frac{p}{q} \mid p, q \in \mathbb{Z},\ q \neq 0\}$
$\mathbb{R}$−the set of all real numbers
$\mathbb{C}$−the set of all complex numbers.

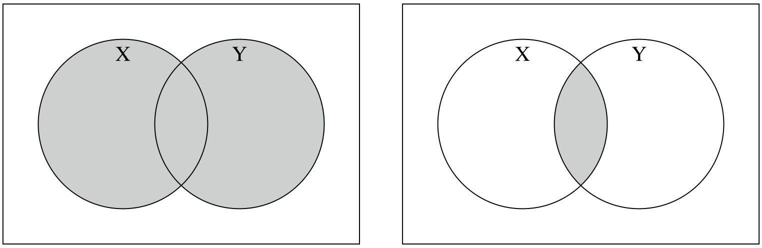
Usually, in a particular context, we have to deal with the elements and subsets of a basic set which is relevant to that particular context. This basic set is called the "Universal Set" and is denoted by $\mathcal{U}$. For example, while studying the number system, we are interested in the set of natural numbers, $\mathbb{N}$, and its subsets such as the set of all prime numbers, the set of all odd numbers, and so forth. In this case $\mathbb{N}$ is the universal set. A null set, often known as an empty set, is another fundamental object in set theory. It is a set with no elements, which means it has no objects or members. In set notation, the null set is commonly represented by Φ or $\{\}$ (an empty pair of curly braces).

Definition 1.3 (*Cardinality*) The cardinality of a set X is the number of elements in X. A set X can be finite or infinite depending on the number of elements in X. Cardinality of X is denoted by $|X|$.

Example 1.2 All the sets mentioned in Example 1.1 are infinite sets. The set of letters in the English alphabet is a finite set.

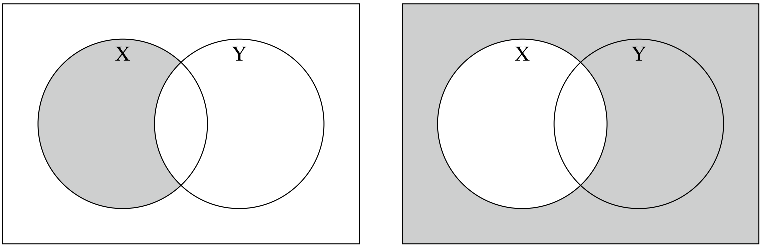
Set Operations

Set operations are fundamental mathematical methods for constructing, manipulating, and analyzing sets. They enable the combination, comparison, and modification of sets in order to acquire insights and solve various mathematical and real-life problems. Union (combining items from several sets), intersection (finding common elements between sets), complement (identifying elements not in a set), and set difference (removing elements from one set based on another) are the fundamental set operations.



(a) (b)

Fig. 1.1 The shaded portions in **a** and **b** represents the union and intersection of the sets X and Y, respectively



(a) (b)

Fig. 1.2 The shaded portion in **a** represents the difference of Y related to X and the shaded portion in **b** represents the complement of a set

Definition 1.4 (*Union and Intersection*) Let X and Y be two sets. The union of X and Y, denoted by $X \cup Y$, is the set of all elements that belong to either X or Y. The intersection of X and Y, denoted by $X \cap Y$, is the set of all elements that belong to both X and Y.

The relationship between sets can be illustrated with the use of diagrams, known as *Venn diagrams*. It was popularized by the famous mathematician *John Venn (1834–1923)*. In a Venn diagram, a rectangle is used to represent the universal set and circles are used to represent its subsets. For example, the union and intersection of two sets are represented in Fig. 1.1.

Definition 1.5 (*Difference of Y related to X*) Let X and Y be two sets. The difference of Y related to X, denoted by $X \setminus Y$, is the set of all elements in X which are not in Y. The difference of a set X related to its universal set $\mathcal{U}$ is called the *complement* of X and is denoted by X^c. That is, $X^c = \mathcal{U} \setminus X$. Keep in mind that $\mathcal{U}^c = \Phi$ and $\Phi^c = \mathcal{U}$ (Fig. 1.2).

Definition 1.6 (*Cartesian Product*) Let X and Y be two sets. The Cartesian product of X and Y, denoted by $X \times Y$, is the set of all ordered pairs (x, y) such that x belongs to X and y belongs to Y. That is, $X \times Y = \{(x, y) \mid x \in X, y \in Y\}$.

Example 1.3 Let $X = \{1, 2, 3\}$ and $Y = \{3, 4, 5\}$. Then the union and intersection of X and Y are $X \cup Y = \{1, 2, 3, 4, 5\}$ and $X \cap Y = \{3\}$, respectively. The difference of Y related to X is $X \setminus Y = \{1, 2\}$, and the Cartesian product of X and Y is $X \times Y = \{(1, 3), (1, 4), (1, 5), (2, 3), (2, 4), (2, 5), (3, 3), (3, 4), (3, 5)\}$.

Remark 1.1 Two sets X and Y are said to be **disjoint**, if their intersection is empty. That is, if $X \cap Y = \Phi$.

We will now try to "connect" elements of distinct sets using the concept, "Relations". A relation between two sets allows for the exploration and quantification of links and relationships between elements of various sets. It essentially acts as a link between elements of another set and elements from another, exposing patterns, dependencies, or correspondences.

Definition 1.7 (*Relation*) A relation R from a non-empty set X to a non-empty set Y is a subset of the Cartesian product $X \times Y$. It is obtained by defining a relationship between the first element and second element (called the *"image"* of first element) of the ordered pairs in $X \times Y$.

The set of all first elements in a relation R is called the domain of the relation R, and the set of all second elements is called the range of R. As we represent sets, a relation may be represented either in the roster form or in the set builder form. In the case of finite sets, a visual representation by an arrow diagram is also possible.

Example 1.4 Consider the sets X and Y from Example 1.3 and their Cartesian product $X \times Y$. Then $R = \{(1, 3), (2, 4), (3, 5)\}$ is a relation between X and Y. The set builder form of the given relation can be given by $R = \{(x, y) \mid y = x + 2, x \in X, y \in Y\}$ (Fig. 1.3).

Remark 1.2 If $|X| = m$ and $|Y| = n$, then $|X \times Y| = mn$ and the number of possible relations from set X to set Y is 2^{mn}.

Definition 1.8 (*Equivalence Relations*) A relation R on a set X is said to be an equivalence relation if and only if the following conditions are satisfied:

(a) $(x, x) \in R$ for all $x \in X$ (*Reflexive*)
(b) $(x, y) \in R$ implies $(y, x) \in R$ (*Symmetric*)
(c) $(x, y) \in R$ and $(y, z) \in R$ implies $(x, z) \in R$ (*Transitive*).

Example 1.5 Consider $\mathbb{N}$ with the relation R, where $(x, y) \in R$ if and only if $x - y$ is divisible by n, where n is a positive integer. We will show that R is an equivalence relation on $\mathbb{N}$. For,

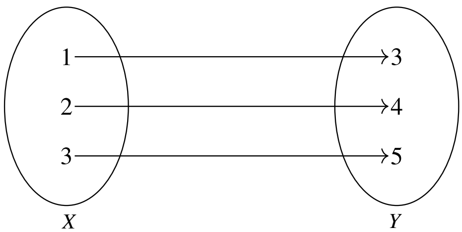


Fig. 1.3 Arrow diagram for R

(a) $(x, x) \in R$ for all $x \in \mathbb{N}$. For, $x - x = 0$ is divisible by n for all $x \in \mathbb{N}$.
(b) $(x, y) \in R$ implies $(y, x) \in R$. For,

$$\begin{aligned}(x, y) \in R &\Rightarrow x - y \text{ is divisible by } n \\ &\Rightarrow -(x - y) \text{ is divisible by } n \\ &\Rightarrow y - x \text{ is divisible by } n \\ &\Rightarrow (y, x) \in R\end{aligned}$$

(c) $(x, y) \in R$ and $(y, z) \in R$ implies $(x, z) \in R$. For,

$$\begin{aligned}(x, y), (y, z) \in R &\Rightarrow x - y \text{ and } y - z \text{ is divisible by } n \\ &\Rightarrow (x - y) + (y - z) \text{ is divisible by } n \\ &\Rightarrow x - z \text{ is divisible by } n \\ &\Rightarrow (x, z) \in R\end{aligned}$$

Thus, R is reflexive, symmetric, and transitive. Hence, R is an equivalence relation.

Example 1.6 Consider the set $X = \{1, 2, 3\}$. Define a relation R on X by $R = \{(1, 1), (2, 2), (1, 2), (2, 3)\}$. Is R an equivalence relation? Clearly, not! We can observe that R is not reflexive as $(3, 3) \notin R$. Also R is not transitive as $(1, 2), (2, 3)$ but $(1, 3) \notin R$. What if we include the elements $(3, 3)$ and $(1, 3)$ to the relation and redefine R as $\tilde{R} = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 3), (1, 3)\}$. Then $\tilde{R}$ is an equivalence relation on X.

Relations define how elements from one set correspond to elements from another, allowing for a broader range of relationships. However, there are specialized relations in which each element in the first set uniquely relates to one element in the second. This connection gives these relations mathematical precision, making them crucial for modeling precise transformations and dependencies in various mathematical disciplines, ranging from algebra to calculus. We refer to such relations as functions.

Functions

Function in mathematics is a rule or an expression that relates how a quantity (dependent variable) varies with respect to another quantity (independent variable) associated with it. They are ubiquitous in mathematics and they serve many purposes.

Definition 1.9 (*Function*) A function f from a set X to a set Y, denoted by $f : X \to Y$, is a relation that assigns to each element $x \in X$ exactly one element $y \in Y$.

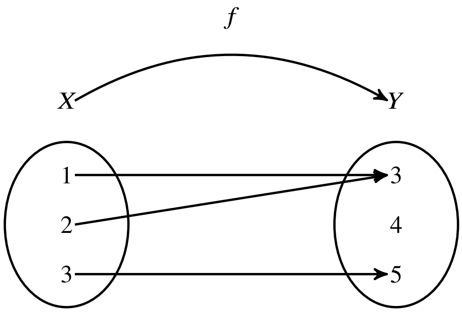
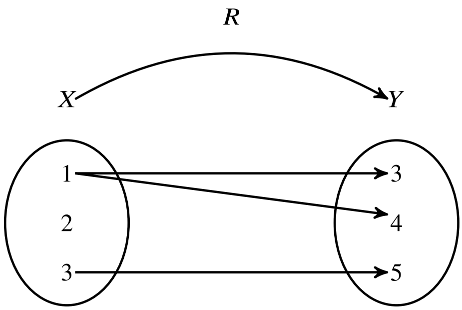


Fig. 1.4 Observe that each element from set X is mapped to exactly one element in set Y. Therefore the given relation is a function. X is called *domain* of f and Y is called the *co-domain* of f. 4 does not belong to the range set of f, as it does not have a pre-image. The range set of f is $\{3, 5\}$

Fig. 1.5 Observe that 1 is mapped to both 3 and 4. Thus $R = \{(1, 3), (1, 4), (2, 3)\}$ is not a function



Then y is called the image of x under f and is denoted by $f(x)$. The set X is called the domain of f and Y is called the co-domain of f. The collection of all images of elements in X is called the range of f.

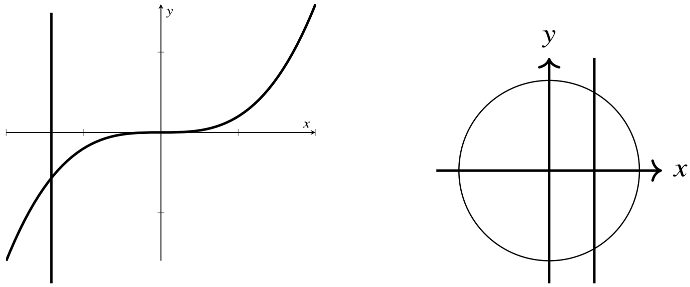
Example 1.7 Consider the sets X and Y from Example 1.3. Define a relation R from the set X to the set Y as $R = \{(1, 3), (2, 3), (3, 5)\}$. Then the relation R is a function from X to Y (Fig. 1.4).

From Definition 1.9, it is clear that any function from a set X to a set Y is a relation from X to Y. But the converse need not be true. Consider the following example.

Example 1.8 Consider the sets X and Y from Example 1.3. Then the relation $R = \{(1, 3), (1, 4), (2, 3)\}$ from the set X to the set Y is not a function as two distinct elements of the set Y are assigned to the element 1 in X (Fig. 1.5).

It would be easier to understand the dependence between the elements if we could geometrically represent a function. As a convention, the visual representation is done by plotting the elements in the domain along the horizontal axis and the corresponding images along the vertical axis.

Definition 1.10 (*Graph of a Function*) Let $f : X \to Y$ be a function. The set $\{(x, f(x)) \in X \times Y \mid x \in X\}$ is called the graph of f.



(a) function (b) not a function

Fig. 1.6 Observe that in the first graph any vertical line drawn in the domain will touch exactly one point of the graph. However, in the second graph it may touch more than one point

Observe that the above-defined set is exactly the same as f, by Definition 1.9. Also keep in mind that not all graphs represent a function. If any vertical line intersects a graph at more than one point, the relation represented by the graph is not a function. This is known as the *vertical line test* (Fig. 1.6).

Definition 1.11 (*One-one function and Onto function*) A function f from a set X to a set Y is called a one-one (injective) function if distinct elements in the domain have distinct images, that is, for every $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ implies $x_1 = x_2$. f is called onto (surjective) if every element of Y is the image of at least one element of X, that is, for every $y \in Y$, $\exists x \in X$ such that $f(x) = y$. A function which is both one-one and onto is called a *bijective* function.

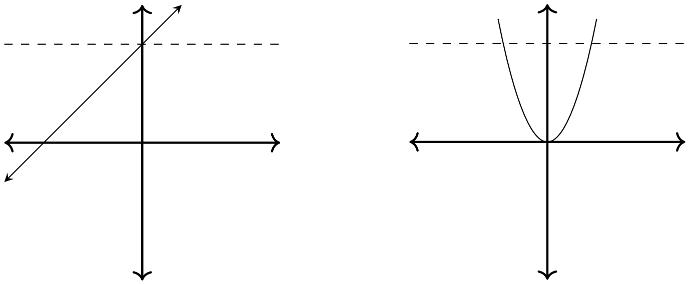
Example 1.9 Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x + 5$ for $x \in \mathbb{R}$. First, we will check whether the function is one-one or not. We will start by assuming $f(x_1) = f(x_2)$ for some $x_1, x_2 \in \mathbb{R}$. Then

$$\begin{aligned} f(x_1) = f(x_2) &\Rightarrow x_1 + 5 = x_2 + 5 \\ &\Rightarrow x_1 = x_2 \end{aligned}$$

Therefore f is one-one. Now to check whether the function is onto, take any $x \in \mathbb{R}$, then $x - 5 \in \mathbb{R}$ with $f(x - 5) = x - 5 + 5 = x$. That is, every element in $\mathbb{R}$ (co-domain) has a pre-image in $\mathbb{R}$ (domain). Thus, f is onto and hence f is a bijective function.

The graph of a function can also be used to check whether a function is one-one. If any horizontal line intersects the graph more than once, then the graph does not represent a one-one function as it implies that two different elements in the domain have the same image. This is known as the *horizontal line test* (Fig. 1.7).

Definition 1.12 (*Composition of two functions*) Let $f : X \to Y$ and $g : Y \to Z$ be any two functions, then the composition $g \circ f$ is a function from X to Z, defined by $(g \circ f)(x) = g(f(x))$ (Fig. 1.8).



(a) $f_1(x) = x+5$ (b) $f_2(x) = x^2$

Fig. 1.7 Consider the graphs of the functions $f_1, f_2{:}\mathbb{R} \to \mathbb{R}$ defined by $f_1(x) = x + 5$ and $f_2(x) = x^2$. Observe that if we draw a horizontal line parallel to the x-axis, it will touch exactly one point on the graph of the function f_1. But on the graph of the function f_2, it touches two points. Then by *horizontal line test*, the first function is one-one whereas the second one is not a one-one function

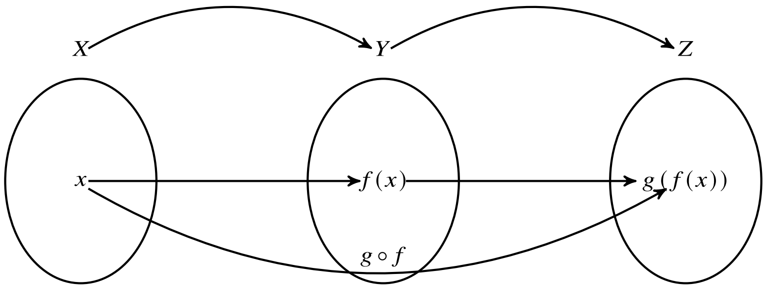


Fig. 1.8 It is clear that the range set of f must be a subset of the domain of g, for the composition function to be defined

Properties

Let $f : X \to Y$, $g : Y \to Z$, and $h : Z \to W$, then

(a) $h \circ (g \circ f) = (h \circ g) \circ f$ (Associative).
(b) If f and g are one-one, then $g \circ f$ is one-one.
(c) If f and g are onto, then $g \circ f$ is onto.

Example 1.10 Consider the functions $f, g : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ and $g(x) = 2x + 1$. Then $(f \circ g)(x) = f(g(x)) = f(2x+1) = (2x+1)^2 = 4x^2 + 4x + 1$ and $(g \circ f)(x) = g(f(x)) = g(x^2) = 2x^2 + 1$. Observe that $f \circ g \neq g \circ f$. Therefore function composition need not necessarily be commutative.

Definition 1.13 (*Inverse of a function*) A function $f : X \to Y$ is said to be invertible if there exists a function $g : Y \to X$ such that $g(f(x)) = x$ for all $x \in X$ and $f(g(y)) = y$ for all $y \in Y$. The inverse function of f is denoted by f^{-1}.

The function f is invertible if and only if f is a bijective function. For, suppose there exists an inverse function g for f. Then

$$f(x_1) = f(x_2) \Rightarrow g(f(x_1)) = g(f(x_2)) \Rightarrow x_1 = x_2$$

That is, f is injective. And $f(g(y)) = y$ for all $y \in Y$ implies that f is onto.

Example 1.11 Consider a function f, defined as in Fig. 1.4. Indeed, from the figure itself, it's evident that function f is not bijective. Thus f is not invertible. Observe that if we define $f(2) = 4$, then f is both one-one and onto. Then define a function, $g : Y \to X$ by $g(3) = 1$, $g(4) = 2$, and $g(5) = 3$. Now $g(f(1)) = g(3) = 1$, $g(f(2)) = g(4) = 2$, and $g(f(3)) = g(5) = 3$. That is, $g(f(x)) = x$ for all $x \in X$. Similarly, we can prove that $f(g(y)) = y$ for all $y \in Y$.

Example 1.12 Consider the function $f(x) = x + 5$, defined as in Example 1.9. We have already shown that the function is bijective. Now, we will find the inverse of f. By definition, we can say that f^{-1} is the function that will undo the operation of f. That is, if a function f maps an element x from set X to y in set Y, its inverse function f^{-1} reverses this mapping, taking y from Y back to x in X. In this case, $X = Y = \mathbb{R}$. If we consider, a $y \in \mathbb{R}$(co-domain), then there exists $x \in \mathbb{R}$(domain) such that $y = x + 5$ (Why?). Then $x = y - 5$. Thus, the function $g(y) = y - 5$ will undo the action of f. We can verify this algebraically as follows:

$$g(f(x)) = g(x + 5) = x + 5 - 5 = x$$

and

$$f(g(x)) = f(x - 5) = x - 5 + 5 = x$$

Thus $f^{-1}(x) = x - 5$.

Example 1.13 Now consider $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. From Fig. 1.7, we can clearly say that f is not bijective. Thus f does not have an inverse in $\mathbb{R}$. But, if we restrict the domain of f to $[0, \infty)$, f is a bijective function. Then the inverse of f is the function $f^{-1}(x) = \sqrt{x}$. For, $g(f(x)) = g(x^2) = \sqrt{x^2} = x$ and $f(g(x)) = f(\sqrt{x}) = (\sqrt{x})^2 = x$.

It is easy to check whether a real function is invertible or not, by just looking at its graph. Consider Fig. 1.9.

Now we will discuss some of the important concepts related to functions defined on the set of all real numbers to itself.

Definition 1.14 (*Continuity at a point*) Let $X \subset \mathbb{R}$ and $f : X \to \mathbb{R}$ be a function. We say that f is continuous at $x_0 \in X$, if given any $\epsilon > 0$ there exists a $\delta > 0$ such that if x is any point in X satisfying $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \epsilon$. Otherwise, f is said to be discontinuous at x_0.

A function is continuous if it is continuous at each point of its domain. In graphical terms, the continuity of a function on the set of all real numbers means that the graph does not have any gaps or breaks. From Fig. 1.7, it is clear that both the functions $f(x) = x + 5$ and $f(x) = x^2$ are continuous. Figure 1.10 gives an example for a discontinuous function.

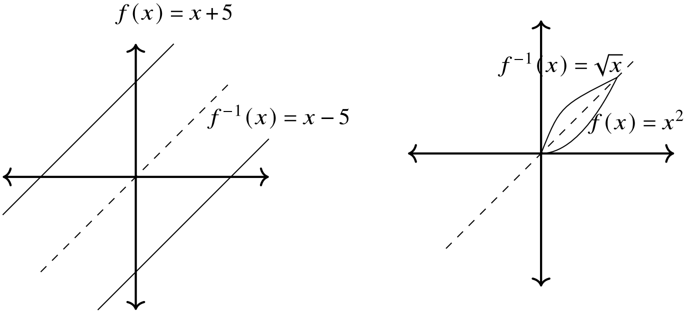
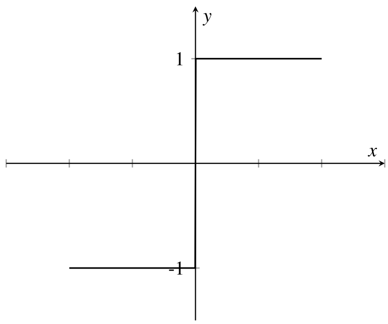


Fig. 1.9 Observe that the graph of $f^{-1}(x)$ is the graph of $f(x)$ reflected about the line $y = x$ (represented by the dotted line)

Fig. 1.10 Consider the signum function, defined by

$$f(x) = \begin{cases} 1, \; if \; x > 0 \\ 0, \; if \; x = 0 \\ -1, \; if \; x < 0 \end{cases}.$$

Clearly, f is not continuous at $x = 0$



Observe that in the definition of continuity of a function at a point, the value of δ depends on both x_0 and ϵ. If δ does not depend on the point x_0, then the continuity is called *uniform continuity*. In other words, a function f is uniformly continuous on a set X, if for every $\epsilon > 0$, there exists $\delta > 0$, such that for every element $x, y \in X$, $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$. Graphically, this means that given any narrow vertical strip of width ϵ on the graph, there exists a corresponding horizontal strip of width δ such that all points in the interval within δ units of each other on the x-axis map to points within ϵ units of each other on the y-axis. Consider the following example.

Example 1.14 Consider the function $f_1(x) = x + 5$. We will show that f_1 is uniformly continuous. For, given any $\epsilon > 0$, choose $\delta = \epsilon$. Then, for any $x, y \in \mathbb{R}$ with $|x - y| < \delta$, we have

$$|f_1(x) - f_1(y)| = |x + 5 - (y + 5)| = |x - y| < \delta = \epsilon$$

Thus $f_1(x) = x + 5$ is uniformly continuous over $\mathbb{R}$. However, the function $f_2(x) = x^2$ is not uniformly continuous on $\mathbb{R}$. Suppose on the contrary that f_2 is uniformly continuous. Fix $\epsilon = 1$. Then, there exists $\delta_0 > 0$, such that for every element $x, y \in \mathbb{R}$, $|f(x) - f(y)| < 1$ whenever $|x - y| < \delta_0$. Now, take $y = x + \frac{\delta_0}{2}$. Then,

$$|f(x) - f(y)| = \left| x^2 - \left(x + \frac{\delta_0}{2} \right)^2 \right| = \left| x\delta_0 + \frac{\delta_0^2}{4} \right| < 1$$

which is a contradiction as x can be chosen arbitrarily.

Now, we will define continuity of a function using the notion of sequences of real numbers.

Definition 1.15 (*Real Sequence*) A real sequence $\{x_n\}$ is a function whose domain is the set $\mathbb{N}$ of natural numbers and co-domain is the set of all real numbers $\mathbb{R}$. In other words, a sequence in $\mathbb{R}$ assigns to each natural number $n = 1, 2, \ldots$ a uniquely determined real number. For example, the function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = \frac{1}{n}$ determines the sequence $\left\{1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$.

Example 1.15 The list of numbers $\{r, r, r, \ldots\}$, where r is any real number, is a sequence called *constant sequence* as we can define a function, $f : \mathbb{N} \to \mathbb{R}$, by $f(n) = r$.

Example 1.16 The list of numbers $\{r, r^2, r^3, \ldots\}$, where r is any real number, is a sequence called *geometric sequence* as we can define a function, $f : \mathbb{N} \to \mathbb{R}$, by $f(n) = r^n$.

Definition 1.16 (*Convergent Sequence*) A real sequence $\{x_n\}$ is said to converge to $x \in \mathbb{R}$, or x is said to be a limit of $\{x_n\}$, denoted by $x_n \to x$ or $\lim\limits_{n\to\infty} x_n = x$, if for every $\epsilon > 0$, there exists a natural number N such that $|x_n - x| < \epsilon$ for all $n \geq N$. Otherwise, we say that $\{x_n\}$ is divergent.

Theorem 1.1 *A real sequence $\{x_n\}$ can have at most one limit.*

Example 1.17 Consider the sequence $\{x_n\}$, where $x_n = \frac{1}{n}$. Clearly, $x_n \to 0$. For, given any $\epsilon > 0$, we have $|x_n - 0| = \left|\frac{1}{n}\right|$. If we take $n > \frac{1}{\epsilon}$, we have $|1/n| < \epsilon$. Thus $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.

Example 1.18 Consider the sequence $\{x_n\}$, defined as in Example 1.15. It is easy to observe that $x_n \to r$ as $|x_n - r| = 0$ for all $n \in \mathbb{N}$.

Example 1.19 Consider the sequence $\{x_n\}$, defined as in Example 1.16. We can observe that the convergence of this sequence depends on the value of r. First of all, by the above example, for $r = 0$ and $r = 1$, $\{x_n\}$ converges to 0 and 1, respectively. Now let $0 < r < 1$. Then $x_n \to 0$. For any $\epsilon > 0$, if we take $N > \frac{ln\ \epsilon}{ln\ r}$ we have $|x_n - 0| = r^n < \epsilon$ for all $n > N$. Similarly, for $-1 < r < 0$, $x_n \to 0$.

Now for $r = -1$, the given sequence becomes $x_n = (-1)^n$. Take $\epsilon = \frac{1}{3}$. Then there does not exist any point $x \in \mathbb{R}$ such that $|x_n - x| < \frac{1}{3}$ as the interval $\left(x - \frac{1}{3}, x + \frac{1}{3}\right)$ must contain both 1 and -1. Therefore $\{x_n\}$ with $x_n = (-1)^n$ does not converge. Similarly, we can prove that the sequence $\{x_n\}$ with $x_n = r^n$ does not converge outside the interval $(-1, 1]$.

As we have discussed convergent sequences, Cauchy sequences must be introduced, which are a specific class of sequences in which the terms become arbitrarily close to each other as the index increases, rather than approaching a single limit.

Definition 1.17 (*Cauchy Sequence*) A real sequence $\{x_n\}$ is said to be a Cauchy sequence, if for any $\epsilon > 0$, there exists a natural number N such that $|x_m - x_n| < \epsilon$ for all $m, n \geq N$.

For a real sequence, the terms convergent sequence and Cauchy sequence do not make any difference. We have the following theorem stating this fact.

Theorem 1.2 *A real sequence* $\{x_n\}$ *is convergent if and only if it is Cauchy.*

However, this may not be true, if we are considering sequences in the set of rational numbers, $\mathbb{Q}$. That is, there exist sequences of rational numbers that are Cauchy but not convergent in $\mathbb{Q}$ (the sequence may not converge to a rational number). For example, consider the sequence $1.41, 1.412, 1.1421, \ldots$. This sequence will converge to $\sqrt{2}$ which is not a rational number (also, see Exercise 13 of this chapter). Now, we will introduce the sequential definition for continuity.

Definition 1.18 (*Sequential Continuity*) A function $f : X \subseteq \mathbb{R} \to \mathbb{R}$ is said to be sequentially continuous at point $x_0 \in X$ if for every $\{x_n\}$ in X with $x_n \to x_0$, we have $f(x_n) \to f(x_0)$. That is if, $\lim\limits_{n\to\infty} x_n = x_0 \Rightarrow \lim\limits_{n\to\infty} f(x_n) = f(x_0)$.

Then, we have the following result which asserts that sequential continuity and continuity of a real function are the same.

Theorem 1.3 *A function* $f : X \subseteq \mathbb{R} \to \mathbb{R}$ *is continuous if and only if it is sequentially continuous.*

Example 1.20 Consider the signum function as defined in Fig. 1.10. We know that f is not continuous at $x = 0$. We can use the definition of sequential continuity to prove this fact. Consider the sequence $\left\{\frac{1}{n}\right\}$. In Example 1.17, we have seen that $\frac{1}{n} \to 0$. However, observe that $f\left(\frac{1}{n}\right) = 1 \to 1 \neq f(0)$. Thus f is not sequentially continuous at 0 and hence f is not continuous at 0.

Now, consider the function $f(x) = x + 5$. We have already seen that f is continuous on $\mathbb{R}$ as its graph does not have any gaps or breaks. Let us check whether f is sequentially continuous or not. Consider any real number $r \in \mathbb{R}$ and a sequence $\{r_n\}$ with $r_n \to r$ as $n \to \infty$. For sequential continuity $f(r_n)$ must converge to $f(r)$. Observe that $f(r_n) = r_n + 5 \to r + 5$ as $n \to \infty$. Thus f is sequential continuous.

Remark 1.3 A set S is said to be countably infinite if there exists a bijective function from $\mathbb{N}$ to S. A set which is empty, finite, or countably infinite is called a countable set. Otherwise it is called uncountable set. For example $\mathbb{Z}$ is countable and $\mathbb{R}$ is uncountable.

Sequence of Functions

Now, we will combine the ideas of functions and sequences discussed so far and define "sequence of functions".

Definition 1.19 (*Sequence of Functions*) Let f_n be real-valued functions defined on an interval $[a, b]$ for each $n \in \mathbb{N}$. Then $\{f_1, f_2, f_3, \ldots\}$ is called a sequence of real-valued functions on $[0, 1]$, and is denoted by $\{f_n\}$.

Example 1.21 For each $n \in \mathbb{N}$, let f_n be defined on $[0, 1]$ by $f_n(x) = x^n$. Then $\{x, x^2, x^3, \ldots\}$ is a sequence of real-valued functions on $[a, b]$.

For a sequence of functions, we have two types of convergences, namely *point-wise convergence* and *uniform convergence*. We will discuss these concepts briefly in this section.

Let $\{f_n\}$ be a sequence of functions on $[a, b]$ and $x_0 \in [a, b]$. Then the sequence of real numbers, $\{f_n(x_0)\}$, may be convergent. In fact, this may be true for all points in $[a, b]$. The limiting values of the sequence of real numbers corresponding to each point $x \in X$ define a function called the limit function or simply the limit of the sequence $\{f_n\}$ of functions on $[a, b]$.

Definition 1.20 (*Point-wise convergence*) Let $\{f_n\}$ be a sequence of real-valued function defined on an interval $[a, b]$. If for each $x \in [a, b]$ and each $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ for all $n > N$, then we say that $\{f_n\}$ converges point-wise to the function f on $[a, b]$ and is denoted by $\lim_{n\to\infty} f_n(x) = f(x),\ \forall\, x \in [a, b]$.

Example 1.22 Let $f_n(x) = x^n$ be defined on $[0, 1]$. By Example 1.19, the limit function $f(x)$ is given by

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases}$$

Let $\epsilon = \frac{1}{2}$. Then for each $x \in [0, 1]$, there exists a positive integer N such that $|f_n(x) - f(x)| < \frac{1}{2}$ for all $n > N$. If $x = 0$, $f(x) = 0$ and $f_n(x) = 0$ for all n. $|f_n(x) - f(x)| < \frac{1}{2}$ is true for all $n > 1$.
If $x = 1$, $f(x) = 1$ and $f_n(x) = 1$ for all n. $|f_n(x) - f(x)| < \frac{1}{2}$ is true for all $n > 1$.
If $x = \frac{3}{4}$, $f(x) = 0$ and $f_n(x) = \left(\frac{3}{4}\right)^n$ for all n. Then

$$|f_n(x) - f(x)| = \left(\frac{3}{4}\right)^n < \frac{1}{2}$$

is true for all $n > 2$.

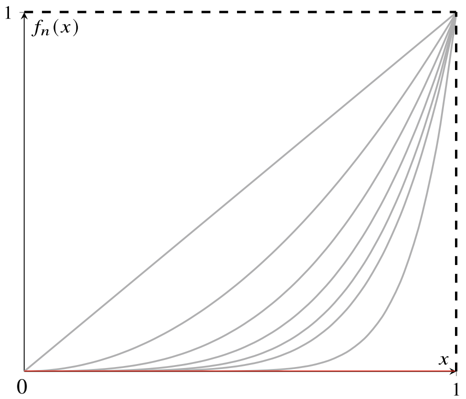


Fig. 1.11 Point-wise convergence of $\{f_n\}$, where $f_n(x) = x^n$, $x \in [0, 1]$

If $x = \frac{9}{10}$, $f(x) = 0$ and $f_n(x) = \left(\frac{9}{10}\right)^n$ for all n. Then

$$|f_n(x) - f(x)| = \left(\frac{9}{10}\right)^n < \frac{1}{2}$$

is true for all $n > 6$ (Fig. 1.11).

Observe that there is no value of N for which $|f_n(x) - f(x)| < \frac{1}{2}$ is true for all $x \in [0, 1]$. N depends on both x and ϵ. But, this is not the case for the following example.

Example 1.23 Consider $f_n(x) = \dfrac{x}{1 + nx}$, $x \geq 0$. Clearly,

$$\lim_{n \to \infty} f_n(x) = f(x) = 0, \ \forall\, x \geq 0$$

Also, we have

$$0 \leq \frac{x}{1 + nx} \leq \frac{x}{nx} = \frac{1}{n}$$

Therefore, $|f_n(x) - f(x)| = |f_n(x)| \leq \frac{1}{n} < \epsilon$ for all $x \geq 0$, provided $N > \frac{1}{\epsilon}$. That is, if $N > \frac{1}{\epsilon}$, then $|f_n(x) - f(x)| < \epsilon$ for all $n > N$ and for all $x \geq 0$. Here N depends only on ϵ. Such type of convergence is called uniform convergence (Fig. 1.12).

Definition 1.21 (*Uniform convergence*) Let $\{f_n\}$ be a real-valued function defined on an interval $[a, b]$. Then $\{f_n\}$ is said to converge uniformly to the function f on $[a, b]$, if for each $\epsilon > 0$, there exists an integer N (dependent on ϵ and independent of x) such that for all $x \in [a, b]$, $|f_n(x) - f(x)| < \epsilon$ for all $n > N$ (Fig. 1.13).

Clearly, we can observe that uniform convergence implies point-wise convergence, but the converse does not hold true always. Also observe that, in Example 1.22, all the functions in $\{f_n\}$ were continuous. However, their point-wise limit was not continuous. In the case of uniform convergence, this is not possible. That is, if $\{f_n\}$ is a sequence of continuous functions and $f_n \to f$ uniformly then f is continuous.

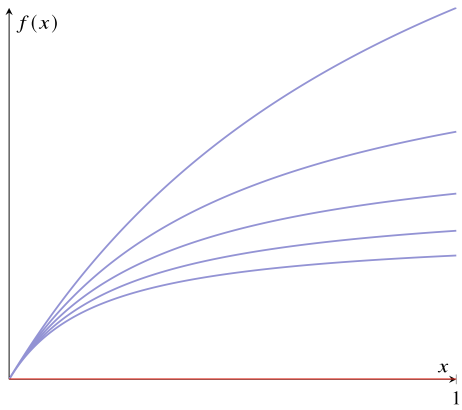


Fig. 1.12 Uniform convergence of $\{f_n\}$, where $f_n(x) = \dfrac{x}{1+nx}$, $x \geq 0$

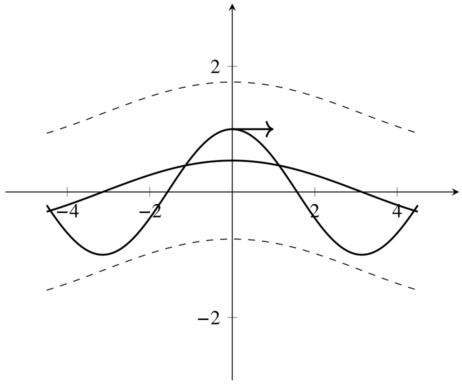


Fig. 1.13 If $\{f_n\}$ converges uniformly to a function f on $[a, b]$, for a given $\epsilon > 0$, there exists a positive integer N such that the graph of $f_n(x)$ for all $n > N$ and for all $x \in [a, b]$ lies between the graphs of $f(x) - \epsilon$ and $f(x) + \epsilon$

1.2 Metric Spaces

In $\mathbb{R}$, we have the notion of usual distance provided by the modulus function, to discuss the ideas like continuity of a function, convergence of a sequence, etc. These concepts can also be extended to a wide range of sets by generalizing the notion of "distance" to these sets by means of a function, called *metric*. A set with such a distance notion defined on it is called as a *metric space*. Consider the following definition.

Definition 1.22 (*Metric Space*) Let X be any non-empty set. A metric (or distance function) on X is a function $d : X \times X \to \mathbb{R}^+$ which satisfies the following properties for all $x, y, z \in X$:

$(M1)$ $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$. (*Non-negativity*)
$(M2)$ $d(x, y) = d(y, x)$. (*Symmetry*)
$(M3)$ $d(x, z) \leq d(x, y) + d(y, z)$. (*Triangle Inequality*)

If d is a metric on X, we say that (X, d) is a metric space.

Example 1.24 Consider the set of all real numbers, $\mathbb{R}$. For $x, y \in \mathbb{R}$, the function defined by

$$d(x, y) = |x - y|$$

is the usual distance between two points on the real line.

$(M1)$ Clearly $d(x, y) = |x - y| \geq 0$ and $d(x, y) = |x - y| = 0$ if and only if $x - y = 0$. That is, if and only if $x = y$.
$(M2)$ $d(x, y) = |x - y| = |y - x| = d(y, x)$
$(M3)$ Also, by the properties of modulus

$$\begin{aligned} d(x, z) &= |x - z| \\ &= |x - y + y - z| \\ &\leq |x - y| + |y - z| \\ &= d(x, y) + d(y, z) \end{aligned}$$

Thus all the conditions for a metric are satisfied and hence $(\mathbb{R}, |.|)$ is a metric space. This metric is known as the *usual metric* or *Euclidean Distance*.

Example 1.25 For any non-empty set X, define a function d by

$$d(x, y) = \begin{cases} 1 \,, & x \neq y \\ 0 \,, & x = y \end{cases}$$

Clearly conditions $(M1)$ and $(M2)$ are satisfied. Now we will check $(M3)$,

Case 1 $x \neq y = z$
Then $d(x, y) = 1, d(x, z) = 1$ and $d(y, z) = 0$
Case 2 $x = y \neq z$
Then $d(x, y) = 0, d(x, z) = 1$ and $d(y, z) = 1$
Case 3 $x = y = z$
Then $d(x, y) = 0, d(x, z) = 0$ and $d(y, z) = 0$
Case 4 $x \neq y \neq z$
Then $d(x, y) = 1, d(x, z) = 1$ and $d(y, z) = 1$.

In all four cases, condition $(M3)$ is clearly satisfied. Hence (X, d) is a metric space for any non-empty set X. The given metric d is known as a *discrete metric*.

Definition 1.23 (*Open Ball*) Let (X, d) be a metric space. For any point $x_0 \in X$ and $\epsilon \in \mathbb{R}^+$,

$$B_\epsilon(x_0) = \{x \in X \mid d(x, x_0) < \epsilon\}$$

is called an *open ball* centered at x_0 with radius ϵ.

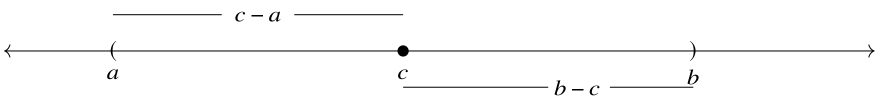


Fig. 1.14 Observe that if we take ϵ less than both $c - a$ and $b - c$, $B_\epsilon(c) \subset (a, b)$

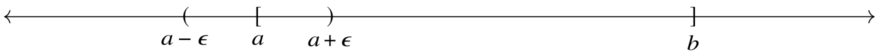


Fig. 1.15 Clearly $B_\epsilon(a) \not\subset [a, b]$ for any $\epsilon > 0$. Also, any open interval containing b is not a subset of $[a, b]$

Definition 1.24 (*Open Set and Closed Set*) Let (X, d) be a metric space. A subset $Y \subseteq X$ is said to be *open* if it contains an open ball about each of its elements. $Y \subseteq X$ is said to be *closed* if its complement Y^c is open.

Example 1.26 Consider the metric space $(\mathbb{R}, |.|)$. Then we can verify that every open interval in the real line is an open set (see Exercise 8 of this chapter). Consider an arbitrary open interval $(a, b) \subset \mathbb{R}$ and choose an arbitrary element $c \in (a, b)$. We have to show that there exists $\epsilon > 0$ such that $B_\epsilon(c) \subset (a, b)$ (Fig. 1.14).

From Fig. 1.14, if we take $\epsilon < min\{c - a, b - c\}$, it is clear that $B_\epsilon(c) \subset (a, b)$ for any $c \in (a, b)$. Similarly, we can prove that the union of open intervals is also an open set in $\mathbb{R}$. But a closed interval $[a, b] \subset \mathbb{R}$ is not an open set as $B_\epsilon(a) \nsubseteq [a, b]$ for any $\epsilon > 0$ (Fig. 1.15).

As $[a, b]^c = (-\infty, a) \cup (b, \infty)$ is an open set, $[a, b]$ is a closed set.

Example 1.27 Every singleton set in a discrete metric space X is an open set. It is obvious from the fact that for any $x \in X$, we have $B_\epsilon(x) = \{x\}$ when $\epsilon < 1$. Also it is interesting to observe that every subset of a discrete metric space is open as every open set can be written as a union of singleton sets. Therefore, every subset of a discrete metric space X is a closed set also.

As we have defined sequences on $\mathbb{R}$, we can define sequences on an arbitrary metric space (X, d) as a function from the set of all natural numbers taking values in X, and we can discuss their convergence based on the distance function d.

Definition 1.25 (*Convergent Sequence*) Sequence $\{x_n\}$ in a metric space (X, d) converges to $x \in X$ if for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $x_n \in B_\epsilon(x)$ for all $n > N$ and x is called the limit of the sequence $\{x_n\}$. We denote this by $x_n \to x$ or $\lim_{n\to\infty} x_n = x$. In other words, we can say that $d(x_n, x) \to 0$ as $n \to \infty$.

Example 1.28 Consider the sequence $\{x_n\}$, where $x_n = r + \frac{1}{n}$, $n \in \mathbb{N}$ in the metric space $(\mathbb{R}, |.|)$ for some $r \in \mathbb{R}$. We will show that $x_n \to r$ in $(\mathbb{R}, |.|)$. For any $\epsilon > 0$, if we take $N > \frac{1}{\epsilon}$

$$d(x_n, r) = \left| r + \frac{1}{n} - r \right| = \left| \frac{1}{n} \right| < \epsilon \ \forall \ n > N$$

That is, $x_n \in B_\epsilon(r)$ for all $n > N$. Therefore $x_n \to r$ in $(\mathbb{R}, |.|)$.

Example 1.29 Let $\{x_n\}$ be a sequence in a metric space (X, d), where d is the discrete metric. We have seen in Example 1.27 that every singleton set in a discrete metric space is open. Therefore for the sequence $\{x_n\}$ to converge to a point $x \in X$, the open set $\{x\}$ must contain almost all terms of the sequence. In other words, a sequence $\{x_n\}$ in a discrete metric space converges if and only if it is of the form $x_1, x_2, \ldots, x_N, x, x, \ldots$. That is, if and only if $\{x_n\}$ is *eventually constant*.

Definition 1.26 (*Cauchy Sequence*) Sequence of points $\{x_n\}$ in a metric space (X, d) is said to be a Cauchy sequence if for every $\epsilon > 0$, there exists an $N_\epsilon \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ for every $m, n > N_\epsilon$.

Theorem 1.4 *In a metric space, every convergent sequence is Cauchy.*

The converse of the above theorem need not be true. That is, there exists metric spaces where every Cauchy sequence may not be convergent.

Example 1.30 Consider the sequence $\{x_n\}$ with $x_n = a + \frac{1}{n}$ in the metric space $((a, b), |.|)$ where (a, b) is any open interval in $\mathbb{R}$. We will show that this sequence is Cauchy but not convergent. For an $\epsilon > 0$, if we choose $N > \frac{2}{\epsilon}$

$$d(x_n, x_m) = \left|\frac{1}{n} - \frac{1}{m}\right| \leq \left|\frac{1}{n}\right| + \left|\frac{1}{m}\right| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \ \forall\, m, n > N$$

That is, the given sequence is a Cauchy sequence. As we have seen in Example 1.28, the given sequence converges to a as $n \to \infty$. As $a \notin (a, b)$, $\{x_n\}$ with $x_n = a + \frac{1}{n}$ is not convergent in $((a, b), |.|)$.

Definition 1.27 (*Complete Metric Space*) A metric space in which every Cauchy sequence is convergent is called a complete metric space.

Example 1.31 By Theorem 1.2, $(\mathbb{R}, |.|)$ is a complete metric space and from Example 1.30, $((a, b), |.|)$ is an incomplete metric space.

Definition 1.28 (*Continuous Function*) Let (X, d_1) and (Y, d_2) be two metric spaces. A function $f : X \to Y$ is said to be continuous at a point $x_0 \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $d_2(f(x), f(x_0)) < \epsilon$ whenever $d_1(x, x_0) < \delta$. f is said to be continuous on X if f is continuous at every point of X.

Theorem 1.5 *Let (X, d_1) and (Y, d_2) be two metric spaces. Then a function $f : X \to Y$ is said to be continuous if and only if the inverse image of any open set of (Y, d_2) is open in (X, d_1).*

The continuity of a function in metric spaces can also be discussed in terms of sequences. Consider the following definition.

Definition 1.29 (*Sequential Continuity*) Let (X, d_1) and (Y, d_2) be two metric spaces. A function $f : X \to Y$ is said to be sequentially continuous at a point $x_0 \in X$ if $\{x_n\}$ is any sequence in (X, d_1) with $x_n \to x_0$, then $f(x_n) \to f(x_0)$ in (Y, d_2).

Theorem 1.6 *Let (X, d_1) and (Y, d_2) be two metric spaces. Then a function $f : X \to Y$ is continuous on X, if and only if it is sequentially continuous.*

1.3 Some Important Algebraic Structures

An algebraic structure consists of a non-empty set together with a collection of operations defined on it satisfying certain conditions or axioms which are defined as per the context under discussion. The operations are of great importance when the resultant obtained by combining two elements in the set belongs to the same set.

Definition 1.30 (*Binary Operation*) Let G be any set. A binary operation $'*'$ on G is a function $* : G \times G \to G$ defined by

$$* (g_1, g_2) = g_1 * g_2$$

Example 1.32 Let $G = \mathbb{R}$, the set of all real numbers, and let $+$ be the usual addition of real numbers. Now $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that $+(a, b) = a + b \in \mathbb{R}$ defines a binary operation. Similarly, the usual multiplication and subtraction of real numbers are also binary operations on $\mathbb{R}$. But as the division of a real number with 0 is not defined, division is not a binary operation.

Definition 1.31 (*Group*) A non-empty set G together with a binary operation $'*'$ is said to be a group, denoted by $(G, *)$, if $'*'$ satisfies the following properties:

(a) $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3 \; \forall \; g_1, g_2, g_3 \in G$ (Associative property)
(b) There exists $e \in G$, such that $e * g = g = g * e \; \forall \; g \in G$ (Existence of Identity)
(c) For each $g \in G$, there exists $g^{-1} \in G$ such that $g * g^{-1} = e = g^{-1} * g$. (Existence of Inverse)

If $'*'$ satisfies $g_1 * g_2 = g_2 * g_1 \; \forall \; g_1, g_2 \in G$ (Commutative property) also, then $(G, *)$ is called an Abelian group.

Example 1.33 Consider $\mathbb{R}$ together with the binary operation $'+'$. Then $\mathbb{R}$ is an Abelian group under the operation $'+'$. For,

(a) Addition is associative over $\mathbb{R}$.
(b) For all $r \in \mathbb{R}$, there exists $0 \in \mathbb{R}$ such that $r + 0 = r = 0 + r$.
(c) For all $r \in \mathbb{R}$, there exists $-r \in \mathbb{R}$ such that $r + (-r) = 0 = (-r) + r$.
(d) Addition is commutative over $\mathbb{R}$.

Similarly, $\mathbb{C}$, the set of all complex numbers, $\mathbb{Q}$, the set of all rational numbers, and $\mathbb{Z}$, the set of all integers together with the binary operation $'+'$ is an Abelian group. But $(\mathbb{R}, .)$ is not a group, where '.' denotes usual multiplication as there does not exist any inverse element for 0.

Example 1.34 Consider $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ under usual multiplication. We can show that $(\mathbb{R}^*, .)$ is an Abelian group. Similarly, we can show that $(\mathbb{Q}^*, .)$ and $(\mathbb{C}^*, .)$ are also Abelian groups where $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$ and $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$. Observe that $\mathbb{Z}^*$ with usual multiplication is not a group as the inverse of every element does not exist in $\mathbb{Z}^*$.

Example 1.35 Consider $\mathbb{R}^+$, the set of all positive real numbers under usual multiplication. We can show that $\left(\mathbb{R}^+, .\right)$ is an Abelian group. Similarly, we can show that $\left(\mathbb{Q}^+, .\right)$ and $\left(\mathbb{C}^+, .\right)$ are also Abelian groups.

Example 1.36 The set $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$, for $n \geq 1$, is a group under the operation *addition modulo n*, denoted by $+_n$. The basic operation is usual addition of elements, which ends by reducing the sum of the elements modulo n, that is, taking the integer remainder when the sum of the elements is divided by n. This group is usually referred to as the *group of integers modulo n*. Consider the following examples:

$+_2$	0	1
0	0	1
1	1	0

$+_3$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$+_4$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

$+_5$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

The above group multiplication table is called Cayley table. A Cayley table, named after the British mathematician *Arthur Cayley (1821–1895)* of the nineteenth century, illustrates the structure of a finite group by arranging all the possible products of all the group's members in a square table resembling an addition or multiplication table.

Example 1.37 A one-one function from a set S onto itself is called a permutation. Consider the set $S = \{1, 2, \ldots, n\}$. Let S_n denote the set of all permutations on S to itself. Then S_n is a non-Abelian group under the operation function composition, called *symmetric group on n letters*. Permutations of finite sets are represented by an explicit listing of each element of the domain and its corresponding image. For example, the elements of S_3 can be listed as $\left\{\rho_0 = \begin{pmatrix} 1\ 2\ 3 \\ 1\ 2\ 3 \end{pmatrix}, \rho_1 = \begin{pmatrix} 1\ 2\ 3 \\ 2\ 3\ 1 \end{pmatrix},\right.$ $\left.\rho_2 = \begin{pmatrix} 1\ 2\ 3 \\ 3\ 1\ 2 \end{pmatrix}, \mu_1 = \begin{pmatrix} 1\ 2\ 3 \\ 1\ 3\ 2 \end{pmatrix}, \mu_2 = \begin{pmatrix} 1\ 2\ 3 \\ 3\ 2\ 1 \end{pmatrix}, \mu_3 = \begin{pmatrix} 1\ 2\ 3 \\ 2\ 1\ 3 \end{pmatrix}\right\}$

Theorem 1.7 *Let* $(G, *)$ *be a group. Then*

(a) the identity element is unique.
(b) each element in G has a unique inverse.

Definition 1.32 (*Subgroup*) A subset H of a group $(G, *)$ is said to be a subgroup of G, if H is a group with respect to the operation $*$ in G. Let $H \leq G$ denote that H is a subgroup of G and $H < G$ denote that H is a subgroup of G, but $H \neq G$.

Example 1.38 We have $(\mathbb{Z}, +) < (\mathbb{Q}, +) < (\mathbb{R}, +)$. But $(\mathbb{Z}_n, +_n)$ is a not a subgroup of $(\mathbb{R}, +)$ even though as sets $\mathbb{Z}_n \subset \mathbb{R}$, as the operations used are different.

Example 1.39 Consider the permutation group S_3. Then $\{\rho_0\}, \{\rho_0, \mu_1\}, \{\rho_0, \mu_2\}, \{\rho_0, \mu_3\}$ and $\{\rho_0, \rho_1, \rho_2\}$ are subgroups of S_3.

Definition 1.33 (*Order of a Group*) Let $(G, *)$ be a group, then the order of G is the number of elements in G.

Example 1.40 Observe that $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, and $(\mathbb{C}, +)$ are groups of order infinity and $(\mathbb{Z}_n, +_n)$ is a group of order n. Also observe that S_n has order $n!$.

Definition 1.34 (*Order of an element*) Let $(G, *)$ be a group, then the order of an element $g \in G$, denoted by $\mathcal{O}(g)$, is the least positive integer n such that $g^n = e$, where e is the identity in G. Clearly, identity element in a group G has order 1.

Example 1.41 Consider the group $(\mathbb{R}, +)$. Then we get that no element other than 0 in $\mathbb{R}$ has finite order. This is because of the fact that repeated addition of a real number will never give us 0.

Example 1.42 Consider a finite group, say $(\mathbb{Z}_4, +_4)$. Then $\mathcal{O}(0) = 1, \mathcal{O}(1) = 4, \mathcal{O}(2) = 2$, and $\mathcal{O}(3) = 4$. It is easy to observe that, in a finite group G, every element has finite order. Consider another example, S_3. Then $\mathcal{O}(\rho_0) = 1, \mathcal{O}(\rho_1) = \mathcal{O}(\rho_2) = 3$, and $\mathcal{O}(\mu_1) = \mathcal{O}(\mu_2) = \mathcal{O}(\mu_3) = 2$.

Remark 1.4 A set G together with a binary operation $'*'$ defined on it is called a Groupoid or Magma. If $'*'$ satisfies associative property also, then $(G, *)$ is called a Semi-group. A semi-group containing an identity element is called a Monoid.

Definition 1.35 (*Group Homomorphism*) Let $(G, *)$ and $\left(G', *'\right)$ be any two groups. A map ϕ from G to G' satisfying $\phi(g_1 * g_2) = \phi(g_1) *' \phi(g_2),\ \forall\ g_1, g_2 \in G$ is called a group homomorphism. If ϕ is one-one and onto, we say that ϕ is an isomorphism or $(G, *)$ and $\left(G', *'\right)$ are isomorphic, denoted by $G \cong G'$.

Definition 1.36 (*Kernel of a Homomorphism*) The kernel of a homomorphism of a group G to a group G' with identity e' is the set of all elements in G which are mapped to e'. That is, $Ker(\phi) = \{g \in G \mid \phi(g) = e'\}$.

Example 1.43 Consider the groups $(\mathbb{R}, +)$ and $(\mathbb{R}^*, .)$. We will show that they are isomorphic. Define $\phi : \mathbb{R} \to \mathbb{R}^*$ by $\phi(x) = e^x$. Then for $x_1, x_2 \in \mathbb{R}$,

$$\phi(x_1 + x_2) = e^{x_1 + x_2} = e^{x_1}.e^{x_2} = \phi(x_1).\phi(x_2)$$

Therefore ϕ is a homomorphism from $\mathbb{R}$ to $\mathbb{R}^*$. Also we can easily verify that ϕ is both one-one and onto. Thus $(\mathbb{R}, +) \cong (\mathbb{R}^*, .)$. Now let us find the Kernel of ϕ. By definition, $Ker(\phi)$ is the set of all elements of the domain which are mapped to the identity element in the co-domain, in this case, 1. Therefore $Ker(\phi) = \{x \in \mathbb{R} \mid \phi(x) = e^x = 1\} = \{0\}$.

Example 1.44 Consider $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +_n)$. Define $\phi : \mathbb{Z} \to \mathbb{Z}_n$ by $\phi(m) = r$, where r is the remainder when m is divided by n. Let us check whether ϕ is a homomorphism or not. Take two elements $m_1, m_2 \in \mathbb{Z}$. By division algorithm, we can write $m_i = q_i n + r_i$ with $0 \le r_i < n$, where $i = 1, 2$ and hence $\phi(m_1) = r_1$ and $\phi(m_2) = r_2$. Observe that $m_1 + m_2 = (q_1 + q_2)n + r_1 + r_2$. Therefore, we can say that $\phi(m_1 + m_2)$ is the remainder when $r_1 + r_2$ is divided by n. That is, $\phi(m_1 + m_2) = r_1 +_n r_2$. Also $\phi(m_1) +_n \phi(m_2) = r_1 +_n r_2$. Thus ϕ is a homomorphism. Now the set of all elements mapped to $0 \in \mathbb{Z}_n$ are integer multiples of n. That is, $Ker(\phi) =< n >$.

Example 1.45 Consider the map $\phi : (\mathbb{R}, +) \to (\mathbb{R}^*, .)$ defined by $\phi(x) = x^2$. Then for $x_1, x_2 \in \mathbb{R}$, we have

$$\phi(x_1 + x_2) = (x_1 + x_2)^2 \neq x_1^2.x_2^2 = \phi(x_1).\phi(x_2)$$

Thus ϕ is not a homomorphism.

Example 1.46 Consider $(\mathbb{R}^*, .)$. Define a map $\phi : \mathbb{R}^* \to \mathbb{R}^*$ by $\phi(x) = |x|$. Then for $x_1, x_2 \in \mathbb{R}^*$, we have

$$\phi(x_1 x_2) = |x_1 x_2| = |x_1||x_2| = \phi(x_1)\phi(x_2)$$

Thus ϕ is a homomorphism from $\mathbb{R}^*$ to itself. Observe that $Ker\ (\phi) = \{x \in \mathbb{R}^* \mid |x| = 1\} = \{-1, 1\}$. Thus ϕ is not one-one (Why?). Also ϕ is not onto as only positive real numbers have pre-images. Therefore ϕ is not an isomorphism.

Theorem 1.8 *Let ϕ be a homomorphism from a group $(G, *)$ to $\left(G', *'\right)$. Then*

(a) if e is the identity element in G, $\phi(e)$ is the identity element in G'.
(b) $Ker\ (\phi)$ is a subgroup of G.
(c) for any $g \in G$, if $\mathcal{O}(g)$ is finite $\mathcal{O}(\phi(g))$ divides $\mathcal{O}(g)$.
(d) for any subgroup H of G, $\phi(H)$ is a subgroup of $\phi(G)$ and if H is Abelian, $\phi(H)$ is also Abelian.

Two algebraic structures $(G, *)$ and $\left(G', *'\right)$ are isomorphic, if there exists a one-one, onto homomorphism from G to G'. But it will be difficult to show that $(G, *)$ and $\left(G', *'\right)$ are not isomorphic, following the definition as it means that there is no one-one homomorphism from G onto G'. It is not possible to check whether such a function exists or not. In such cases, we could use the idea of structural properties of an algebraic structure, which are properties that must be shared by any isomorphic structure. Cardinality is an example for structural property.

Example 1.47 In Remark 1.3, we have seen that $\mathbb{R}$ is an uncountable set and $\mathbb{Z}$ is a countable set. Hence $(\mathbb{R}, +)$ and $(\mathbb{Z}, +)$ are not isomorphic.

Theorem 1.9 (Cyclic subgroup) *Let $(G, *)$ be a group. Then the set $\{g^n \mid g \in G, n \in \mathbb{Z}\}$ is a subgroup of G called cyclic subgroup of G generated by g, denoted by $< g >$.*

If the group $G =< g >$ for some $g \in G$, then G is called a *cyclic group* and g is called a generator of G.

Example 1.48 $(\mathbb{Z}, +)$ is a cyclic group with two generators $\{1, -1\}$.

Example 1.49 $(\mathbb{Z}_n, +_n)$ is a cyclic group. The generators are the elements $m \in \mathbb{Z}_n$ with $gcd(m, n) = 1$, where $gcd(m, n)$ denotes the greatest common divisor for m and n (verify).

Theorem 1.10 *Let $(G, *)$ be a cyclic group with generator g. If $\mathcal{O}(G)$ is finite, then $(G, *) \cong (\mathbb{Z}_n, +_n)$ and if $\mathcal{O}(G)$ is infinite, then $(G, *) \cong (\mathbb{Z}, +)$.*

Example 1.50 By Example 1.47, $(\mathbb{R}, +)$ is not a cyclic group.

Definition 1.37 (*Coset*) Let $(G, *)$ be a group and H be a non-trivial subgroup of G. Then $gH = \{g * h \mid h \in H\}$ is called left coset of H in G containing g and $Hg = \{h * g \mid h \in H\}$ is called right coset of H in G containing g.

Example 1.51 Consider $(\mathbb{Z}_8, +_8)$ and the subgroup $H = \{0, 2, 4, 6\}$ of $\mathbb{Z}_8$. Then

$$0H = \{0, 2, 4, 6\} = 2H = 4H = 6H$$

$$1H = \{1, 3, 5, 7\} = 3H = 5H = 7H$$

Also observe that as $(\mathbb{Z}_8, +_8)$ is an Abelian group, the left and right cosets of each element coincide.

Example 1.52 Consider the subgroup $H = \{\rho_0, \mu_1\}$ in S_3. Then

$$\rho_0 H = \{\rho_0, \mu_1\} = \mu_1 H$$

$$\rho_1 H = \{\rho_1, \mu_3\} = \mu_3 H$$

$$\rho_2 H = \{\rho_2, \mu_2\} = \mu_2 H$$

are the distinct left cosets of H in G and

$$H\rho_0 = \{\rho_0, \mu_1\} = H\mu_1$$

$$H\rho_1 = \{\rho_1, \mu_2\} = H\mu_2$$

$$H\rho_2 = \{\rho_2, \mu_3\} = H\mu_3$$

are the distinct right cosets of H in G

Theorem 1.11 (Lagrange's Theorem) *Let G be a finite group and H be a subgroup of G, then $\mathcal{O}(H)$ divides $\mathcal{O}(G)$. Moreover, the number of distinct left/right cosets of H in G is* $\dfrac{\mathcal{O}(G)}{\mathcal{O}(H)}$.

Example 1.53 In Example 1.51, $H = \{0, 2, 4, 6\}$ and $G = \mathbb{Z}_8$. We have $\mathcal{O}(H) = 4$ and $\mathcal{O}(G) = 8$. Clearly, $\mathcal{O}(H)$ divides $\mathcal{O}(G)$ and the number of distinct left/right cosets of H in G is $\dfrac{\mathcal{O}(G)}{\mathcal{O}(H)} = 2$

Example 1.54 In Example 1.52, $H = \{\rho_0, \mu_1\}$ and $G = S_3$. We have $\mathcal{O}(H) = 2$ and $\mathcal{O}(G) = 6$. Clearly, $\mathcal{O}(H)$ divides $\mathcal{O}(G)$ and the number of distinct left/right cosets of H in G is $\dfrac{\mathcal{O}(G)}{\mathcal{O}(H)} = 3$.

Definition 1.38 (*Normal Subgroup*) A subgroup H of G is called a normal subgroup of G if $gH = Hg$ for all $g \in G$.

Example 1.55 From Example 1.51, $H = \{0, 2, 4, 6\}$ is a normal subgroup of $(\mathbb{Z}_8, +_8)$. In fact, every subgroup of an Abelian group is a normal subgroup (verify).

Example 1.56 From Example 1.52, $H = \{\rho_0, \mu_1\}$ is not a normal subgroup of S_3.

Theorem 1.12 (Factor Group) *Let* $(G, *)$ *be a group and* H *be a normal subgroup. Then the set* $G/H = \{gH \mid g \in G\}$ *is a group under the operation* $*'$*, where* $*'$ *is defined by* $(g_1H) *' (g_2H) = (g_1 * g_2)H$.

Example 1.57 In Example 1.55 we have seen that $H = \{0, 2, 4, 6\}$ is a normal subgroup of $(\mathbb{Z}_8, +_8)$. From Example 1.51, $G/H = \{0H, 1H\}$. Then G/H is a group, with the operation $*'$ defined as $(0H) *' (0H) = (0H)$, $(0H) *' (1H) = (1H) *' (0H) = (1H)$, and $(1H) *' (1H) = (0H)$.

Example 1.58 Consider the group $(\mathbb{Z}, +)$. Clearly $3\mathbb{Z} = \{\ldots, -6, -3, 0, 3, 6\}$ is a normal subgroup of $\mathbb{Z}$. Then $G/H = \{0\,(3\mathbb{Z})\,, 1\,(3\mathbb{Z})\,, 2\,(3\mathbb{Z})\}$ is a group, with the operation $*$ defined as $0\,(3\mathbb{Z}) *' 0\,(3\mathbb{Z}) = 0\,(3\mathbb{Z})\,, 0\,(3\mathbb{Z}) *' 1\,(3\mathbb{Z}) = 1\,(3\mathbb{Z}) *' 0\,(3\mathbb{Z}) = 1\,(3\mathbb{Z})\,, 0\,(3\mathbb{Z}) *' 2\,(3\mathbb{Z}) = 2\,(3\mathbb{Z}) *' 0\,(3\mathbb{Z}) = 2\,(3\mathbb{Z})\,, 2\,(3\mathbb{Z}) *' 1\,(3\mathbb{Z}) = 0\,(3\mathbb{Z})\,, 1\,(3\mathbb{Z}) *' 2\,(3\mathbb{Z}) = 0\,(3\mathbb{Z})\,, 1\,(3\mathbb{Z}) *' 1\,(3\mathbb{Z}) = 0\,(3\mathbb{Z})$ and $2\,(3\mathbb{Z}) *' 2\,(3\mathbb{Z}) = 1\,(3\mathbb{Z})$.

Theorem 1.13 (First Isomorphism Theorem) *Let* ϕ *be a homomorphism from a group* G *to a group* G'*. Then the mapping* $\Psi : G/Ker\,(\phi) \to G'$ *given by* $\Psi\,(gKer(\phi)) = \phi\,(g)$ *is an isomorphism. That is,* $G/Ker\,(\phi) \cong \phi\,(G)$.

Example 1.59 In Example 1.44, we have seen that $\phi\,(m) = m \; mod \; n$ is a homomorphism from $(\mathbb{Z}, +)$ and $(\mathbb{Z}_n, +_n)$ with $Ker\,(\phi) =< n >$. Therefore by Theorem 1.13, $\mathbb{Z}/ < n > \cong \mathbb{Z}_n$.

Definition 1.39 (*Ring*) A non-empty set R together with two operations $'+'$ and $'.'$, known as addition and multiplication, respectively, is called a ring (denoted by $\langle \mathcal{R}, +, . \rangle$) if the following conditions are satisfied:

(a) $(\mathcal{R}, +)$ is an Abelian group.
(b) $(\mathcal{R}, .)$ is a semi-group.
(c) For all $r_1, r_2, r_3 \in \mathcal{R}$

$$r_1.(r_2 + r_3) = r_1.r_2 + r_1.r_3 \text{ (left distributive law)}$$

$$(r_1 + r_2).r_3 = r_1.r_3 + r_2.r_3 \text{ (right distributive law)}$$

If there exists a non-zero element $1 \in \mathcal{R}$ such that for every element $r \in \mathcal{R}, r.1 = r = 1.r$, then $\langle \mathcal{R}, +, . \rangle$ is called a ring with unity and if multiplication is also commutative, then the ring is called a commutative ring.

Example 1.60 The set of all real numbers under usual addition and multiplication is a commutative ring with unity. From Example 1.33, we have $(\mathbb{R}, +)$ is an Abelian group. Clearly, the usual multiplication $'.'$ is closed, associative, and commutative over $\mathbb{R}$. Also $1 \in \mathbb{R}$ acts as unity and the distributive laws are satisfied. Similarly $\langle \mathbb{C}, +, . \rangle$,$\langle \mathbb{Q}, +, . \rangle$, and $\langle \mathbb{Z}, +, . \rangle$ are commutative rings with unity.

Example 1.61 The set $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$, for $n \geq 1$, under the operations addition and multiplication modulo n (taking the integer remainder when the product is divided by n) is a ring with unity 1.

Definition 1.40 (*Sub-Ring*) A sub-ring of a ring $\mathcal{R}$ is a subset of the $\mathcal{R}$ that is a ring under the induced operations from $\mathcal{R}$.

Example 1.62 Clearly $\langle \mathbb{Q}, +, . \rangle$ is a sub-ring of $\langle \mathbb{Q}, +, . \rangle$. Also $\langle \mathbb{Q}, +, . \rangle$ is a sub-ring of $\langle \mathbb{R}, +, . \rangle$ which is again a sub-ring of $\langle \mathbb{C}, +, . \rangle$

Example 1.63 $\mathbb{Z}_n$, for $n \geq 1$, is a ring under the operation *addition modulo n* and *multiplication modulo n* (denoted by $\times_n$). The basic operation in $\times_n$ is multiplication, which ends by reducing the result modulo n; that is, taking the integer remainder when the result is divided by n as in $+_n$.

Definition 1.41 (*Division Ring*) Let $\langle \mathcal{R}, +, . \rangle$ be a ring with unity $'1'$. An element $r \in \mathcal{R}$ is a unit of $\mathcal{R}$ if it has multiplicative inverse in $\mathcal{R}$. That is, if there exists an element $r^{-1} \in \mathcal{R}$ such that $r.r^{-1} = 1 = r^{-1}.r$. If every non-zero element in $\mathcal{R}$ is a unit, then $\mathcal{R}$ is called a division ring or skew-field.

Example 1.64 $\langle \mathbb{R}, +, . \rangle$ is a division ring as for any $r(\neq 0) \in \mathbb{R}$, there exists $\frac{1}{r} \in \mathbb{R}$ such that $r \cdot \frac{1}{r} = 1 = \frac{1}{r} \cdot r$.

Theorem 1.14 *An element $m \in \mathbb{Z}_n$ is a unit if and only if $gcd(m, n) = 1$.*

Corollary 1.1 *$\mathbb{Z}_n$ is a division ring only if n is a prime.*

Definition 1.42 (*Field*) A field is a commutative division ring. In other words, $\langle \mathcal{R}, +, . \rangle$ is a field if the following conditions are satisfied:

(a) $(\mathcal{R}, +)$ is an Abelian group.
(b) $(\mathcal{R} \setminus \{0\}, .)$ is an Abelian group.

Example 1.65 The set of all real numbers $\mathbb{R}$ under usual addition and multiplication is a field. Similarly, the set of all complex numbers $\mathbb{C}$ and the set of all rational numbers $\mathbb{Q}$ under usual addition and multiplication are fields.

Example 1.66 From Corollary 1.1, the set $\mathbb{Z}_n$ is a field under the operations addition and multiplication modulo n, if and only if n is a prime (Why?). Clearly, $\langle \mathbb{Z}_n, +_n, \times_n \rangle$ is an example for a finite field.

Example 1.67 The set of all integers $\mathbb{Z}$ under usual addition and multiplication is not a field as it is not a division ring. But $\mathbb{Z}$ is a commutative ring with unity.

Definition 1.43 (*Sub-Field*) A sub-field of a field is a subset of the field that is a field under the induced operations from the field.

Example 1.68 Clearly $\langle \mathbb{Q}, +, . \rangle$ is a sub-field of $\langle \mathbb{R}, +, . \rangle$ which is again a sub-field of $\langle \mathbb{C}, +, . \rangle$.

1.4 Polynomials

Polynomials are a type of mathematical expression built by combining variables by the operations addition, subtraction, and multiplication. They are an important tool in mathematics as many mathematical problems can be encoded into polynomial equations. In this section, we will discuss some of the important properties of polynomials in one variable.

Definition 1.44 (*Ring of polynomials*) Let $\mathbb{K}$ be a field. Consider the set

$$\mathbb{K}[x] = \{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n \mid a_i \in \mathbb{K}, n \in \mathbb{Z}^+\}$$

$a_i \in \mathbb{K}$ are called coefficients of the polynomial, and the order of the highest power of x with non-zero coefficient is called the degree of the polynomial. For $f(x) = a_0 + a_1 x + \cdots + a_n x^n$, $g(x) = b_0 + b_1 x + \cdots + b_m x^m \in \mathbb{K}[x]$, define

$$f(x) + g(x) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_{k-1} + b_{k-1})x^{k-1} + (a_k + b_k)x^k$$

where $k = max(m, n)$, $a_i = 0$ for $i > n$ and $b_i = 0$ for $i > m$. Also

$$f(x)g(x) = c_0 + c_1 x + \cdots + c_{m+n-1}x^{m+n-1} + c_{m+n}x^{m+n}$$

where $c_k = a_k b_0 + a_{k-1}b_1 + \cdots + a_1 b_{k-1} + a_0 b_k$ for $k = 0, 1, \ldots, m+n$. Then $\mathbb{K}[x]$ forms a ring with respect to the operations defined above, called the ring of polynomials over $\mathbb{K}$ in the indeterminate x.

Remark 1.5 If the coefficient of the highest power of x is the multiplicative identity of $\mathbb{K}$, then the polynomial is called a monic polynomial. Two elements in $\mathbb{K}[x]$ are equal if and only they have the same coefficients for all powers of x.

Theorem 1.15 (Division Algorithm) *Let $\mathbb{K}$ be a field and let $f(x), g(x) \in \mathbb{K}[x]$ with $g(x) \neq 0$. Then there exists unique polynomials $q(x), r(x) \in \mathbb{K}[x]$ such that $f(x) = g(x)q(x) + r(x)$ and either $r(x) = 0$ or $deg[r(x)] < deg[g(x)]$. If $r(x) = 0$ we have $f(x) = g(x)q(x)$ and we say that $g(x)$ is a factor $f(x)$.*

Theorem 1.16 *Let $\mathbb{K}$ be a field and let $f(x), g(x) \in \mathbb{K}[x]$. The greatest common divisor of $f(x)$ and $g(x)$, denoted by $(f(x), g(x))$, is the unique monic polynomial $r(x) \in \mathbb{K}[x]$ such that*

1. *$r(x)$ is a factor of both $f(x)$ and $g(x)$.*
2. *if $q(x) \in \mathbb{K}[x]$ is a factor of both $f(x)$ and $g(x)$, then $r(x)$ is a factor of $q(x)$.*

Moreover, there exists polynomials $l(x), m(x) \in \mathbb{K}[x]$ *such that*

$$r(x) = l(x)f(x) + m(x)g(x)$$

Remark 1.6 If $(f(x), g(x)) = 1$, then we say that $f(x), g(x) \in \mathbb{K}[x]$ are relatively prime.

Definition 1.45 (*Zero of a polynomial*) Let $f(x) \in \mathbb{K}[x]$; an element $\mu \in \mathbb{K}$ is called a zero (or a root) of $f(x)$ if $f(\mu) = 0$.

Theorem 1.17 (Factor Theorem) *Let* $\mathbb{K}$ *be a field and* $f(x) \in \mathbb{K}[x]$. *Then* $\mu \in \mathbb{K}$ *is a zero of* $f(x)$ *if and only if* $x - \mu$ *is a factor of* $f(x)$.

Definition 1.46 (*Algebraically Closed Field*) A field $\mathbb{K}$ is said to be an algebraically closed field, if every non-constant polynomial in $\mathbb{K}[x]$ has a root in $\mathbb{K}$.

Theorem 1.18 (Fundamental Theorem of Algebra) *The field of complex numbers is algebraically closed. In other words, every non-constant polynomial in* $\mathbb{C}[x]$ *has at least one root in* $\mathbb{C}$.

From the above theorem, we can infer that every polynomial of degree n in $\mathbb{C}[x]$ has exactly n roots in $\mathbb{C}$.

Example 1.69 Consider $x^2 + 1 \in \mathbb{R}[x]$. As the given polynomial has no root in $\mathbb{R}$, the field of real numbers is not algebraically closed, whereas if we consider $x^2 + 1$ as a polynomial in $\mathbb{C}[x]$, it has roots in $\mathbb{C}$.

Remark 1.7 (*Vieta's Formula*) Let $f(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{K}[x]$ with roots $x_1, x_2, \ldots, x_n$, then

$$x_1 + x_2 + \cdots + x_n = -\frac{a_{n-1}}{a_n}$$

$$x_1x_2 \cdots x_n = (-1)^n \frac{a_0}{a_n}$$

It is named after the French mathematician *Francois Viete (1540–1603).*

1.5 Matrices

A matrix in mathematics is a rectangular arrangement of numbers, symbols, or functions in rows and columns. They are of great importance in mathematics and are widely used in linear algebra to study linear transformations which will be discussed in later chapters.

Definition 1.47 An $m \times n$ matrix A over a field $\mathbb{K}$ is a rectangular array of m rows and n columns of entries from $\mathbb{K}$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Such a matrix, written as $A = \left(a_{ij}\right)$, where $1 \le i \le m$, $1 \le j \le n$ is said to be of *size* (or order) $m \times n$. Two matrices are considered to be equal if they have the same size and same corresponding entries in all positions. $\mathbb{M}_{m\times n}(\mathbb{K})$ denotes the set of all $m \times n$ matrices with entries from $\mathbb{K}$.

Matrix Operations

Let us discuss some of the important operations that are used in the collection of all matrices.

Definition 1.48 (*Matrix Addition*) Let $A = \left(a_{ij}\right)$ and $B = \left(b_{ij}\right)$, where $1 \le i \le m$, $1 \le j \le n$ be any two elements of $\mathbb{M}_{m\times n}(\mathbb{K})$. Then $A + B = \left(a_{ij} + b_{ij}\right) \in \mathbb{M}_{m\times n}(\mathbb{K})$. Two matrices must have an equal number of rows and columns to be added.

Properties
For any matrices A, B and $C \in \mathbb{M}_{m\times n}(\mathbb{K})$

1. $A + B = B + A$. (Commutativity)
2. $A + (B + C) = (A + B) + C$. (Associativity)
3. There exists a matrix $O \in \mathbb{M}_{m\times n}(\mathbb{K})$ with all entries 0 such that $A + O = A$. (Existence of Identity)
4. There exists a matrix $-A$ such that $A + (-A) = O$. (Existence of Inverse)

Remark 1.8 $\mathbb{M}_{m\times n}(\mathbb{K})$ with matrix addition defined on it forms an Abelian group.

Definition 1.49 (*Matrix Multiplication*) Let $A = \left(a_{ij}\right)_{m\times n}$ and $B = \left(b_{ij}\right)_{n\times p}$. Then their product $AB \in \mathbb{M}_{m\times p}$ and its (i, j)th entry is given by

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

For AB to make sense, the number of columns of A must equal the number of rows of B. Then we say that the size of matrices A and B are compatible for multiplication.

Properties

For any matrices A, B and $C \in \mathbb{M}_{n\times n}(\mathbb{K})$

1. $A(BC) = (AB)C$ (Associativity)
2. $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$. (Distributive laws)

Remark 1.9 1. Matrix multiplication need not be commutative. For example, if $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 4 & 5 \\ 6 & 0 & 8 \end{pmatrix}$ then $AB = \begin{pmatrix} -3 & 4 & -3 \\ 12 & 0 & 16 \end{pmatrix}$. Note that BA is undefined. It need not be commutative even if BA is defined. For example, if $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 4 \\ 6 & 0 \end{pmatrix}$ then $AB = \begin{pmatrix} -3 & 4 \\ 12 & 0 \end{pmatrix}$ and $AB = \begin{pmatrix} 3 & 5 \\ 6 & -6 \end{pmatrix}$.

2. The set of all invertible matrices over the field $\mathbb{K}$ under matrix multiplication forms a non-Abelian group, denoted by $GL_n(\mathbb{K})$. Also observe that $\mathbb{M}_{n\times n}(\mathbb{K})$ forms a ring under the operations matrix addition and multiplication.

Definition 1.50 (*Scalar Multiplication*) Let $A = \left[a_{ij}\right] \in \mathbb{M}_{m\times n}(\mathbb{K})$ and $\lambda \in \mathbb{K}$, then $\lambda A = \left[\lambda a_{ij}\right] \in \mathbb{M}_{m\times n}(\mathbb{K})$.

Properties

For any matrices $A, B \in \mathbb{M}_{m\times n}(\mathbb{K})$ and $\lambda, \mu \in \mathbb{K}$

1. $\lambda(A + B) = \lambda A + \lambda B$
2. $(\lambda + \mu)A = \lambda A + \mu A$
3. $\lambda(\mu A) = (\lambda\mu)A$
4. $A(\lambda B) = \lambda(AB) = (\lambda A)B$.

Definition 1.51 (*Transpose of a matrix*) The transpose of an $m \times n$ matrix $A = \left[a_{ij}\right]$ is the $n \times m$ matrix (denoted by A^T), given by $A^T = \left[a_{ji}\right]$.

Properties

Let A and B be matrices of appropriate order, then

1. $\left(A^T\right)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$
4. $(kA)^T = kA^T$.

Definition 1.52 (*Conjugate transpose of a matrix*) The conjugate transpose of an $m \times n$ matrix $A = \left[a_{ij}\right]$ is the $n \times m$ matrix (denoted by A^*) given by $A^* = \left[\overline{a_{ji}}\right]$ where bar denotes complex conjugation (if $a_{ij} = c + id$, then $\overline{a_{ij}} = c - id$).

Properties

Let A and B be matrices of appropriate orders and λ be a scalar, then

1. $(A^*)^* = A$
2. $(A + B)^* = A^* + B^*$
3. $(AB)^* = B^* A^*$
4. $(\lambda A)^* = \overline{\lambda} A^*$, where $\overline{\lambda}$ is the conjugate of λ.

Definition 1.53 (*Trace of a matrix*) Let $A = [a_{ij}]$ be an $n \times n$ matrix. The trace of A, denoted by $tr(A)$, is the sum of diagonal entries; that is $tr(A) = \sum_{i=1}^{n} a_{ii}$.

Properties

For any $n \times n$ matrices $A, B, C,$ *and* D and $\lambda \in \mathbb{R}$, we have the following properties:

1. Trace is a linear function.
   $tr(A + B) = tr(A) + tr(B)$
   $tr(\lambda A) = \lambda \, tr(A)$
2. $tr(A^T) = tr(A)$ and $tr(A^*) = \overline{(trA)}$
3. $tr(AB) = tr(BA)$
4. $tr(ABCD) = tr(DABC) = tr(CDAB) = tr(BCDA)$
5. $tr(ABC) \neq tr(ACB)$ in general.
6. $tr(AB) \neq tr(A).tr(B)$ in general.

Definition 1.54 (*Determinant of a matrix*) For each square matrix A with entries in $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$), we can associate a single element of $\mathbb{K}$ called determinant of A, denoted by $det\ (A)$.
If A is a 1×1 matrix, i.e., $A = [a_{11}]$, then its determinant is defined by $det(A) = a_{11}$.
If A is a 2×2 matrix, say $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, then its determinant is defined by

$$det(A) = a_{11}a_{22} - a_{21}a_{12}$$

The determinant for a square matrix with higher dimension n may be defined inductively as follows:

$$det\ (A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

for a fixed j, where M_{ij} is the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by deleting ith row and jth column, called *minor* of the element a_{ij}.

Properties

Let A and B be any $n \times n$ matrices and λ be any scalar, then

1. $det\ (I_n) = 1$, where I_n is the $n \times n$ identity matrix.
2. $det\ (A^T) = det\ (A)$ and $det\ (A^*) = \overline{det\ (A)}$.
3. $det\ (AB) = det\ (A)\ det\ (B)$.
4. $det\ (\lambda A) = \lambda^n\ det\ (A)$.
5. If B is a matrix obtained from A by multiplying one row (or column) by a scalar λ, then $det\ (B) = \lambda\ det\ (A)$.
6. If B is a matrix obtained from A by interchanging any two rows (or columns) of A then $det\ (B) = -\ det\ (A)$.
7. If two rows of a matrix are identical then the matrix has determinant zero.
8. If B is a matrix obtained from A by adding λ times one row (or column) of A to another row (or column) of A, then $det\ (B) = det\ (A)$.

Remark 1.10 An $n \times n$ matrix with determinant zero is called *singular matrix*, otherwise it is called a *non-singular matrix.*

Definition 1.55 (*Adjoint of a Matrix*) The adjoint of a matrix $A = \left[a_{ij}\right]_{n\times n}$ (denoted by $adj\ (A)$) is the transpose of the co-factor matrix, where co-factor matrix of $A = \left[a_{ij}\right]_{n\times n}$ is $\left[(-1)^{i+j}M_{ij}\right]_{n\times n}$, where M_{ij} is the determinant of the $(n-1)\times(n-1)$ matrix obtained from A by deleting ith row and jth column, called minor of the ijth element.

Properties
Let A and B be any $n \times n$ matrices, then

1. $adj(I_n) = I_n$
2. $adj(AB) = adj(B)\ adj(A)$
3. $adj(kA) = k^{n-1}adj(A)$
4. $adj(A^m) = (adj(A))^m$
5. $adj(A^T) = (adj(A))^T$
6. $A\ adj(A) = det(A)\ I = adj(A)\ A$
7. $det\ (adj(A)) = (det(A))^{n-1}$
8. $adj\ (adj(A)) = (det(A))^{n-2}\ A.$

Definition 1.56 (*Inverse of a matrix*) The inverse of a square matrix $A_{n\times n}$ if it exists is the matrix $A^{-1}_{n\times n}$ such that $AA^{-1} = I_n = A^{-1}A$ and is given by $A^{-1} = \frac{1}{det(A)}adj(A)$.

Properties
Let A and B be any $n \times n$ matrices and λ be any scalar, then

1. The inverse of a matrix if it exists is unique.
2. A is invertible if and only if $det\ A \neq 0$.
3. $\left(A^{-1}\right)^{-1} = A$.
4. $(kA)^{-1} = k^{-1}A^{-1}$, where $k \neq 0$ is any scalar.
5. $det\ (A^{-1}) = \frac{1}{det\ (A)}$.
6. $(AB)^{-1} = B^{-1}A^{-1}$.
7. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$.

Remark 1.11 1. There are matrices for which $AB = I$ but $BA \neq I$. For example take
$A = \begin{bmatrix}1 & 2\end{bmatrix}$ and $B = \begin{bmatrix}1\\0\end{bmatrix}$. Then $AB = \begin{bmatrix}1\end{bmatrix} = I$ and $BA = \begin{bmatrix}1 & 2\\0 & 0\end{bmatrix} \neq I$.
2. If $A = \begin{bmatrix}a & b\\c & d\end{bmatrix}$ is invertible, then A^{-1} is given by $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix}d & -b\\-c & a\end{bmatrix}$.
3. Set of all $n \times n$ non-singular matrices with entries from the field $\mathbb{K}$ under matrix multiplication forms a non-Abelian group called *general linear group*, and is denoted by $GL_n\ (\mathbb{K})$.

    1. For any matrices $A, B \in GL_n\ (\mathbb{K})$, $AB \in GL_n\ (\mathbb{K})$ ($det(A), det(B) \neq 0 \Rightarrow det(AB) \neq 0$). (Closure property)

2. Matrix multiplication is associative.
3. $I_n \in GL_n(\mathbb{K})$ acts as identity matrix.
4. For each $A \in GL_n(\mathbb{K})$, we have $det(A) \neq 0$ and hence A^{-1} exists. Also, $det\ (A^{-1}) = \frac{1}{det\ (A)}$, and thus $A^{-1} \in GL_n(\mathbb{K})$.

Definition 1.57 (*Rank of a matrix*) The rank of a matrix is the order of the highest order sub-matrix having non-zero determinant.

Properties

1. Let A be an $m \times n$ matrix. Then $Rank(A) \leq min\{m, n\}$.
2. Only zero matrix has rank zero.
3. A square matrix $A_{n\times n}$ is invertible if and only if $Rank(A) = n$.
4. *Sylvester's Inequality:* If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then

$$Rank(A) + Rank(B) - n \leq Rank(AB) \leq min\{Rank(A), Rank(B)\}$$

   This result is named after the famous English mathematician *James Joseph Sylvester (1814–1897).*
5. *Frobenius Inequality:* Let A, B, and C be any matrices such that AB, BC, and ABC exists, then

$$Rank(AB) + Rank(BC) \leq Rank(ABC) + Rank(B)$$

   This result is named after the famous German mathematician *Ferdinand Georg Frobenius (1849–1917).*
6. Rank is sub-additive. That is, $Rank(A + B) \leq Rank(A) + Rank(B)$.
7. $Rank(A) = Rank(A^T) = Rank(A^T A)$.
8. $Rank(kA) = Rank(A)$ if $k \neq 0$.

Definition 1.58 (*Block Matrix*) A block matrix or a partitioned matrix is a matrix that is defined using smaller matrices called blocks.

Example 1.70 Consider $X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}_{5\times5}$ where $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}_{2\times2}$, $B = \begin{bmatrix} 2 & 1 & 3 \\ 6 & 2 & 7 \end{bmatrix}_{2\times3}$, $C = \begin{bmatrix} 1 & 0 \\ 5 & 2 \\ 7 & 3 \end{bmatrix}_{3\times2}$, and $D = \begin{bmatrix} 1 & 9 & 8 \\ 4 & 2 & 1 \\ 7 & 0 & 1 \end{bmatrix}_{3\times3}$.

Properties

1. Let $X = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$ where $A_{n\times n}$, $B_{n\times m}$, $C_{m\times n}$, and $D_{m\times m}$ are matrices.
   If A is invertible, then

$$det\ (X) = (det\ (A))\left(det\ (D - CA^{-1}B)\right)$$

Definition 1.59 (*Block Diagonal Matrix*) A block diagonal matrix is a block matrix which is a square matrix such that all blocks except the diagonal ones are zero.

Properties

1. Consider a block diagonal matrix of the form
$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n \end{bmatrix}$$, where each $A_i' s$ is a square matrix. Then
   (a) $det(A) = det(A_1)det(A_2) \cdots det(A_n)$
   (b) $Tr(A) = Tr(A_1) + Tr(A_2) + \cdots + Tr(A_n)$
   (c) $Rank(A) = Rank(A_1) + Rank(A_2) + \cdots + Rank(A_n)$.

Definition 1.60 (*Elementary Operations*) There are three kinds of elementary matrix operations:

(1) Interchanging two rows (or columns).
(2) Multiplying each element in a row (or column) by a non-zero number.
(3) Multiplying a row (or column) by a non-zero number and adding the result to another row (or column).

When these operations are performed on rows, they are called *elementary row operations;* and when they are performed on columns, they are called *elementary column operations.*

Definition 1.61 (*Equivalent matrices*) Two matrices A and B are said to be row(column) equivalent if there is a sequence of elementary row(column) operations that transforms A into B and is denoted by $A \sim B$.

Definition 1.62 (*Row Echelon form of a matrix*) A matrix is said to be in row echelon form when it satisfies the following conditions:

(a) Each leading entry (the first non-zero entry in a row) is in a column to the right of the leading entry in the previous row.
(b) Rows with all zero elements, if any, are below rows having a non-zero element.

If the matrix also satisfies the condition

(c) The first non-zero element in each row, called the leading entry or pivot, is 1.

Then the matrix is in *reduced row echelon form.*

Example 1.71 Consider the matrix $A = \begin{bmatrix} 3 & 2 & 1 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 6 & 11 & 12 \end{bmatrix}$. Now

$$
\begin{aligned}
A = \begin{bmatrix} 3 & 2 & 1 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 6 & 11 & 12 \end{bmatrix} & \qquad R_1 \leftrightarrow R_2 \\
\sim \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \\ 1 & 6 & 11 & 12 \end{bmatrix} & \qquad \begin{matrix} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 - R_1 \end{matrix} \\
\sim \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -8 \\ 0 & 4 & 8 & 8 \end{bmatrix} & \qquad R_3 \to R_3 + R_2 \\
\sim \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -4 & -8 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix} & \qquad R_2 \to -\tfrac{1}{4} R_2 \\
\sim \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix} = B &
\end{aligned}
$$

Then B is called the reduced row echelon form of A.

Remark 1.12 1. A matrix is equivalent to any of its row echelon form and reduced row echelon form. The reduced row echelon form of A is unique.
2. The rank of a matrix is equal to the number of non-zero rows in its row echelon form. For example, the matrix $A = \begin{bmatrix} 3 & 2 & 1 & 4 \\ 1 & 2 & 3 & 4 \\ 1 & 6 & 11 & 12 \end{bmatrix}$ has rank 2 as it is equivalent to $B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, which is in the row echelon form.

1.6 Euclidean Space $\mathbb{R}^n$

In a mathematical environment, Euclidean space is a geometric concept that contains all conceivable positions and locations. It provides the theoretical framework for many other mathematical fields, including classical geometry. We can use well-defined connections and rules to describe points, lines, angles, and distances inside this space. It acts as a foundational tool and gives a framework for comprehending spatial relationships. Any point in $\mathbb{R}^n$ is a list of n real numbers, denoted as $v = (v_1, v_2, \ldots, v_n)$. For convenience, we may use this list as a matrix with one column or one row called *column vector* and *row vector*, respectively. In the physical world, a vector is a quantity which has both magnitude and direction, which can be easily visualized when we work on $\mathbb{R}^2$ or $\mathbb{R}^3$.

Vectors in $\mathbb{R}^2$

Algebraically, a vector in $\mathbb{R}^2$ is simply an ordered pair of real numbers. That is $\mathbb{R}^2 = \{(v_1, v_2) \mid v_1, v_2 \in \mathbb{R}\}$. Two vectors (u_1, u_2) and (v_1, v_2) are equal if and only if the corresponding components are equal. That is, if and only if $u_1 = v_1$ and $u_2 = v_2$. Now we can define some operations on $\mathbb{R}^2$.

Definition 1.63 (*Vector Addition*) The sum of two vectors $u = (u_1, u_2)$ and $v = (v_1, v_2)$, denoted by $u + v$, is given by $u + v = (u_1 + v_1, u_2 + v_2) \in \mathbb{R}^2$.

Properties
Let $u = (u_1, u_2)$, $v = (v_1, v_2)$, $w = (w_1, w_2) \in \mathbb{R}^2$. Then

1. $u + v = (u_1 + v_1, u_2 + v_2) = (v_1 + u_1, v_2 + u_2) = v + u$. (Commutative)
2. $u + (v + w) = (u_1 + (v_1 + w_1), u_2 + (v_2 + w_2)) = ((u_1 + v_1) + w_1, (u_2 + v_2) + w_2) = (u + v) + w$. (Associative)
3. There exists $\mathbf{0} = (0, 0)$ such that $v + \mathbf{0} = v$ for all v. (Existence of identity element)
4. For each $v \in \mathbb{R}^2$, there exists $-v = (-v_1, -v_2) \in \mathbb{R}^2$ such that $v + (-v) = \mathbf{0}$. (Existence of inverse)

Remark 1.13 The set $\mathbb{R}^2$ with vector addition forms an Abelian group.

Definition 1.64 (*Scalar Multiplication*) Let $v = (v_1, v_2) \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$, then $\lambda v = (\lambda v_1, \lambda v_2) \in \mathbb{R}^2$.

Properties
Let $u = (u_1, u_2)$, $v = (v_1, v_2) \in \mathbb{R}^2$ and $\lambda, \mu \in \mathbb{R}$. Then

1. $\lambda(u + v) = (\lambda(u_1 + v_1), \lambda(u_2 + v_2)) = \lambda(u_1, u_2) + \lambda(v_1, v_2) = \lambda u + \lambda v$
2. $(\lambda + \mu)v = ((\lambda + \mu)v_1, (\lambda + \mu)v_2) = \lambda(v_1, v_2) + \mu(v_1, v_2) = \lambda v + \mu v$
3. $\lambda(\mu v) = (\lambda\mu)v = \mu(\lambda v)$.

From the above properties, it is clear that $0v = 0$ for any $v \in V$ and $0 \in \mathbb{R}$. Also, $(-1)v = -v$ for any $v \in V$ and $-1 \in \mathbb{R}$.

The Geometric Notion of Vectors in $\mathbb{R}^2$

Corresponding to every vector in $\mathbb{R}^2$, there exists a point in the Cartesian plane, and each point in the Cartesian plane represents a vector in $\mathbb{R}^2$. But the representation of vectors in $\mathbb{R}^2$ as points of Cartesian plane does not provide much information about the operations like vector addition and scalar multiplication. So it is better to represent a vector in $\mathbb{R}^2$ as a directed line segment which begins at the origin and ends at the point. Such a visualization of a vector v is called position vector of v. Then as in the physical world, the vector possess both magnitude and direction. However, to represent a vector in $\mathbb{R}^2$, the directed line segment need not start from the origin;

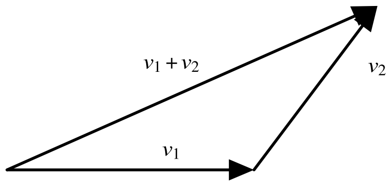


Fig. 1.16 Triangle law of vector addition

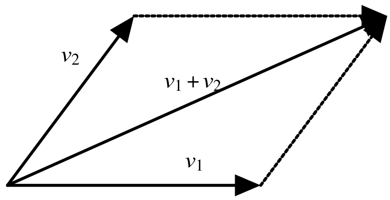


Fig. 1.17 Parallelogram law of vector addition

it may start at some point in $\mathbb{R}^2$, but the magnitude and direction cannot vary. For convenience, the directed line segment is considered to be starting from the origin.

Theorem 1.19 (Triangle Law of Vector Addition) *If two vectors are represented in magnitude and direction by the two sides of a triangle, taken in order, then their sum is represented in magnitude and direction by the third side of the triangle, taken in the reverse order (Fig. 1.16).*

Theorem 1.20 (Parallelogram Law of vector Addition) *If two vectors are represented in magnitude and direction by the two adjacent sides of a parallelogram, then their sum is represented in magnitude and direction by the diagonal of the parallelogram through their common point (Fig. 1.17).*

These ideas of vectors and vector operations in $\mathbb{R}^2$ can be extended to general Euclidean space $\mathbb{R}^n$.

1.7 System of Linear Equations

Solving simultaneous linear equations is one among the central problems in algebra. In this section, we will get to know some of the methods that are used to solve the system of linear equations. Let us start by discussing the solution of a system having n equations in n unknowns. Consider the basic problem with $n = 1$, i.e., consider an equation of the form, $ax = b$. We know that there are three possible numerical realizations for this equation:

(1) $a \neq 0$: In this case, we know that the equation have a unique solution, which is $x = \frac{b}{a}$.
(2) $a, b = 0$: Any numerical value for x will be a solution for this equation. That is, there are infinite number of solutions.

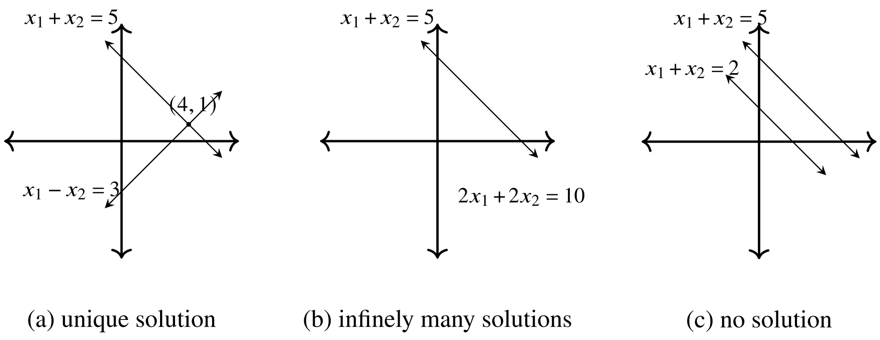


Fig. 1.18 Observe that in **a**, the lines $x_1 + x_2 = 5$ and $x_1 - x_2 = 3$ have a unique intersection point (4, 1), in **b** both the equations $x_1 + x_2 = 5$ and $2x_1 + 2x_2 = 10$ represent the same line and in **c**, the lines $x_1 + x_2 = 5$ and $x_1 + x_2 = 2$ are parallel to each other

(3) $a = 0, b \neq 0$: Then it is clear that no numerical value of x would satisfy the equation. That is, the system has no solutions.

Now consider a set of two equations in 2 unknowns x_1 and x_2:

$$a_1x_1 + a_2x_2 = b_1$$

$$a_3x_1 + a_4x_2 = b_2$$

We know that these equations represent two lines on a plane and solution of this system, if it exists, are the intersecting points of these two lines. If the lines are intersecting, either there will be a unique intersection point or there will be an infinite number of intersection points and if the lines are non-intersecting, they must be parallel to each other. Thus, here also, there are only three possibilities. The possibilities will be the same in the case of a system of n equations with n unknowns. The three possibilities are demonstrated in the Fig. 1.18.

Now that we have seen the possibilities for the number of solutions of a system of equations, we have to find a method to solve a system of linear equations. Consider a system of n equations in n unknowns $x_1, x_2, \ldots, x_n$ given by

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\vdots \qquad \vdots \qquad\quad \vdots \qquad &\;\;\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}$$

The system can be written in the form $Ax = b$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

The matrix A is called the *coefficient matrix*. A method to solve this system is given by *Gabriel Cramer (1704–1752)*, using the determinants of the coefficient matrix and matrices obtained from it by replacing one column by the column vector of right-hand sides of the equations. *Cramer's rule* states that if $x = (x_1, x_2, \ldots, x_n)$ is a solution of the system, $x_i = \frac{det(A_i)}{det(A)}$, $i = 1, 2, \ldots, n$, where A_i is the matrix obtained by replacing the ith column of A by the column vector b. Observe that this rule is applicable only if $det(A) \neq 0$. For example, consider the equations $x_1 + x_2 = 5$ and $x_1 - x_2 = 3$. The system can be expressed in the form,

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

As $det(A) = -2 \neq 0$, we have

$$x = \frac{det\left(\begin{bmatrix} 5 & 1 \\ 3 & -1 \end{bmatrix}\right)}{det\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)} = 4 \text{ and } y = \frac{det\left(\begin{bmatrix} 1 & 5 \\ 1 & 3 \end{bmatrix}\right)}{det\left(\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)} = 1$$

As we can see, *Cramer's rule* is applicable only if the determinant of A is non-zero. Even if the determinant of A is non-zero, this rule may cause computational difficulties for higher values of n. Also it cannot be applied to a system of m equations in n unknowns. Another method to find the solution of a system of equations is *elimination*, in which multiples of one equation is added or subtracted to other equations so as to remove the unknowns from the equations till only one equation in one by unknown remains, which can be solved easily. We can use the value of this unknown to find the value of the remaining ones.

Consider a system of m equations in n unknowns $x_1, x_2, \ldots, x_n$ given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

The system can be written in the form $Ax = b$, where $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$. The matrix A is called the *coefficient matrix*, and the matrix $[A \mid b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{bmatrix}$ is called the *augmented matrix* of the system. If $b = 0$, then the system is called a *homogeneous system.* Otherwise, it is called *non-homogeneous system.* A system is said to be *consistent*, if it has a solution. Otherwise, it is called *inconsistent.* We will see that a homogeneous system is always consistent, whereas a non-homogeneous system can be inconsistent (as given in Fig. 1.18c).

Gauss Elimination Method

Consider a system of equations given by $Ax = b$. We can solve the system using the following method called *Gauss elimination method*, named after the famous German mathematician *Carl Friedrich Gauss (1777–1855).*

1. Construct the augmented matrix for the given system of equations.
2. Use elementary row operations to transform the augmented matrix to its row echelon form.
3. The system
   - is consistent if and only if $Rank\,[A \mid b] = Rank(A)$.
     - ◇ has unique solution if and only if $Rank\,[A \mid b] = Rank(A) = n$.
     - ◇ has an infinite number of solutions if $Rank\ [A \mid b] = Rank(A) = r < n$.
   - is inconsistent if and only if $Rank\,[A \mid b] \neq Rank\ (A)$.
4. If the system is consistent, write and solve the new set of equations corresponding to the row echelon form of the augmented matrix.

If reduced row echelon form is used, the method is called *Gauss–Jordan* method.

Remark 1.14 A homogeneous system $Ax = 0$ is always consistent (since $Rank\,[A \mid 0] = Rank(A)$ always). The system

- has a unique solution if $Rank(A) = n$.
- has infinite number of solutions if and only if $Rank(A) = r < n$.

Example 1.72 Consider the system of equations

$$2x_1 + 3x_2 + 5x_3 = 9$$

$$7x_1 + 3x_2 - 2x_3 = 8$$

$$2x_1 + 3x_2 + \lambda_1 x_3 = \lambda_2$$

where λ_1 and λ_2 are some real numbers.
The above system can be written in the matrix form $Ax = b$ as

$$\begin{bmatrix} 2 & 3 & 5 \\ 7 & 3 & -2 \\ 2 & 3 & \lambda_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \\ \lambda_2 \end{bmatrix}$$

Now the augmented matrix $[A \mid b]$ is given by

$$[A \mid b] = \begin{bmatrix} 2 & 3 & 5 & 9 \\ 7 & 3 & -2 & 8 \\ 2 & 3 & \lambda_1 & \lambda_2 \end{bmatrix} \qquad \begin{matrix} R_2 \to R_2 - \frac{7}{2}R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 2 & 3 & 5 & 9 \\ 0 & \frac{-15}{2} & \frac{-39}{2} & \frac{-47}{2} \\ 0 & 0 & \lambda_1 - 5 & \lambda_2 - 9 \end{bmatrix}$$

As the first two rows in the reduced form are non-zero, both $Rank(A)$ and $Rank\,[A \mid b]$ are greater than or equal to 2.

- ◇ The system has unique solution if and only if $Rank\,[A \mid b] = Rank(A) = 3$. That is, if $\lambda_1 \neq 5$ and for any arbitrary values λ_2.
- ◇ The system has an infinite number of solutions if $Rank\,[A \mid b] = Rank(A) < 3$. If $\lambda_1 = 5$ and $\lambda_2 = 9$, we have $Rank\,[A \mid b] = Rank(A) = 2 < 3$.
- ◇ The system has no solution when $Rank\,[A \mid b] \neq Rank(A)$. That is, if $\lambda_1 = 5$ and $\lambda_2 \neq 9$.

If $b = 0$ in the above system, then

- ◇ The homogeneous system has a unique solution if and only if $Rank(A) = 3$. That is, if $\lambda_1 \neq 5$ the given system has only the zero vector as solution.
- ◇ If $\lambda_1 = 5$, then $Rank(A) = 2 < 3$ and hence the given system has an infinite number of solutions.

As we have identified the values of λ_1 and λ_2 for which the given system is consistent, let us try to compute the solutions of the given system for some particular values of λ_1 and λ_2. Take $\lambda_1 = 1$ and $\lambda_2 = 9$. Then,

$$[A \mid b] \sim \begin{bmatrix} 2 & 3 & 5 & 9 \\ 0 & \frac{-15}{2} & \frac{-39}{2} & \frac{-47}{2} \\ 0 & 0 & -4 & 0 \end{bmatrix}$$

That is, the given system is reduced to the following equivalent form:

$$\begin{aligned} 2x_1 + 3x_2 + 5x_3 &= 9 \\ \frac{15}{2}x_2 + \frac{39}{2}x_3 &= \frac{47}{2} \\ -4x_3 &= 0 \end{aligned}$$

Thus, we have $x = \begin{bmatrix} \frac{-1}{5} \\ \frac{47}{15} \\ 0 \end{bmatrix}$ as the unique solution for the given system. Similarly, if we take $\lambda_1 = 5$ and $\lambda_2 = 9$, we can show that set of all solutions of the given system is $\left\{(x_1, x_2, x_3) \mid x_3 \in \mathbb{R}, x_1 = \frac{14x_3-2}{10} \; and \; x_2 = \frac{47-39x_3}{15}\right\}$ (Verify!).

Remark 1.15 If the coefficient matrix A is an $n \times n$ non-singular matrix, then the system $Ax = b$ has a unique solution $x = A^{-1}b$.

LU Decomposition

The LU decomposition method consists of factorizing A into a product of two triangular matrices

$$A = LU$$

where L is the lower triangular and U is the upper triangular. We use the *Doolittle method* to convert A into the form $A = LU$, where L and U are as mentioned above. We initialize this process by setting $A = IA$ and use Gaussian elimination procedure to achieve the desired form. The pivot element is identified in each column during this procedure, and if necessary, the rows are switched. We update the entries of both I and A on the right-hand side in accordance with each column, using row operations to remove elements below the main diagonal and multipliers to generate L. We get a lower triangular matrix L with ones on its principal diagonals and an upper triangular matrix U after iterating over all the columns. This decomposition allows us to reduce the solution of the system $Ax = b$ to solving two triangular systems $Ly = b$ and $Ux = y$. Generally, there are many such factorizations. If L is required to have all diagonal elements equal to 1, then the decomposition, when it exists, is unique. This method was introduced by the Polish mathematician *Tadeusz Julian Banachiewicz (1882–1954)*.

Example 1.73 Consider the system of equations

$$2x_1 - x_2 + 3x_3 = 9$$

$$4x_1 + 2x_2 + x_3 = 9$$

$$-6x_1 - x_2 + 2x_3 = 12$$

The above system can be written in the matrix form $Ax = b$ as

$$\begin{bmatrix} 2 & -1 & 3 \\ 4 & 2 & 1 \\ -6 & -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \\ 12 \end{bmatrix}$$

Consider the coefficient matrix A. We will use elementary row transformations to convert A into the form LU. We have

$$A = \begin{bmatrix} 2 & -1 & 3 \\ 4 & 2 & 1 \\ -6 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 \\ 4 & 2 & 1 \\ -6 & -1 & 2 \end{bmatrix} \quad \begin{matrix} R_2 \to R_2 - (2)R_1 \\ R_3 \to R_3 - (-3)R_1 \end{matrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & -5 \\ 0 & -4 & 11 \end{bmatrix} \quad R_3 \to R_3 - (-1)R_1$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & -5 \\ 0 & 0 & 6 \end{bmatrix} = LU$$

Now $Ly = b$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 9 \\ 9 \\ 12 \end{bmatrix}$$

Solving the system, we get $y_1 = 9$, $y_2 = -9$, and $y_3 = 30$. Now consider the system $Ux = y$

$$\begin{bmatrix} 2 & -1 & 3 \\ 0 & 4 & -5 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 9 \\ -9 \\ 30 \end{bmatrix}$$

Solving the system, we get $x_1 = -1$, $x_2 = 4$, and $x_3 = 5$.

Theorem 1.21 *If y and z are two distinct solutions of $Ax = b$, then $\lambda y + \mu z$ is also a solution of $Ax = b$, for any scalars $\lambda, \mu \in \mathbb{K}$ with $\lambda + \mu = 1$. If $b = 0$, $\lambda y + \mu z$ is a solution of $Ax = 0$, for any scalars $\lambda, \mu \in \mathbb{K}$.*

Proof Suppose that $b \neq 0$ and y and z are two given solutions of $Ax = b$, then $Ay = b$ and $Az = b$. Let $\lambda, \mu \in \mathbb{K}$ be such that $\lambda + \mu = 1$. Then

$$A(\lambda y + \mu z) = \lambda Ay + \mu Az = \lambda b + \mu b = (\lambda + \mu)b = b$$

Now let $b = 0$. If y and z are two given solutions of $Ax = 0$, then $Ay = 0$ and $Az = 0$. Then

$$A(\lambda y + \mu z) = \lambda Ay + \mu Az = 0$$

Hence the proof.

1.8 Exercises

1. For any sets A and B, show that
   (a) $A \cap B \subseteq A$, $B \subseteq A \cup B$.
   (b) $A \subseteq B$ if and only if $A \cap B = A$.
2. Consider the relation $R = \{(0, 1), (0, 2), (1, 2)\}$ on $X = \{0, 1, 2\}$. Check whether R is an equivalence relation.
3. Let $f : X \to Y$ and $g : Y \to Z$ be any two functions. Then show that
   (a) if f and g are one-one, then $g \circ f$ is one-one.
   (b) if f and g are onto, then $g \circ f$ is onto.
4. Check whether the following functions are bijective or not.
   (a) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2 + 1$
   (b) $f : [0, \pi] \to [-1, 1]$ defined by $f(x) = sin\ x$
   (c) $f : \mathbb{R}^* \to \mathbb{R}^*$ defined by $f(x) = \frac{1}{x}$
   (d) $f : \mathbb{C} \to \mathbb{C}$ defined by $f(z) = \overline{z}$.
5. Let $\lambda_i, \mu_i \in \mathbb{K}$, $i \in \mathbb{N}$. Then show that
   (a) for $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$, we have

   $$\sum_{i=1}^{\infty} |\lambda_i \mu_i| \leq \left(\sum_{i=1}^{\infty} |\lambda_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^{\infty} |\mu_i|^q \right)^{\frac{1}{q}}$$

   (b) for $1 < p < \infty$, we have

   $$\left(\sum_{i=1}^{\infty} |\lambda_i + \mu_i|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |\lambda_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |\mu_i|^p \right)^{\frac{1}{p}}$$

   These inequalities are called *Holder's inequality* and *Minkowski's inequality*, respectively.
6. For $1 < p < \infty$, consider the following collections of sequences.

$$l^p = \left\{ v = (v_1, v_2, \ldots) \mid v_i \in \mathbb{K} \text{ and } \sum_{i=1}^{\infty} |v_i|^p < \infty \right\}$$

and

$$l^\infty = \left\{ v = (v_1, v_2, \ldots) \mid v_i \in \mathbb{K} \text{ and } \sup_{i \in \mathbb{N}} |v_i| < \infty \right\}$$

Show that for $u = (u_1, u_2, \ldots), v = (v_1, v_2, \ldots) \in l^p$

$$d_p(u, v) = \left(\sum_{i=1}^{\infty} |u_i - v_i|^p \right)^{\frac{1}{p}}$$

defines a metric on l^p and for $u = (u_1, u_2, \ldots), v = (v_1, v_2, \ldots) \in l^\infty$,

$$d_\infty(u, v) = \sup_{i \in \mathbb{N}} |u_i - v_i|$$

defines a metric on l^∞.

7. Let X be a metric space with respect to the metrics d_1 and d_2. Then show that each of the following:

   (a) $d(x, y) = d_1(x, y) + d_2(x, y)$
   (b) $d(x, y) = \dfrac{d_1(x, y)}{1 + d_1(x, y)}$
   (c) $d(x, y) = max\{d_1(x, y) + d_2(x, y)\}$

   also defines a metric on X.
8. Let (X, d) be a metric space. Show that

   (a) union of any number of open sets is open.
   (b) finite intersection of open sets is open.

   Also give an example to show that arbitrary intersection of open sets need not necessarily be open.
9. Show that a set is closed if and only if it contains all its limit points.
10. Show that $\left(l^p, d_p\right)$ and (l^∞, d_∞) are complete metric spaces.
11. Show that a closed subspace of a complete metric space is complete.
12. Prove that if a sequence of continuous functions on $[a, b]$ converges on $[a, b]$ and the convergence is uniform on $[a, b]$, then the limit function f is continuous on $[a, b]$.
13. Let $x \in \mathbb{R}$. Show that the sequence $\{x_n\}$, where $x_n = \frac{\lfloor nx \rfloor}{n}$, is a rational sequence that converges to x. ($\lfloor x \rfloor$ denotes the greatest integer less than or equal to x.)
14. Let $(G, *)$ be a group. Then show that

    (a) the identity element in G is unique.
    (b) each element in G has a unique inverse.

15. **Center of a group**: Let $(G, *)$ be group. The center of G, denoted by $\mathcal{Z}(G)$, is the set of all elements of G that commute with every other element of G.
    (a) Show that $\mathcal{Z}(G)$ is a subgroup of G.
    (b) Show that $\mathcal{Z}(G) = G$ for an Abelian group.
    (c) Find the center of $GL_2(\mathbb{K})$ and S_3.
16. Find the order of the following elements in $GL_2(\mathbb{K})$
    (a) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
    (b) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.
17. Let $\phi : (G, *) \to (G', *')$ be a homomorphism. Then, prove the following:
    (a) if e is the identity element in G, $\phi(e)$ is the identity element in G'.
    (b) $Ker\ (\phi)$ is a subgroup of G.
    (c) for any $g \in G$, if $\mathcal{O}(g)$ is finite $\mathcal{O}(\phi(g))$ divides $\mathcal{O}(g)$.
    (d) for any subgroup H of G, $\phi(H)$ is a subgroup of $\phi(G)$ and if H is Abelian, $\phi(H)$ is also Abelian.
18. Consider $\phi : GL_n(\mathbb{K}) \to (\mathbb{R}^*, .)$, defined by $\phi(A) = det(A)$.
    (a) Show that ϕ is a homomorphism.
    (b) Find $Ker\ (\phi)$.
19. Show that every cyclic group is Abelian.
20. Find the normal subgroups of S_3.
21. Prove that $\langle \mathbb{Q}, +, . \rangle$, $\langle \mathbb{R}, +, . \rangle$, and $\langle \mathbb{C}, +, . \rangle$ are fields with respect to the given algebraic operations. Also show that $\langle \mathbb{Z}, +, . \rangle$ is not a field.
22. Give an example of a finite field.
23. Show that $\mathbb{K}[x] = \{a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + a_n x^n \mid a_i \in \mathbb{K}, n \in \mathbb{Z}^+\}$ forms a ring with respect to the operations defined in Definition 1.44.
24. Prove *the Fundamental Theorem of Algebra*.
25. Show that the set of all $n \times n$ matrices with entries in $\mathbb{K}$, denoted by $M_n(\mathbb{K})$ with matrix addition and scalar multiplication, forms a ring with unity.
26. Find the rank of the matrix $A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 4 & 5 & 3 & 6 \\ 0 & 1 & 2 & -1 \end{bmatrix}$ using row reduced echelon form.
27. Show that the set of all solutions of a homogeneous system of equations forms a group with respect to coordinate-wise addition and scalar multiplication.
28. Consider the system of equations

$$2x_1 + x_2 + 3x_3 = 9$$

$$3x_1 + 2x_2 + 5x_3 = 15$$

$$4x_1 - 2x_2 + 7x_3 = 16$$

Solve the above system of equations using

(a) Gauss Elimination method
(b) LU Decomposition method.

Is it possible to solve this system using *Cramer's rule*? If yes, find the solution using *Cramer's rule*.

Solved Questions related to this chapter are provided in Chap. 8.

Chapter 2
Vector Spaces

This chapter explores one of the fundamental topics in linear algebra. It starts by defining vector spaces, highlighting their importance as mathematical structures with essential qualities such as closure under addition and scalar multiplication. Subspaces are introduced as vector space subsets with their vector space features, followed by an in-depth analysis of linear dependence and independence of vectors, which are critical for constructing bases. The ideas of span and basis are emphasized as critical tools for understanding the structure of vector spaces, with dimension serving as a quantitative measure of their complexity. Finally, the chapter looks into vector space sums and the particular case of the direct sum, providing a more in-depth understanding of vector space composition.

2.1 Introduction

In Chap. 1, we have called an element of Euclidean space $\mathbb{R}^n$ a "vector". From this chapter onwards, we will be using the term "vector" with a broader meaning. An element of a *vector space* is called a vector. Roughly speaking, a vector space is a collection of objects which are closed under vector addition and scalar multiplication and are subjected to some reasonable rules. The rules are chosen so that we can manipulate the vectors algebraically. We can also consider a vector space as a generalization of the Euclidean space. In this chapter, we will be discussing vector spaces in detail.

Definition 2.1 (*Vector space*) A vector space or linear space V over a field $\mathbb{K}$ is a non-empty set together with two operations called vector addition (denoted by '$+$') and scalar multiplication (as the elements of $\mathbb{K}$ are called scalars) satisfying certain conditions:

$(V1)$ $v_1 + v_2 \in V$ for all $v_1, v_2 \in V$.
$(V2)$ $\lambda v \in V$ for all $\lambda \in \mathbb{K}$ and $v \in V$.

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_2

$(V1)$ and $(V2)$ respectively imply that V is closed under both vector addition and scalar multiplication. The following properties are familiar as we have seen these in Chap. 1, associated with another algebraic structure, *Group*.

- $(V3)$ $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ for all $v_1, v_2, v_3 \in V$.
- $(V4)$ there is an element $\mathbf{0} \in V$ such that $v + \mathbf{0} = v$ for all $v \in V$.
- $(V5)$ for each $v \in V$ there exists an element $-v \in V$ such that $v + (-v) = \mathbf{0}$.

Thus we can say that V under vector addition must be a group. Now $(V6)$ imply that $(V, +)$ is not just any group, it must be an Abelian group.

- $(V6)$ $v_1 + v_2 = v_2 + v_1$ for all $v_1, v_2 \in V$.

Along with closure properties and $(V, +)$ being an Abelian group, the following properties also must be satisfied for V to be a vector space over the field $\mathbb{K}$ under the given operations.

- $(V7)$ $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$ for all $\lambda \in \mathbb{K}$ and $v_1, v_2 \in V$.
- $(V8)$ $(\lambda + \mu)v = \lambda v + \mu v$ for all $\lambda, \mu \in \mathbb{K}$ and $v \in V$.
- $(V9)$ $(\lambda\mu)v = \lambda(\mu v)$ for all $\lambda, \mu \in \mathbb{K}$ and $v \in V$.
- $(V10)$ $1v = v$ for all $v \in V$.

Now let us get familiar with some of the important vector spaces that we will see throughout this book. Let us start with a basic one.

Example 2.1 Consider V as the set of all real numbers, $\mathbb{R}$ under usual addition as vector addition and usual multiplication as scalar multiplication, the scalars being taken from the field $\mathbb{R}$ itself. In Chap. 1, we have seen that $(\mathbb{R}, +)$ is an Abelian group. Scalar multiplication in this case is the usual multiplication of real numbers, which is closed. Properties $(V7) - (V10)$ are easy to verify. Thus $\mathbb{R}$ over $\mathbb{R}$ is a vector space. Similarly, we can show that $\mathbb{C}$ over $\mathbb{C}$ is a vector space. What about $\mathbb{C}$ over $\mathbb{R}$ and $\mathbb{R}$ over $\mathbb{C}$?

Example 2.2 Let $\mathbb{K}$ be any field. Then $\mathbb{K}^n$ is a vector space over $\mathbb{K}$, where n is a positive integer and

$$\mathbb{K}^n = \left\{ \left(x_1, x_2, \ldots, x_n\right) \mid x_1, x_2, \ldots, x_n \in \mathbb{K} \right\}$$

Addition and scalar multiplication are defined component-wise as we have seen in the previous chapter:

$$\left(x_1, x_2, \ldots, x_n\right) + \left(y_1, y_2, \ldots, y_n\right) = \left(x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n\right)$$

$$\lambda \left(x_1, x_2, \ldots, x_n\right) = \left(\lambda x_1, \lambda x_2, \ldots, \lambda x_n\right), \ \lambda \in \mathbb{K}$$

In particular, $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ and $\mathbb{C}^n$ is a vector space over $\mathbb{C}$ (Verify). Is $\mathbb{R}^n$ a vector space over $\mathbb{C}$?

Example 2.3 The collection of all $m \times n$ matrices, $\mathbb{M}_{m\times n}(\mathbb{K})$, with the usual matrix addition and scalar multiplication is a vector space over $\mathbb{K}$.

Example 2.4 If $\mathbb{F}$ is a sub-field of a field $\mathbb{K}$, then $\mathbb{K}$ is a vector space over $\mathbb{F}$, with addition and multiplication just being the operations in $\mathbb{K}$. Thus, in particular, $\mathbb{C}$ is a vector space over $\mathbb{R}$ and $\mathbb{R}$ is a vector space over $\mathbb{Q}$.

Example 2.5 Let $\mathbb{P}_n[a, b]$ denote the set of all polynomials of degree less than or equal to n defined on $[a, b]$ with coefficients from the field $\mathbb{K}$. For $p, q \in \mathbb{P}_n[a, b]$, and $\lambda \in \mathbb{K}$ the addition and scalar multiplication are defined by

$$(p + q)(x) = p(x) + q(x) = (a_n + b_n)x^n + \cdots + (a_1 + b_1)x + (a_0 + b_0)$$

where $p(x) = a_n x^n + \cdots + a_1 x + a_0$ and $q(x) = b_n x^n + \cdots + b_1 x + b_0$ and

$$(\lambda p)(x) = \lambda\,(p(x)) = (\lambda a_n)x^n + \cdots + (\lambda a_1)x + (\lambda a_0)$$

$\mathbb{P}_n[a, b]$ along with zero polynomial forms a vector space over $\mathbb{K}$. Denote by $\mathbb{P}[a, b]$ the collection of all polynomials defined on $[a, b]$ with coefficients from $\mathbb{K}$. Then $\mathbb{P}[a, b]$ is a vector space over $\mathbb{K}$ with respect to the above operations for polynomials.

Example 2.6 Let $C[a, b]$ denote the set of all real-valued continuous functions on the interval $[a, b]$. If f and g are continuous functions on $[a, b]$, then the vector addition and scalar multiplication are defined by

$$(f + g)(x) = f(x) + g(x) \text{ and } (\lambda f)(x) = \lambda f(x)$$

where $\lambda \in \mathbb{R}$. Then $C[a, b]$ is a vector space with respect to the above operations over the field $\mathbb{R}$.

Example 2.7 Let $\mathbb{K}$ be any field. Let V consist of all sequences $\{a_n\}$ in $\mathbb{K}$ that have only a finite number of non-zero terms a_n. If $\{a_n\}$ and $\{b_n\}$ are in V and $\lambda \in \mathbb{K}$, define

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \text{ and } \lambda\{a_n\} = \{\lambda a_n\}$$

With the above operations V forms a vector space over $\mathbb{K}$.

Example 2.8 $V = \{0\}$ over the field $\mathbb{K}$ is a vector space called the *zero space*.

Now, we will establish some of the basic properties of vector spaces.

Theorem 2.1 *Let V be a vector space over a field $\mathbb{K}$. Then the following statements are true.*

(a) $0v = \mathbf{0}$ *for each* $v \in V$.
(b) $\lambda\mathbf{0} = \mathbf{0}$ *for each* $\lambda \in \mathbb{K}$.
(c) *For* $v \in V$ *and* $\lambda \in \mathbb{K}$, *if* $\lambda v = \mathbf{0}$, *then either* $\lambda = 0$ *or* $v = \mathbf{0}$.
(d) *If* v_1, v_2, *and* v_3 *are vectors in a vector space* V *such that* $v_1 + v_3 = v_2 + v_3$, *then* $v_2 = v_3$.
(e) $(-\lambda)v = -(\lambda v) = \lambda(-v)$ *for each* $\lambda \in \mathbb{K}$ *and each* $v \in V$.

Proof (a) For $v \in V$, by $(V2)$, $0v \in V$. By $(V5)$, for $0v \in V$ there exists $(-0v)$ such that $0v + (-0v) = \mathbf{0}$. And by using $(V8)$,

$$0v = (0+0)v = 0v + 0v \Rightarrow 0v = \mathbf{0}$$

(b) For $\lambda \in \mathbb{K}$ by $(V2)$ $\lambda\mathbf{0} \in V$. By $(V5)$, for $\lambda\mathbf{0} \in V$ there exists $(-\lambda\mathbf{0})$ such that $\lambda\mathbf{0} + (-\lambda\mathbf{0}) = \mathbf{0}$. And by using $(V7)$,

$$\lambda\mathbf{0} = \lambda(\mathbf{0}+\mathbf{0}) = \lambda\mathbf{0} + \lambda\mathbf{0} \Rightarrow \lambda\mathbf{0} = \mathbf{0}$$

(c) Let $\lambda v = \mathbf{0}$. From (1), if $\lambda = 0$, then $\lambda v = \mathbf{0}$. Now suppose that $\lambda \neq 0$, then there exists $\frac{1}{\lambda} \in \mathbb{K}$ and $\frac{1}{\lambda}(\lambda v) = \frac{1}{\lambda}\mathbf{0} \Rightarrow v = \mathbf{0}$.

(d) Suppose that $v_1, v_2, v_3 \in V$ be such that $v_1 + v_3 = v_2 + v_3$. Since $v_3 \in V$, by $(V5)$ there exists $-v_3 \in V$ such that $v_3 + (-v_3) = \mathbf{0}$. Then

$$\begin{aligned} v_1 + v_3 = v_2 + v_3 &\Rightarrow (v_1 + v_3) + (-v_3) = (v_2 + v_3) + (-v_3) \\ &\Rightarrow v_1 + (v_3 + (-v_3)) = v_2 + (v_3 + (-v_3)) \text{ (using } (V3)) \\ &\Rightarrow v_1 = v_2 \text{ (using } (V6)) \end{aligned}$$

(e) By $(V5)$, we have $\lambda v + (-(\lambda v)) = \mathbf{0}$. Also $\lambda v + (-\lambda)v = (\lambda + (-\lambda))v = 0$. By the uniqueness of additive inverse, this implies that $(-\lambda)v = -(\lambda v)$. In particular, $(-1)v = -v$. Now by $(V9)$,

$$\lambda(-v) = \lambda[(-1)v] = [\lambda(-1)]v = (-\lambda)v$$

From the next section, we will use 0 for zero vector, instead of $\mathbf{0}$.

2.2 Subspaces

For vector spaces, there may exist subsets which themselves are vector spaces under the same operations as defined in the parent space. Such subsets of a vector space are called subspaces. We will define the subspace of a vector space as follows.

Definition 2.2 (*Subspace*) A subset W of a vector space V over a field $\mathbb{K}$ is called a *subspace* of V if W is a vector space over $\mathbb{K}$ with the operations of addition and scalar multiplication defined on V.

If V is a vector space, then V and $\{0\}$ are subspaces of V called *trivial subspaces*. The latter is also called the *zero subspace* of V. A subspace W of V is called a proper subspace if $V \neq W$. Otherwise it is called an improper subspace (if it exists). Can you find any subspaces for the vector space $\mathbb{R}$ over $\mathbb{R}$ other than $\mathbb{R}$ and $\{0\}$? By definition, a subspace is a vector space in its own right. To check whether a subset is a subspace, we don't have to verify all the conditions $(V1) - (V10)$. The following theorem gives the set of conditions that are to be verified.

Theorem 2.2 *Let V be a vector space over a field $\mathbb{K}$. A subset W of V is a subspace if and only if the following three conditions hold for the operations defined in V.*

(a) $0 \in W$.
(b) $w_1 + w_2 \in W$ *whenever* $w_1, w_2 \in W$.
(c) $\lambda w \in W$ *whenever* $\lambda \in \mathbb{K}$ *and* $w \in W$.

Proof Suppose that W is a subspace of V. Then W is a vector space with the operation addition and scalar multiplication defined on V. Therefore (b) and (c) are satisfied. And by the uniqueness of identity element in a vector space $0 \in W$.

Conversely suppose that the conditions (a), (b), and (c) are satisfied. We have to show that W is a vector space with the operations defined on V. Since W is a subset of the vector space V, the conditions $(V3)$, $(V5) - (V10)$ are automatically satisfied by the elements in W. Therefore W is a subspace of V.

Certainly, we can observe that Condition (a) in the above theorem need not be checked separately, as it can be obtained from Condition (c) with $\lambda = 0$. But Condition (a) can be used to identify subsets which are not subspaces as shown in the following example.

Example 2.9 Let $V = \mathbb{R}^2 = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\}$. We have seen that $\mathbb{R}^2$ is a vector space over $\mathbb{R}$. Consider $W_1 = \{(x_1, x_2) \mid x_1 + x_2 = 0\}$ and $W_2 = \{(x_1, x_2) \mid x_1 + x_2 = 1\}$. Then W_1 is a subspace of V. For,

(a) Clearly, the additive identity $(0, 0)$ is in W_1.
(b) Take two elements $(x_1, x_2), (y_1, y_2) \in W_1$. Then $x_1 + x_2 = 0$ and $y_1 + y_2 = 0$. This implies that $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2) \in W_1$ as $x_1 + x_2 + y_1 + y_2 = 0$.
(c) Take $(x_1, x_2) \in W_1$ and $\lambda \in \mathbb{R}$. Then $x_1 + x_2 = 0$. This implies that $\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2) \in W_1$ as $\lambda x_1 + \lambda x_2 = \lambda(x_1 + x_2) = 0$.

But W_2 is not a subspace of $\mathbb{R}^2$ as zero vector does not belong to W_2. Now let us discuss the geometry of W_1 and W_2 a bit. W_1 and W_2 represent two lines on the plane as shown in the figure (Fig. 2.1).

Later, we will see that the only non-trivial proper subspaces of $\mathbb{R}^2$ are straight lines passing through origin.

Example 2.10 Let $V = \mathbb{M}_{n \times n}(\mathbb{K})$ and $W = \left\{A \in \mathbb{M}_{n \times n}(\mathbb{K}) \mid A^T = A\right\}$. That is, W is the set of all $n \times n$ symmetric matrices over $\mathbb{K}$. We will check whether the conditions in Theorem 2.2 are satisfied or not.

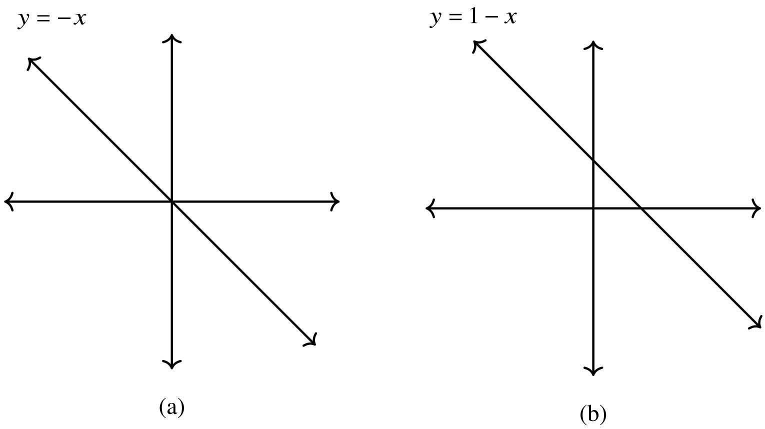


Fig. 2.1 Observe that W_1 depicted in (**a**) (straight line passing through origin) is a subspace and W_2 depicted in (**b**) (straight line not passing through origin) is not a subspace

(a) The zero matrix is equal to its transpose and hence belongs to W.
(b) By the properties of symmetric matrices, the sum of two symmetric matrices is again a symmetric matrix. That is, $A + B \in W$ whenever $A, B \in W$.
(c) Also $\lambda A \in W$ whenever $A \in W$ and $\lambda \in \mathbb{K}$, since $(\lambda A)^T = \lambda A^T = \lambda A$ as $A^T = A$.

Therefore, the set of all $n \times n$ symmetric matrices over $\mathbb{K}$ is a subspace of $\mathbb{M}_{n\times n}(\mathbb{K})$. What about the set of all $n \times n$ skew-symmetric matrices over $\mathbb{K}$?

Example 2.11 Let $V = \mathbb{P}_2[a, b]$. Consider $W = \{p \in \mathbb{P}_2[a, b] \mid p(0) = 0\}$.

(a) Since $p(0) = 0$ for zero polynomial, zero polynomial belongs to W.
(b) Take $p, q \in W$, then $p(0) = q(0) = 0$ and hence $(p + q)(0) = p(0) + q(0) = 0$. Thus $p + q \in W$ whenever $p, q \in W$.
(c) Let $p \in W$ and $\lambda \in \mathbb{R}$, then $(\lambda p)(0) = \lambda p(0) = 0$. That is, $\lambda p \in W$ whenever $p \in W$ and $\lambda \in \mathbb{R}$.

Therefore $\{p \in \mathbb{P}_2[a, b] \mid p(0) = 0\}$ is a subspace of $\mathbb{P}_2[a, b]$. Now, consider the subset $\tilde{W} = \{p \in \mathbb{P}_2[a, b] \mid p(0) = 1\}$. Is it a subspace of $\mathbb{P}_2[a, b]$? It is not!! (Why?)

Remark 2.1 To check whether a subset of a vector space is a subspace, we verify only the closure properties of vector addition and scalar multiplication in the given set. Therefore Theorem 2.2 can also be stated as follows:

- A subset W of a vector space V is a subspace of V if and only if $\lambda w_1 + \mu w_2 \in W$, whenever $w_1, w_2 \in W$ and $\lambda, \mu \in \mathbb{K}$
- A subset W of a vector space V is a subspace of V if and only if $\lambda w_1 + w_2 \in W$, whenever $w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$.

Example 2.12 In the previous chapter, we have seen that the collection of all solutions to the system $Ax = 0$ satisfies the conditions in Remark 2.1 where $A \in \mathbb{M}_{m\times n}(\mathbb{K})$ and hence they form a subspace of $\mathbb{K}^n$. That is, the solutions of a homogeneous system form a vector space under the operations defined on $\mathbb{K}^n$. But the solutions of a non-homogeneous system does not form a vector space as zero vector is never a solution for a non-homogeneous system.

The next theorem gives a method to construct new subspaces from known subspaces.

Theorem 2.3 *Let W_1 and W_2 be two subspaces of a vector space V over a field $\mathbb{K}$, then their intersection $W_1 \cap W_2 = \{w \mid w \in W_1$ and $w \in W_2\}$ is a subspace of V.*

Proof Since W_1 and W_2 are subspaces of V, $0 \in W_1$ and $0 \in W_2$. Therefore $0 \in W_1 \cap W_2$. Now let $v, w \in W_1 \cap W_2$, then

$$
\begin{aligned}
v, w \in W_1 \cap W_2 &\Rightarrow v, w \in W_1 \text{ and } v, w \in W_2\\
&\Rightarrow v + w \in W_1 \text{ and } v + w \in W_2 \text{ as } W_1 \text{ and } W_2 \text{ are subspaces}\\
&\Rightarrow v + w \in W_1 \cap W_2
\end{aligned}
$$

For $\lambda \in \mathbb{K}$ and $w \in W_1 \cap W_2$,

$$
\begin{aligned}
w \in W_1 \cap W_2 &\Rightarrow w \in W_1 \text{ and } w \in W_2\\
&\Rightarrow \lambda w \in W_1 \text{ and } \lambda w \in W_2 \text{ as } W_1 \text{ and } W_2 \text{ are subspaces}\\
&\Rightarrow \lambda w \in W_1 \cap W_2
\end{aligned}
$$

Therefore $W_1 \cap W_2$ is a subspace of V.

The above result can be extended to any number of subspaces. As we have shown that the intersection of subspaces is again a subspace, it is natural to ask whether the union of subspaces is again a subspace. It is clear that the union of two subspaces need not be a subspace of V (Fig. 2.2).

The following theorem gives a scenario in which union of two subspaces of a vector space is again a subspace of the same.

Theorem 2.4 *Let V be a vector space over the field $\mathbb{K}$ and let W_1 and W_2 be subspaces of V. Then $W_1 \cup W_2$ is a subspace of V if and only if either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.*

Proof Let W_1 and W_2 be subspaces of V. Suppose that either $W_2 \subseteq W_1$ or $W_1 \subseteq W_2$. Then $W_1 \cup W_2$ is either W_1 or W_2. In either cases, $W_1 \cup W_2$ is a subspace of V. Conversely, suppose that $W_1 \cup W_2$ is a subspace of V, $W_1 \nsubseteq W_2$ and $W_2 \nsubseteq W_1$. Then there exists at least one element $w_1 \in W_1$ such that $w_1 \notin W_2$ and $w_2 \in W_2$ such that $w_2 \notin W_1$. As $W_1, W_2 \subseteq W_1 \cup W_2$ both $w_1, w_2 \in W_1 \cup W_2$. Since $W_1 \cup W_2$ is a subspace of V, $w_1 + w_2 \in W_1 \cup W_2$. Then either $w_1 + w_2 \in W_1$ or $w_1 + w_2 \in W_2$.

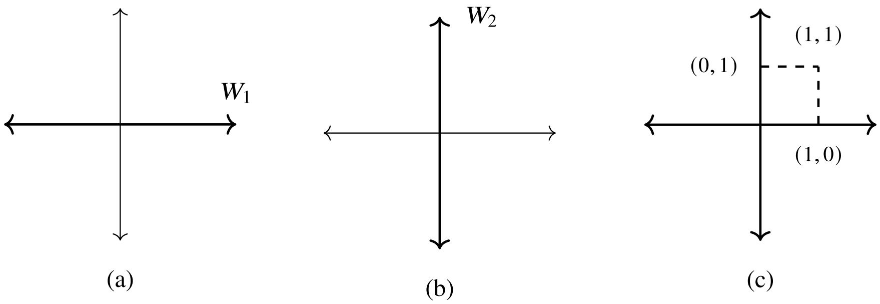


Fig. 2.2 Consider $V = \mathbb{R}^2$, take $W_1 = x - axis$ and $W_2 = y - axis$ (depicted as (**a**) and (**b**) respectively). Then W_1 and W_2 are subspaces of V but $W_1 \cup W_2$ is not a subspace of V, since $(1, 0) \in W_1,\ (0, 1) \in W_2$, but $(1, 0) + (0, 1) = (1, 1) \notin W_1 \cup W_2$, as we can observe from (**c**)

Suppose $w_1 + w_2 \in W_1$. Since $w_1 \in W_1$ and W_1 is a subspace, $-w_1 \in W_1$ and hence $(-w_1) + w_1 + w_2 = (-w_1 + w_1) + w_2 = w_2 \in W_1$ which is a contradiction. Now suppose $w_1 + w_2 \in W_2$. Since $w_2 \in W_2$ and W_2 is a subspace, $-w_2 \in W_2$ and hence $w_1 + w_2 + (-w_2) = w_1 + (w_2 - w_2) = w_1 \in W_2$ which is again a contradiction. Therefore our assumption is wrong. That is, $W_1 \cup W_2$ is a subspace of V if and only if either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Example 2.13 Let V be the vector space $\mathbb{R}^3$ over $\mathbb{R}$. Consider $W_1 = \left\{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\right\}$ and $W_2 = \{(0, x_2, 0) \mid x_2 \in \mathbb{R}\}$. Clearly, $W_1 \cup W_2 = W_1$ is a subspace. Observe that $W_2 \subset W_1$.

2.3 Linear Dependence and Independence

Let V be a vector space over a field $\mathbb{K}$. Let $v_1, v_2, \ldots, v_n \in V$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$. Then the vector

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

is called a *linear combination* of the vectors and the scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ are called the *coefficients* of the linear combination. If all the coefficients are zero, then $v = 0$, which is trivial. Now suppose that there exists a non-trivial representation for 0, that is, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ not all zero such that a linear combination of the given vectors equals zero . Then we say that the vectors $v_1, v_2, \ldots, v_n$ are *linearly dependent*. In other words, the vectors $v_1, v_2, \ldots, v_n$ are linearly dependent if and only if there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ not all zero such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

The vectors $v_1, v_2, \ldots, v_n$ are *linearly independent* if they are not linearly dependent. That is,

$$\text{if } \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0, \text{ then } \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$$

Clearly, any subset of a vector space V containing zero vector is linearly dependent as 0 can be written as $0 = 1.0$. Since $\lambda v = 0$ implies either $\lambda = 0$ or $v = 0$, any singleton subset of V containing a non-zero vector is linearly independent.

Example 2.14 Consider the vector space $V=\mathbb{R}^2$ and the subset $S_1 = \{(1, 0), (1, 1)\}$. To check whether S_1 is linearly dependent or not, consider a linear combination of vectors in S_1 equals zero for some scalars λ_1 and λ_2. Then

$$\begin{aligned}\lambda_1(1, 0) + \lambda_2(1, 1) = (0, 0) &\Rightarrow (\lambda_1 + \lambda_2, \lambda_2) = (0, 0)\\ &\Rightarrow \lambda_1 + \lambda_2 = 0, \lambda_2 = 0\\ &\Rightarrow \lambda_1 = 0, \lambda_2 = 0\end{aligned}$$

That is, there does not exist non-trivial representation for zero vector in $\mathbb{R}^2$ using vectors of S_1. Thus S_1 is linearly independent. Note that $(1, 0)$ cannot be obtained by scaling $(1, 1)$ or vice verse.

Now consider a subset $S_2 = \{(1, 0), (2, 0)\}$ of $\mathbb{R}^2$ and a linear combination of the vectors in S_2 equals zero. Then

$$\begin{aligned}\lambda_1(1, 0) + \lambda_2(2, 0) = (0, 0) &\Rightarrow (\lambda_1 + \lambda_2, \lambda_2) = (0, 0)\\ &\Rightarrow \lambda_1 + 2\lambda_2 = 0\end{aligned}$$

Then there are infinitely many possibilities for λ_1 and λ_2. For example, $\lambda_1 = 2$ and $\lambda_2 = -1$ is one such possibility. Clearly, $2(1, 0) + (-1)(2, 0) = (0, 0)$. Thus the zero vector in $\mathbb{R}^2$ has a non-trivial representation using the vectors of S_2. Thus S_2 is linearly dependent. Note that $(2, 2) = 2(1, 1)$ is a scaled version of $(1, 1)$ (Fig. 2.3).

Using the above geometrical idea, try to characterize the linearly independent sets in $\mathbb{R}$ and $\mathbb{R}^2$. Also observe that the equation, $\lambda_1(1, 0) + \lambda_2(1, 1) = (\lambda_1 + \lambda_2, \lambda_2) = (0, 0)$ formed by vectors in S_1, from the above example, can be written in the form of a system of homogeneous equation as $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. We have seen in Section 1.7 that a $n \times n$ homogeneous system $Ax = 0$ has a non-trivial solution when the coefficient matrix A has rank less than n. In this case, $rank\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right) = 2$. Therefore the system does not have a non-trivial solution. That is, $\lambda_1 = \lambda_2 = 0$. Now, for vectors in S_2, observe that the coefficient matrix $A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ has rank 1, which implies that there exists a non-trivial representation for the zero vector. Using this idea, can we say something about the linear dependency/independency of a collection of vectors in $\mathbb{R}^2$? Is it possible to generalize this idea to $\mathbb{R}^n$? Think!!!

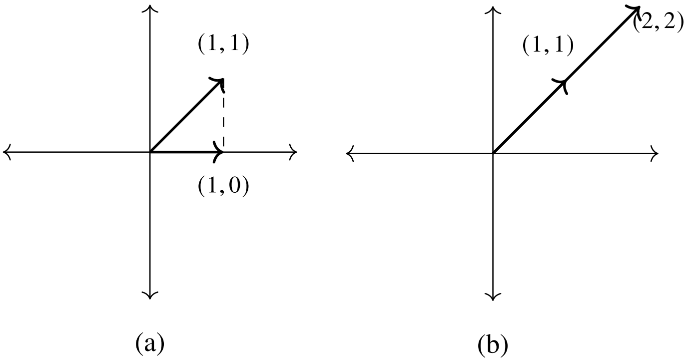


Fig. 2.3 Examples for **a** linearly independent vectors in $\mathbb{R}^2$ and **b** linearly dependent vectors in $\mathbb{R}^2$. Observe that the linearly independent vectors lie on two distinct straight lines passing through origin and the linearly dependent vectors lie on the same line passing through origin. We will soon prove that a set of two vectors is linearly dependent if and only if one vector is a scalar multiple of the other

Remark 2.2 We can say that the number of linearly independent vectors in a collection S of m vectors of $\mathbb{K}^n$ is the rank of the $n \times m$ matrix A formed by the vectors in S as columns. As the rank of a matrix and its transpose is the same, we may redefine the rank of a matrix as the number of linearly independent rows or columns of that matrix.

Example 2.15 Consider the vector space $V = \mathbb{P}_2[a, b]$ and the subset $S_1=\{1, x, x^2\}$. Now, for $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{K}$,

$$\lambda_1.1 + \lambda_2 x + \lambda_3 x^2 = 0 \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$$

Thus S_1 is linearly independent.

Now consider the subset $S_2 = \{1 - x, 1 + x^2, 3 - 2x + x^2\}$ of $\mathbb{P}_2[a, b]$. As

$$2(1 - x) + 1(1 + x^2) = 3 - 2x + x^2$$

S_2 is linearly dependent.

As we have seen in the previous example, consider the matrices $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 1 & 3 \\ -1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}$. Is there any relation between the rank of these matrices and the linear dependency/independency of vectors in S_1 and S_2 given in Example 2.15?

The following results are some of the important consequences of definitions of linear dependence and independence.

Theorem 2.5 *Let V be a vector space over a field $\mathbb{K}$ and $W = \{w_1, w_2, \ldots, w_n\}$ be a subset of V, where $n \geq 2$. Then W is linearly dependent if and only if at least one vector in W can be written as a linear combination of the remaining vectors in W.*

Proof Suppose that W is linearly dependent. Then there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, not all zero such that

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n = 0$$

Without loss of generality, assume that $\lambda_1 \neq 0$. Then since $\lambda_1 \in \mathbb{K}$, $\frac{1}{\lambda_1} \in \mathbb{K}$ and hence

$$w_1 = -\frac{\lambda_2}{\lambda_1} w_2 - \frac{\lambda_3}{\lambda_1} w_3 - \cdots - \frac{\lambda_n}{\lambda_1} w_n$$

Conversely suppose that one vector in W can be written as a linear combination of the remaining vectors in W. Without loss of generality, take $w_1 = \lambda_2 w_2 + \cdots + \lambda_n w_n$. Then $w_1 - \lambda_2 w_2 + \cdots + \lambda_n w_n = 0$. That is, there exists a non-trivial representation for zero. Therefore W is linearly dependent.

Corollary 2.1 *A subset of a vector space V containing two non-zero vectors is linearly dependent if and only if one vector is a scalar multiple of the other.*

Proof Suppose that $\{v_1, v_2\} \subseteq V$ be linearly dependent. Then there exists scalars $\lambda_1, \lambda_2 \in \mathbb{K}$ not both zero such that $\lambda_1 v_1 + \lambda_2 v_2 = 0$. Without loss of generality, let $\lambda_1 \neq 0$. Then $v_1 = -\frac{\lambda_2}{\lambda_1} v_2$. The converse part is trivial.

Theorem 2.6 *Let V be a vector space over a field $\mathbb{K}$, and let $W_1 \subseteq W_2 \subseteq V$. If W_1 is linearly dependent, then W_2 is linearly dependent and if W_2 is linearly independent, then W_1 is linearly independent.*

Proof Suppose that W_1 is linearly dependent and $W_1 \subseteq W_2$. Then there exists $v_1, v_2, \ldots, v_n \in W_1$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, not all 0 such that $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$. Since $W_1 \subseteq W_2$, W_2 is linearly dependent.

Now suppose that W_2 is linearly independent. Then from above W_1 is linearly independent. For if W_1 is linearly dependent, then W_2 must be linearly dependent.

Thus we can say that any super set of a linearly dependent set is linearly dependent and any subset of a linearly independent set is linearly independent.

2.4 Basis and Dimension

In this section, we will study the basic building blocks of vector spaces known as basis. A basis of a vector space is a subset of the vector space which can be used to uniquely represent each vector in the given space. We will start by the following definition.

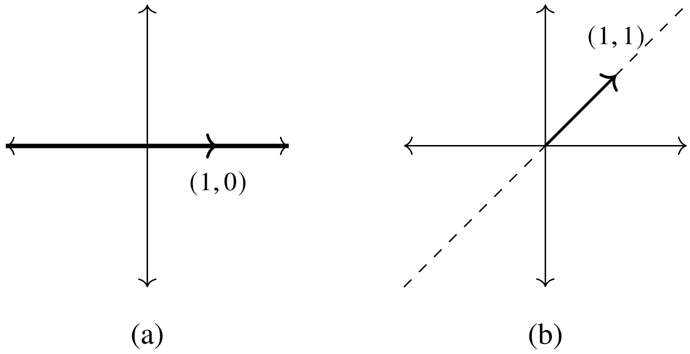


Fig. 2.4 Observe that both $Span(S_1)$ and $Span(S_2)$ are straight lines passing through origin

Definition 2.3 (*Span of a set*) Let $S = \{v_1, v_2, \ldots, v_n\}$ be a subset of a vector space V. Then the span of S, denoted by $span(S)$, is the set consisting of all linear combinations of the vectors in S. That is,

$$span(S) = \{\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \mid \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}\}$$

For convenience, we define $span\{\phi\} = \{0\}$. A subset S of a vector space V spans (or generates) V if $span(S) = V$. If there exists a finite subset S of V such that $span(S) = V$, then V is called finite-dimensional vector space. Otherwise it is called infinite-dimensional vector space.

Example 2.16 Consider $S_1 = \{(1, 0)\}$ and $S_2 = \{(1, 1)\}$ in $\mathbb{R}^2$. Then (Fig. 2.4)

$$Span(S_1) = \{\lambda(1, 0) \mid \lambda \in \mathbb{R}\} = \{(\lambda, 0) \mid \lambda \in \mathbb{R}\} = x - \text{axis}$$

and

$$Span(S_2) = \{\lambda(1, 1) \mid \lambda \in \mathbb{R}\} = \{(\lambda, \lambda) \mid \lambda \in \mathbb{R}\}$$

In fact, span of any non-zero vector of the form $(x_1, 0)$ in $\mathbb{R}^2$ will be the $x-$axis and span of any non-zero vector of the form (x_1, x_1) in $\mathbb{R}^2$ will be the line $y = x$. In general, we can say that span of any single non-zero vector in $\mathbb{R}^2$ will be a straight line passing through that vector and the origin. This can be generalized to $\mathbb{R}^n$ also. Now consider the set $S_3 = \{(1, 0, 0), (0, 1, 0)\}$ in $\mathbb{R}^3$. Then (Fig. 2.5)

$$\begin{aligned} Span(S_3) &= \{\lambda_1(1, 0, 0) + \lambda_2(0, 1, 0) \mid \lambda_1, \lambda_2 \in \mathbb{R}\} \\ &= \{(\lambda_1, \lambda_2, 0) \mid \lambda_1, \lambda_2 \in \mathbb{R}\} = x - y \text{ plane} \end{aligned}$$

Theorem 2.7 *Let V be a vector space over a field $\mathbb{K}$. Let $S = \{v_1, v_2, \ldots, v_n\}$ be a subset of V, then $span(S)$ is a subspace of V and any subspace of V that contains S must also contain span (S).*

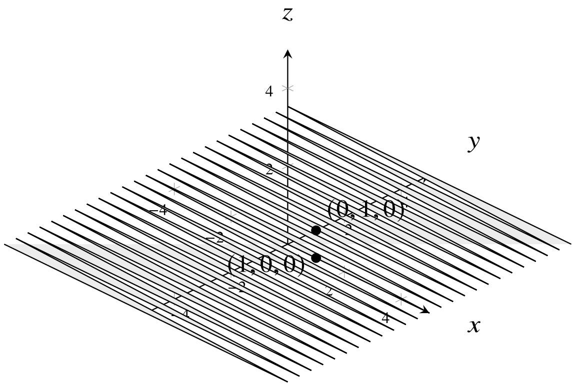


Fig. 2.5 The span of $S_3 = \{(1, 0, 0), (0, 1, 0)\}$ in $\mathbb{R}^3$ is the entire $x - y$ plane

Proof Clearly, $0 = 0v_1 + 0v_2 + \cdots + 0v_n \in span(S)$. Let $v, w \in span(S)$. Then there exists $\lambda_1, \lambda_2, \ldots, \lambda_n, \mu_1, \mu_2, \ldots, \mu_n \in \mathbb{K}$ such that $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ and $v = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n$. Then

$$u + v = (\lambda_1 + \mu_1)v_1 + (\lambda_2 + \mu_2)v_2 + \cdots + (\lambda_n + \mu_n)v_n \in span(S)$$

and for $\mu \in \mathbb{K}$,

$$\mu v = \mu(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = (\mu\lambda_1)v_1 + (\mu\lambda_2)v_2 + \cdots + (\mu\lambda_n)v_n \in span(S)$$

Therefore $span(S)$ is a subspace of V. Now let W be any subspace of V containing $S = \{v_1, v_2, \ldots, v_n\}$. Then for any scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$, as W is a subspace of V, $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in W$. That is, $span(S) \subseteq W$.

Remark 2.3 Consider a matrix $A \in \mathbb{M}_{m\times n}(\mathbb{K})$. We can view each row(column) as a vector in $\mathbb{K}^n(\mathbb{K}^m)$. The span of the row vectors of A is called *row space* of A and the span of the column vectors of A is called *column space* of A.

Definition 2.4 (*Basis*) Let V be a vector space over a field $\mathbb{K}$. If a set $B \subseteq V$ is linearly independent and $span(B) = V$, then B is called a basis for V. If the basis has some specific order, then it is called an ordered basis.

Theorem 2.8 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and S=$\{v_1, v_2, \ldots, v_n\}$ spans V. Then S can be reduced to a basis B of V.*

Proof Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and $S = \{v_1, v_2, \ldots, v_n\}$ spans V. Let $S_\sigma = \{v_{\sigma_1}, v_{\sigma_2}, \ldots, v_{\sigma_k}\}$ denote the set of all non-zero elements of S. Now, we will construct a linearly independent set B from S, with $span(B) = S$. Pick the element $v_{\sigma_1} \in S_\sigma$ to B. If $v_{\sigma_2} = \lambda v_{\sigma_1}$, for some $\lambda \in \mathbb{K}$, then $v_{\sigma_2} \notin B$, otherwise $v_{\sigma_2} \in B$. Now consider $v_{\sigma_3} \in S_\sigma$. If $v_{\sigma_3} = \lambda_1 v_{\sigma_1} + \lambda_2 v_{\sigma_2}$ for some $\lambda_1, \lambda_2 \in \mathbb{K}$, then $v_{\sigma_3} \notin B$, otherwise $v_{\sigma_3} \in B$. Proceeding like this, after σ_k steps we will get a linearly independent set with $span(B) = V$.

Corollary 2.2 *Every finite-dimensional vector space V has a basis.*

Proof Let V be a finite-dimensional vector space. Then there exists a finite subset S of V with $span(S) = V$. Then by S can be reduced to a basis.

Example 2.17 Consider the set

$$B = \{e_1 = (1, 0, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, 0, 0, \dots, 1)\}$$

in $\mathbb{K}^n$ over $\mathbb{K}$. We will show that B is a basis for $\mathbb{K}^n$. Let us consider an element $a = (a_1, a_2, \dots, a_n) \in \mathbb{K}^n$ arbitrarily, then we have $a = a_1e_1 + a_2e_2 + \cdots + a_ne_n$. That is, every element in $\mathbb{K}^n$ can be written as a linear combination of elements in B with coefficients from $\mathbb{K}$. Thus B spans $\mathbb{K}^n$ over $\mathbb{K}$. Also

$$a_1e_1 + a_2e_2 + \cdots + a_ne_n = 0 \Rightarrow a_1 = a_2 = \cdots = a_n = 0$$

That is, B is linearly independent. Therefore B is a basis for $\mathbb{K}^n$ over $\mathbb{K}$ and is called the standard ordered basis for $\mathbb{K}^n$ over $\mathbb{K}$.

Example 2.18 Consider the set $B = \left\{E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ in $\mathbb{M}_{2\times2}(\mathbb{K})$ over the field $\mathbb{K}$. Consider an element $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{M}_{2\times2}(\mathbb{K})$. Then

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

That is, B spans $\mathbb{M}_{2\times2}(\mathbb{K})$ over the field $\mathbb{K}$. Also

$$\lambda_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \lambda_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \lambda_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \lambda_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$\Rightarrow \begin{bmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. That is, B is linearly independent. Therefore B is a basis for $\mathbb{M}_{2\times2}(\mathbb{K})$ over $\mathbb{K}$.

Example 2.19 Consider the set $B = \{1, x, \dots, x^n\}$ in $\mathbb{P}_n[a, b]$ over $\mathbb{R}$. Then B is linearly independent as

$$\lambda_1.0 + \lambda_1 x + \cdots + \lambda_n x^n = 0 \Rightarrow \lambda_0 = \lambda_1 = \cdots = \lambda_n = 0$$

and clearly B spans $\mathbb{P}_n[a, b]$. Therefore B is a basis for $\mathbb{P}_n[a, b]$ over $\mathbb{R}$

Example 2.20 Now consider a subset $S = \{(1, 1, 2), (2, 1, 1), (3, 2, 3), (-1, 0, 1)\}$ of $\mathbb{R}^3$ over $\mathbb{R}$. We know that $span(S)$ is a subspace of $\mathbb{R}^3$. Can you find a basis for $span(S)$? To find a basis for $span(S)$, we have to find a linearly independent subset $\tilde{S}$ of $\mathbb{R}^3$ such that $span(S) = span(\tilde{S})$. We may observe that the $span(S)$ is the same as the row space of the matrix $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \\ 3 & 2 & 3 \\ -1 & 0 & 1 \end{bmatrix}$. Thus to find a basis for $span(S)$, it is enough to find the linearly independent rows of A. We can reduce A to the row reduced form as $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & -1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. From this we can say that the set $\tilde{S} = \{(1, 1, 2), (0, -1, -3)\}$ forms a basis for $span(S)$.

Theorem 2.9 *Let V be a vector space over a field $\mathbb{K}$. If $B = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then any $v \in V$ can be uniquely expressed as a linear combination of vectors in B.*

Proof Let B be a basis of V and $v \in V$. Suppose that v can be expressed as a linear combination of vectors in B as

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

and as

$$v = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n$$

where $\lambda_i, \mu_i \in \mathbb{K}$ for all $i = 1, 2, \ldots, n$. Subtracting the second expression from first, we get

$$0 = (\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 + \cdots + (\lambda_n - \mu_n)v_n$$

Since B is linearly independent, this implies that $\lambda_i - \mu_i = 0$ for all $i = 1, 2, \ldots, n$. That is, $\lambda_i = \mu_i$ for all $i = 1, 2, \ldots, n$.

Theorem 2.10 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and B be a basis of V. Then basis is a minimal spanning set in V. That is, if B is a basis of V, there does not exist a proper subset of B that spans V.*

Proof Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and $B = \left\{v_1, v_2, \ldots, v_n\right\}$ be a basis of V. Let S be a proper subset of B that spans V. Since $S \subset B$ and $S \neq B$, there exists at least one element v such that $v \in B$ and $v \notin S$. Rearrange the elements of B so that the first k elements are also elements of S and the remaining $n - k$ elements belong to B only. Now take any element $v_{k+i} \in B$ where $i \in \{1, 2, \ldots n - k\}$. Since $span(S) = V$ and $v_{k+i} \in V$, there exists $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ such that $v_{k+i} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k$. This can also be written as $v_{k+i} = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k + 0v_{k+1} + \cdots + 0v_n$. Also as $v_{k+i} \in B$, v_{k+i} can be represented as a linear combination of elements of B by taking 1 as the coefficient to v_{k+i} and 0 as the coefficient for all elements in B other than v_{k+i}. This is a contradiction to the fact that representation for any element with respect to a basis must be unique.

Theorem 2.11 *Let V be a finite-dimensional vector space and S be a minimal spanning set of V, then S is a basis.*

Proof Let $S = \{v_1, v_2, \ldots, v_n\}$ be a minimal spanning set of V. To prove that S is a basis, it is enough to show that S is linearly independent. Suppose that it is linearly dependent, then by Theorem 2.5, at least one element say $v_i \in S$ can be written as a linear combination of the remaining vectors. Then $S \setminus \{v_i\}$ is a spanning set for V. This is a contradiction to the fact that S is a minimal spanning set.

Theorem 2.12 *Let V be a vector space over a field $\mathbb{K}$ and $B = \{v_1, v_2, \ldots, v_n\}$ be a basis of V. Let $W = \{w_1, w_2, \ldots, w_m\}$ be a linearly independent set in V, then $m \leq n$.*

Proof Since $B = \{v_1, v_2, \ldots, v_n\}$ is a basis of V, B spans V and B is linearly independent. Since $w_1 \in V$, by the previous theorem w_1 has a unique representation using the vectors in B, say

$$w_1 = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \tag{2.1}$$

Now we can express one of the v_i, say v_k, in terms of w_1 and the remaining v_i's. That is,

$$v_k = \mu w_1 + \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_n v_n \tag{2.2}$$

where $\mu = \frac{-1}{\lambda_k}$ and $\mu_j = \frac{-\lambda_j}{\lambda_k}$, $j \neq k$.

Now we will show that the set $B_1 = \{w_1, v_1, v_2, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n\}$ obtained by replacing v_k by w_1 is a basis for V. That is, we will prove that B_1 is linearly independent and B_1 spans V. Suppose that they are linearly dependent. Then by Theorem 2.5 at least one of the vectors in B_1 can be written as a linear combination of the remaining vectors. Since (2.1) is the unique representation for w_1, we cannot express w_1 in terms of $v_1, v_2, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n$. Therefore some $v_l \in B_1$ can be written as a linear combination of the remaining vectors in B_1. That is, there exist scalars $\alpha, \alpha_1, \ldots, \alpha_{l-1}, \alpha_{l+1}, \ldots, \alpha_{k-1}, \alpha_{k+1}, \ldots, \alpha_n \in \mathbb{K}$ such that

$$v_l = \alpha w_1 + \alpha_1 v_1 + \cdots + \alpha_{l-1} v_{l-1} + \alpha_{l+1} v_{l+1} + \cdots + \alpha_{k-1} v_{k-1} + \alpha_{k+1} v_{k+1} + \cdots + \alpha_n v_n$$

Now substituting (2.1) in the above equation we get that v_l can be expressed as a linear combination of vectors in B, which is a contradiction as B is linearly independent. Therefore B_1 is linearly independent. Since v_k can be expressed as in (2.2), $span(B_1) = span(B) = V$. Therefore B_1 is a basis of V. We repeat this process by replacing some $v_j \in B_1$, by w_2, and so on.

Now if $m \le n$, $B_m = \{w_1, w_2, \ldots, w_m, v_{i_1}, v_{i_2}, \ldots, v_{i_{m-n}}\}$ is a basis for V. If $m > n$, $B_n = \{w_1, w_2, \ldots, w_n\}$ is a basis for V. Then $w_{n+1} \in W$ can be written as a linear combination of vectors in B_n, which is a contradiction to the fact that W is linearly independent. Therefore $m \le n$.

Basis of a vector space is not unique. For example, consider $\mathbb{R}^2$. Clearly $B_1 = \{(1, 0), (0, 1)\}$ is a basis for $\mathbb{R}^2$ as any vector $(x_1, x_2) \in \mathbb{R}^2$ can be written as $(x_1, x_2) = x_1(1, 0) + x_2(0, 1), x_1, x_2 \in \mathbb{R}$ and B_1 is linearly independent. Now consider the set $B_\lambda = \{(1, 0), (0, \lambda)\}$. Then B_λ is a basis for $\mathbb{R}^2$ for any $\lambda \neq 0 \in \mathbb{R}$ as any vector $(x_1, x_2) \in \mathbb{R}^2$ can be written as $(x_1, x_2) = x_1(1, 0) + \frac{x_2}{\lambda}(0, \lambda), x_1, x_2 \in \mathbb{R}$ and B_λ is linearly independent for any $\lambda \neq 0 \in \mathbb{R}$. The following corollary shows that any two bases for a vector space have the same cardinality.

Corollary 2.3 *For a finite-dimensional vector space V over $\mathbb{K}$, any two bases for V have the same cardinality.*

Proof Let $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_m\}$ be any two bases for V. Consider B_1 as a basis and B_2 as a linearly independent set, then by the above theorem, $m \le n$. Now consider B_2 as a basis and B_1 as a linearly independent set, then $n \le m$. Therefore $m = n$.

Corollary 2.4 *Let V be a vector space over a field $\mathbb{K}$ and B be a basis of V. Then basis is a maximal linearly independent set in V. That is, if B is a basis of V, there does not exist a linearly independent set S such that $B \subset S \subset V$.*

Proof Let B be a basis of V and S be a linearly independent set in V. By the Theorem 2.12, the cardinality of S is less than or equal to cardinality of B. Therefore there does not exist a linearly independent set S such that $B \subset S \subset V$.

In the above corollary, we have shown that every basis is a maximal linearly independent set. Now we will prove that the converse is also true.

Theorem 2.13 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Let $S = \{v_1, v_2, \ldots, v_n\}$ be a maximal linearly independent set in V, then S is a basis.*

Proof Let S be a maximal linearly independent set in V. To show that S is a basis, it is enough to prove that $span(S) = V$. Suppose that this is not true. Then there exists a non-zero vector $v \in V$ such that $v \notin span(S)$. Now consider the set $S_1 = S \cup \{v\}$. We will show that S_1 is linearly independent, which will be a contradiction to the fact that S is maximal. Now let $\lambda, \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ be such that $\lambda v + \lambda_1 v_1 + \lambda_1 v_2 + \cdots + \lambda_n v_n = 0$. If $\lambda = 0$, then as S is linearly independent $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. If $\lambda \neq 0$, as $v \notin span(S)$, the expression $\lambda v + \lambda_1 v_1 + \lambda_1 v_2 + \cdots + \lambda_n v_n = 0$ is not possible. Therefore S_1 is linearly independent.

Theorem 2.14 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and S be a linearly independent subset of V. Then S can be extended to a basis.*

Proof Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Let $B=\{v_1, v_2, \ldots, v_n\}$ be a basis of V. Let S be a linearly independent subset of V. Now $S \cup B$ is a spanning set of V. By Theorem 2.8, it can be reduced to a basis. If $|S| = n$, then by Theorem 2.12, S is a maximal linearly independent set and hence a basis. Suppose that $|S| < n$. Then take the vector $v_1 \in B$. If $v_1 \notin span(S)$, then $S_1 = S \cup \{v_1\}$ is a linearly independent set. If $|S_1| = n$, then as above S_1 is a basis. If $v_1 \in span(S)$, discard v_1. Then choose $v_2 \in V$ and proceed in the same way. By repeating this process, we obtain a basis for V which is an extension of S.

The following theorem summarizes the results from Theorems 2.10–2.14.

Theorem 2.15 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and $B = \{v_1, v_2, \ldots, v_n\}$. Then the following are equivalent:*

(a) B is a basis of V.
(b) B is a minimal spanning set.
(c) B is a maximal linearly independent set.

In Corollary 2.3, we have seen that any basis for a vector space has the same cardinality. Therefore, we can uniquely define a quantity to express the cardinality of a basis for a vector space.

Definition 2.5 (*Dimension*) Let V be a vector space over a field $\mathbb{K}$ and B be basis of V. The number of elements of B is called dimension of V. It is denoted by $dim(V)$. For convenience, the dimension of $\{0\}$ is defined as 0.

Example 2.21 From Example 2.17, it is easy to observe that $\mathbb{K}^n$ over $\mathbb{K}$ has dimension n.

Example 2.22 From Example 2.18, $\mathbb{M}_{2\times 2}(\mathbb{K})$ over $\mathbb{K}$ has dimension 4. In general, $\mathbb{M}_{n\times n}(\mathbb{K})$ over $\mathbb{K}$ has dimension n^2.

Example 2.23 From Example 2.19, $\mathbb{P}_n[a, b]$ over $\mathbb{R}$ has dimension $n + 1$.

What about the dimension of $\mathbb{P}[a, b]$? Does there exist a finite set which is linearly independent and spans $\mathbb{P}[a, b]$? If such a finite set does not exist, such vector spaces are called infinite-dimensional vector spaces. Can you give another example for an infinite-dimensional vector space? What about $C[a, b]$? Now, the following remark discusses some interesting facts about the importance of field $\mathbb{K}$, while considering a vector space $V(\mathbb{K})$.

Remark 2.4 One set can be a vector space over different fields and their dimension may vary with the field under consideration. For example $\mathbb{C}$ = the set of all complex numbers is a vector space over both the fields $\mathbb{R}$ and $\mathbb{C}$. Since every element $a + bi \in \mathbb{C}$ can be written as

$$a + bi = (a + bi)1$$

where $a + bi \in \mathbb{C}$ (field under consideration) and $1 \in \mathbb{C}$ (set under consideration), $\{1\}$ is a basis for $\mathbb{C}(\mathbb{C})$ and $dim_{\mathbb{C}}(\mathbb{C}) = 1$. If $\mathbb{R}$ is the field under consideration, then $a + bi \in \mathbb{C}$ can be written as

$$a + bi = a(1) + b(i)$$

where $a, b \in \mathbb{R}$ and $1, i \in \mathbb{C}$. Therefore $\{1, i\}$ is a basis for $\mathbb{C}$ $(\mathbb{R})$ and $dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Theorem 2.16 *Let V be a finite-dimensional vector space, then*

(a) Every spanning set of vectors in V with cardinality the same as that of $dim(V)$ is a basis of V.
(b) Every linearly independent set of vectors in V with cardinality the same as that of $dim(V)$ is a basis of V.

Proof (a) Let V be a finite-dimensional vector space with $dim(V) = n$. Then by Corollary 2.3, any basis of V have cardinality n. Let S be subset of V with $span(S) = V$ and $|S| = n$. By Theorem 2.8 any spanning set can be reduced to a basis. Therefore S is a basis for V.
(b) Let V be a finite-dimensional vector space with $dim(V) = n$. Let S be a linearly independent subset of V with $|S| = n$. By Theorem 2.14 any linearly independent set S can be extended to a basis. Therefore S is a basis for V.

Theorem 2.17 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Let W be a subspace V. Then W is finite-dimensional and $dim(W) \leq dim(V)$. Moreover, if $dim(W) = dim(V)$, then $V = W$.*

Proof Let W be a subspace of V. Then W is a vector space with the operations defined on B. Consider a basis B for W. Then B is a linearly independent set in V. Then by Theorem 2.12, $dim(W) \leq dim(V)$. If $dim(W) = dim(V)$, then by the previous theorem, B is a basis for V also and hence $V = W$.

Example 2.24 Consider the vector space $\mathbb{R}^2$ over $\mathbb{R}$. Let W be a subspace of $\mathbb{R}^2$. Since $dim(\mathbb{R}^2) = 2$ the only possible dimensions for W are 0, 1, and 2. If $dim(W) = 0$, then $W = \{0\}$ and if $dim(W) = 2$, then $W = \mathbb{R}^2$. Now let $dim(W) = 1$. Then W is spanned by some non-zero vector. Therefore W is given by $W = \{\lambda v \mid \lambda \in \mathbb{R}\}$ for some $v \neq 0 \in \mathbb{R}^2$. That is, W is a line passing through origin. Hence the only subspaces of $\mathbb{R}^2$ are the zero space, lines passing through origin, and $\mathbb{R}^2$ itself. Similarly, the only subspaces of $\mathbb{R}^3$ are the zero space, lines passing through origin, planes passing through origin, and $\mathbb{R}^3$ itself.

2.5 Sum and Direct Sum

In the previous section, we have seen that the union of two subspaces need not necessarily be a subspace. Therefore analogous to union of subsets in set theory, we define a new concept called the sum of subspaces and analogous to disjoint union of subsets we introduce direct sums.

Theorem 2.18 *Let $W_1, W_2, \ldots, W_n$ be subspaces of a vector space over a field $\mathbb{K}$, then their sum $W_1 + W_2 + \cdots + W_n = \{w_1 + w_2 + \cdots + w_n \mid w_i \in W_i\}$ is a subspace of V and it is the smallest subspace of V containing $W_1, W_2, \ldots, W_n$.*

Proof Since $W_1, W_2, \ldots, W_n$ are subspaces of V, $0 \in W_i$ for all $i = 1, 2, \ldots, n$. Then

$$0 = 0 + 0 + \cdots + 0 \in W_1 + W_2 + \cdots + W_n$$

Now let $v, w \in W_1 + W_2 + \cdots + W_n$ and $\lambda \in \mathbb{K}$, then $v = v_1 + v_2 + \cdots + v_n$ and $w = w_1 + w_2 + \cdots + w_n$ where $v_i, w_i \in W_i$ for all $i = 1, 2, \ldots, n$. As each W_i is a subspace of V, $v_i + w_i \in W_i$ and $\lambda v_i \in W_i$ for all $i = 1, 2, \ldots, n$. Hence

$$v + w = \sum_{i=1}^{n} (v_i + w_i) \in W_1 + W_2 + \cdots + W_n$$

and

$$\lambda v = \sum_{i=1}^{n} \lambda v_i \in W_1 + W_2 + \cdots + W_n$$

Therefore $W_1 + W_2 + \cdots + W_n$ is a subspace of V. Since $w_i \in W_i$ can be written as $w_i = 0 + \cdots + 0 + w_i + 0 + \cdots + 0 \in W_1 + W_2 + \cdots + W_n$, $W_1 + W_2 + \cdots + W_n$ contains each W_i. Now to prove that $W_1 + W_2 + \cdots + W_n$ is the smallest subspace containing $W_1, W_2, \ldots, W_n$, we will show that any subspace of V containing $W_1, W_2, \ldots, W_n$ contains $W_1 + W_2 + \cdots + W_n$. Let W be any subspace containing $W_1, W_2, \ldots, W_n$. Let $w = w_1 + w_2 + \cdots + w_n \in W_1 + W_2 + \cdots + W_n$ where $w_i \in W_i$ for all $i = 1, 2, \ldots, n$. Since W is a subspace of V and W contains $W_1, W_2, \ldots, W_n$, $w \in W$.

Example 2.25 Let $V = \mathbb{R}^2$. Consider $W_1 = \{(x_1, x_2) \mid x_1 = x_2, x_1, x_2 \in \mathbb{R}\}$ and $W_2 = \{(x_1, x_2) \mid x_1 = -x_2, x_1, x_2 \in \mathbb{R}\}$. Then W_1 and W_2 are subspaces of V (Fig. 2.6).

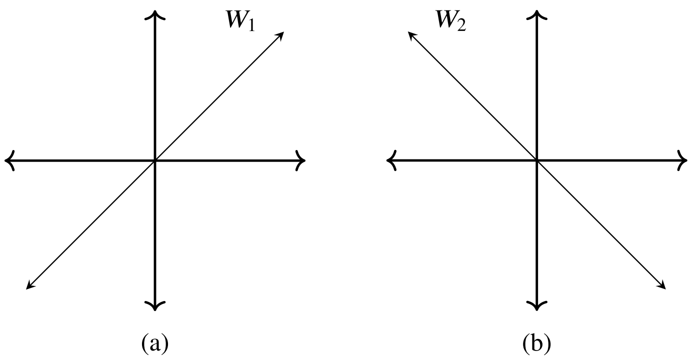


Fig. 2.6 Observe that both W_1 and W_2 depicted in (**a**) and (**b**) respectively are straight lines passing through origin and hence are subspaces of $\mathbb{R}^2$

Any vector $(x_1, x_2) \in \mathbb{R}^2$ can be written as a linear combination of elements of W_1 and W_2 as follows:

$$(x_1, x_2) = \left(\frac{x_1 + x_2}{2}, \frac{x_1 + x_2}{2}\right) + \left(\frac{x_1 - x_2}{2}, \frac{x_2 - x_1}{2}\right) \in W_1 + W_2$$

As $W_1 + W_2$ is a subspace of $\mathbb{R}^2$, this implies that $W_1 + W_2 = \mathbb{R}^2$. Also observe that the representation of any vector as the sum of elements in W_1 and W_2 is unique here.

Example 2.26 Let $V = \mathbb{M}_{2\times 2}(\mathbb{R})$. Consider

$$W_1 = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} \mid a_{11}, a_{12}, a_{22} \in \mathbb{R} \right\}$$

and

$$W_2 = \left\{ \begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix} \mid a_{11}, a_{21}, a_{22} \in \mathbb{R} \right\}$$

Then W_1 and W_2 are subspaces of V (Verify). Also any vector in $\mathbb{M}_{2\times 2}(\mathbb{R})$ can be expressed as a sum of elements in W_1 and W_2. But here this expression is not unique. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix} \in W_1 + W_2$$

and

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix} \in W_1 + W_2$$

If the elements can be expressed uniquely, then it has particular importance and is called *direct sum*. That is, the sum $W_1 + W_2$ is called *direct sum* denoted by $W_1 \oplus W_2$ if every element $w \in W_1 + W_2$ can be uniquely written as $w = w_1 + w_2$, where $w_1 \in W_1$ and $w_2 \in W_2$. That is, if $w = v_1 + v_2$, where $v_1 \in W_1$ and $v_2 \in W_2$, then $v_1 = w_1$ and $v_2 = w_2$.

Definition 2.6 (*Direct sum*) Let V be a vector space over a field $\mathbb{K}$ and $W_1, W_2, \ldots, W_n$ be subspaces of V. If every element in V can be uniquely represented as a sum of elements in $W_1, W_2, \ldots, W_n$, then V is called the direct sum of $W_1, W_2, \ldots, W_n$ and is denoted by $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$.

Suppose we have a vector space V over a field $\mathbb{K}$ and subspaces $W_1, W_2, \ldots, W_n$ of V. Then it is not easy to check whether every element in V has a unique representation as the sum of elements of $W_1, W_2, \ldots, W_n$. The following theorem provides a solution for this.

Theorem 2.19 *Let V be a vector space over a field $\mathbb{K}$ and $W_1, W_2, \ldots, W_n$ be subspaces of V. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$ if and only if the following conditions are satisfied:*

(a) $V = W_1 + W_2 + \cdots + W_n$
(b) *zero vector has only the trivial representation.*

Proof Let $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$. Then by the definition of direct sum both (a) and (b) hold. Conversely, suppose that both (a) and (b) hold. Let $v \in V$ have two representations namely,

$$v = v_1 + v_2 + \cdots + v_n \tag{2.3}$$

and

$$v = w_1 + w_2 + \cdots + w_n \tag{2.4}$$

where $v_i, w_i \in W_i$ for all $i = 1, 2, \ldots, n$. Then subtracting (2) from (1) gives

$$0 = (v_1 - w_1) + (v_2 - w_2) + \cdots + (v_n - w_n)$$

and as zero has trivial representation only, $v_i - w_i = 0$ for all $i = 1, 2, \ldots, n$ which implies $v_i = w_i$ for all $i = 1, 2, \ldots, n$. That is, every vector has a unique representation. Therefore $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$.

Example 2.27 Consider $V = \mathbb{R}^2$ and take W_1 and W_2 as in Example 2.25. Then $V = W_1 \oplus W_2$. We already know that $V = W_1 + W_2$. It is enough to prove that the zero vector has only the trivial representation. Let $(x_1, x_1) \in W_1$ and $(x_2, -x_2) \in W_2$ be such that $(x_1, x_1) + (x_2, -x_2) = (0, 0)$. This implies that $(x_1 + x_2, x_1 - x_2) = (0, 0)$ and hence $x_1 = x_2 = 0$. Thus zero vector has only the trivial representation.

The following theorem gives a necessary and sufficient condition to check whether the sum of two subspaces is a direct sum or not.

Theorem 2.20 *Let V be a vector space over a field $\mathbb{K}$. Let W_1 and W_2 be two subspaces of V, then $V = W_1 \oplus W_2$ if and only if $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$.*

Proof Let $V = W_1 \oplus W_2$, then by the definition of direct sum $V = W_1 + W_2$. If $w \in W_1 \cap W_2$, then

$$w \in W_1 \cap W_2 \Rightarrow w \in W_1 \textit{ and } w \in W_2 \Rightarrow -w \in W_2$$

Now $0 = w + (-w) \in W_1 + W_2$. Since $V = W_1 \oplus W_2$, this implies that $w = 0$. That is, $W_1 \cap W_2 = \{0\}$.

Conversely, suppose that $V = W_1 + W_2$ and $W_1 \cap W_2 = \{0\}$. Let $0 = w_1 + w_2$ where $w_1 \in W_1$ and $w_2 \in W_2$ be a non-trivial representation of the zero vector. Now $0 = w_1 + w_2 \Rightarrow -w_1 = w_2 \in W_1$, since W_1 is a subspace. As $W_1 \cap W_2 = \{0\}$, this implies that $w_1 = w_2 = 0$.

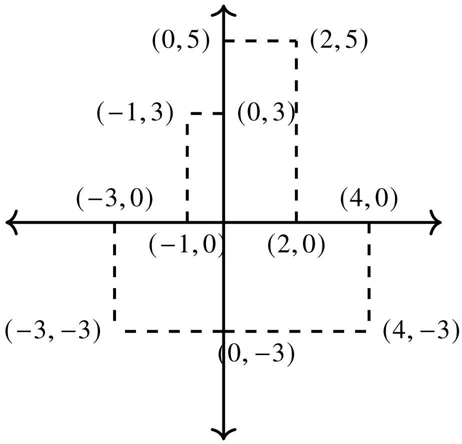


Fig. 2.7 Observe that any vector in $\mathbb{R}^2$ can be written as a sum of elements of W_1 and W_2. Also observe that $W_1 \cap W_2 = \{0\}$

Example 2.28 Let $V = \mathbb{P}_3[a, b]$. Let

$$W_1 = \{a_0 + a_2x^2 \mid a_0, a_2 \in \mathbb{R}\}$$

and

$$W_2 = \{a_1x + a_3x^3 \mid a_1, a_3 \in \mathbb{R}\}$$

Any element in $\mathbb{P}_3[a, b]$ is of the form $a_0 + a_1x + a_2x^2 + a_3x^3$. Then clearly $\mathbb{P}_3[a, b] = W_1 + W_2$. Also $W_1 \cap W_2 = \{0\}$, as polynomials in W_1 and W_2 have different orders. Therefore $\mathbb{P}_3[a, b] = W_1 \oplus W_2$.

Example 2.29 Let $V = \mathbb{R}^2$. Let $W_1 = \{(x_1, 0) \mid x_1 \in \mathbb{R}\}$ and $W_2 = \{(0, x_2) \mid x_2 \in \mathbb{R}\}$.

Then any vector $(x_1, x_2) \in \mathbb{R}^2$ can be written as $(x_1, x_2) = (x_1, 0) + (0, x_2) \in W_1 + W_2$. Since $W_1 + W_2$ is a subspace of $\mathbb{R}^2$, we get $V = W_1 + W_2$. Also $W_1 \cap W_2 = \{0\}$. Therefore $\mathbb{R}^2 = W_1 \oplus W_2$ (Fig. 2.7).

The examples discussed deal with subspaces of finite dimensional vector spaces. Now let us give you an example from an infinite-dimensional vector space.

Example 2.30 Let $V = C[a, b]$. Take

$$W_1 = \{f(x) \mid f(-x) = -f(x)\}$$

and

$$W_2 = \{f(x) \mid f(-x) = f(x)\}$$

W_1 and W_2 are respectively the collection of all odd functions and even functions. (Verify that they are subspaces of $C[a, b]$.) Now, for any $f \in C[a, b]$, consider $f_1(x) = \frac{f(x)-f(-x)}{2}$ and $f_2(x) = \frac{f(x)+f(-x)}{2}$. We have

$$f_1(-x) = \frac{f(-x) - f(-(-x))}{2} = \frac{-(f(x) - f(-x))}{2} = -f_1(x)$$

and

$$f_2(-x) = \frac{f(-x) + f(-(-x))}{2} = \frac{f(-x) + f(x)}{2} = f_2(x)$$

Thus $f_1 \in W_1$ and $f_2 \in W_2$. Clearly, $f = f_1 + f_2$ and hence $C[a, b] = W_1 + W_2$. Also observe that $W_1 \cap W_2 = \{0\}$. For if $f \in W_1 \cap W_2$, $f(-x) = -f(x) = f(x) \forall x \in [a, b]$. This gives $f(x) = 0$ for all $x \in [a, b]$. Thus we can conclude that $C[a, b] = W_1 \oplus W_2$.

Observe that the above proposition discusses the case of two subspaces only. When asking about a possible direct sum with more than two subspaces, it is not enough to check that the intersection of any two of the subspaces is $\{0\}$. For example, consider the subspaces of $\mathbb{R}^3$ given by $W_1 = \{(x_1, 0, 0) \mid x_1 \in \mathbb{R}\}$, $W_2 = \{(0, x_2, x_3) \mid x_2, x_3 \in \mathbb{R}\}$, $W_3 = \{(x_1, x_1, 0) \mid x_1 \in \mathbb{R}\}$. Clearly, $\mathbb{R}^3 = W_1 + W_2 + W_3$ and $W_1 \cap W_2 = W_1 \cap W_3 = W_2 \cap W_3 = \{0\}$ (verify). But $\mathbb{R}^3 \neq W_1 \oplus W_2 \oplus W_3$ as $(0, 0, 0) = (0, 0, 0) + (0, 0, 0) + (0, 0, 0)$ and $(0, 0, 0) = (1, 0, 0) + (0, 1, 0) + (-1, -1, 0)$.

Now we will discuss the dimension of the sum of two subspaces of a finite-dimensional vector space.

Theorem 2.21 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and W_1, W_2 be two subspaces of V, then*

$$dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2)$$

Proof Let W_1, W_2 be two subspaces of finite-dimensional vector space V. Then $W_1 \cap W_2$ is also a subspace of V. Let $\{u_1, u_2, \ldots, u_l\}$ be a basis for $W_1 \cap W_2$. Since $W_1 \subseteq W_1 \cap W_2$, $\{u_1, u_2, \ldots, u_l\}$ is a linearly independent set in W_1, and hence it can be extended to a basis $\{u_1, u_2, \ldots, u_l, v_1, v_2, \ldots, v_m\}$ of W_1. Similarly, let $\{u_1, u_2, \ldots, u_l, w_1, w_2, \ldots, w_n\}$ be a basis of W_2. Clearly

$$B = \{u_1, u_2, \ldots, u_l, , v_1, v_2, \ldots, v_m, w_1, w_2, \ldots, w_n\}$$

is a spanning set of $W_1 + W_2$. Now will show that B is a basis for $W_1 + W_2$. It is enough to show that B is linearly independent. Let $\lambda_1, \ldots, \lambda_l, \mu_1, \ldots, \mu_m, \xi_1, \ldots, \xi_n \in \mathbb{K}$ be such that

$$\lambda_1 u_1 + \cdots + \lambda_l u_l + \mu_1 v_1 + \cdots + \mu_m v_m + \xi_1 w_1 + \cdots + \xi_n w_n = 0 \tag{2.5}$$

This implies

$$\xi_1 w_1 + \cdots + \xi_n w_n = -\lambda_1 u_1 - \cdots - \lambda_l u_l - \mu_1 v_1 - \cdots - \mu_m v_m \in W_1 \cap W_2$$

as $\{u_1, u_2, \ldots, u_l, v_1, v_2, \ldots, v_m\}$ is basis for W_1 and $\{w_1, w_2, \ldots, w_n\} \subseteq W_2$. Now $\{u_1, u_2, \ldots, u_l\}$ is a basis for $W_1 \cap W_2$ implying there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_l \in \mathbb{K}$ such that

$$\xi_1 w_1 + \cdots + \xi_n w_n = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_l u_l$$

Since $\{u_1, u_2, \ldots, u_l, w_1, w_2, \ldots, w_n\}$ is a basis for W_2, the above equation implies that $\xi_1 = \cdots = \xi_n = \alpha_1 = \cdots = \alpha_l = 0$. Then (3) changes to $\lambda_1 u_1 + \cdots + \lambda_l u_l + \mu_1 v_1 + \cdots + \mu_m v_m = 0$. Since $\{u_1, u_2, \ldots, u_l, v_1, v_2, \ldots, v_m\}$ is a basis of W_1, this implies that $\lambda_1 = \cdots = \lambda_l = \mu_1 = \cdots = \mu_m = 0$. That is, B is linearly independent. Thus we have shown that B is a basis for $W_1 + W_2$. Now

$$\begin{aligned} dim(W_1 + W_2) &= l + m + n \\ &= (l + m) + (l + n) - l \\ &= dim(W_1) + dim(W_2) - dim(W_1 \cap W_2) \end{aligned}$$

Example 2.31 Consider the vector space $\mathbb{M}_{2\times 2}(\mathbb{R})$ over the field $\mathbb{R}$. Let

$$W_1 = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \mid a_{11}, a_{12}, a_{22} \in \mathbb{R} \right\}$$

and

$$W_2 = \left\{ \begin{bmatrix} a_{11} & -a_{12} \\ a_{12} & 0 \end{bmatrix} \mid a_{11}, a_{12} \in \mathbb{R} \right\}$$

Verify that W_1 and W_2 are subspaces of $\mathbb{M}_{2\times 2}(\mathbb{R})$. Since $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ is a basis for W_1, $dim(W_1) = 3$ and as $\left\{ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$ is a basis for W_2, $dim(W_2) = 2$. Now

$$W_1 \cap W_2 = \left\{ \begin{bmatrix} a_{11} & 0 \\ 0 & 0 \end{bmatrix} \mid a_{11} \in \mathbb{R} \right\}$$

Since $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$ is a basis for $W_1 \cap W_2$, $dim(W_1 \cap W_2) = 1$. Thus

$$dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2) = 4$$

Hence $W_1 + W_2 = \mathbb{M}_{2\times 2}(\mathbb{R})$.

Example 2.32 Consider the vector space $\mathbb{P}_4[a, b]$. Let

$$W_1 = \{\lambda_0 + \lambda_2 x^2 + \lambda_4 x^4 \mid \lambda_0, \lambda_2, \lambda_4 \in \mathbb{R}\}$$

and

$$W_2 = \{\lambda_1 x + \lambda_3 x^3 \mid \lambda_1, \lambda_3 \in \mathbb{R}\}$$

Since $\{1, x^2, x^4\}$ is a basis for W_1, $dim(W_1) = 3$ and as $\{x, x^3\}$ is a basis for W_2, $dim(W_2) = 2$. Clearly $dim(W_1 \cap W_2) = 0$ (How?) and hence

$$dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2) = 5$$

As $W_1 + W_2 = \mathbb{P}_4[a, b]$ and $W_1 \cap W_2 = \{0\}$, we have $\mathbb{P}_4[a, b] = W_1 \oplus W_2$.

Theorem 2.22 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Let $W_1, W_2, \ldots, W_n$ be subspaces of V, such that $V = W_1 + W_2 + \cdots + W_n$ and $dim(V) = dim(W_1) + dim(W_2) + \cdots + dim(W_n)$. Then $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$.*

Proof Let V be a finite-dimensional vector space with $W_1, W_2, \ldots, W_n$ as subspaces of V. Consider a basis B_i for each $i = 1, 2, \ldots, n$ and let $B = \cup_{i=1}^{n} B_i$. Since $V = W_1 + W_2 + \cdots + W_n$, B spans V. Now suppose that B is linearly dependent. Then at least one of the vectors can be written as a linear combination of other vectors. Then $dim(V) < dim(W_1) + dim(W_2) + \cdots + dim(W_n)$, which is a contradiction. Therefore B is linearly independent and hence B is a basis of V. Now let $0 = w_1 + w_2 + \cdots + w_n$ where $w_i \in W_i$. Since B_i is a basis for W_i, each $w_i \in W_i$ can be expressed uniquely as a sum of elements in B_i. i.e., 0 can be written as a linear combination of elements of B. As B is a basis for V, this implies that the coefficients are zero. That is, $w_i = 0$ for all $i = 1, 2, \ldots, n$. Therefore $V = W_1 \oplus W_2 \oplus \cdots \oplus W_n$.

2.6 Exercises

1. Show that the collections given in Examples 2.2–2.7 are vector spaces with respect to the given operations.
2. Consider the vector space $\mathbb{R}^2$ with usual addition and multiplication over $\mathbb{R}$. Give an example for a subset of $\mathbb{R}^2$ which is
   (a) closed under addition but not closed under scalar multiplication.
   (b) closed under scalar multiplication but not closed under addition.
3. Does $\mathbb{R}^2$ over $\mathbb{R}$ with operations defined by
$$(x_1, x_2) + (y_1, y_2) = (x_1 + x_2, y_1 + y_2) \; and \; \lambda(x_1, x_2) = (\lambda x_1, 0)$$
   form a vector space?
4. Check whether the following vectors are linearly dependent or not.
   (a) $\{(1, 2), (2, 1)\}$ in $\mathbb{R}^2$ over $\mathbb{R}$.
   (b) $\{(1, 2, 1), (2, 1, 1), (1, 1, 2)\}$ in $\mathbb{R}^3$ over $\mathbb{R}$.
   (c) $\{(i, -i), (-1, 1)\}$ in $\mathbb{C}^2$ over $\mathbb{R}$.
   (d) $\{(i, -i), (-1, 1)\}$ in $\mathbb{C}^2$ over $\mathbb{C}$.

(e) $\{1+x, 1+x^2\}$ in $\mathbb{P}_2[a,b]$ over $\mathbb{R}$.
(f) $\{2, x-2, 1+x+x^2, x^3-x^2\}$ in $\mathbb{P}_3[a,b]$ over $\mathbb{R}$.
(g) $\left\{\begin{bmatrix}1 & 2\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\2 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\2 & 1\end{bmatrix}, \begin{bmatrix}1 & 1\\2 & 0\end{bmatrix}\right\}$ in $\mathbb{M}_2(\mathbb{R})$ over $\mathbb{R}$.
(h) $\left\{\begin{bmatrix}1 & -1\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}, \begin{bmatrix}3 & -1\\2 & 1\end{bmatrix}\right\}$ in $\mathbb{M}_2(\mathbb{R})$ over $\mathbb{R}$.

5. Let $\{v_1, v_2\}$ be a linearly independent subset of a vector space V over a field $\mathbb{K}$. Then show that $\{v_1+v_2, v_1-v_2\}$ is linearly independent only if characteristic of $\mathbb{K}$ is not equal to 2.
6. Check whether the following subsets of $\mathbb{R}^2$ are subspaces of $\mathbb{R}^2$ over $\mathbb{R}$. If yes, find its dimension and write down a basis.

(a) $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 = 1\}$
(b) $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 0\}$
(c) $\{(x_1, x_2) \in \mathbb{R}^2 \mid \frac{x_1}{x_2} = 1\}$
(d) $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1, x_2 \leq 0\}$
(e) $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 0\}$.

7. Check whether the following subsets of $\mathbb{M}_2(\mathbb{K})$ are subspaces of $\mathbb{M}_2(\mathbb{K})$ over $\mathbb{K}$. If yes, find its dimension and write down a basis.

(a) $\left\{\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} \in \mathbb{M}_2(\mathbb{K}) \mid a_{11} + a_{12} = 0\right\}$
(b) $\left\{\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} \in \mathbb{M}_2(\mathbb{K}) \mid a_{11} + a_{12} = 1\right\}$
(c) $\{A \in \mathbb{M}_2(\mathbb{K}) \mid det(A) = 0\}$
(d) $\{A \in \mathbb{M}_2(\mathbb{K}) \mid det(A) \neq 0\}$
(e) $\left\{\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} \in \mathbb{M}_n(\mathbb{K}) \mid a_{11} = a_{22}\right\}$.

8. Check whether the following subsets of $\mathbb{P}_2(\mathbb{R})$ are subspaces of $\mathbb{P}_2(\mathbb{R})$ over $\mathbb{R}$. If yes, find its dimension and write down a basis.

(a) $\{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid p(0) = 0\}$
(b) $\{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid p(0) = 1\}$
(c) $\{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid p(0) = p(1) = 0\}$
(d) $\{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid p(x) \geq 0\}$
(e) $\{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid p(x) = p(-x)\}$.

9. State whether the following statements are true or false.

(a) A non-trivial vector space over the fields $\mathbb{R}$ or $\mathbb{C}$ always has an infinite number of elements.
(b) The set of all rational numbers $\mathbb{Q}$ is a vector space over $\mathbb{R}$ under usual addition and multiplication.
(c) $\{(x_1, x_2) \mid x_1^2 + x_2^2 = 0, x_1, x_2 \in \mathbb{C}\}$ is a subspace of $\mathbb{C}^2$ over $\mathbb{C}$.

(d) There exists a non-trivial subspace of $\mathbb{R}$ over $\mathbb{R}$ under usual addition and multiplication.
(e) $\{(i, 1), (-1, i)\}$ is a linearly independent set in $\mathbb{C}^2$ over $\mathbb{C}$.
(f) If W_1, W_2, W are subspaces of a vector space V such that $W_1 + W = W_2 + W$, then $W_1 = W_2$.
(g) If W_1, W_2 are subspaces of $\mathbb{R}^7$ with $dim\,(W_1) = 4$ and $dim\,(W_2) = 4$, then $dim\,(W_1 \cap W_2) = 1$.

10. Show that $\mathbb{R}$ with usual addition and multiplication over $\mathbb{Q}$ is an infinite-dimensional vector space. (**Hint:** Use the fact that π is a transcendental number.)
11. Find the row space and column space of $\begin{bmatrix} 1 & -1 \\ 2 & 0 \end{bmatrix}$.
12. Show that the rows of a 3×2 matrix are linearly dependent.
13. Show that the columns of a 3×5 matrix are linearly dependent.
14. Which of the following collection of vectors span $\mathbb{R}^2$ over $\mathbb{R}$?

(a) $\{(1, 1)\}$
(b) $\{(1, 2), (0, 4)\}$
(c) $\{(0, 0), (1, -1), (3, 2)\}$
(d) $\{(2, 4), (4, 8)\}$
(e) $\{(3, 2), (1, 4), (4, 6)\}$.

15. Which of the following collection of vectors span $\mathbb{R}^3$ over $\mathbb{R}$?

(a) $\{(1, 1, 0), (0, 1, 1)\}$
(b) $\{(0, 2, 0), (1, 0, 0), (1, 2, 0)\}$
(c) $\{(0, 0, -1), (0, 1, -1), (-1, 1, -1)\}$
(d) $\{(0, 4, 2), (0, 8, 4), (1, 12, 6)\}$
(e) $\{(1, 3, 2), (1, 2, 3), (3, 2, 1), (2, 1, 3)\}$.

16. Which of the following collection of vectors span $\mathbb{P}^2[a, b]$ over $\mathbb{R}$?

(a) $\{x^2 + 1, x^2 + x, x + 1\}$
(b) $\{x + 1, x - 1, x^2 - 1\}$
(c) $\{x^2 + x + 1, 2x - 1\}$
(d) $\{2x^2 - x + 1, x^2 + x, 2x - 3, x^2 - 5\}$
(e) $\{x + 1, 2x + 2, x^2 + x\}$.

17. Let W_1, W_2 be subsets of a vector space V over the field $\mathbb{K}$. Show that

(a) $span\,(W_1 \cap W_2) \subseteq span\,(W_1) \cap span\,(W_2)$.
(b) $span\,(W_1) \cup span\,(W_2) \subseteq span\,(W_1 \cup W_2)$.

Does the converse hold in both a) and b)?

18. Let W_1, W_2 be subspaces of a vector space V over the field $\mathbb{K}$. Show that $span\,(W_1 + W_2) = span\,(W_1) + span\,(W_2)$.
19. Let $V_1 = \{v_1, v_2, \ldots, v_n\}$, $V_2 = \{v_1, v_2, \ldots, v_n, v\}$ be subsets of a vector space V. Then $span\,(V_1) = span\,(V_2)$ if and only if $v \in span\,(V_1)$.

20. Check whether the given collection of vectors form a basis for corresponding vector spaces.

    (a) $\{(2, 1, 1), (1, 2, 1), (1, 1, 2)\}$ for $\mathbb{R}^3$ over $\mathbb{R}$.
    (b) $\{1, x-1, (x-1)^2\}$ for $\mathbb{P}_2(\mathbb{R})$ over $\mathbb{R}$.
    (c) $\{1, x^2-1, 2x^2+5\}$ for $\mathbb{P}_2(\mathbb{R})$ over $\mathbb{R}$.
    (d) $\left\{\begin{bmatrix}1 & 0\\ 0 & -2\end{bmatrix}, \begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 2 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right\}$ for $\mathbb{M}_2(\mathbb{R})$ over $\mathbb{R}$.

21. Determine which of the given subsets forms a basis for $\mathbb{R}^3$ over $\mathbb{R}$. Express the vector $(1, 2, 3)$ as a linear combination of the vectors in each subset that is a basis.

    (a) $\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$
    (b) $\{(1, 2, 1), (2, 1, 1), (1, 1, 2)\}$
    (c) $\{(2, 3, 1), (1, -2, 0), (1, 5, 1)\}$.

22. Check whether the sets given in Questions $14-16$ form a basis for the respective vector spaces. If not, find the dimension of their span.
23. Find the dimension of span of the following collection of vectors:

    (a) $\{(1, -2), (-2, 4)\}$ in $\mathbb{R}^2$ over $\mathbb{R}$.
    (b) $\{(-2, 3), (1, 2), (5, 6)\}$ in $\mathbb{R}^2$ over $\mathbb{R}$.
    (c) $\{(0, 3, 1), (-1, 2, 3), (2, 3, 0), (-1, 2, 4)\}$ in $\mathbb{R}^3$ over $\mathbb{R}$.
    (d) $\{1+x, x^2+x+1\}$ in $\mathbb{P}_2[a, b]$ over $\mathbb{R}$.
    (e) $\{1-x, x^2, 2x^2+x-1\}$ in $\mathbb{P}_2[a, b]$ over $\mathbb{R}$.
    (f) $\left\{\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}, \begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 1 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\ -1 & 1\end{bmatrix}\right\}$ in $\mathbb{M}_2(\mathbb{R})$ over $\mathbb{R}$.
    (g) $\left\{\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}\right\}$ in $\mathbb{M}_2(\mathbb{R})$ over $\mathbb{R}$.

    Also, find a basis for the linear space spanned by the vectors.
24. Consider two subspaces of $\mathbb{R}^4$ given by

$$W_1 = \{(x_1, x_2, 2x_1, x_1 + x_2) \in \mathbb{R}^4 \mid x_1, x_2 \in \mathbb{R}\}$$

    and

$$W_2 = \{(x_1, 2x_1, x_2, x_1 - x_2) \in \mathbb{R}^4 \mid x_1, x_2 \in \mathbb{R}\}$$

    Find

    (a) $W_1 + W_2$ and $W_1 \cap W_2$.
    (b) $dim\,(W_1 + W_2)$ and $dim\,(W_1 \cap W_2)$.

25. Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and W_1 be a subspace of V. Prove that there exists a subspace W_2 of V such that $V = W_1 \oplus W_2$.
26. Let V be a vector space over a field $\mathbb{K}$ and $W_1, W_2, \ldots, W_n$ be subspaces of V with $W_i \cap W_j = \{0\}\ \forall\, i \neq j$ and $W_1 + W_2 + \cdots + W_n = V$. Is the sum a direct sum?

27. Let $W_1 = \{A \in \mathbb{M}_n(\mathbb{K}) \mid A_{ij} = 0 \ \forall i \geq j\}$, $W_2 = \{A \in \mathbb{M}_n(\mathbb{K}) \mid A_{ij} = 0 \ \forall i \leq j\}$, and $W_3 = \{A \in \mathbb{M}_n(\mathbb{K}) \mid A_{ij} = 0 \ \forall\, i \neq j\}$. Then show that $\mathbb{M}_n(\mathbb{K}) = W_1 \oplus W_2 \oplus W_3$.
28. Let
$$W_1 = \{A \in \mathbb{M}_n(\mathbb{K}) \mid A^T = A\}$$
and
$$W_2 = \{A \in \mathbb{M}_n(\mathbb{K}) \mid A^T = -A\}$$
Then show that $\mathbb{M}_n(\mathbb{K}) = W_1 \oplus W_2$.

Solved Questions related to this chapter are provided in Chap. 9.

Chapter 3
Linear Transformations

In this chapter, we delve deeply into a key concept of linear transformations which map between vector spaces in linear algebra. It begins with the definition and key properties of linear transformations, emphasizing their significance as they preserve vector space operations. Examples such as the differential operator, which maps a function to its derivative, and the integral operator, which maps a function to its integral, are discussed. Both of these exhibit linearity qualities that are essential in calculus and mathematical analysis. The concepts of range spaces and null spaces are presented, providing an insight into the possible outputs and dependencies of linear transformations. The relationship between linear transformations and matrices is illustrated, showing how matrices can be used to represent these transformations and for ease of computations. It covers fundamental ideas like projection, rotation, reflection, shear, and other transformations that provide helpful insights into the geometric manipulation of vectors and shapes in two-dimensional spaces. Invertible linear transformations and isomorphism of vector spaces are discussed. The chapter also deals with the concept of changing coordinate bases, shedding light on how different bases can affect the representation of vectors and linear transformations. We further go into linear functionals, emphasizing their importance in dual spaces.

3.1 Introduction

In this chapter, we will be discussing functions on vector spaces that preserve the structure. This gives us an important class of functions called linear transformations. Vaguely, we can say that a linear transformation is a function between two vector spaces that preserve algebraic operations. When we discuss linear transformations from $\mathbb{R}^2$ to itself, we could see that the transformation has some interesting geometric properties. The term *"transformation"* just indicates that it transforms the input vector to give us an output vector and the term *"linear"* suggests that

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_3

(a) all lines must remain lines, without getting curved and
(b) the origin must remain fixed. That is, the image of the origin must be the origin itself.

These two ideas can be used to verify whether a function is linear or not when a transformation is defined from $\mathbb{R}^2$ to itself.

To say geometrically that a function is a linear transformation, we must be able to say that conditions (a) and (b) as mentioned above must be satisfied, which is a tedious task. Also, we need a valid strategy to check whether a function defined on an arbitrary vector space is a linear transformation or not. To tackle such situations, we have the following definition for a linear transformation.

Definition 3.1 (*Linear Transformations*) Let V and W be vector spaces over the field $\mathbb{K}$. A linear transformation T from V into W is a function such that

(a) $T(v_1 + v_2) = T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$ and
(b) $T(\lambda v) = \lambda T(v)$ for all $v \in V$ and $\lambda \in \mathbb{K}$.

(a) is called the additive property and (b) is called the homogeneity property, both of which can be written together as $T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2)$ for $v_1, v_2 \in V$ and $\lambda_1, \lambda_2 \in \mathbb{K}$. This compact form is called the *principle of superposition.* A linear transformation from V to itself is called a *linear operator.*

Now that we have put forward algebraic conditions to check whether a function on a vector space is a linear transformation or not, we can look into some of the important examples. Many of the important functions that we use in pure and applied mathematics like differentiation and integration in calculus, rotations, reflections, and projection in geometry are in fact linear transformations.

Example 3.1 Let $V = \mathbb{R}^2$ over the field $\mathbb{R}$. Define $T : V \to V$ by

$$T(x_1, y_1) = (2x_1 + 3y_1, 5x_1)$$

Let $\lambda \in \mathbb{R}$ and $v_1 = (x_1, y_1), v_2 = (x_2, y_2) \in V$. Then $\lambda v_1 + v_2 = (\lambda x_1 + x_2, \lambda y_1 + y_2)$, and

$$\begin{aligned} T(\lambda v_1 + v_2) &= T(\lambda x_1 + x_2, \lambda y_1 + y_2) \\ &= (2(\lambda x_1 + x_2) + 3(\lambda y_1 + y_2), 5(\lambda x_1 + x_2)) \\ &= (\lambda(2x_1 + 3y_1) + (2x_2 + 3y_2), \lambda(5x_1) + 5x_2) \\ &= \lambda\,(2x_1 + 3y_1, 5x_1) + (2x_2 + 3y_2, 5x_2) \\ &= \lambda T(v_1) + T(v_2) \end{aligned}$$

Therefore T is a linear operator on V.

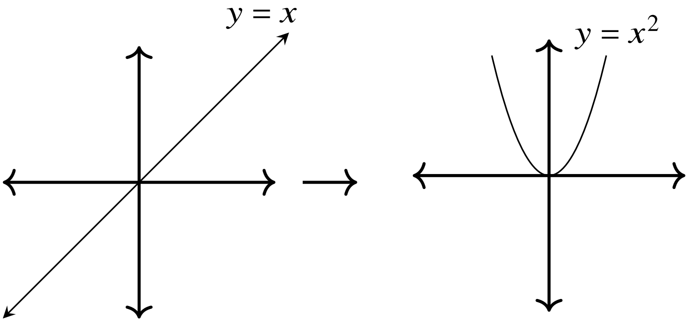


Fig. 3.1 Observe that the line $y = x$ is transformed to a curve $y = x^2$. Thus the above transformation is not linear

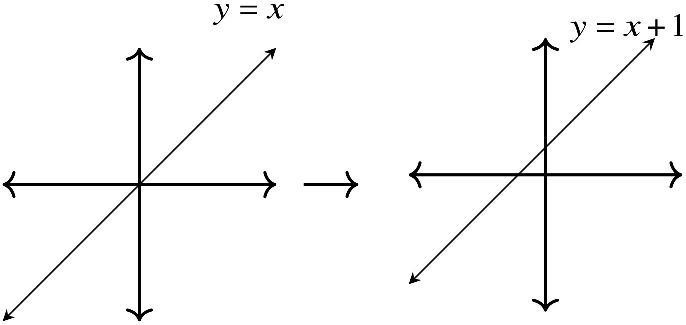


Fig. 3.2 Observe that the line $y = x$ is transformed to a line $y = x + 1$. But the origin is not mapped onto itself. Thus the above transformation is not linear

Example 3.2 From Fig. 3.1, it is clear that the function $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_1(x_1, x_2) = (x_1, x_1^2)$ is not a linear transformation. Now let us check the linearity of T_1 using Definition 3.1. That is, we have to check whether T_1 satisfies the *principle of superposition* for all vectors in $\mathbb{R}^2$ or not. Observe that

$$T_1(1, 1) + T_1(-1, 1) = (0, 2) \neq T_1((1, 1) + (-1, 1)) = T_1(0, 2) = (0, 0)$$

Thus, our assertion that T_1 is not a linear transformation is confirmed. Similarly, we can check whether the function $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_2(x_1, x_2) = (x_1, x_1 + 1)$ is a linear transformation or not. As

$$T_2(1, 1) + T_2(-1, 1) = (0, 2) \neq T_2((1, 1) + (-1, 1)) = T_2(0, 2) = (0, 1)$$

by Definition 3.1, T_2 is not a linear transformation, which we have already observed from Fig. 3.2.

Example 3.3 Let $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ over the field $\mathbb{R}$. Define $T : V \to W$ by $T(v) = Av$, where $A \in M_{m\times n}(\mathbb{R})$ is a fixed matrix. Then, for $v_1, v_2 \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$,

$$T(\lambda v_1 + v_2) = A(\lambda v_1 + v_2) = \lambda A(v_1) + A(v_2) = \lambda T(v_1) + T(v_2)$$

Therefore T is a linear transformation from V to W.

Example 3.4 Let $V = \mathbb{P}_n(\mathbb{R})$ and $W = \mathbb{P}_{n-1}(\mathbb{R})$ over the field $\mathbb{R}$. Define $T : V \to W$ by $T(p(x)) = \frac{d}{dx}(p(x))$. Then T is a linear transformation from V to W. For $p(x), q(x) \in \mathbb{P}_n(\mathbb{R})$ and $\lambda \in \mathbb{R}$,

$$\begin{aligned} T\left((\lambda p + q)(x)\right) &= \frac{d}{dx}\left((\lambda p + q)(x)\right) \\ &= \frac{d}{dx}\left(\lambda p(x)\right) + \frac{d}{dx}\left(q(x)\right) \\ &= \lambda \frac{d}{dx}\left(p(x)\right) + \frac{d}{dx}\left(q(x)\right) \\ &= \lambda T(p(x)) + T(q(x)) \end{aligned}$$

Therefore T is a linear transformation from V to W.

Example 3.5 Let $V = \mathbb{P}_{n-1}(\mathbb{R})$ and $W = \mathbb{P}_n(\mathbb{R})$ over the field $\mathbb{R}$. Define $T : V \to W$ by $(Tp)(x) = \int_0^x p(t)dt$. Then T is a linear transformation from V to W. For $p(x), q(x) \in \mathbb{P}_{n-1}(\mathbb{R})$ and $\lambda \in \mathbb{R}$,

$$\left(T(\lambda p + q)\right)(x) = \int_0^x (\lambda p + q)(t)dt = \lambda \int_0^x p(t)dt + \int_0^x q(t)dt = \lambda (Tp)(x) + (Tq)(x)$$

Therefore T is a linear transformation from V to W.

Example 3.6 Let $V = \mathbb{M}_{m\times n}(\mathbb{K})$ and $W = \mathbb{M}_{n\times m}(\mathbb{K})$ over the field $\mathbb{K}$. Define $T : V \to W$ by $T(A) = A^T$. Then for $A, B \in \mathbb{M}_{m\times n}$ and $\lambda \in \mathbb{K}$

$$T(\lambda A + B) = (\lambda A + B)^T = \lambda A^T + B^T = \lambda T(A) + T(B)$$

Therefore T is a linear transformation from V to W.

Example 3.7 Let V and W be any two arbitrary vector spaces over a field $\mathbb{K}$. The linear transformation $I : V \to V$ defined by $I(v) = v$ for all $v \in V$ is called the identity transformation. The linear transformation $\mathcal{O} : V \to W$ defined by $\mathcal{O}(v) = 0$ for all $v \in V$ is called the zero transformation.

Now let us discuss some of the important properties of linear transformations.

Theorem 3.1 *Let V and W be vector spaces over the field $\mathbb{K}$ and $T : V \to W$ be a function*

(a) If T is linear, then $T(0) = 0$.
(b) T is linear if and only if $T(\lambda v_1 + v_2) = \lambda T(v_1) + T(v_2)$ for all $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$.
(c) If T is linear, then $T(v_1 - v_2) = T(v_1) - T(v_2)$ for all $v_1, v_2 \in V$.

(d) *T is linear if and only if, for* $v_1, v_2, \ldots, v_n \in V$ *and* $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$*, we have*

$$T\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i T(v_i)$$

(e) *Let* $\{v_1, v_2, \ldots, v_n\}$ *be a linearly dependent set in* V*, then* $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ *is a linearly dependent set in* W*.*

Proof Suppose that $T : V \to W$ is linear, then

$$T(0) = T(0+0) = T(0) + T(0) \Rightarrow T(0) = 0$$

Proof of (b), (c), and (d) are trivial from the definition of a linear transformation. Now let $\{v_1, v_2, \ldots, v_n\}$ be a linearly dependent set in V, then there exists at least one vector in V, say v_i such that

$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n$$

where $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n \in \mathbb{K}$ and are not all zero. Then

$$\begin{aligned} T(v_i) &= T\left(\lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n\right) \\ &= \lambda_1 T(v_1) + \cdots + \lambda_{i-1} T(v_{i-1}) + \lambda_{i+1} T(v_{i+1}) + \cdots + \lambda_n T(v_n) \end{aligned}$$

Therefore $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a linearly dependent set in W.

Observe that Theorem 3.1(a) gives a necessary condition, not a sufficient one. That is, $T(0) = 0$ need not imply that T is a linear transformation. For example, consider $T : \mathbb{R} \to \mathbb{R}$ such that $T(v) = v^2$. Clearly, $T(0) = 0$. But

$$T(v_1 + v_2) = (v_1 + v_2)^2 = v_1^2 + v_2^2 + 2v_1 v_2 \neq v_1^2 + v_2^2 = T(v_1) + T(v_2)$$

That is, T is not linear. But if $T(0) \neq 0$, then we can say that T is not linear. (b) part of the above theorem is used to prove that a given function is a linear transformation.

As linear transformations preserve linear combinations, to describe a linear transformation on a vector space V, it is enough to identify the images of the basis vectors of the domain under the linear transformation.

Theorem 3.2 *Let* V *be a finite-dimensional vector space over the field* $\mathbb{K}$ *with basis* $\{v_1, v_2, \ldots, v_n\}$*. Let* W *be a vector space over the same field* $\mathbb{K}$ *and* $\{w_1, w_2, \ldots, w_n\}$ *be an arbitrary set of vectors in* W*, then there exists exactly one linear transformation* $T : V \to W$ *such that* $T(v_i) = w_i$*, where* $i = 1, 2, \ldots, n$*.*

Proof Let V be a finite-dimensional vector space over the field $\mathbb{K}$ with basis $\{v_1, v_2, \ldots, v_n\}$. Then for each $v \in V$, there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ such that $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Now define

$$T(v) = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n$$

for each $v \in V$. Then T is well defined and $T(v_i) = w_i$, for all $i = 1, 2, \ldots, n$. Now we have to prove that T is a linear transformation. Take $u, v \in V$. Then $u = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n$ and $v = \xi_1 v_1 + \xi_2 v_2 + \cdots + \xi_n v_n$ for some scalars $\mu_1, \mu_2, \ldots, \mu_n, \xi_1, \xi_2, \ldots, \xi_n \in \mathbb{K}$. Now for $\lambda \in \mathbb{K}$,

$$\begin{aligned} T(\lambda u + v) &= T\left(\lambda\left(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n\right) + \xi_1 v_1 + \xi_2 v_2 + \cdots + \xi_n v_n\right) \\ &= T\left(\left(\lambda\mu_1 + \xi_1\right) v_1 + \left(\lambda\mu_2 + \xi_2\right) v_2 + \cdots + \left(\lambda\mu_n + \xi_n\right) v_n\right) \\ &= \left(\lambda\mu_1 + \xi_1\right) w_1 + \left(\lambda\mu_2 + \xi_2\right) w_2 + \cdots + \left(\lambda\mu_n + \xi_n\right) w_n \\ &= \lambda\left(\mu_1 w_1 + \mu_2 w_2 + \cdots + \mu_n w_n\right) + \xi_1 w_1 + \xi_2 w_2 + \cdots + \xi_n w_n \\ &= \lambda T(u) + T(v) \end{aligned}$$

Therefore T is a linear transformation. Now suppose that there exists another linear transformation $\tilde{T} : V \to W$ such that $\tilde{T}(v_i) = w_i$, where $i = 1, 2, \ldots, n$. Since $\tilde{T}$ is linear, for each vector $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in V$,

$$\begin{aligned} \tilde{T}(v) &= \tilde{T}\left(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n\right) \\ &= \lambda_1 \tilde{T}(v_1) + \lambda_2 \tilde{T}(v_2) + \cdots + \lambda_n \tilde{T}(v_n) \\ &= \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_n w_n \\ &= T(v) \end{aligned}$$

That is, there exists exactly one linear transformation $T : V \to W$ such that $T(v_i) = w_i$, for all $i = 1, 2, \ldots, n$.

Example 3.8 Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Let $T : V \to W$ be a linear transformation such that $T(1, 0) = (2, 1, 0)$ and $T(0, 1) = (1, 0, 3)$. Since $\{(1, 0), (0, 1)\}$ is a basis for $\mathbb{R}^2$, from Theorem 3.2, there exists only one such T. To find T, take $(x_1, y_1) \in \mathbb{R}^2$. Since

$$(x_1, y_1) = x_1(1, 0) + y_1(0, 1)$$

we get

$$\begin{aligned} T(x_1, y_1) &= x_1 T(1, 0) + y_1 T(0, 1) \\ &= x_1(2, 1, 0) + y_1(1, 0, 3) \\ &= (2x_1 + y_1, x_1, 3y_1) \end{aligned}$$

In Example 3.3, we have seen that $T(v) = Av$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, where $A \in M_{m\times n}(\mathbb{R})$ is a fixed matrix. In particular, for $n = 2$ and $m = 3$, if we take $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}$, we can say that $v \to Av$ is a linear transformation. We can represent this linear transformation by $\tilde{T}$, i.e., $\tilde{T}(v) = Av$. What is the speciality of this transformation? Is there any similarity between the transformations $\tilde{T}$ and the transformation T defined in Example 3.8?

We can observe that

$$\tilde{T}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}2 & 1\\1 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}2\\1\\0\end{bmatrix} \quad \textit{and} \quad \tilde{T}\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}2 & 1\\1 & 0\\0 & 3\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1\\0\\3\end{bmatrix}$$

Then by Theorem 3.2, we can say that T and $\tilde{T}$ are the same. This is an interesting fact, isn't it? This relation between the set of linear transformations from an n dimensional space to an m dimensional space and the set of all $m \times n$ real matrices is worth exploring, and we will be studying this relation in detail in this chapter.

3.2 Range Space and Null Space

Considering that a linear transformation from V to W is a function that preserves structure, a linear transformation has two significant sets associated with it: the null set and the range set. In fact, they are subspaces of V and W respectively. The range space of a linear transformation consists of all possible output vectors that can be obtained by applying the transformation to input vectors. It represents the span of the transformed vectors in the co-domain. On the other hand, the null space of a linear transformation comprises all input vectors that are mapped to the zero vector in the co-domain, forming a subspace of the domain. Together, these spaces provide valuable insights into the behavior and properties of the linear transformation. In this section, we will discuss in detail these subspaces and some of the important results associated.

Definition 3.2 (*Range set and Null set*) Let V and W be vector spaces over the field $\mathbb{K}$, and let $T : V \to W$ be linear, then range set of T, denoted by $\mathcal{R}(T)$, is a subset of W consisting of all images of vectors in V under T. That is,

$$\mathcal{R}(T) = \{T(v) \mid v \in V\}$$

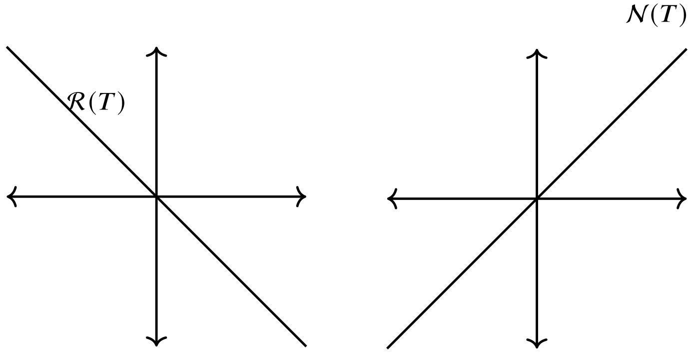
and the null set or kernel of T, denoted by $\mathcal{N}(T)$, is the set of all vectors $v \in V$ such that $T(v) = 0$. That is,

$$\mathcal{N}(T) = \{v \in V \mid T(v) = 0\}$$

Consider the following example. We can observe some interesting facts about the range set and null set associated with a linear transformation.

Example 3.9 Consider a function $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T(x_1, x_2) = (x_1 - x_2, x_2 - x_1)$$



(a) (b)

Fig. 3.3 Both range set and null set of T; plotted as (**a**) and (**b**) respectively are straight lines passing through origin. Thus both $\mathcal{R}(T)$ and $\mathcal{N}(T)$ are subspaces of $\mathbb{R}^2$ for the given T

Clearly T is a linear transformation (Verify). Now let us find the range set of T. By definition,

$$\mathcal{R}(T) = \left\{T(x_1, x_2) \mid (x_1, x_2) \in \mathbb{R}^2\right\} = \{(x_1 - x_2, x_2 - x_1) \mid x_1, x_2 \in \mathbb{R}\}$$

which is the straight line $y = -x$. And the null set of T is given by

$$\begin{aligned}\mathcal{N}(T) &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid T(x_1, x_2) = 0\right\} \\ &= \left\{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - x_2, x_2 - x_1) = (0, 0)\right\}\end{aligned}$$

which is the straight line $y = x$ (Fig. 3.3).

One of the interesting facts to observe here is that both $\mathcal{R}(T)$ and $\mathcal{N}(T)$ are subspaces of $\mathbb{R}^2$ for the given T, as we have seen in Fig. 3.3. Will this be true for every linear transformation $T : V \to W$, where V and W are any two arbitrary vector spaces? That is, will $\mathcal{R}(T)$ be a subspace of W and $\mathcal{N}(T)$ be a subspace of V? The following theorem will give us an answer.

Another interesting fact to observe here is that the lines $y = x$ and $y = -x$ are perpendicular to each other. So far, we haven't defined the tools to analyze this fact. We will study this interesting observation in detail in Chap. 5.

In the next theorem we will prove that if T is a linear transformation between two vector spaces V and W, then $\mathcal{R}(T)$ and $\mathcal{N}(T)$ are subspaces of W and V, respectively.

Theorem 3.3 *Let V and W be vector spaces over the field $\mathbb{K}$, and let $T : V \to W$ be a linear transformation. Then $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are subspaces of V and W respectively.*

Proof Let $T : V \to W$ be a linear transformation. We have $\mathcal{N}(T) = \{v \in V \mid T(v) = 0\}$. Clearly, $\mathcal{N}(T) \subseteq V$. Now for $v_1, v_2 \in \mathcal{N}(T)$ and $\lambda \in \mathbb{K}$, we have

$$T(\lambda v_1 + v_2) = \lambda T(v_1) + T(v_2) = 0$$

Therefore $\lambda v_1 + v_2 \in \mathcal{N}(T)$ for all $v_1, v_2 \in \mathcal{N}(T)$ and $\lambda \in \mathbb{K}$. Hence $\mathcal{N}(T)$ is a subspace of V.

Also we have $\mathcal{R}(T) = \{T(v) \mid v \in V\}$. Clearly, $\mathcal{R}(T) \subseteq W$. As $\mathcal{R}(T)$ is range space, for $w_1, w_2 \in \mathcal{R}(T)$ there exists $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Since $v_1, v_2 \in V$ and V is a vector space over the field $\mathbb{K}$, $\lambda v_1 + v_2 \in V$, where $\lambda \in \mathbb{K}$. Then

$$T(\lambda v_1 + v_2) = \lambda T(v_1) + (v_2) = \lambda w_1 + w_2 \in \mathcal{R}(T)$$

Hence $\mathcal{R}(T)$ is a subspace of W.

If $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are finite dimensional, the dimensions of $\mathcal{N}(T)$ and $\mathcal{R}(T)$ are called *Nullity*(T) (read as nullity of T) and *Rank*(T) (read as rank of T), respectively. Now, let T be a linear transformation from a finite-dimensional vector space V to a vector space W. From Theorem 3.2, it is clear that, if we know the images of basis elements of V, it is easy to find $\mathcal{R}(T)$. If $B = \{v_1, v_2, \ldots, v_n\}$ is a basis for V, then

$$\mathcal{R}(T) = span\ (T(B)) = span\{T(v_1), T(v_2), \ldots, T(v_n)\}$$

We can also conclude that, if $dim(V) = n$, then $Rank(T) \leq n$.

Remark 3.1 Let A be an $m \times n$ matrix with entries from the field $\mathbb{K}$. We have seen that the space spanned by the rows of A is a subspace of $\mathbb{K}^n$ called row space of A and the space spanned by the columns of A is a subspace of $\mathbb{K}^m$, called the column space of A. The dimensions of the row space and column space are called the *row rank* and *column rank* of A, respectively. We will later show that *row rank*(A)=*column rank*(A) for any $m \times n$ matrix A. The column space of a matrix A is also known as the image or range of A, denoted by $Im(A)$ or $\mathcal{R}(A) Col(A)$. That is,

$$Im(A) = \{Ax \mid x \in \mathbb{K}^n\}$$

and the null space or kernel of A, denoted by $Ker\ A$ or $\mathcal{N}(A)$, is given by

$$Ker(A) = \{x \in \mathbb{K}^n \mid Ax = 0\}$$

Now let us discuss an interesting example. The relation between linear transformations and matrices will become more evident in the following one.

Example 3.10 Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ in Example 3.8 defined by $T(x_1, y_1) = (2x_1 + y_1, x_1, 3y_1)$. Then the range space of T is given by

$$\mathbb{R}(T) = span\ (T(B)) = span\{T(1, 0), T(0, 1)\} = span\{(2, 1, 0), (1, 0, 3)\}$$

As the vectors (2, 1, 0) and (1, 0, 3) are linearly independent, $Rank(T) = 2$. Now the null space of T is

$$\mathcal{N}(T) = \{x \in \mathbb{R}^2 \mid T(x) = 0\} = \{x = (x_1, y_1) \in \mathbb{R}^2 \mid (2x_1 + y_1, x_1, 3y_1) = (0, 0, 0)\} = \{(0, 0)\}$$

Thus $Nullity(T) = 0$. Now, consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}$. We can observe that the column space of A is the span of $\{(2, 1, 0), (1, 0, 3)\}$ and kernel of A is $\{(0, 0)\}$. Does there exist any relation between the range space of a matrix and the range space of the corresponding linear transformation?

Example 3.11 Let $V = \mathbb{P}_3(\mathbb{R})$ and $W = \mathbb{P}_2(\mathbb{R})$. Define $T : V \to W$ by $T(p(x)) = \frac{d}{dx}(p(x))$. Consider the basis $B = \{1, x, x^2, x^3\}$ for V. Then the range space of T is given by

$$\mathcal{R}(T) = span\,(T(B)) = span\{T(1), T(x), T(x^2), T(x^3)\} = span\{1, 2x, 3x^2\}$$

Therefore $Rank(T) = 3$ and

$$\mathcal{N}(T) = \left\{p(x) \in \mathbb{P}_3(\mathbb{R}) \mid \frac{d}{dx}(p(x)) = 0\right\} = \{\text{constant polynomials}\}$$

Hence $Nullity(T) = 1$.

Example 3.12 Let $V = \mathbb{M}_{m\times n}(\mathbb{K})$ and $W = \mathbb{M}_{n\times m}(\mathbb{K})$ over the field $\mathbb{K}$. Define $T : V \to W$ by $T(A) = A^T$. In Example 3.6, we have shown that T is a linear transformation. Since for each $A \in \mathbb{M}_{n\times m}(\mathbb{K})$, there exists $A^T \in \mathbb{M}_{m\times n}(\mathbb{K})$ such that $T\left(A^T\right) = \left(A^T\right)^T = A$, $\mathcal{R}(T) = W$ and hence

$$Rank(T) = dim(W) = mn$$

Now

$$T(A) = 0 \Rightarrow A^T = 0 \Rightarrow A = 0$$

Therefore $\mathcal{N}(T) = \{0\}$ and hence $Nullity(T) = 0$.

Example 3.13 Let V and W be any two arbitrary vector spaces over a field $\mathbb{K}$. Consider the identity transformation and zero transformation defined as in Example 3.7. Then $\mathcal{R}(I) = V$ and $\mathcal{N}(I) = \{0\}$. Also $\mathcal{R}(O) = \{0\}$ and $\mathcal{N}(O) = V$.

Observe that in each of the above examples, if you consider the sum of rank and nullity of respective linear transformations, you will get the dimension of V. Is this true in general? Now, we will prove one of the important results in the theory of linear transformations on a finite-dimensional vector space which will answer this question.

Theorem 3.4 *(*Rank-Nullity Theorem*) Let V be a finite-dimensional vector space over a field $\mathbb{K}$ and let W be a vector space over the same field $\mathbb{K}$. Let $T : V \to W$ be a linear transformation, then*

$$Nullity(T) + Rank(T) = dim(V)$$

Proof Let V and W be a finite-dimensional vector spaces over the field $\mathbb{K}$, and let T be a linear transformation from V to W. Let $\mathcal{N}(T)$ be the null space of T. As $\mathcal{N}(T)$ is a subspace of V and V is finite dimensional, by Theorem 2.17, $\mathcal{N}(T)$ is finite dimensional and hence it has a finite basis, say $\{v_1, v_2, \ldots, v_k\}$. Since $\{v_1, v_2, \ldots, v_k\}$ is a linearly independent set in V, then by Theorem 2.14, it can be extended to a basis $B = \{v_1, v_2, \ldots, v_n\}$ of V. We know that

$$\mathcal{R}(T) = span\{T(v_1), T(v_2), \ldots, T(v_n)\}$$

But as $T(v_1) = T(v_2) = \cdots = T(v_k) = 0$,

$$\mathcal{R}(T) = span\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$$

Now we will prove that $\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ is linearly independent. Let $\lambda_{k+1}, \lambda_{k+2}, \ldots, \lambda_n \in \mathbb{K}$ be such that

$$\lambda_{k+1}T(v_{k+1}) + \lambda_{k+2}T(v_{k+2}) + \cdots + \lambda_n T(v_n) = 0$$

which implies

$$T\,(\lambda_{k+1}v_{k+1} + \lambda_{k+2}v_{k+2} + \cdots + \lambda_n v_n) = 0$$

That is, $\lambda_{k+1}v_{k+1} + \lambda_{k+2}v_{k+2} + \cdots + \lambda_n v_n \in \mathcal{N}(T)$. Since $\{v_1, v_2, \ldots, v_k\}$ is a basis of $\mathcal{N}(T)$, there exists $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ such that

$$\lambda_{k+1}v_{k+1} + \lambda_{k+2}v_{k+2} + \cdots + \lambda_n v_n = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k$$

Since B is a basis for V, this implies that $\lambda_i = 0$ for all $i = 1, 2, \ldots, n$. Therefore $\{T(v_{k+1}), T(v_{k+2}), \ldots, T(v_n)\}$ is linearly independent and hence is a basis of $\mathcal{R}(T)$. Therefore $Rank(T) = n - k$. Now

$$Nullity(T) + Rank(T) = k + n - k = n = dim(V)$$

Hence the proof.

We can verify *Rank-Nullity Theorem* for the linear transformations given in Examples 3.10–3.13.

Corollary 3.1 *Let V and W be finite-dimensional vector spaces over a field $\mathbb{K}$ with $dim(V) < dim(W)$, then no linear transformation $T : V \to W$ is onto.*

Proof Let V and W be finite-dimensional vector spaces over a field $\mathbb{K}$ with $dim(V) < dim(W)$. By *Rank-Nullity theorem*, we have

$$Rank(T) = dim(V) - Nullity(T) \leq dim(V) < dim(W)$$

Therefore T cannot be onto.

Interestingly, the one-oneness of a linear transformation is closely related to its null space. The relation is explained in the following theorem.

Theorem 3.5 *Let V and W be vector spaces over a field $\mathbb{K}$ and let $T : V \to W$ be a linear transformation. Then T is one-one if and only if $\mathcal{N}(T) = \{0\}$.*

Proof Let V and W be vector spaces over a field $\mathbb{K}$ and let $T : V \to W$ be a linear transformation. Suppose that T is one-one. That is, $T(v_1) = T(v_2) \Rightarrow v_1 = v_2$ for all $v_1, v_2 \in V$. Let $v \in \mathcal{N}(T)$. Then $T(v) = 0 = T(0) \Rightarrow v = 0$.

Conversely suppose that $\mathcal{N}(T) = \{0\}$. Now $T(v_1) = T(v_2) \Rightarrow T(v_1 - v_2) = 0 \Rightarrow v_1 - v_2 \in \mathcal{N}(T) = \{0\} \Rightarrow v_1 = v_2$. Hence, T is one-one. Therefore T is one-one if and only if $\mathcal{N}(T) = \{0\}$.

This gives an easy way to check whether a linear transformation is one-one or not. Also consider the following corollary, which is an immediate consequence of *Rank-Nullity Theorem* and the above theorem.

Corollary 3.2 *Let V and W be finite-dimensional subspaces over a field $\mathbb{K}$ with $dim(V) > dim(W)$, then no linear transformation $T : V \to W$ is one-one.*

Proof Let V and W be finite-dimensional subspaces over a field $\mathbb{K}$ with $dim(V) > dim(W)$. Since $\mathcal{R}(T)$ is a subspace of W, we have $Rank(T) \leq dim(W)$. Now by *Rank-Nullity theorem*

$$\begin{aligned} Nullity(T) &= dim(V) - Rank(T) \\ &\geq dim(V) - dim(W) \\ &> 0 \end{aligned}$$

Therefore T cannot be one-one.

Example 3.14 Observe that, by Theorem 3.5, the linear transformations defined in Examples 3.10 and 3.12 are one-one, whereas the linear transformation defined in Example 3.11 is not one-one. Also observe that the identity transformation is one-one, but zero transformation is not one-one.

Example 3.15 Let $V = \mathbb{M}_{2\times 2}(\mathbb{R})$ and $W = \mathbb{P}_2[a, b]$. For a matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathbb{M}_{2\times 2}(\mathbb{R})$, define

$$T(A) = a_{11} + (a_{12} + a_{21})\,x + a_{22}x^2$$

Consider the standard ordered basis, $\{E_{11}, E_{12}, E_{21}, E_{22}\}$ of $\mathbb{M}_{2\times 2}(\mathbb{R})$ as defined in Example 2.18. Then, we have

$$T(E_{11}) = 1 \quad T(E_{12}) = T(E_{21}) = x \quad T(E_{22}) = x^2$$

Therefore $\mathcal{R}(T) = span\{1, x, x^2\} = \mathbb{P}_2[a, b]$ and hence $Rank(T) = 3$. Then by *Rank - Nullity Theorem*, $Nullity(T) = 4 - 3 = 1$. Thus T is not one-one.

We have seen that if the dimension of the domain V is greater than the dimension of the co-domain W, there does not exist a one-one linear transformation from V to W. We have also seen that if the dimension of V is less than dimension of W, there does not exist an onto linear map from V to W. But, if the linear transformation is defined between two vector spaces of equal dimension, which is finite, there is no need to distinguish between one-one functions and onto functions. The following theorem states this fact.

Theorem 3.6 *Let V and W be vector spaces over a field $\mathbb{K}$ with equal dimension(finite), and let $T : V \to W$ be a linear transformation. Then T is one-one if and only if T is onto.*

Proof Let V and W be vector spaces over a field $\mathbb{K}$ with $dim(V) = dim(W)$, and let $T : V \to W$ be a linear transformation. Suppose that T is one-one. Then by the above theorem $\mathcal{N}(T) = \{0\}$ and hence $Nullity(T) = 0$. By *Rank-Nullity theorem*, $Rank(T) = dim(V) = dim(W)$. That is, $\mathcal{R}(T)$ is a subspace of V with dimension same as that of W. Therefore $\mathcal{R}(T) = W$ and hence T is onto.

Conversely, suppose that T is onto. Then $\mathcal{R}(T) = W$ and hence $Rank(T) = dim(W) = dim(V)$. Again by *Rank-Nullity theorem*, $\mathcal{N}(T) = \{0\}$ and hence $Nullity(T) = 0$. Hence, T is one-one.

We have seen that a linear transformation maps linearly dependent sets to linearly dependent sets. But for a linearly independent set, this is not necessarily true. For example, let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x_1, y_1) = (x_1, 0)$. Consider the standard ordered basis, $B = \{(1, 0), (0, 1)\}$, for $\mathbb{R}^2$. Then $T(1, 0) = (1, 0)$ and $T(0, 1) = (0, 0)$. Clearly, $T(B) = \{(1, 0), (0, 0)\}$ is a linearly dependent set. Thus a linear transformation may map a linearly independent set to a linearly dependent set.

Theorem 3.7 *Let V and W be vector spaces over a field $\mathbb{K}$. Let $T : V \to W$ be a linear transformation which is one-one. Then a subset S of V is linearly independent if and only if $T(S)$ is linearly independent.*

Proof Let V and W be vector spaces over a field $\mathbb{K}$ and $T : V \to W$ be a one-one linear transformation. Then $\mathcal{N}(T) = \{0\}$. Consider $S = \{v_1, v_2, \ldots, v_n\} \subseteq V$. Suppose that S is linearly independent. Let $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ be such that

$$\lambda_1 T(v_1) + \lambda_2 T(v_2) + \cdots + \lambda_n T(v_n) = 0$$

which implies

$$T(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = 0$$

That is, $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in \mathcal{N}(T) = \{0\}$. Since S is linearly independent, this implies that $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Conversely, suppose that $T(S) = \{T(v_1), T(v_2), \ldots, T(v_n)\}$ is linearly independent. If S is linearly dependent, then there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ (not all zero) such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

Then

$$T(\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n) = T(0) = 0$$

That is, $\lambda_1 T(v_1) + \lambda_2 T(v_2) + \cdots + \lambda_n T(v_n) = 0$, which is a contradiction, since $T(S)$ is linearly independent. Therefore S is linearly independent.

Now, we will prove that the column rank of a matrix is equal to its row rank and we can simply call it the rank of the matrix.

Theorem 3.8 *If $A \in \mathbb{M}_{n\times n}(\mathbb{K})$, then $column\ rank(A) = row\ rank(A)$.*

Proof Define $T : \mathbb{R}^n \to \mathbb{R}^m$ by $T(v) = Av$. In Example 3.3, we have seen that T is a linear transformation. Then, the range set of T is the set of all $b \in \mathbb{R}^m$ such that $Ax = b$, where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. If $A_1, A_2, \ldots, A_n$ denote the columns of A, we can write

$$Ax = x_1 A_1 + x_2 A_2 + \cdots + x_n A_n$$

This implies that the range space of T is spanned by the columns of A. In other words, $\mathcal{R}(T) = Im(A)$. Thus,

$$Rank(T) = column\ rank(A)$$

Also, observe that $\mathcal{N}(T) = Ker(A)$. Then by *Rank-Nullity Theorem*,

$$dim\ (Ker(A)) + column\ rank(A) = n$$

From Sect. 1.7, we have

$$dim\ (Ker(A)) = n - row\ rank(A)$$

Thus, we have $column\ rank(A) = row\ rank(A)$.

Earlier, we have observed some intriguing similarities between the matrix $\begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}$ and the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x_1, x_2) = (2x_1 + y_1, x_1, 3y_1)$, right? But we didn't have any clue how a relation can be drawn between this matrix and the linear transformation T. The next section will give us a good idea regarding this relation.

3.3 Matrix Representation of a Linear Transformation

Let V be an n dimensional vector space over a field $\mathbb{K}$ and W be an m dimensional vector space over $\mathbb{K}$. Let $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_m\}$ be bases of V and W, respectively. Now for each $v \in V$, there exists a unique set of scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$$

Then the matrix $[v]_{B_1} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}_{n \times 1}$ is called the coordinate representation of the vector v with respect to the basis B_1. Let T be a linear transformation from V to W. Now $T(v_1), T(v_2), \ldots, T(v_n)$ are all vectors in W and each can be expressed as a linear combination of basis vectors in B_2. In particular,

$$T(v_1) = a_{11} w_1 + a_{21} w_2 + \cdots + a_{m1} w_m$$

for some scalars $a_{11}, a_{21}, \ldots, a_{m1}$. In general

$$T(v_j) = a_{1j} w_1 + a_{2j} w_2 + \cdots + a_{mj} w_m = \sum_{i=1}^{m} a_{ij} w_i \tag{3.1}$$

for some scalars $a_{1j}, a_{2j}, \ldots, a_{mj}$ $(j = 1, 2, \ldots, m)$. Then the *matrix representation* of T is

$$[T]_{B_1}^{B_2} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}_{m \times n}$$

Now let $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \in V$. Then

$$\begin{aligned}
T(v) &= \lambda_1 T(v_1) + \lambda_2 T(v_2) + \cdots + \lambda_n T(v_n) \\
&= \lambda_1 \left(\sum_{i=1}^{m} a_{i1} w_i \right) + \lambda_2 \left(\sum_{i=1}^{m} a_{i2} w_i \right) + \cdots + \lambda_n \left(\sum_{i=1}^{m} a_{in} w_i \right) \\
&= \sum_{i=1}^{m} (a_{i1}\lambda_1 + a_{i2}\lambda_2 + \cdots + a_{in}\lambda_n) w_i
\end{aligned}$$

Therefore

$$\begin{aligned}
[Tv]_{B_2} &= \begin{bmatrix} a_{11}\lambda_1 + a_{12}\lambda_2 + \cdots + a_{1n}\lambda_n \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \cdots + a_{2n}\lambda_n \\ \vdots \\ a_{m1}\lambda_1 + a_{m2}\lambda_2 + \cdots + a_{mn}\lambda_n \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \\
&= [T]_{B_1}^{B_2} [v]_{B_1}
\end{aligned}$$

Example 3.16 Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ in Example 3.8 defined by $T(x_1, y_1) = (2x_1 + y_1, x_1, 3y_1)$. Consider the standard ordered bases $B_1 = \{(1, 0), (0, 1)\}$ and $B_2 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Then

$$\begin{aligned}
T(1, 0) = (2, 1, 0) = 2(1, 0, 0) + 1(0, 1, 0) + 0(0, 0, 1) \\
T(0, 1) = (1, 0, 3) = 1(1, 0, 0) + 0(0, 1, 0) + 3(0, 0, 1)
\end{aligned}$$

Therefore the matrix of the linear transformation is

$$[T]_{B_1}^{B_2} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Consider the element $v_1 = (2, 5) \in \mathbb{R}^2$. Then $T(v_1) = (9, 2, 15)$. Consider the coordinate representation for both v_1 and $T(v_1)$. We have $[v_1]_{B_1} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ and $[T(v_1)]_{B_2} = \begin{bmatrix} 9 \\ 2 \\ 15 \end{bmatrix}$. Then $[T]_{B_1}^{B_2} [v_1]_{B_1} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 9 \\ 2 \\ 15 \end{bmatrix} = [T(v_1)]_{B_2}$. Similarly, for $v_2 = (1, -1) \in \mathbb{R}^2$, we have $T(v_2) = (1, 1, -3)$. The coordinate representations for both

v_2 and $T(v_2)$ are $[v_2]_{B_1} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $[T(v_2)]_{B_2} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}$, respectively. Then

$$[T]_{B_1}^{B_2} [v_2]_{B_1} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} = [T(v_2)]_{B_2}.$$

Remark 3.2 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then the matrix representation of T is given by $[T]_{B_1}^{B_2} = \begin{bmatrix} | & & | \\ T(e_1) & \dots & T(e_n) \\ | & & | \end{bmatrix}$ where B_1 and B_2 are the standard ordered bases for $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Example 3.17 Let $V = \mathbb{P}_3(\mathbb{R})$ and $W = \mathbb{P}_2(\mathbb{R})$. Define $T : V \to W$ by $(Tp)(x) = \frac{d}{dx}(p(x))$. Consider the bases $B_1 = \{1, x, x^2, x^3\}$ and $B_2 = \{1, x, x^2\}$ for V and W, respectively:

$$\begin{aligned} T(1) &= 0 = 0.1 + 0x + 0x^2 \\ T(x) &= 1 = 1.1 + 0x + 0x^2 \\ T(x^2) &= 2x = 0.1 + 2x + 0x^2 \\ T(x^3) &= 3x^2 = 0.1 + 0x + 3x^2 \end{aligned}$$

Therefore the matrix of the linear transformation is

$$[T]_{B_1}^{B_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Now consider $v = 3x^3 + 2x^2 \in V$. Then $T(v) = 9x^2 + 4x$. Consider the coordinate representation for both v and $T(v)$. We have $[v]_{B_1} = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 3 \end{bmatrix}$ and $[T(v)]_{B_2} = \begin{bmatrix} 0 \\ 4 \\ 9 \end{bmatrix}$.

Then $[T]_{B_1}^{B_2} [v]_{B_1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ 9 \end{bmatrix} = [T(v)]_{B_2}$.

Example 3.18 Let $V = \mathbb{R}^2$ and $W = \mathbb{P}_2[a, b]$. Define $T : V \to W$ by $T(\alpha_1, \alpha_2) = 2\alpha_1 x^2 + (\alpha_1 + \alpha_2)x + 3\alpha_2$ (verify that T is a linear transformation). Consider the bases $B_1 = \{(1, 1), (1, -1)\}$ and $B_2 = \{1, x, x^2\}$ for V and W, respectively:

$$\begin{aligned} T(1, 1) &= 2x^2 + 2x + 3 = 3.1 + 2x + 2x^2 \\ T(1, -1) &= 2x^2 - 3 = (-3).1 + 0x + 2x^2 \end{aligned}$$

Therefore the matrix of the linear transformation is

$$[T]_{B_1}^{B_2} = \begin{bmatrix} 3 & -3 \\ 2 & 0 \\ 2 & 2 \end{bmatrix}$$

Consider $v = (5, 6) \in \mathbb{R}^2$. Then $T(v) = 10x^2 + 11x + 18$. Consider the coordinate representation for both v and $T(v)$. We have $[v]_{B_1} = \begin{bmatrix} \frac{11}{2} \\ \frac{-1}{2} \end{bmatrix}$ and $[T(v)]_{B_2} = \begin{bmatrix} 18 \\ 11 \\ 10 \end{bmatrix}$.
Then $[T]_{B_1}^{B_2} [v]_{B_1} = \begin{bmatrix} 3 & -3 \\ 2 & 0 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \frac{11}{2} \\ \frac{-1}{2} \end{bmatrix} = \begin{bmatrix} 18 \\ 11 \\ 10 \end{bmatrix} = [T(v)]_{B_2}$.

Now that we have got a flavor of the relation between matrices and linear transformations, let us discuss a bit the geometry of linear transformations. It will be easy to visualize the transformation, if it is defined on $\mathbb{R}^2$ or $\mathbb{R}^3$.

Geometry of Linear Transformations on $\mathbb{R}^2$

In the realm of linear transformations in $\mathbb{R}^2$, geometry plays a central role, serving as the primary framework for understanding how these transformations reshape the fundamental properties of points, lines, and shapes within the two-dimensional plane. For instance, rotations can change the orientation of objects, scaling can stretch or shrink them, and reflections can flip them across lines of symmetry. Linear transformations can also introduce shearing effects or map points to new locations entirely. Understanding the geometry of these transformations is crucial for applications in various fields of science and engineering, as it enables us to model and manipulate objects and phenomena in a mathematically rigorous manner. In this section, we will discuss the geometrical aspects of linear transformations in $\mathbb{R}^2$.

1. **Projection**: Consider the linear transformations $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T_1(x_1, x_2) = (x_1, 0) \text{ and } T_2(x_1, x_2) = (0, x_2)$$

The matrix representation of T_1 and T_2 with respect to standard ordered basis B is

$$[T_1]_B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } [T_2]_B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then T_1 is called projection to $x-$ axis and T_2 is called projection to $y-$ axis (Figs. 3.4).

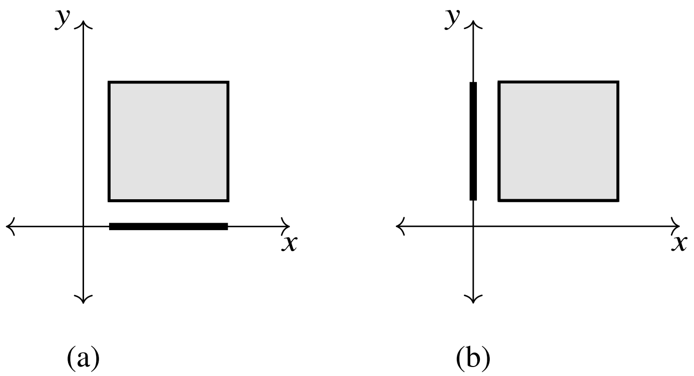


Fig. 3.4 Image of the shaded area under **a** the map T_1 and **b** the map T_2 respectively

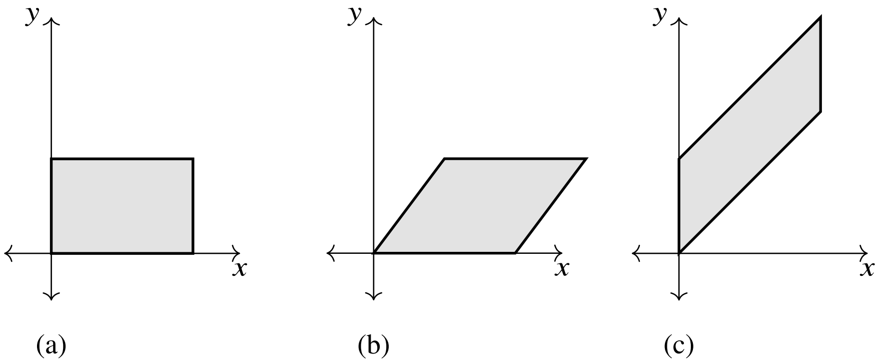


Fig. 3.5 **a** shaded area **b** image under the map T_3 and **c** image under the map $T_4(\lambda = 1$ in both cases)

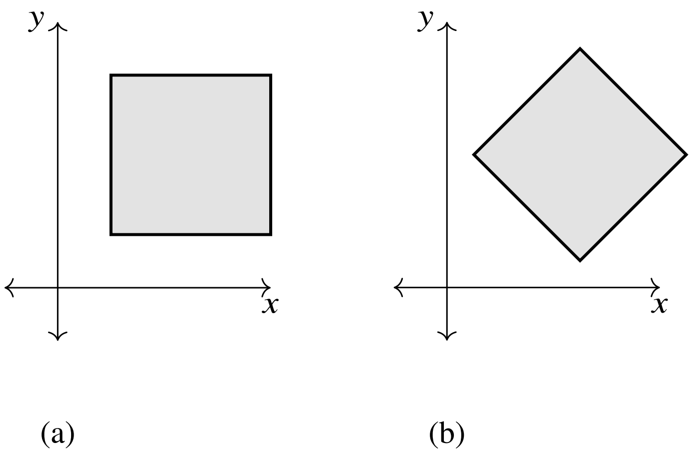


Fig. 3.6 **a** shaded area and **b** image of the shaded area when $\theta = 45°$

2. **Shear**: Consider the linear transformation $T_3, T_4 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T_3(x_1, x_2) = (x_1 + \lambda x_2, x_2) \text{ and} T_4(x_1, x_2) = (x_1, x_1 + \lambda x_2)$$

where $\lambda \in \mathbb{R}$. Then T_1 and T_2 are called horizontal shear and vertical shear, respectively (Fig. 3.5).

3. **Rotation**: Consider the linear transformation defined by

$$T_\theta(x_1, x_2) = (x_1 cos\ \theta - x_2 sin\ \theta, x_1 sin\ \theta + x_2 cos\ \theta)$$

Then T_θ rotates every vector (x_1, x_2) counter clockwise by $\theta°$. The matrix representation of T_θ is the matrix $\begin{bmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{bmatrix}$ (Fig. 3.6).

Remark 3.3 a. $\begin{bmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{bmatrix}^n = \begin{bmatrix} cos\ n\theta & -sin\ n\theta \\ sin\ n\theta & cos\ n\theta \end{bmatrix}$ for every positive integer $n \in \mathbb{Z}$.

b. The basic rotation matrices in $\mathbb{R}^3$ which rotates the vectors by an angle θ about the x, y, z axes in the clockwise directions are

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\ \theta & -sin\ \theta \\ 0 & sin\ \theta & cos\ \theta \end{bmatrix}, \quad R_y = \begin{bmatrix} cos\ \theta & 0 & sin\ \theta \\ 0 & 1 & 0 \\ -sin\ \theta & 0 & cos\ \theta \end{bmatrix},$$

$$R_z = \begin{bmatrix} cos\ \theta & -sin\ \theta & 0 \\ sin\ \theta & cos\ \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

4. **Reflection**: Consider the linear transformation defined by

$$\tilde{T_\theta}(x_1, x_2) = (x_1 cos\ 2\theta + x_2 sin\ 2\theta, x_1 sin\ 2\theta - x_2 cos\ 2\theta)$$

Then $\tilde{T_\theta}$ reflects every vector (x_1, x_2) with respect to a line which makes an angle $\theta°$ in the positive direction of $x-$axis. The matrix representation of $\tilde{T_\theta}$ is the matrix $\begin{bmatrix} cos\ 2\theta & sin\ 2\theta \\ sin\ 2\theta & -cos\ 2\theta \end{bmatrix}$. For example the reflection matrix with respect to x - axis is given by $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and the reflection matrix with respect to y - axis is given by $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ (Fig. 3.7).

5. **Scaling and Contraction**: Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = \lambda(x_1, x_2)$, where $\lambda \in \mathbb{R}$. Then T is called scaling if $\lambda > 1$ and is contraction if $0 < \lambda < 1$ (Fig. 3.8).

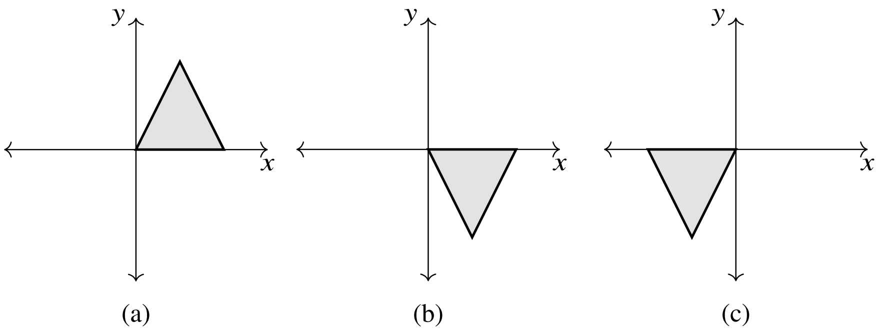


Fig. 3.7 **a** shaded area **b** reflection of the shaded area with respect to x-axis and **c** reflection of the shaded area with respect to y-axis

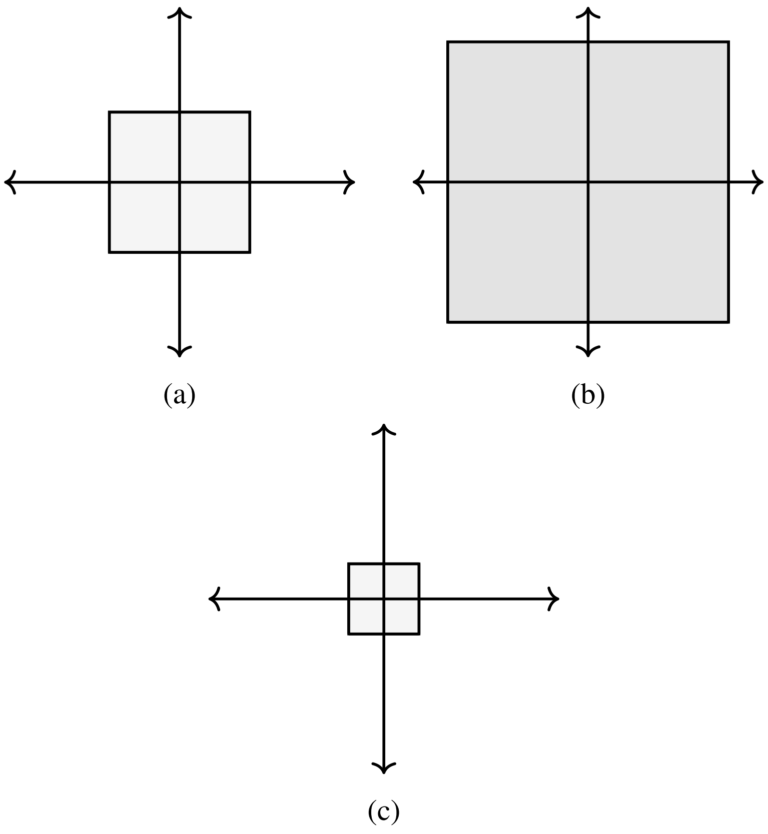


Fig. 3.8 **a** shaded area **b** image of shaded portion under T when $\lambda = 2$ and **c** image of shaded portion under T when $\lambda = \frac{1}{2}$

3.4 Algebra of Linear Transformations

In this section, we will study vector space structure inherited by the set of all linear transformations on a vector space in detail. We will define addition and scalar multiplication in the set of all linear transformations from V to W as in the following theorem. Later, we will also prove that the set of all linear transformations from V to W forms a vector space with respect to these operations.

Theorem 3.9 *Let V and W be vector spaces over the field $\mathbb{K}$. Let T_1 and T_2 be linear transformations from V into W. The function $(T_1 + T_2)$ defined by*

$$(T_1 + T_2)(v) = T_1(v) + T_2(v)$$

is a linear transformation from V into W. If $\xi \in \mathbb{K}$, the function (ξT) defined by $(\xi T)(v) = \xi\,(T(v))$ is a linear transformation from V into W.

Proof Let $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$. Then since T_1 and T_2 are linear transformations from V into W,

$$\begin{aligned}(T_1 + T_2)(\lambda v_1 + v_2) &= T_1(\lambda v_1 + v_2) + T_2(\lambda v_1 + v_2)\\ &= \lambda T_1(v_1) + T_1(v_2) + \lambda T_2(v_1) + T_2(v_2)\\ &= \lambda(T_1(v_1) + T_2(v_1)) + T_1(v_2) + T_2(v_2)\\ &= \lambda(T_1 + T_2)(v_1) + (T_1 + T_2)(v_2)\end{aligned}$$

Therefore $(T_1 + T_2)$ is a linear transformation. Now for any linear transformation T from V into W and $\xi \in \mathbb{K}$,

$$\begin{aligned}(\xi T)(\lambda v_1 + v_2) &= \xi\,(T(\lambda v_1 + v_2))\\ &= \xi\,(\lambda T(v_1) + T(v_2))\\ &= (\xi\lambda)T(v_1) + \xi T(v_2)\\ &= \lambda(\xi T)(v_1) + (\xi T)(v_2)\end{aligned}$$

Therefore ξT is a linear transformation from V into W.

We have shown that a linear transformation can be represented by a matrix. Now let us discuss the relation between the matrices of the linear transformation $T_1 + T_2$ with the matrices of the linear transformation T_1 and T_2. The following theorem also discusses the relation between the matrices of the linear transformations ξT and T, where $\xi \in \mathbb{K}$. We are slowly establishing a relationship between the collection of all linear transformations from an n dimensional vector space V to an m dimensional vector space W and $\mathbb{M}_{m\times n}(\mathbb{K})$.

Theorem 3.10 *Let V and W be finite-dimensional vector spaces over the field $\mathbb{K}$ with ordered bases B_1 and B_2 respectively, and let $T_1, T_2 : V \to W$ be linear transformations. Then*

(a) $[T_1 + T_2]_{B_1}^{B_2} = [T_1]_{B_1}^{B_2} + [T_2]_{B_1}^{B_2}$ *and*
(b) $[\lambda T]_{B_1}^{B_2} = \lambda\, [T]_{B_1}^{B_2}$ *for all* $\lambda \in \mathbb{K}$.

Proof Let V and W be finite-dimensional vector spaces over the field $\mathbb{K}$ with ordered bases $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_m\}$, respectively.

(a) Let $T_1, T_2 : V \to W$ be linear transformations. Then there exist scalars a_{ij} and b_{ij}, where $1 \le i \le m$ and $1 \le j \le n$ such that $T_1(v_j) = \sum_{i=1}^m a_{ij} w_i$ and $T_2(v_j) = \sum_{i=1}^m b_{ij} w_i$. Then

$$(T_1 + T_2)(v_j) = T_1(v_j) + T_2(v_j) = \sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i = \sum_{i=1}^{m} \left(a_{ij} + b_{ij}\right) w_i$$

Therefore

$$\begin{aligned}
[T_1 + T_2]_{B_1}^{B_2} &= \begin{bmatrix} a_{11} + b_{11} & \ldots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \ldots & a_{mn} + b_{mn} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & \ldots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \ldots & b_{mn} \end{bmatrix} \\
&= [T_1]_{B_1}^{B_2} + [T_2]_{B_1}^{B_2}
\end{aligned}$$

(b) Now consider $\lambda T : V \to W$ for $\lambda \in \mathbb{K}$. Since $(\lambda T)(v_j) = \lambda\left(T(v_j)\right)$ and T is a linear transformation from V to W, there exist scalars a_{ij}, where $1 \le i \le m$ and $1 \le j \le n$ such that

$$(\lambda T)(v_j) = \lambda\left(T(v_j)\right) = \lambda \left(\sum_{i=1}^{m} a_{ij} w_i\right) = \sum_{i=1}^{m} \lambda a_{ij} w_i$$

Then

$$[\lambda T]_{B_1}^{B_2} = \begin{bmatrix} \lambda a_{11} & \ldots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \ldots & \lambda a_{mn} \end{bmatrix} = \lambda \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} = \lambda\, [T]_{B_1}^{B_2}$$

Example 3.19 Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Let $T_1, T_2 : V \to W$ be defined by $T_1(x_1, x_2) = (x_2, x_1, 0)$ and $T_2(x_1, x_2) = (x_1, x_2, x_1 + x_2)$ (verify that T_1 and T_2 are linear transformations). Consider the basis $B_1 = \{(1, 0), (0, 1)\}$ for V and $B_2 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ for W. Since

$$T_1(1,0) = (0,1,0) = 0(1,0,0) + 1(0,1,0) + 0(0,0,1)$$

$$T_1(0,1) = (1,0,0) = 1(1,0,0) + 0(0,1,0) + 1(0,0,1)$$

the matrix of T_1 with respect to B_1 and B_2 is $[T_1]_{B_1}^{B_2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$. Since

$$T_2(1,0) = (1,0,1) = 1(1,0,0) + 0(0,1,0) + 1(0,0,1)$$

$$T_2(0,1) = (0,1,1) = 0(1,0,0) + 1(0,1,0) + 1(0,0,1)$$

the matrix of T_2 with respect to B_1 and B_2 is $[T_2]_{B_1}^{B_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$. Now $(T_1 + T_2)(x_1, x_2) = (x_1 + x_2, x_1 + x_2, x_1 + x_2)$. Since

$$(T_1 + T_2)(1,0) = (1,1,1) = 1(1,0,0) + 1(0,1,0) + 1(0,0,1)$$

$$(T_1 + T_2)(0,1) = (1,1,1) = 1(1,0,0) + 1(0,1,0) + 1(0,0,1)$$

the matrix of $(T_1 + T_2)$ with respect to B_1 and B_2 is

$$[T_1 + T_2]_{B_1}^{B_2} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} = [T_1]_{B_1}^{B_2} + [T_2]_{B_1}^{B_2}$$

We have seen that $\mathbb{M}_{m\times n}(\mathbb{K})$ is vector space over the field $\mathbb{K}$ with matrix addition and scalar multiplication. The above theorem relates the matrix addition and scalar multiplication with addition of linear transformations and multiplication of a linear transformation by a scalar. Now we will prove that the collection of all linear transformations from V to W is also a vector space.

Theorem 3.11 *Let V and W be vector spaces over a field $\mathbb{K}$. Then the set of all linear transformations from V to W, denoted by $\mathcal{L}(V, W)$, is a vector space with respect to the addition and scalar multiplication defined as in Theorem 3.9.*

Proof Let $T_1, T_2 \in \mathcal{L}(V, W)$ and $\xi \in \mathbb{K}$. Then by the above theorem $(V1)$ and $(V2)$ are satisfied.

(V3) For any $v \in V$,

$$\begin{aligned} ((T_1 + T_2) + T_3)(v) &= (T_1 + T_2)(v) + T_3(v) \\ &= T_1(v) + T_2(v) + T_3(v) \\ &= T_1(v) + (T_2 + T_3)(v) \\ &= (T_1 + (T_2 + T_3))(v) \end{aligned}$$

That is, $(T_1 + T_2) + T_3 = T_1 + (T_2 + T_3)$.

(V4) The linear transformation $\hat{O}$ acts as the identity element in $\mathcal{L}(V, W)$. For any $T \in \mathcal{L}(V, W)$

$$(\hat{O} + T)(v) = \hat{O}(v) + T(v) = T(v)$$

That is, $\hat{O} + T = T$.

(V5) For any $T \in \mathcal{L}(V, W)$, take $\xi = -1 \in \mathbb{K}$, then by $(V2)$, $-T \in \mathcal{L}(V, W)$ and

$$(T + (-T))(v) = T(v) + ((-T)(v)) = T(v) - T(v) = 0 = \hat{O}(v)$$

That is, $T + (-T) = \hat{O}$.

(V6) For any $v \in V$,

$$(T_1 + T_2)(v) = T_1(v) + T_2(v) = T_2(v) + T_1(v) = (T_2 + T_1)(v)$$

That is, $T_1 + T_2 = T_2 + T_1$.

(V7) For $T_1, T_2 \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{K}$,

$$\lambda [(T_1 + T_2)(v)] = \lambda [T_1(v) + T_2(v)] = \lambda T_1(v) + \lambda T_2(v) = [\lambda T_1 + \lambda T_2](v)$$

That is, $\lambda(T_1 + T_2) = \lambda T_1 + \lambda T_2$.

(V8) For $\lambda, \mu \in \mathbb{K}$ and $T \in \mathcal{L}(V, W)$,

$$[(\lambda + \mu) T](v) = (\lambda + \mu)(T(v)) = \lambda(T(v)) + \mu(T(v)) = (\lambda T)(v) + (\mu T)(v) = (\lambda T + \mu T)(v)$$

That is, $(\lambda + \mu) T = \lambda T + \mu T$.

(V9) For $\lambda, \mu \in \mathbb{K}$ and $T \in \mathcal{L}(V, W)$,

$$[(\lambda\mu) T](v) = (\lambda\mu)(T(v)) = \lambda [(\mu T)](v)$$

That is, $(\lambda\mu) T = \lambda(\mu T)$

(V10) Now $(1T)(v) = 1(T(v)) = T(v) \Rightarrow (1T) = 1(T)$.

Thus conditions $(V1) - (V10)$ are satisfied. Therefore $\mathcal{L}(V, W)$ is a vector space over the field $\mathbb{K}$.

Theorem 3.12 *Let U, V, and W be vector spaces over the field $\mathbb{K}$. Let T_1 be a linear transformation from U into V and T_2 a linear transformation from V into W. Then the composed function T_2T_1 defined by $(T_2T_1)(u) = T_2(T_1(u))$ is a linear transformation from U to W.*

Proof Let $u_1, u_2 \in U$ and $\lambda \in \mathbb{K}$. Then

$$
\begin{aligned}
(T_2T_1)(\lambda u_1 + u_2) &= T_2\left(T_1(\lambda u_1 + u_2)\right) \\
&= T_2\left(\lambda T_1(u_1) + T_2(u_2)\right) \\
&= \lambda\left[T_2\left(T_1(u_1)\right)\right] + T_2\left(T_1(u_2)\right) \\
&= \lambda(T_2T_1)(u_1) + (T_2T_1)(u_2)
\end{aligned}
$$

Therefore T_2T_1 is a linear transformation from V to W.

Now we will prove that the matrix of the composition of two linear transformations is analogous to the product of matrices of the transformations.

Theorem 3.13 *Let U, V, and W be finite-dimensional vector spaces with ordered bases B_1, B_2, and B_3, respectively. Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Then*

$$[T_2T_1]_{B_1}^{B_3} = [T_2]_{B_2}^{B_3}\,[T_1]_{B_1}^{B_2}$$

Proof Let U, V, and W be vector spaces over a field $\mathbb{K}$ with bases $B_1 = \{u_1, u_2, \ldots, u_m\}$, $B_2 = \{v_1, v_2, \ldots, v_n\}$, and $B_3 = \{w_1, w_2, \ldots, w_p\}$, respectively. Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Then there exist scalars a_{ij}, where $1 \le i \le n$ and $1 \le j \le m$ such that $T_1(u_j) = \sum_{i=1}^{n} a_{ij} v_i$. Now

$$(T_2T_1)(u_j) = T_2\big(T_1(u_j)\big) = T_2\left(\sum_{i=1}^{n} a_{ij} v_i\right) = \sum_{i=1}^{n} a_{ij} T_2(v_i)$$

Now there exists scalars b_{ki}, where $1 \le k \le p$ such that $T_2(v_i) = \sum_{k=1}^{p} b_{ki} w_k$. Therefore

$$(T_2T_1)(u_j) = \sum_{i=1}^{n} a_{ij} T_2(v_i) = \sum_{i=1}^{n} a_{ij}\left(\sum_{k=1}^{p} b_{ki} w_k\right) = \sum_{k=1}^{p}\left(\sum_{i=1}^{n} b_{ki} a_{ij}\right) w_k = \sum_{k=1}^{p} c_{kj} w_k$$

where $c_{kj} = \sum_{i=1}^{n} b_{ki} a_{ij}$. Therefore

$$
\begin{aligned}
[T_2T_1]_{B_1}^{B_3} &= \begin{bmatrix} \sum_{i=1}^{n} b_{1i}a_{i1} & \sum_{i=1}^{n} b_{1i}a_{i2} & \ldots & \sum_{i=1}^{n} b_{1i}a_{in} \\ \sum_{i=1}^{n} b_{2i}a_{i1} & \sum_{i=1}^{n} b_{2i}a_{i2} & \ldots & \sum_{i=1}^{n} b_{2i}a_{in} \\ \vdots & \vdots & \ldots & \vdots \\ \sum_{i=1}^{n} b_{pi}a_{i1} & \sum_{i=1}^{n} b_{pi}a_{i2} & \ldots & \sum_{i=1}^{n} b_{pi}a_{in} \end{bmatrix} \\
&= \begin{bmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ b_{p1} & b_{p2} & \ldots & b_{pn} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ldots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix} \\
&= [T_2]_{B_2}^{B_3}\,[T_1]_{B_1}^{B_2}
\end{aligned}
$$

Example 3.20 Let $T_1, T_2 : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T_1(x_1, x_2, x_3) = (0, x_1, x_2)$ and $T_2(x_1, x_2, x_3) = (x_2, x_3, x_1)$ (verify that T_1 and T_2 are linear transformations). Consider the standard ordered basis B for $\mathbb{R}^3$. Then $[T_1]_B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $[T_2]_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Now $T_2T_1 : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$(T_2T_1)(x_1, x_2, x_3) = T_2\,(T_1(x_1, x_2, x_3)) = T_2\,(0, x_1, x_2) = (x_1, x_2, 0)$$

Then $[T_2T_1]_B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = [T_2]_B\,[T_1]_B$. And $T_1T_2 : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$(T_1T_2)(x_1, x_2, x_3) = T_1\,(T_2(x_1, x_2, x_3)) = T_1\,(x_2, x_3, x_1) = (0, x_2, x_3)$$

Also $[T_1T_2]_B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = [T_1]_B\,[T_2]_B$.

Example 3.21 Let $T_1 : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ and $T_2 : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the linear transformations defined by

$$(T_1p)\,(x) = \int_0^x p(t)dt \quad \text{and} \quad (T_2p)\,(x) = \frac{d}{dx}(p(x))(x)$$

Consider the bases $B_1 = \{1, x, x^2\}$ and $B_2 = \{1, x, x^2, x^3\}$. Then $[T_1]_{B_1}^{B_2} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$ and $[T_2]_{B_2}^{B_1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$. Now T_2T_1 is a linear transformation from $\mathbb{P}_2(\mathbb{R})$ to $\mathbb{P}_2(\mathbb{R})$. From calculus, we know that $T_2T_1 = I$. Also $[T_2]_{B_2}^{B_1}\,[T_1]_{B_1}^{B_2} = I_3$.

Now we will prove some properties of linear transformations which are analogous to properties satisfied by matrices.

Theorem 3.14 *Let V be a vector space over the field $\mathbb{K}$. Let T, T_1, and T_2 be linear operators on V, and let λ be an element of $\mathbb{K}$. Then*

(a) $IT = TI = T$.
(b) $T(T_1 + T_2) = TT_1 + TT_2$ *and* $(T_1 + T_2)T = T_1T + T_2T$.
(c) $\lambda(T_1T_2) = (\lambda T_1)T_2 = T_1(\lambda T_2)$.

Proof (a) For any $v \in V$, we have

$$(IT)(v) = I(T(v)) = T(v) = T(I(v)) = (TI)(v)$$

Hence, $IT = TI = T$.

(b) For any $v \in V$,

$$\begin{aligned}[T(T_1 + T_2)](v) &= T[(T_1 + T_2)(v)] \\ &= T[T_1(v) + T_2(v)] \\ &= T(T_1(v)) + T(T_2(v)) \\ &= (TT_1)(v) + (TT_2)(v)\end{aligned}$$

Hence, $T(T_1 + T_2) = TT_1 + TT_2$. Also

$$\begin{aligned}[(T_1 + T_2)T](v) &= (T_1 + T_2)(T(v)) \\ &= T_1(T(v)) + T_2(T(v)) \\ &= (T_1T)(v) + (T_2T)(v)\end{aligned}$$

Hence, $(T_1 + T_2)T = T_1T + T_2T$.

(c) For any $v \in V$,

$$[\lambda(T_1T_2)](v) = \lambda[(T_1T_2)(v)] = \lambda[T_1(T_2(v))] = (\lambda T_1)(T_2(v)) = [(\lambda T_1)T_2](v)$$

Hence, $\lambda(T_1T_2) = (\lambda T_1)T_2$. Also

$$T_1[(\lambda T_2)(v)] = T_1[\lambda(T_2(v))] = \lambda[T_1(T_2(v))] = \lambda[(T_1T_2)(v)] = [\lambda(T_1T_2)](v)$$

Hence, $\lambda(T_1T_2) = T_1(\lambda T_2)$.

3.5 Invertible Linear Transformations

Think of a magical device that can transform an object in space into any shape and rotate, squash, or stretch. This device is an illustration of a linear transformation. Now imagine that you stretched out a square into a rectangle with this machine. You would need a second magical device that works opposite to the first one if you wished to return to the initial square. All the spinning, stretching, and squashing that the first machine did can be undone by this amazing device. For these magical devices, the inverse of a linear transformation functions as the "undo" button. Being able to return to your starting point makes it an effective tool in mathematics and science, particularly when working with transformations in the realm of matrices and vectors (Fig. 3.9).

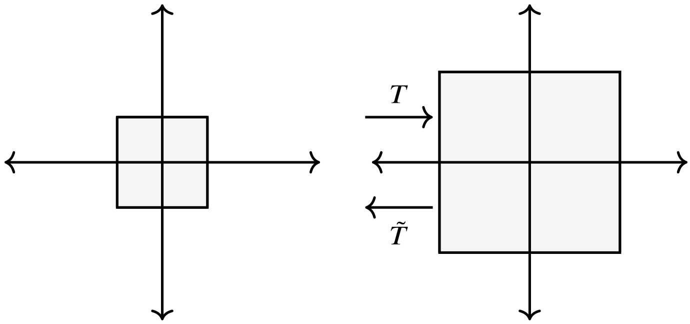


Fig. 3.9 Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = 2(x_1, x_2)$. If we define $\tilde{T}(x_1, x_2) = \frac{1}{2}(x_1, x_2)$, then we can "undo" the action of T and *vice versa*

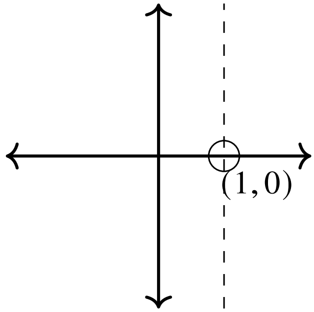


Fig. 3.10 We can see that all the points on the line $x = 1$ are mapped onto the point $(1, 0)$. Observe that reversing this action will not produce a function as $(1, 0)$ should be mapped onto all the points on the line $x = 1$

Now, one crucial question that comes to our mind is whether it is possible to reverse the action of every linear transformation. For example, consider the projection map onto the x-axis (Fig.3.10).

From the above figure, it is clear that the transformation has to be one-one in order to discuss its inverse. Also, if T is onto, we can take W as the domain of our inverse function. Then, we have the following definition for the inverse of a linear transformation.

Definition 3.3 (*Inverse*) Let V and W be vector spaces over the field $\mathbb{K}$. Let T be a linear transformation from V into W. A function $\tilde{T}$ from W into V is said to be an inverse of T if $\tilde{T}T = I_V$, the identity function on V and $T\tilde{T} = I_W$, the identity function on W. Furthermore, T is invertible if and only if T is both one-one and onto. The inverse of a linear transformation T is denoted by T^{-1}.

Example 3.22 Let $V = \mathbb{R}^2$ with basis $B = \{(1, 0), (0, 1)\}$. Define $T, \tilde{T} : V \to V$ by $T(x_1, x_2) = (x_1, x_1 + x_2)$ and $\tilde{T}(x_1, x_2) = (x_1, x_2 - x_1)$. Clearly both T and $\tilde{T}$ are linear transformations. Now

$$T\tilde{T}(x_1, x_2) = T(x_1, x_2 - x_1) = (x_1, x_2)$$

$$\tilde{T}T(x_1, x_2) = \tilde{T}(x_1, x_1 + x_2) = (x_1, x_2)$$

Thus, $\tilde{T}T = T\tilde{T} = I$. Therefore T is an invertible linear transformation on $\mathbb{R}^2$ with $T^{-1} = \tilde{T}$.

Now that we have defined the inverse of a linear transformation, other important questions to be answered are that if the inverse of a linear transformation exists, will it be linear and unique?

Theorem 3.15 *Let V and W be vector spaces over a field $\mathbb{K}$, and let $T : V \to W$ be an invertible linear transformation. Then $T^{-1} : W \to V$ is linear and unique.*

Proof Let $w_1, w_2 \in W$. Since T is both one-one and onto, there exist unique vectors $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Then $T^{-1}(w_1) = v_1$ and $T^{-1}(w_2) = v_2$. Now for $\lambda \in \mathbb{K}$, $\lambda w_1 + w_2 \in W$, with $T(\lambda v_1 + v_2) = \lambda w_1 + w_2$. Then

$$T^{-1}(\lambda w_1 + w_2) = \lambda v_1 + v_2 = \lambda T^{-1}(w_1) + T^{-1}(w_2)$$

Therefore T^{-1} is a linear transformation from W to V. Now suppose that there exists two functions $\tilde{T}_1, \tilde{T}_2 : W \to V$ such that $T\tilde{T}_1 = I_W = T\tilde{T}_2$ and $\tilde{T}_1 T = I_V = \tilde{T}_2 T$. Now

$$\tilde{T}_1 = \tilde{T}_1 I_W = \tilde{T}_1(T\tilde{T}_2) = (\tilde{T}_1 T)\tilde{T}_2 = I_V \tilde{T}_2 = \tilde{T}_2$$

That is, inverse of a linear transformation, if it exists, is unique.

The idea of inverse of a linear transformation can be used to identify the similarity between vector spaces. For example, a vector $(x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ can be observed as a 2×2 matrix $\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$ in $\mathbb{M}_{2\times 2}(\mathbb{R})$ as their vector addition and scalar multiplication can be associated in an identical manner. Such vector spaces are said to be isomorphic. An isomorphism of vector spaces is similar to discovering a link between two universes that allows them to be interpreted as identical despite their differences in appearance. This correspondence is more than just matching elements; it is a unique relationship in which operations such as addition and scalar multiplication in one space completely mirror those in the other. Isomorphisms, in essence, indicate that these seemingly separate vector spaces are fundamentally the same in terms of their underlying algebraic features, making them a valuable notion for reducing complex issues and integrating diverse areas of mathematics and science.

Definition 3.4 (*Isomorphism*) Let V and W be vector spaces over the field $\mathbb{K}$, then V is said to be isomorphic to W if there exists an invertible linear transformation from V to W. That is, if there exists a one-one and onto linear transformation from V to W.

Example 3.23 Let $V = \mathbb{P}_2(\mathbb{R})$ and $W = \mathbb{R}^3$ over the field $\mathbb{R}$. Define $T : V \to W$ by

$$T(ax^2 + bx + c) = (a, b, c)$$

Let $p(x) = a_1x^2 + b_1x + c_1, q(x) = a_2x^2 + b_2x + c_2 \in \mathbb{P}_2(\mathbb{R})$ and $\lambda \in \mathbb{R}$. Now $\lambda p(x) + q(x) = (\lambda a_1 + a_2)x^2 + (\lambda b_1 + b_2)x + (\lambda c_1 + c_2)$ and

$$\begin{aligned} T(\lambda p(x) + q(x)) &= (\lambda a_1 + a_2, \lambda b_1 + b_2, \lambda c_1 + c_2) \\ &= \lambda(a_1, b_1, c_1) + (a_2, b_2, c_2) = \lambda T(p(x)) + T(q(x)) \end{aligned}$$

Therefore T is a Linear transformation from V to W. Now

$$\begin{aligned} \mathcal{N}(T) &= \{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid T(p(x)) = (0, 0, 0)\} \\ &= \{p(x) \in \mathbb{P}_2(\mathbb{R}) \mid (a_1, b_1, c_1) = (0, 0, 0)\} = \{0\} \end{aligned}$$

Since $\mathcal{N}(T) = \{0\}$, T is one-one. Since V and W are of equal dimension, T is onto. Hence, T is an isomorphism from $\mathbb{P}_2(\mathbb{R})$ to $\mathbb{R}^3$.

Now we will prove that any two finite-dimensional vector spaces with equal dimension are isomorphic.

Theorem 3.16 *Let V and W be finite-dimensional vector spaces over a field $\mathbb{K}$, then V is isomorphic to W if and only if $dim(V) = dim(W)$.*

Proof Let V and W be vector spaces over a field $\mathbb{K}$. Suppose that V is isomorphic to W. Then there exists an invertible linear transformation T from V to W. Since T is one-one, by Theorem 3.5, $Nullity(T) = 0$ and as T is onto $\mathcal{R}(T) = W$. Then by *Rank-Nullity theorem*,

$$dim(V) = Rank(T) = dim(W)$$

Conversely, suppose that $dim(V) = dim(W)$. Let $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_n\}$ be bases for V and W, respectively. By Theorem 3.2, there exists a linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, \ldots, n$. Then

$$\mathcal{R}(T) = span\{T(v_1), T(v_2), \ldots T(v_n)\} = \{w_1, w_2, \ldots, w_n\} = W$$

Therefore T is onto. Since $dim(V) = dim(W)$, T is also one-one. Therefore T is an isomorphism.

Corollary 3.3 *Let V be a vector space over the field $\mathbb{K}$ with $dim(V) = n$. Then V is isomorphic to $\mathbb{K}^n$ over $\mathbb{K}$.*

Theorem 3.17 *Let V and W be vector spaces over a field $\mathbb{K}$ with bases B_1 and B_2, respectively. Let $T : V \to W$ be a linear transformation. Then T is invertible if and only if $[T]_{B_1}^{B_2}$ is invertible. Furthermore,*

$$\left[T^{-1}\right]_{B_2}^{B_1} = \left([T]_{B_1}^{B_2}\right)^{-1}$$

Proof Let V and W be vector spaces over a field $\mathbb{K}$, and let $T : V \to W$ be an invertible linear transformation. Then there exists a linear transformation $T^{-1} : W \to V$ such that $TT^{-1} = I_W$, the identity transformation on W and $T^{-1}T = I_V$, the identity transformation on V. Also by the above theorem, we have $dim(V) = dim(W)$ and hence $[T]_{B_1}^{B_2}$ is an $n \times n$ matrix. Therefore

$$I_n = [I_V]_{B_1} = \left[T^{-1}T\right]_{B_1} = \left[T^{-1}\right]_{B_2}^{B_1} [T]_{B_1}^{B_2}$$

Also

$$I_n = [I_W]_{B_2} = \left[TT^{-1}\right]_{B_2} = [T]_{B_1}^{B_2} \left[T^{-1}\right]_{B_2}^{B_1}$$

Therefore $\left[T^{-1}\right]_{B_2}^{B_1} = \left([T]_{B_1}^{B_2}\right)^{-1}$. Now suppose that $A = [T]_{B_1}^{B_2}$ is invertible. Then there exists a matrix $B = \left[b_{ij}\right]$ such that $AB = I_n = BA$. Let $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_n\}$ be bases for V and W, respectively. Now define $\tilde{T} : W \to V$ such that $\tilde{T}(w_j) = \sum_{i=1}^{n} b_{ij} v_i$ for $i = 1, 2, \ldots n$. Then $\left[\tilde{T}\right]_{B_2}^{B_1} = B$. Therefore

$$\left[\tilde{T}T\right]_{B_1} = \left[\tilde{T}\right]_{B_2}^{B_1} [T]_{B_1}^{B_2} = BA = I_n = [I_V]_{B_1}$$

and

$$\left[T\tilde{T}\right]_{B_1} = [T]_{B_1}^{B_2} \left[\tilde{T}\right]_{B_2}^{B_1} = AB = I_n = [I_W]_{B_1}$$

Hence, T is invertible with $T^{-1} = \tilde{T}$.

Example 3.24 Consider Example 3.22. Now consider the matrix representation of both T and $\tilde{T}$:

$$[T]_B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad and \quad [\tilde{T}]_B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

Clearly $\left[\tilde{T}\right]_B = ([T]_B)^{-1}$.

We have seen that corresponding to every linear transformation from an n dimensional vector space V to an m dimensional vector space W, there exists a matrix representation $[T]_{B_1}^{B_2}$, where B_1 and B_2 are bases of V and W, respectively. Now we will prove that the space of all linear transformations from V to W is identical to the space of all $m \times n$ matrices.

Theorem 3.18 *Let V and W be vector spaces over $\mathbb{K}$ with bases $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_m\}$, respectively. For each linear transformation T from V into W, there is an $m \times n$ matrix $[T]_{B_1}^{B_2}$ with entries in $\mathbb{K}$ such that $[Tv]_{B_2} = [T]_{B_1}^{B_2}[v]_{B_1}$ for all $v \in V$. Then the function $\Phi : \mathcal{L}(V, W) \to \mathbb{M}_{m\times n}(\mathbb{K})$, defined by $\Phi(T) = [T]_{B_1}^{B_2}$, is an isomorphism.*

Proof Let V and W be vector spaces over $\mathbb{K}$ with bases $B_1 = \{v_1, v_2, \ldots, v_n\}$ and $B_2 = \{w_1, w_2, \ldots, w_m\}$, respectively. Define $\Phi : \mathcal{L}(V, W) \to \mathbb{M}_{m\times n}(\mathbb{K})$ by $\Phi(T) = [T]_{B_1}^{B_2}$. By Theorem 3.10, Φ is a linear transformation. Now we have to prove that Φ is both one-one and onto. It is enough to show that, for any matrix $A \in \mathbb{M}_{m\times n}(\mathbb{K})$, there exists a unique $T \in \mathcal{L}(V, W)$ such that $\Phi(T) = A$. Let $M = \left[a_{ij}\right] \in \mathbb{M}_{m\times n}(\mathbb{K})$. Define $T : V \to W$ by $T(v_j) = \sum_{i=1}^{m} a_{ij} w_i$ where $1 \le j \le n$. Then $[T]_{B_1}^{B_2} = A$, and hence $\Phi(T) = A$. That is, Φ is onto. By Theorem 3.2, such a T is unique. Hence Φ is one-one. Therefore $\mathcal{L}(V, W)$ is isomorphic to $\mathbb{M}_{m\times n}(\mathbb{K})$.

Thus, the space of all linear transformations from an n dimensional space V to an m dimensional space W is a vector space of dimension mn.

3.6 Change of Coordinate Matrix

We know that basis of a vector space is not unique. The coordinate representation of a vector depends on the basis that we are choosing, i.e., the same vector can have different representations when we are choosing different bases. Therefore the same linear transformation can be represented by different matrices depending upon the basis. In this section, we will discuss the relation between these representations.

Theorem 3.19 *Let V be a finite-dimensional vector space over a field $\mathbb{K}$ with ordered bases B_1 and B_2, and let $P = [I_V]_{B_1}^{B_2}$ where I_V is the identity transformation on V. Then*

(a) P is invertible.
(b) For any $v \in V$, $[v]_{B_2} = P\,[v]_{B_1}$.
(c) If T is a linear operator on V, then $[T]_{B_2} = P^{-1}\,[T]_{B_1}\,P$.

P is called the change of coordinate matrix as P changes B_1 coordinates into B_2 coordinates and P^{-1} changes B_2 coordinates into B_1 coordinates.

Proof (a) Since I_V is invertible, by Theorem 3.17, P is invertible.
(b) For any $v \in V$,

$$[v]_{B_2} = [I_V(v)]_{B_2} = [I_V]_{B_1}^{B_2}\,[v]_{B_1} = P\,[v]_{B_1}$$

(c) Since $I_V T = T = T I_V$, by Theorem 3.13, we have

$$P\,[T]_{B_1} = [I_V]_{B_1}^{B_2}\,[T]_{B_1}^{B_1} = [I_V T]_{B_1}^{B_2} = [T I_V]_{B_1}^{B_2} = [T]_{B_2}^{B_2}\,[I_V]_{B_1}^{B_2} = [T]_{B_2}\,P$$

Since P is invertible, $[T]_{B_2} = P^{-1}\,[T]_{B_1}\,P$.

Example 3.25 Let $V = \mathbb{R}^2$. Consider the bases $B_1 = \{(1, 0), (0, 1)\}$, $B_2 = \{(1, 1), (1, -1)\}$ for V and $v = (5, 2)$. Then $[v]_{B_1} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $[v]_{B_2} = \begin{bmatrix} \frac{7}{2} \\ \frac{3}{2} \end{bmatrix}$. Since

$$(1,0) = \frac{1}{2}(1,1) + \frac{1}{2}(1,-1)$$

$$(0,1) = \frac{1}{2}(1,1) + \left(-\frac{1}{2}\right)(1,-1)$$

the change of coordinate matrix $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$ and $P^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Clearly

$$P\,[v]_{B_1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{7}{2} \\ \frac{3}{2} \end{bmatrix} = [v]_{B_2}$$

and

$$P^{-1}\,[v]_{B_2} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \frac{7}{2} \\ \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} = [v]_{B_1}$$

Now consider the linear transformation $T : V \to V$ defined by

$$T(x_1, x_2) = (x_2, x_1)$$

Since $T(1,0) = (0,1) = 0(1,0) + 1(0,1)$ and $T(0,1) = (1,0) = 1(1,0) + 0(0,1)$. The matrix of T with respect to B_1 is $[T]_{B_1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Since $T(1,1) = (1,1) = 1(1,1) + 0(1,-1)$ and $T(1,-1) = (-1,1) = 0(1,1) + (-1)(1,-1)$ The matrix of T with respect to B_2 is $[T]_{B_2} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Now

$$P^{-1}\,[T]_{B_1}\,P = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = [T]_{B_2}$$

3.7 Linear Functionals and Dual Space

So far, we have discussed linear transformations between vector spaces. In this section, we will discuss linear transformations which are defined from a vector space to the field associated with it.

Definition 3.5 (*Linear Functionals*) Let V be a vector space over the field $\mathbb{K}$. A function $f : V \to \mathbb{K}$ is said to be a linear functional, if

$$f(\lambda v_1 + v_2) = \lambda f(v_1) + f(v_2)$$

for all $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$. The set of all linear functionals on V forms a vector space, called the **dual space of** V, and is denoted by V^*.

Example 3.26 Let V be a vector space over the field $\mathbb{K}$. Clearly, the map f defined by $f(v) = 0,\ \forall\, v \in V$ is a linear functional on V.

Example 3.27 Let $V = \mathbb{M}_{n\times n}(\mathbb{R})$. Define a map $f : \mathbb{M}_{n\times n}(\mathbb{R}) \to \mathbb{R}$ by $f(A) = Tr(A)$. As we have already seen that Tr is a linear function, Tr is a linear functional on $\mathbb{M}_{n\times n}(\mathbb{R})$.

Example 3.28 Let $V = \mathbb{P}_n[a, b]$. Define a map $f : \mathbb{P}_n[a, b] \to \mathbb{R}$ by $f(p) = p(0)$. As

$$f(\lambda p + q) = (\lambda p + q)(0) = \lambda p(0) + q(0) = \lambda f(p) + f(q)$$

for all $p, q \in \mathbb{P}_n[a, b]$ and $\lambda \in \mathbb{R}$, f is a linear functional on $\mathbb{P}_n[a, b]$.

Example 3.29 Let $V = C[a, b]$. Define a map $f : C[a, b] \to \mathbb{R}$ defined by $f(p) = \int_a^b p(x)dx$. As

$$f(\lambda p + q) = \int_a^b (\lambda p + q)(x)dx = \lambda \int_a^b p(x)dx + \int_a^b q(x)dx = \lambda f(p) + f(q)$$

for all $p, q \in C[a, b]$ and $\lambda \in \mathbb{R}$, f is a linear functional on $C[a, b]$.

We already know that the set of all linear functionals on a vector space V forms a vector space, called the *dual space* of V. If V is a finite-dimensional vector space, we can get a rather explicit description of the dual space. Consider the following theorem.

Theorem 3.20 *Let V be a finite-dimensional vector space over the field $\mathbb{K}$, and let $B = \{v_1, v_2, \ldots, v_n\}$ be a basis of V. Then there exists a unique basis $B^* = \{f_1, f_2, \ldots f_n\}$ for V^*, where f_i is given by*

$$f_i(v_j) = \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j \end{cases}$$

Then for each linear functional f on V, we have $f = \sum_{i=1}^n f(v_i) f_i$ and for each vector $v \in V$, we have $v = \sum_{i=1}^n f_i(v) v_i$.

Proof First we will prove that $B^* = \{f_1, f_2, \ldots f_n\}$, where f_i is given by

$$f_i(v_j) = \begin{cases} 1, & if\ i = j \\ 0, & if\ i \neq j \end{cases}$$

is linearly independent. Let $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ be such that

$$\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n = 0$$

i.e., $(\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n)(v) = 0$ for all $v \in V$. Then,

$$(\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_n f_n)(v_i) = 0 \Rightarrow \lambda_i = 0,\ \forall\, i = 1, 2 \ldots, n$$

Thus $\{f_1, f_2, \ldots f_n\}$ is linearly independent. Also by Theorem 3.18, we have $dim(V) = dim(V^*)$. Hence $B^* = \{f_1, f_2, \ldots f_n\}$ is a basis for V^*, called the **dual basis** of B. By definition itself B^* is unique. Now, for any linear functional $f \in V^*$, there exists $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ such that $f = \sum_{i=1}^{n} \lambda_i f_i$. Then

$$f(v_j) = \sum_{i=1}^{n} \lambda_i f_i(v_j) = \lambda_j,\ \forall\, j = 1, 2, \ldots, n$$

Thus for each linear functional f on V, we have $f = \sum_{i=1}^{n} f(v_i) f_i$. Similarly, for each $v \in V$, there exist scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ such that $v = \sum_{i=1}^{n} \lambda_i v_i$. Then

$$f_j(v) = f_j\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i f_j(v_i) = \lambda_j,\ \forall\, j = 1, 2, \ldots, n$$

Hence for any $v \in V$, we have $v = \sum_{i=1}^{n} f_i(v) v_i$.

Consider the following example for a better understanding of the above theorem.

Example 3.30 Let $V = \mathbb{R}^2$ and consider a basis $B = \{(1, 2), (2, 2)\}$. Now, let us find the dual basis $B^* = \{f_1, f_2\}$ corresponding to V. By definition, $f_1(1, 2) = 1$ and $f_1(2, 2) = 0$. Then

$$f_1(1, 2) = 1 \Leftrightarrow f_1\,(1(1, 0) + 2(0, 1)) = f_1(1, 0) + 2 f_1(0, 1) = 1$$

$$f_1(2, 2) = 0 \Leftrightarrow f_1\,(2(1, 0) + 2(0, 1)) = 2 f_1(1, 0) + 2 f_1(0, 1) = 0$$

This implies that $f_1(1, 0) = -1$ and $f_1(0, 1) = 1$. Thus, we have

$$f_1(x_1, x_2) = x_1 f_1(1, 0) + x_2 f_1(0, 1) = x_2 - x_1$$

Similarly, we have ,$f_2(1, 2) = 0$ and $f_2(2, 2) = 1$. Then

$$f_2(1, 2) = 0 \Leftrightarrow f_2\,(1(1, 0) + 2(0, 1)) = f_2(1, 0) + 2 f_2(0, 1) = 0$$

$$f_2(2, 2) = 1 \Leftrightarrow f_2\,(2(1, 0) + 2(0, 1)) = 2 f_2(1, 0) + 2 f_2(0, 1) = 1$$

This implies that $f_2(1, 0) = 1$ and $f_2(0, 1) = \frac{-1}{2}$. Thus, we have

$$f_2(x_1, x_2) = x_1 f_2(1, 0) + x_2 f_2(0, 1) = x_1 - \frac{x_2}{2}$$

Thus $B^* = \{x_2 - x_1, x_1 - \frac{x_2}{2}\}$ forms the dual basis corresponding to $B = \{(1, 2), (2, 2)\}$. Now consider a linear functional $f(x_1, x_2) = 2x_1 - 3x_2$ on $\mathbb{R}^2$. Then we can write f as a linear combination of the dual basis elements. Observe that

$$f(1, 2)f_1 + f(2, 2)f_2 = -4(x_2 - x_1) - 2\left(x_1 - \frac{x_2}{2}\right) = 2x_1 - 3x_2 = f(x_1, x_2)$$

Also, the coordinates of a vector relative to the basis can be obtained using the dual basis. For example, $(3, 4) \in \mathbb{R}^2$,

$$f_1(3, 4)(1, 2) + f_2(3, 4)(2, 2) = 1(1, 2) + 1(2, 2) = (3, 4)$$

The above theorem gives a good description of the dual basis B^* corresponding to a basis B of V. If $B = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for a vector space V, $B^* = \{f_1, f_2, \ldots f_n\}$ is the dual basis, where f_i is the function which assigns to each vector v in V the ith coordinate of v relative to the ordered basis B. Thus if $f \in V^*$, and we have $f(v_i) = \mu_i$, then for $v = \sum_{i=1}^{n} \lambda_i v_i$, we have $f(v) = \sum_{i=1}^{n} \lambda_i \mu_i$. In other words, if $B = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for V, and describe each vector in V by its coordinates $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ relative to B, then every linear functional on V has the form $f(v) = \sum_{i=1}^{n} \lambda_i \mu_i$.

Example 3.31 Consider the basis $\{(3, 2), (1, 1)\}$ for $V = \mathbb{R}^2$. Then for any $(x_1, x_2) \in \mathbb{R}^2$ as

$$(x_1, x_2) = (x_1 - x_2)(3, 2) + (-2x_1 + 3x_2)(1, 1)$$

we get $f_1(x_1, x_2) = x_1 - x_2$ and $f(x_1, x_2) = -2x_1 + 3x_2$. Thus $\{x_1 - x_2, -2x_1 + 3x_2\}$ is the dual basis corresponding to the basis $\{(3, 2), (1, 1)\}$.

Example 3.32 Consider the basis $\{1, 1 + x, x + x^2\}$ for $V = \mathbb{P}_2[a, b]$. Then for any $a_0 + a_1x + a_2x^2 \in \mathbb{R}^2$ as

$$a_0 + a_1x + a_2x^2 = (a_0 - a_1 + a_2)1 + (a_1 - a_2)(1 + x) + a_2(x + x^2)$$

we get $f_1(a_0 + a_1x + a_2x^2) = a_0 - a_1 + a_2$, $f_2(a_0 + a_1x + a_2x^2) = a_1 - a_2$ and $f_3(a_0 + a_1x + a_2x^2) = a_2$.

Now, let us discuss the range space and null space of a linear functional. If f is a non-zero linear functional, then the range space of f is the scalar field itself. Then, by *Rank-Nullity Theorem*, we can say that $Nullity(f) = n - 1$, if V is an n dimensional space. In a vector space of dimension n, a subspace of dimension $n - 1$ is called a *hyperspace*. In fact we can say that every hyperspace is the null space of a linear functional (see Exercise 26).

3.8 Exercises

1. Check whether which of the following functions defines a linear transformation from $\mathbb{R}^2$ over $\mathbb{R}$ to itself:

   (a) $T(x_1, x_2) = (x_1 + 1, x_2 + 1)$
   (b) $T(x_1, x_2) = (3x_1, 7x_2)$
   (c) $T(x_1, x_2) = (sin\, x, 0)$
   (d) $T(x_1, x_2) = (x_1, x_2^2)$
   (e) $T(x_1, x_2) = (2x_1, \frac{1}{2}x_2)$
   (f) $T(x_1, x_2) = (x_1, x_1x_2)$.

2. Check which of the following functions define a linear transformation:

   (a) $T : \mathbb{R}^3 \to \mathbb{P}_2[a, b]$ defined by $T(\alpha_1, \alpha_2, \alpha_3) = \alpha_1 + (\alpha_1 + \alpha_2)\, x + (\alpha_1 + \alpha_2 + \alpha_3)\, x^2$.
   (b) $T : \mathbb{P}(\mathbb{R}) \to \mathbb{P}(\mathbb{R})$ defined by $(Tp)\,(x) = x^2 p(x) + \frac{d}{dx}(p(x))$.
   (c) $T : \mathbb{R}^4 \to \mathbb{M}_2\,(\mathbb{R})$ defined by $T(x_1, x_2, x_3, x_4) = \begin{bmatrix} 1 + x_1 & x_1 + x_2 \\ x_2 + x_3 & x_3 + x_4 \end{bmatrix}$.
   (d) $T : \mathbb{P}_2[a, b] \to \mathbb{M}_2\,(\mathbb{R})$ defined by $(Tp)\,(x) = \begin{bmatrix} p(0) & p(1) + p(2) \\ 0 & p(3) \end{bmatrix}$.
   (e) $T : \mathbb{M}_2\,(\mathbb{R}) \to \mathbb{M}_2\,(\mathbb{R})$ defined by $T(A) = A + I$, where $I \in \mathbb{M}_2\,(\mathbb{R})$ is the identity matrix.
   (f) Fix $A \in \mathbb{M}_n\,(\mathbb{R})$ and $b \neq\in \mathbb{R}^n$. Define $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(x) = Ax + b$.
   (g) Fix $A \in \mathbb{M}_n\,(\mathbb{R})$. Define $T : \mathbb{M}_n\,(\mathbb{R}) \to \mathbb{M}_n\,(\mathbb{R})$ by $T(B) = A^{-1}BA$.

3. Show that $T : \mathbb{M}_n\,(\mathbb{R}) \to \mathbb{R}$ defined by $T(A) = Tr(A)$ is a linear transformation. What about $T(A) = det(A)$? Here $Tr(A)$ and $det(A)$ denote the trace of A and determinant of A, respectively.
4. Let V be a one-dimensional vector space over a field $\mathbb{K}$. Show that every linear transformation $T : V \to V$ is of the form $T(v) = \lambda v$ for some $\lambda \in \mathbb{K}$.
5. Let $T_1 : V \to W$ be a linear map and $T_2 : V \to W$ be a non-linear map. Then what about $T_1 + T_2$? Is it always non-linear?
6. Find $T : \mathbb{R}^3 \to \mathbb{R}^3$, if $T(1, 0, 0) = (1, 0, 1), T(1, 1, 0) = (0, 2, 1)$, and $T(1, 1, 1) = (0, 0, 1)$. Is T unique?
7. Find the range space and null space of the following linear transformations:

   (a) $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3) = (x_3, 0, x_1)$.
   (b) $T : \mathbb{P}_2[a, b] \to \mathbb{M}_2\,(\mathbb{R})$ defined by $T(a_0 + a_1x + a_2x^2) = \begin{bmatrix} a_0 & a_1 \\ a_2 & 0 \end{bmatrix}$.
   (c) $T : \mathbb{M}_2\,(\mathbb{R}) \to \mathbb{M}_2\,(\mathbb{R})$ defined by $T(A) = A - A^T$.

8. Find a linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^3$ for which the null space is spanned by $v_1 = (1, 1, 1, 1)$, $v_2 = (1, 0, 0, 1)$, and the range space of T is spanned by $w_1 = (1, 1, 0)$ and $w_2 = (1, 0, 1)$, if it exists.
9. Show that if $A \in \mathbb{M}_{m\times n}(\mathbb{K})$, then row rank of A= column rank of A.

10. Let V be a vector space over a field $\mathbb{K}$. Does there exist linear transformations T_1, T_2 on V with

    (a) $\mathcal{R}(T_1) = \mathcal{R}(T_2)$ and $\mathcal{N}(T_1) = \mathcal{N}(T_2)$
    (b) $\mathcal{R}(T_1) = \mathcal{N}(T_2)$ and $\mathcal{N}(T_1) = \mathcal{R}(T_2)$.

11. Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ be defined by

$$T(x_1, x_2, x_3, x_4) = (x_1 - x_2, x_2 - x_3, x_3 - x_4, 0)$$

    Then

    (a) verify that T is a linear transformation on $\mathbb{R}^4$.
    (b) find $\mathcal{R}(T)$ and $\mathcal{N}(T)$.
    (c) verify *Rank-Nullity* theorem.
    (d) Is T invertible?

12. Check whether the following statements are true or false:

    (a) There exists a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T(1, 0) = (1, 2)$, $T(1, 1) = (0, 3)$ and $T(2, 3) = (2, 5)$.
    (b) $T : \mathbb{P}_{n-1}(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ defined by $(Tp)(x) = \int_0^x p(t)dt$ is onto.
    (c) $T : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_{n-1}(\mathbb{R})$ defined by $(Tp)(x) = \frac{d}{dx}(p(x))$ is one-one.
    (d) There exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with $\mathcal{R}(T) = \mathcal{N}(T)$ if and only if n is even.
    (e) Let $T : V \to V$ be a linear transformation with $\mathcal{R}(T) \cap \mathcal{N}(T) = \{0\}$, then $V = \mathcal{R}(T) \oplus \mathcal{N}(T)$.
    (f) Let $T : V \to V$ be a linear transformation on a finite-dimensional vector space V and $A \in \mathbb{M}_n(\mathbb{K})$ be its matrix representation. Then A is unique.
    (g) Let V and W be finite-dimensional vector spaces over the field $\mathbb{K}$. Then an invertible linear transformation $T : V \to W$ maps a basis of V to a basis of W.
    (h) Let $T_1, T_2 : V \to W$ be isomorphisms from V to W. Then $T_1 + T_2$ is also an isomorphism from V to W.

13. Let V be a finite-dimensional space and $T_1, T_2 : V \to V$ be linear transformations such that $T_1 T_2 = I$. Then show that $T_2 = T_1^{-1}$. Also show that there exist linear transformations T_1, T_2 such that $T_1 T_2 = I$ and $T_2 T_1 \neq I$, if V is infinite-dimensional.
14. Let V and W be finite-dimensional vector spaces over the field $\mathbb{K}$, and let $T_1 : V \to W$ be a linear transformation. Then

    (a) T_1 is one-one if and only if there exists $T_2 : W \to V$ such that $T_2 T_1$ is the identity map on V.
    (b) T_1 is onto if and only if there exists $T_2 : W \to V$ such that $T_1 T_2$ is the identity map on W.

15. Let V be a vector space over $\mathbb{K}$ and $T : V \to V$ be any linear transformation. Show that

    (a) $\mathcal{R}(T) \cap \mathcal{N}(T) = \{0\}$ if and only if $\mathcal{N}\left(T^2\right) \subseteq \mathcal{N}(T)$.
    (b) T^2 is the zero transformation if and only if $\mathcal{R}(\mathcal{T}) \subseteq \mathcal{N}(T)$.

16. Consider the linear transformations $T_1, T_2, T_3 : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T_1(x_1, x_2, x_3) = (x_1 + x_2, x_1 + x_3), T_2(x_1, x_2, x_3) = (2x_1, x_1 + x_2 + x_3)$$

and

$$T_3(x_1, x_2, x_3) = (x_1, x_2)$$

Check whether $\{T_1, T_2, T_3\}$ is linearly independent in $\mathcal{L}\left(\mathbb{R}^3, \mathbb{R}^2\right)$.

17. Let $T_1 : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_{n-1}(\mathbb{R})$ be defined by $(T_1 p)(x) = \frac{d}{dx}(p(x))$ and $T_2 : \mathbb{P}_{n-1}(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ be defined by $(T_2 p)(x) = xp(x)$. Show that $T_1T_2 - T_2T_1 = I$, where I is the identity operator on $\mathbb{P}_{n-1}(\mathbb{R})$.
18. Fix $A \in \mathbb{M}_n(\mathbb{R})$. Define $T : \mathbb{R}^n \to \mathbb{R}^n$ by $T(v) = Av$. Describe a situation where T becomes an isomorphism, if it exists.
19. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and $A \in \mathbb{M}_n(\mathbb{R})$ be its matrix representation. then what is the relation between $\mathcal{R}(T)$ and $Im(A)$, where $Im(A)$ denotes the column space of A.
20. Let $T : \mathbb{R}^3 \to \mathbb{P}_2[a, b]$ be defined by

$$T(\alpha_1, \alpha_2, \alpha_3) = (\alpha_2 + \alpha_3)x + (\alpha_1 + \alpha_3)x^2$$

Find $[T]_{B_1}^{B_2}$ where $B_1 = \{(2, 0, 1), (1, 2, 0), (0, 1, 2)\}$ and $B_2 = \{1, 1 + x, (1 + x)^2\}$.

21. Let $T : \mathbb{M}_2(\mathbb{R}) \to \mathbb{M}_2(\mathbb{R})$ be defined by $T(A) = \frac{1}{2}\left(A + A^T\right)$. Find $[T]_B$ where $B = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$.
22. Let V and W be vector spaces over $\mathbb{K}$ with $dim(V) = dim(W)$. Show that there exist bases, B_1 of V and B_2 of W such that $[T]_{B_1}^{B_2}$ is a diagonal matrix.
23. Let $B_1 = \{1, x, x^2\}$, $B_2 = \{1, 1 + 2x, 1 + 2x + 3x^2\}$, and $B_1 = \{1, 2 + x, 1 + x^2\}$ be three bases for $\mathbb{P}_2[a, b]$. Then find

    (a) change of basis matrix from B_1 to B_2.
    (b) change of basis matrix from B_1 to B_3.
    (c) change of basis matrix from B_2 to B_3.

    Find the relation between them and generalize, if possible.

24. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x_1, x_2) = (x_1 + x_2, 2x_2)$. Consider the bases $B_1 = \{(1, 1), (1, -1)\}$ and $B_2 = \{(1, 2), (0, 1)\}$ of $\mathbb{R}^2$. Then find

    (a) the matrix representation of T with respect to B_1.
    (b) the matrix representation of T with respect to B_2.

(c) change of basis matrix from B_1 to B_2.
(d) What is the relation between $[T]_{B_1}$ and $[T]_{B_2}$?

25. Show that every linear functional on $\mathbb{K}^n$ is of the form

$$f(x_1, x_2, \ldots, x_n) = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$$

for some scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$.
26. Let $T : V \to \mathbb{K}$ be a linear transformation. Show that for an element $v \in V$ with $v \notin \mathcal{N}(T)$, we have $V = span\{v\} \oplus \mathcal{N}(T)$.
27. Consider the basis $B = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ for $\mathbb{R}^3$. Find the dual basis B^* corresponding to B.
28. Let $V = \mathbb{P}_2[a, b]$. Consider 3 distinct real numbers $\lambda_1, \lambda_2, \lambda_3$, and define $f_i(p) = p(\lambda_i)$. Then

(a) Show that $B^* = \{f_1, f_2, f_3\}$ forms a dual basis for V^*.
(b) Find the ordered basis B of V corresponding to B^*.

29. Let V and W be vector spaces over the field $\mathbb{K}$ with respective dual spaces V^* and W^*. Let $T : V \to W$ be a linear transformation.

(a) Show that the map $\tilde{T} : W^* \to V^*$ defined by $\tilde{T}g = g \circ T$ is a linear map. (The map $\tilde{T}$ is called the *Transpose of T*.)
(b) Suppose that $\{f_1, f_2, \ldots, f_n\}$ in V^* is a dual basis corresponding to the basis $\{v_1, v_2, \ldots, v_n\}$ in V and that $\{g_1, g_2, \ldots, g_n\}$ in V^* is a dual basis corresponding to the basis $\{w_1, w_2, \ldots, w_n\}$ in W. If A is the matrix representation of T, then show that A^T is the matrix representation of $\tilde{T}$ with respect to the above dual bases.

30. Find the matrix representation of transpose of the linear transformation, $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by

$$T(x_1, x_2, x_3) = (2x_1 - x_2 + 2x_3, x_1 + 2x_2 - x_3)$$

where the bases of $\mathbb{R}^3$ and $\mathbb{R}^2$ are, respectively, $\{(2, 0, 1), (1, 2, 0), (0, 1, 0)\}$ and $\{(3, 2), (1, 1)\}$.

Solved Questions related to this chapter are provided in Chap. 9.

Chapter 4
Eigenvalues and Eigenvectors

In this chapter, we explore the foundational concepts of eigenvalues and eigenvectors, providing a deep understanding of their definition, properties, and far-reaching applications of linear algebra. Eigenvalues and eigenvectors are introduced as crucial properties of square matrices. Eigenvalues represent the scaling factors by which eigenvectors are stretched or compressed when the matrix operates on them. Matrix similarity is discussed as a fundamental concept, highlighting how similar matrices share the same eigenvalues. We delve into the importance of diagonalization, where a matrix is transformed into a diagonal matrix using its eigenvectors. This process simplifies matrix exponentiation and powers, which are crucial for solving differential equations and modeling dynamical systems and their stability analysis. The chapter provides a thorough grasp of when diagonalization is possible by examining the necessary and sufficient conditions for a matrix to be diagonalizable. To deal with non-diagonalizable matrices, generalized eigenvectors are introduced, leading to the Jordan Canonical Form notion. This form aids in the analysis of complicated systems by providing insight into the structure of non-diagonalizable matrices. From this point onwards, for convenience λ is used both as a variable and as a scalar. The usage is evident from the context.

4.1 Eigenvalues and Eigenvectors

Consider a homogeneous linear system of differential equations of the form

$$\begin{cases} \frac{dy_1}{dx} &= a_{11}y_1 + a_{12}y_2 \\ \frac{dy_2}{dx} &= a_{21}y_1 + a_{22}y_2 \end{cases} \tag{4.1}$$

where $a_{11}, a_{12}, a_{21}, a_{22}$ are real constants. We shall represent the above system in an alternate form using matrices, as follows:

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_4

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix} \tag{4.2}$$

If we name the matrices in Eq. (4.2) as,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \text{ and } \frac{dy}{dx} = \begin{bmatrix} \frac{dy_1}{dx} \\ \frac{dy_2}{dx} \end{bmatrix},$$

Equation (4.2) can be written in a compact form as

$$Ay = \frac{dy}{dx} \tag{4.3}$$

We seek non-trivial solutions of the form

$$y_1 = \mu_1 e^{\lambda x} \text{ and } y_2 = \mu_2 e^{\lambda x}$$

for the system (4.1), where μ_1, μ_2 and λ are constants. That is, we need a solution of the form

$$y = ve^{\lambda x}, \text{ where } v = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

Substituting $y = ve^{\lambda x}$ in Eq. (4.3), we get

$$\lambda v e^{\lambda x} = Ave^{\lambda x}$$

which then reduces to

$$\lambda v = Av \tag{4.4}$$

That is, we need to find a non-zero vector v and scalar λ such that Eq. (4.4) is satisfied. In many practical cases, numerous mathematical problems can be formulated in the form of Eq. (4.4). Such problems are called *Eigenvalue problems.* We may rewrite Eq. (4.4) as

$$(A - \lambda I)v = 0 \tag{4.5}$$

The matrix $(A - \lambda I)$ is called the *characteristic matrix* of A and the Eq. (4.5) is the *characteristic matrix equation*. From Sect. 1.7, we know that Eq. (4.5) has a non-zero solution if and only if

$$det(A - \lambda I) = 0 \tag{4.6}$$

Observe that $det(A - \lambda I)$ will be a polynomial of order n in λ, if A is an $n \times n$ matrix, and is referred to as *characteristic polynomial* of A. Equation (4.6) is known as *characteristic equation* of A. From Theorem 1.18, we know that over an algebraically closed field, such an equation will have n solutions/roots. The roots are called *characteristic values* or *eigenvalues* of A. If we denote the eigenvalues of A by $\lambda_1, \lambda_2, \ldots, \lambda_n$, we have $det(A - \lambda_i I) = 0$ for each $i = 1, 2, \ldots, n$. Then for each λ_i there exists non-zero vectors v_i satisfying $(A - \lambda_i I)v_i = 0$. Such vectors v_i are called *characteristic vectors or eigenvectors* of A associated with the eigenvalue λ_i. Consequently, we have $Av = \lambda v$. Thus we can have the following definition.

Definition 4.1 (*Eigenvalues and Eigenvectors*) Let A be an $n \times n$ matrix with entries from the field $\mathbb{C}$. A non-zero vector $v \in \mathbb{C}^n$ is said to be an eigenvector of A, if there exists $\lambda \in \mathbb{C}$ such that $Av = \lambda v$. The scalar λ is called an eigenvalue of A. In other words, $\lambda \in \mathbb{C}$ is an eigenvalue of an $n \times n$ matrix A if there exists a non-zero vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$.

Consider the following example for a better understanding of the ideas that we have discussed above.

Example 4.1 Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Then the characteristic polynomial of A is

$$det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3)$$

and $det(A - \lambda I) = 0$ implies that $\lambda_1 = -1$ and $\lambda_2 = 3$. Now to find the eigenvector associated with the eigenvalue $\lambda_1 = -1$, we have to find a non-zero vector $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ satisfying $(A + I)v = 0$.

$$\begin{aligned} (A + I)v = 0 &\Rightarrow \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow v_1 + v_2 = 0 \end{aligned}$$

Thus any non-zero vector from the set $W_1 = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 + v_2 = 0\}$ is an eigenvector of A corresponding to the eigenvalue $\lambda_1 = -1$. In particular, we can say that $v = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue -1. Similarly, for $\lambda_2 = 3$

$$\begin{aligned} (A - 3I)v = 0 &\Rightarrow \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Rightarrow -v_1 + v_2 = 0 \end{aligned}$$

Thus any non-zero vector from the set $W_2 = \{(v_1, v_2) \in \mathbb{R}^2 \mid -v_1 + v_2 = 0\}$ is an eigenvector of A corresponding to the eigenvalue $\lambda_2 = 3$. In particular, we can say that $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 3. Now let us plot the sets W_1 and W_2.

Clearly both W_1 and W_2 are subspaces of $\mathbb{R}^2$. Another interesting fact to observe is that the vector $(1, -1)$ spans W_1 and $(1, 1)$ spans W_2. Is this true in general?

Remark 4.1 The geometrical significance of eigenvalues and eigenvectors of a matrix are of great importance in matrix theory and linear algebra. In Example 4.1, we have seen that the eigenvalue -1 changed the direction of the eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and the eigenvalue 3 stretched the vector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, three times. If the matrix A has a real eigenvalue λ, Definition 4.1 means that the eigenvector of A associated with λ is a vector that experiences a change in sign or magnitude or both and λ is the amount of stretch or shrink, the eigenvector is subjected to by the action of A.

We know that a non-zero vector v is an eigenvector of a matrix A corresponding to the eigenvalue λ if and only if v is a solution of the matrix equation of $(A - \lambda I)v = 0$. That is, if and only if $v \in \mathcal{N}(A - \lambda I)$, the null space of the matrix $A - \lambda I$. This justifies our observation in the Example 4.1 that the sets W_1 and W_2 are subspaces of $\mathbb{R}^2$. The ideas that we have discussed so far can be summarized as follows to characterize the eigenvalues of a square matrix A.

Theorem 4.1 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $\lambda \in \mathbb{K}$, then the following are equivalent;*

(a) *λ is an eigenvalue of A.*
(b) $\mathcal{N}(A - \lambda I) \neq \{0\}$
(c) $det(A - \lambda I) = 0$.

Proof $(a) \Rightarrow (b)$ Let A be an $n \times n$ matrix with entries from a field $\mathbb{K}$. Let $\lambda \in \mathbb{K}$ be an eigenvalue of A. Then there exists a vector $v \neq 0 \in \mathbb{K}^n$ such that $Av = \lambda v$. Now

$$Av = \lambda v \Leftrightarrow (A - \lambda I)v = 0 \Leftrightarrow v \in \mathcal{N}(A - \lambda I)$$

Therefore $\mathcal{N}(A - \lambda I) \neq \{0\}$.
$(b) \Rightarrow (c)$ Suppose that $\mathcal{N}(A - \lambda I) \neq \{0\}$. Let $v \neq 0 \in \mathcal{N}(A - \lambda I)$. Then $(A - \lambda I)v = 0$. That is, the homogeneous system of equations $(A - \lambda I)v = 0$ has a non-trivial solution. This is true only if $det(A - \lambda I) = 0$.
$(c) \Rightarrow (a)$ Now suppose that $det(A - \lambda I) = 0$. This implies that the homogeneous system of equations $(A - \lambda I)x = 0$ has a non-trivial solution, say $v \in \mathbb{R}^n$. That is, $(A - \lambda I)v = 0$. This implies that there exists a non-zero vector $v \in \mathbb{R}^n$ such that $Av = \lambda v$. Therefore λ is an eigenvalue of A.

Thus we have seen that the collection of all eigenvectors of a square matrix A associated with an eigenvalue λ forms a vector space called as *eigenspace* of A corresponding to the eigenvalue λ. The dimension of the eigenspace associated with λ is called as the *geometric multiplicity* of λ. By Rank–Nullity Theorem, *geometric multiplicity* of an eigenvalue λ of A is given by $n - Rank(A - \lambda I)$. Another term related to an eigenvalue λ of a matrix A is its algebraic multiplicity. The *algebraic multiplicity* of an eigenvalue λ of A is defined as the number of times λ appears as a root of the characteristic polynomial.

Example 4.2 In Example 4.1, we have seen that the matrix $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ has two eigenvalues -1 and 3. As the characteristic polynomial of A is $(\lambda + 1)(\lambda - 3)$, the algebraic multiplicity of both -1 and 3 (denoted by $AM(-1)$ and $AM(3)$, respectively) is 1. Also, we have observed that

$$\mathcal{N}(A + I) = \{(v_1, v_2) \in \mathbb{R}^2 \mid v_1 + v_2 = 0\} = span\{(-1, 1)\}$$

and

$$\mathcal{N}(A - 3I) = \{(v_1, v_2) \in \mathbb{R}^2 \mid -v_1 + v_2 = 0\} = span\{(1, 1)\}$$

Therefore the geometric multiplicity of both -1 and 3 (denoted by $GM(-1)$ and $GM(3)$ respectively) is also 1.

Now pick one eigenvector from each of the eigenspaces $\mathcal{N}(A + I)$ and $\mathcal{N}(A - 3I)$ of the matrix $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Let us pick $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ from $\mathcal{N}(A + I)$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ from $\mathcal{N}(A - 3I)$. We can clearly notice that these vectors are linearly independent. Our next theorem generalizes this fact. That is, we will prove that the eigenvectors of a matrix A corresponding to its distinct eigenvalues will be linearly independent.

Theorem 4.2 *Let $A \in \mathbb{M}_{n \times n}(\mathbb{K})$ and let $v_1, v_2, \ldots, v_m \in \mathbb{K}^n$ be eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{K}$ respectively , then $\{v_1, v_2, \ldots, v_m\}$ is linearly independent.*

Proof Let k be the smallest positive integer such that $v_1, v_2, \ldots, v_k$ are linearly independent. If $k = m$, then there is nothing to prove. Now let $k < m$. Then $\{v_1, v_2, \ldots, v_{k+1}\} \subset \{v_1, v_2, \ldots, v_m\}$ is linearly dependent. Hence, there exists scalars $\mu_1, \mu_2, \ldots, \mu_k$ such that

$$v_{k+1} = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k$$

Multiplying by A on both sides, we get

$$\begin{aligned} Av_{k+1} &= A(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k) \\ &= \mu_1 A(v_1) + \mu_2 A(v_2) + \cdots + \mu_k A(v_k) \\ &= \mu_1 \lambda_1 v_1 + \mu_2 \lambda_2 v_2 + \cdots + \mu_k \lambda_k v_k \end{aligned}$$

Since v_{k+1} is a an eigenvector of A corresponding to the eigenvalue λ_{k+1}, we have

$$\begin{aligned} Av_{k+1} &= \lambda_{k+1} v_{k+1} \\ &= \lambda_{k+1}(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k) \\ &= \mu_1 \lambda_{k+1} v_1 + \mu_2 \lambda_{k+1} v_2 + \cdots + \mu_k \lambda_{k+1} v_k \end{aligned}$$

From the above two equations, we get

$$\mu_1(\lambda_1 - \lambda_{k+1})v_1 + \mu_2(\lambda_2 - \lambda_{k+1})v_1 + \cdots + \mu_k(\lambda_k - \lambda_{k+1})v_k = 0$$

Since $v_1, v_2, \ldots, v_k$ are linearly independent, we get

$$\mu_i(\lambda_i - \lambda_{k+1}) = 0 \; for \; all \; i = 1, 2, \ldots, k$$

Now as $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ are distinct, we get $\mu_i = 0$ for all $i = 1, 2, \ldots, k$. This implies that v_{k+1} is the zero vector which is a contradiction.

Corollary 4.1 *An $n \times n$ matrix A can have at most n distinct eigenvalues.*

Proof Suppose that A have $n + 1$ distinct eigenvalues. Then as eigenvectors corresponding to distinct eigenvalues are linearly independent, A has $n + 1$ linearly independent eigenvectors which is a contradiction since $\mathbb{K}^n$ is of dimension n.

Example 4.3 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$$

Then

$$det\;(A - \lambda I) = 0 \Rightarrow \begin{vmatrix} 1-\lambda & 2 & 3 \\ 0 & 1-\lambda & 2 \\ 0 & 0 & 7-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)^2\,(7-\lambda) = 0$$

Thus $(1 - \lambda)^2(7 - \lambda) = 0$ is the characteristic equation of A and hence the eigenvalues of A are $1, 1, 7$. Since 1 appears two times as a root of the characteristic equation, the algebraic multiplicity of 1 is 2 and algebraic multiplicity of 7 is 1. That is, $AM(1) = 2$ and $AM(7) = 1$.

Now let us find the eigenvectors corresponding to the eigenvalue, $\lambda_1 = 1$.

$$(A - I)v = 0 \Rightarrow \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow v_2 = v_3 = 0$$

Therefore

$$\mathcal{N}(A - I) = \{(v_1, 0, 0) \in \mathbb{R}^3 \mid v_1 \in \mathbb{R}\} = span\{(1, 0, 0)\}$$

Hence geometric multiplicity of $\lambda_1 = 1$ is 1. That is, $GM(1) = 1$. This fact can also be verified by *Rank-Nullity Theorem*, as

$$n - Rank(A - \lambda_1 I) = 3 - Rank(A - I) = 3 - 2 = 1$$

Now for $\lambda_2 = 7$,

$$(A - 7I)v = 0 \Rightarrow \begin{bmatrix} -6 & 2 & 3 \\ 0 & -6 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow v_1 = \frac{11}{6} v_2, v_3 = 3v_2$$

Therefore

$$\mathcal{N}(A - 7I) = \left\{ \left(\frac{11}{6} v_2, v_2, 3v_2 \right) \in \mathbb{R}^3 \mid v_2 \in \mathbb{R} \right\} = span \left\{ \left(\frac{11}{6}, 1, 3 \right) \right\}$$

Hence geometric multiplicity of $\lambda_2 = 7$ is 1. That is, $GM(7) = 1$. Verification by *Rank-Nullity Theorem* is as follows,

$$n - Rank(A - \lambda I) = 3 - Rank(A - 7I) = 3 - 2 = 1$$

Also observe that the eigenvectors corresponding to 1 and 7 are linearly independent.

An intriguing fact to keep in mind from the above example is that the geometric multiplicity and algebraic multiplicity of every eigenvalue need not be the same. Matrices having $AM(\lambda) = GM(\lambda)$ for all eigenvalues λ are of greater importance in Mathematics. We will study about such matrices later in this chapter. Now let us give a definite form for the characteristic polynomial of an $n \times n$ matrix.

Theorem 4.3 *The characteristic polynomial of an $n \times n$ matrix A is a polynomial of degree n and is of the form*

$$det(A - \lambda I) = (-1)^n \left[\lambda^n + \mu_{n-1}\lambda^{n-1} + \cdots + \mu_1 \lambda + \mu_0 \right]$$

where $\mu_0, \mu_1, \ldots, \mu_n \in \mathbb{K}$.

Proof We prove this by induction on n. Suppose that $n = 1$, then A is of the form $A = [a_{11}]$, where $a_{11} \in \mathbb{K}$. Then for $\lambda \in \mathbb{K}$, $det(A - \lambda I) = 0$ implies that $a_{11} - \lambda = (-1)(\lambda - a_{11}) = 0$. This implies that the result is true for $n = 1$. Now assume that the result is true for $n - 1$. That is, the characteristic polynomial of an $(n-1) \times (n-1)$ matrix is a polynomial of degree $n - 1$ and is of the form $(-1)^{n-1} \left[\lambda^{n-1} + \xi_{n-2}\lambda^{n-2} + \cdots + \xi_1 \lambda + \xi_0 \right]$ where $\xi_0, \xi_1, \ldots, \xi_{n-2} \in \mathbb{K}$. Now consider an $n \times n$ matrix $A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix}$. Then $A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} - \lambda \end{bmatrix}$. Now calculate the determinant of $A - \lambda I$ by expanding the matrix along a column or row. In either case, some $(-1)(\lambda - a_{ii})$ is multiplied with the determinant of an $(n-1) \times (n-1)$ matrix, which is a polynomial of degree $n - 1$ and is of the form $(-1)^{n-1} \left[\lambda^{n-1} + \xi_{n-2}\lambda^{n-2} + \cdots + \xi_1 \lambda + \xi_0 \right]$ where $\xi_0, \xi_1, \ldots, \xi_{n-2} \in \mathbb{K}$ by our induction hypothesis. Therefore the characteristic polynomial of an $n \times n$ matrix A is a polynomial of degree n with leading coefficient $(-1)^n$.

Corollary 4.2 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then the product of eigenvalues of A is equal to $det(A)$ and the sum of the eigenvalues of A is equal to $Tr\ (A)$.*

Proof From the above theorem, for an $n \times n$ matrix A, the characteristic polynomial $det(A - \lambda I)$ is of the form $(-1)^n \left[\lambda^n + \mu_{n-1}\lambda^{n-1} + \cdots + \mu_1\lambda + \mu_0\right]$. Since the roots of the characteristic polynomial are eigenvalues, from *Vieta's Formula*, we know that the product of eigenvalues is equal to the constant term in the polynomial and the sum of the eigenvalues is equal to the coefficient of λ^{n-1}. Therefore product of eigenvalues of $A = (-1)^n\mu_0$ and the sum of the eigenvalues of $A = (-1)^n\mu_{n-1}$. Also $det(A) = det(A - 0I) = (-1)^n\mu_0$. Therefore the product of eigenvalues of an $n \times n$ matrix A is equal to $det(A)$. Now expanding $det(A - \lambda I)$ we get that $Tr(A) = (-1)^n\mu_n$. Therefore the sum of the eigenvalues of a matrix A is equal to $Tr\ (A)$.

From the above corollary, we can conclude that $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is singular if and only if 0 is an eigenvalue of A. For, A is singular implies that $det(A) = 0$. Since $det(A)$ is the product of eigenvalues of A, atleast one of the eigenvalues of A must be 0. Conversely, if 0 is an eigenvalue of A, then the product of eigenvalues of $A = det(A) = 0$.

Example 4.4 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

from Example 4.1. We have seen the eigenvalues of A are -1 and 3. Clearly, we can observe that sum of eigenvalues of $A = 2 = Tr(A)$ and product of eigenvalues of $A = -3 = det(A)$.

Example 4.5 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$$

given in Example 4.3. The eigenvalues of A are 1, 1 and 7. Clearly, sum of eigenvalues of $A = 9 = Tr(A)$ and product of eigenvalues of $A = 7 = det(A)$.

Remark 4.2 The characteristic equation of a 2×2 matrix is of the form

$$\lambda^2 - tr(A)\lambda + det(A) = 0$$

and the characteristic equation of a 3×3 matrix is of the form

$$-\lambda^3 + tr(A)\lambda^2 - [M_{11} + M_{22} + M_{33}]\lambda + det(A) = 0$$

where M_{11}, M_{22} and M_{33} are the minors of the diagonal elements.

If we know the eigenvalues of an $n \times n$ matrix A, we could find the eigenvalues of some matrices associated or related with A. The following theorem shows that, if A is an invertible matrix, then the eigenvalues of A^{-1} are the multiplicative inverses of eigenvalues of A.

Theorem 4.4 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be a nonsingular matrix. If λ is an eigenvalue of A, then λ^{-1} is an eigenvalue of A^{-1}.*

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be a nonsingular matrix. Then by Corollary 4.2, all eigenvalues of A are non-zero. If λ is an eigenvalue of A, there exists a non-zero vector $v \in \mathbb{K}^n$ such that $Av = \lambda v$. Multiplying both sides with A^{-1} we get, $A^{-1}(Av) = A^{-1}(\lambda v)$. That is, $v = \lambda A^{-1} v$ which implies that $A^{-1}v = \lambda^{-1}v$. Therefore λ^{-1} is an eigenvalue of A^{-1} with eigenvector v.

Likewise we can compute the eigenvalues of powers of A, if we know the eigenvalues of A.

Theorem 4.5 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Let λ be an eigenvalue of A with an eigenvector v, then λ^m is an eigenvalue of A^m with eigenvector v, for any positive integer m.*

Proof Let λ be an eigenvalue of A with an eigenvector v. Then $Av = \lambda v$. We have to show that $A^m v = \lambda^m v$ for any positive integer m. Clearly this is true for $m = 1$. Now assume that the result is true for $m - 1$. i.e, $A^{m-1}v = \lambda^{m-1}v$. Now

$$A^m v = A\left(A^{m-1}v\right) = A\left(\lambda^{m-1}v\right) = \lambda^{m-1}A(v) = \lambda^m v$$

Hence, λ^m is an eigenvalue of A^m with eigenvector v, for any positive integer m.

Using the eigenvalues of an $n \times n$ matrix A, we have characterized the eigenvalues of A^m, where m is a positive integer and A^{-1}, when A is invertible. Now consider a polynomial of degree m, given by

$$p(x) = a_0 + a_1 x + \cdots + a_m x^m$$

If we evaluate this polynomial with $x = A$,

$$p(A) = a_0 I + a_1 A + \cdots + a_m A^m$$

we get a *matrix polynomial*. Again, using the eigenvalues of A, we can compute the eigenvalues of $p(A)$. Consider the following theorem.

Theorem 4.6 *Let $p(x) = a_0 + a_1x + \cdots + a_m x^m \in \mathbb{K}[x]$ be a polynomial of degree m, where $\mathbb{K}$ is an algebraically closed field. If λ is an eigenvalue of $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ with an eigenvector v, then $p(\lambda)$ is an eigenvalue of $p(A) = a_0I + a_1A + \cdots + a_mA^m$ with the eigenvector v. Conversely, if $m \geq 1$ and if μ is an eigenvalue of $p(A)$, then there is some eigenvalue λ of A such that $p(\lambda) = \mu$.*

Proof Let λ be an eigenvalue of A with eigenvector v. Then $Av = \lambda v$. Now

$$\begin{aligned} p(A)v &= (a_0I + a_1A + \cdots + a_mA^m)v \\ &= a_0Iv + a_1Av + \cdots + a_mA^mv \\ &= a_0v + a_1\lambda v + \cdots + a_m\lambda^m v \\ &= (a_0 + a_1\lambda + \cdots + a_m\lambda^m)v \\ &= p(\lambda)v \end{aligned}$$

Hence, $p(\lambda)$ is an eigenvalue of $p(A)$ with the eigenvector v.

Conversely for $m \geq 1$, if μ is an eigenvalue of $p(A)$, then there exists a non-zero vector $v \in V$ such that $(p(A) - \mu I)v = 0$. Then $det(p(A) - \mu I) = 0$. Since $\mathbb{K}$ is an algebraically closed field, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{K}$ such that

$$p(x) - \mu = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_m)$$

Then

$$p(A) - \mu I = (A - \lambda_1 I)(A - \lambda_2 I)\cdots(A - \lambda_m I)$$

and as $det(p(A) - \mu I) = 0$, $det(A - \lambda_i I) = 0$ for atleast one i. This implies that λ_i is an eigenvalue of A. Also $p(\lambda_i) - \mu = 0$. Hence, λ_i is an eigenvalue of A and $p(\lambda_i) = \mu$.

Thus Theorem 4.5 can be considered as a special case of Theorem 4.6. Consider the following example.

Example 4.6 Let A be the matrix

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

from Example 4.1. Consider a polynomial $q(\lambda) = \lambda^2 + 2\lambda + 1$. Then

$$q(A) = A^2 + 2A + I = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} + \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}$$

Observe that the characteristic equation of $q(A)$ is $x^2 - 16x = 0$ and hence the eigenvalues of $q(A)$ are 0 and 16. By Theorem 4.6, the eigenvalues of $q(A)$ must be of the form $q(\lambda)$, where λ is an eigenvalue of A. We have already seen that the eigenvalues of A are -1 and 3. Note that $q(-1) = 0$ and $q(3) = 16$.

Now let us consider the characteristic polynomial of A, given by

$$p(\lambda) = \lambda^2 - 2\lambda - 3$$

It will be interesting to observe that

$$p(A) = A^2 - 2A - 3I = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} - \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

That is, we have

$$A^2 - 2A = 3I \tag{4.7}$$

As all the eigenvalues of A are non-zero, A is invertible. Multiplying by A^{-1} on both sides of Eq. (4.7), we have $3A^{-1} = A - 2I$. Therefore

$$A^{-1} = \frac{1}{3}(A - 2I) = \frac{1}{3}\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$$

That is, we can write A^{-1} in terms of A, when A is invertible. Will this be possible always? We will discuss this in the next theorem.

Observe that the characteristic equation of A^{-1} is $\lambda^2 + \frac{2}{3}\lambda - \frac{1}{3}$ and hence the eigenvalues of A^{-1} are -1 and $\frac{1}{3}$. Note that the eigenvalues of A^{-1} are the multiplicative inverses of the eigenvalues of A.

If we compute $p(A)$ for the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$ given in Example 4.3, we will get $p(A) = 0$, where $p(A)$ is the characteristic polynomial of A. This is interesting, right? In the next theorem, we will prove that this will be true for every square matrix. That is, *every square matrix satisfies its characteristic equation.* This is one of the most important theorems in matrix theory, named after famous mathematicians, *Arthur Cayley (1821–1895)* and *William R. Hamilton (1805–1865).*

Theorem 4.7 (Cayley–Hamilton Theorem) *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$, and let*

$$p(\lambda) = det(A - \lambda I) = (-1)^n\left[\lambda^n + \mu_{n-1}\lambda^{n-1} + \cdots + \mu_1\lambda + \mu_0\right]$$

be the characteristic polynomial of A, then A satisfies its own characteristic polynomial. That is,

$$p(A) = (-1)^n\left[A^n + \mu_{n-1}A^{n-1} + \cdots + \mu_1 A + \mu_0 I\right] = 0$$

Proof We have

$$(A - \lambda I)adj(A - \lambda I) = det(A - \lambda I)I = p(\lambda)I \tag{4.8}$$

where $adj(A - \lambda I) = \left[p_{ij}(\lambda)\right]_{n\times n}$, $p_{ij}(\lambda)$ is a polynomial of degree $n - 1$ in λ and $1 \leq i, j \leq n$. Therefore, we can represent $adj(A - \lambda I)$ as

$$adj(A - \lambda I) = A_0 + A_1\lambda + \cdots + A_{n-1}\lambda^{n-1}$$

where $A_0, A_1, \ldots A_{n-1} \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then from Eq. 4.8,

$$\begin{aligned}(-1)^n \left[\lambda^n + \mu_{n-1}\lambda^{n-1} + \cdots + \mu_1\lambda + \mu_0\right] I &= (A - \lambda I)\left(A_0 + A_1\lambda + \cdots + A_{n-1}\lambda^{n-1}\right)\\ &= AA_0 + (AA_1 - A_0)\lambda + \cdots + (-A_{n-1})\lambda^n\end{aligned}$$

Now comparing the coefficients of powers of λ, we get

$$\begin{aligned}AA_0 &= (-1)^n\mu_0 I\\ AA_1 - A_0 &= (-1)^n\mu_1 I\\ &\vdots\\ AA_{n-1} - A_{n-2} &= (-1)^n\mu_{n-1} I\\ -A_{n-1} &= (-1)^n I\end{aligned}$$

Multiplying these equations on the left by $I, A, \ldots, A^n$ respectively , we get

$$\begin{aligned}AA_0 &= (-1)^n\mu_0 I\\ A^2A_1 - AA_0 &= (-1)^n\mu_1 A\\ &\vdots\\ A^nA_{n-1} - A^{n-1}A_{n-2} &= (-1)^n\mu_{n-1}A_{n-1}\\ -A^nA_{n-1} &= (-1)^nA_n\end{aligned}$$

By adding these equations, we get

$$p(A) = (-1)^n\left[A^n + \mu_{n-1}A^{n-1} + \cdots + \mu_1A + \mu_0 I\right] = 0$$

That is, A satisfies its characteristic equation.

Example 4.7 Consider the matrix

$$A = \begin{bmatrix} 2 & 0 & 0\\ 1 & 2 & 1\\ -1 & 0 & 1\end{bmatrix}$$

Then the characteristic equation of A is

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = (2-\lambda)^2(1-\lambda) = 0$$

Observe that

$$A^3 - 5A^2 + 8A - 4I = \begin{bmatrix} 8 & 0 & 0 \\ 7 & 8 & 7 \\ -7 & 0 & 1 \end{bmatrix} - 5\begin{bmatrix} 4 & 0 & 0 \\ 3 & 4 & 3 \\ -3 & 0 & 1 \end{bmatrix} + 8\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} - 4\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0$$

That is, A satisfies its characteristic equation. Since $det(A) \neq 0$, A is invertible. Now

$$A^3 - 5A^2 + 8A - 4I = 0 \Rightarrow 4I = A^3 - 5A^2 + 8A$$

Multiplying by A^{-1} on both sides

$$A^{-1} = \frac{1}{4}[A^2 - 5A + 8I]$$

This is an important application of *Cayley–Hamilton theorem*. We can also see that

$$A^2 - 3A + 2I = \begin{bmatrix} 4 & 0 & 0 \\ 3 & 4 & 3 \\ -3 & 0 & 1 \end{bmatrix} - 3\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0$$

i.e., A also satisfies the polynomial equation $\lambda^2 - 3\lambda + 2 = 0$.

From this example, we get that for a matrix A, there are polynomials $p(\lambda)$ other than the characteristic polynomial of A for which $p(A) = 0$.

Definition 4.2 (*Annihilating polynomial*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. If for $f(\lambda) \in \mathbb{K}[\lambda]$, we have $f(A) = 0$, then $f(\lambda)$ is called an annihilating polynomial of A.

Definition 4.3 (*Minimal polynomial*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$, then the minimal polynomial of A is the least degree monic polynomial $q(\lambda) \in \mathbb{K}[\lambda]$ such that $q(A) = 0$.

Clearly, minimal polynomial and characteristic polynomial of a matrix A are annihilating polynomials of A. The following theorems discusses the relation between these polynomials.

Theorem 4.8 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. If $p(\lambda) \in \mathbb{K}[\lambda]$ is an annihilating polynomial of A, then the minimal polynomial divides $p(\lambda)$.*

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $p(\lambda) \in \mathbb{K}[\lambda]$ be such that $p(A) = 0$. Let $q(\lambda)$ be the minimal polynomial of A. By division algorithm for polynomials, there exists polynomials $m(\lambda)$ and $r(\lambda)$ such that

$$p(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$

where $deq(r(\lambda)) < deq(q(\lambda))$. Then $r(\lambda) = p(\lambda) - m(\lambda)q(\lambda)$. This implies that $r(A) = p(A) - m(A)q(A) = 0$ which is contradiction, since $q(\lambda)$ is the least degree polynomial that is satisfied by the matrix A. Therefore $r(\lambda) = 0$ and hence $p(\lambda) = m(\lambda)q(\lambda)$.

Corollary 4.3 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then minimal polynomial of A divides characteristic polynomial of A.*

Proof Let $p(\lambda)$ be the characteristic polynomial of A. By *Cayley–Hamilton theorem*, $p(\lambda)$ is an annihilating polynomial of A and hence the result follows.

Theorem 4.9 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then the minimal polynomial of A and the characteristic polynomial of A have same roots except for multiplicities.*

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. $p(\lambda)$ and $q(\lambda)$ be the characteristic and minimal polynomial of A respectively. Since minimal polynomial divides characteristic polynomial, there exists $m(\lambda) \in \mathbb{K}[\lambda]$ such that $p(\lambda) = m(\lambda)q(\lambda)$. Let $\lambda_1 \in \mathbb{K}$ be such that $q(\lambda_1) = 0$. Then $p(\lambda_1) = m(\lambda_1)q(\lambda_1) = 0$. Hence, λ_1 is a root of $p(\lambda)$.

Now let λ_1 be a root of $p(\lambda)$. Then there exists an eigenvector $v \neq 0$ such that $Av = \lambda_1 v$. Since $q(\lambda)$ is the minimal polynomial of A, we have $q(A)v = 0$. Then by Theorem 4.6, $q(\lambda_1)v = 0$. Then as $v \neq 0$, $q(\lambda_1) = 0$. Hence, λ_1 is a root of $q(\lambda)$.

Remark 4.3 Let $p(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$ be the characteristic polynomial of a matrix A, then the minimal polynomial is of the form $q(\lambda) = (\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_k)^{m_k}$ where $m_i \leq n_i$ for all $i = 1, 2, \ldots, k$.

Example 4.8 Consider the matrix

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

from Example 4.7. We have seen that the characteristic polynomial of A is $(2-\lambda)^2(1-\lambda)$. Then by the above remark, the minimal polynomial has two possibilities, $(2-\lambda)(1-\lambda)$ and $(2-\lambda)^2(1-\lambda)$. Since $A^2 - 3A + 2I = (2I - A)(I - A) = 0$, the minimal polynomial of A is $(2-\lambda)(1-\lambda)$.

Example 4.9 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

from Example 4.1. The characteristic polynomial of A is $(\lambda + 1)(\lambda - 3)$. As the characteristic polynomial and minimal polynomial must have the same roots the minimal polynomial of A is also $(\lambda + 1)(\lambda - 3)$.

Example 4.10 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$$

from Example 4.3. The characteristic polynomial of A is $(1-\lambda)^2(7-\lambda)$. Then there are two possibilities for the minimal polynomial of A, which are $(1-\lambda)(7-\lambda)$ and $(1-\lambda)^2(7-\lambda)$ it self. As we can see that $(I-A)(7I-A) \neq 0$, $(1-\lambda)^2(7-\lambda)$ is the minimal polynomial of A.

4.2 Diagonalization

A matrix $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is called a diagonal matrix if each of its non diagonal elements is zero. A diagonal matrix with all its main diagonal entries equal is called a scalar matrix. For a diagonal matrix D, the eigenvalues are precisely its diagonal entries and e_i is an eigenvector of D with eigenvalue d_{ii}, where d_{ii} denotes the ith diagonal entry of D and e_i is the ith element in the standard ordered basis for $\mathbb{K}^n$. These properties of diagonal matrices can be used in many applications. In this section, we will be discussing whether every square matrix can be made similar to a diagonal matrix.

Remark 4.4 The collection of all diagonal matrices, denoted by $\mathcal{D}$, under matrix addition forms an Abelian group where the zero matrix acts as the identity and inverse of each element A is $-A$. But under matrix multiplication, $\mathcal{D}$ does not form a group, as a diagonal matrix is invertible if and only if all its diagonal entries are non-zero.

Definition 4.4 (*Similar matrices*) Two $n \times n$ matrices A and B are said to be *similar* if there exists an invertible matrix P such that $P^{-1}AP = B$.

Example 4.11 Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \text{ and } B = \begin{bmatrix} 3 & 0 \\ 6 & -1 \end{bmatrix}$$

As for $P = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$, $P^{-1}AP = B$ implies that A and B are similar. We know that the characteristic polynomial of A is $(\lambda+1)(\lambda-3)$. Observe that the characteristic equation of B is also $(\lambda+1)(\lambda-3)$. Thus A and B have same eigenvalues. Will this be true always? The next theorem will give us the answer.

Theorem 4.10 *Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ be similar matrices. Then they have same characteristic polynomials.*

Proof Suppose that $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ are similar matrices. Then there exists an invertible matrix P such that $P^{-1}AP = B$. The characteristic polynomial of B is given by $det(B - \lambda I)$ where λ is an indeterminate. Now,

$$\begin{aligned} det(B - \lambda I) &= det(P^{-1}AP - \lambda P^{-1}P) \\ &= det\left[P^{-1}(A - \lambda I)\, P\right] \\ &= det(P^{-1})det(A - \lambda I)det\,(P) \\ &= det(A - \lambda I) \end{aligned}$$

That is, A and B have the same characteristic polynomial.

Corollary 4.4 *Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ be similar matrices, then they have same trace and determinant.*

Proof Since similar matrices have same characteristic polynomial, they have same eigenvalues with same algebraic multiplicities. Then by Corollary 4.2, they have same trace and determinant.

Theorem 4.11 *Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ be similar matrices. Then they have same rank and nullity.*

Proof Suppose that $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ are similar matrices. Then, there exist an invertible matrix P such that $P^{-1}AP = B$. Let $\{v_1, v_2, \ldots, v_k\}$ be a basis for $\mathcal{N}(B)$, where $k \le n$. Now for $i = 1, 2, \ldots, n$,

$$0 = Bv_i = P^{-1}APv_i \Rightarrow A(Pv_i) = 0$$

Hence, $\{Pv_1, Pv_2, \ldots, Pv_k\}$ is a subset of $\mathcal{N}(A)$. We will prove that this set will form a basis of $\mathcal{N}(A)$. Suppose that there exists scalars $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{K}$ such that

$$\mu_1 Pv_1 + \mu_2 Pv_2 + \cdots + \mu_k Pv_k = 0$$

which implies that

$$P(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k) = 0$$

Since P is invertible, we have

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k = 0$$

As $\{v_1, v_2, \ldots, v_k\}$ is a basis, the set is linearly independent. This implies that $\mu_1 = \mu_2 = \cdots = \mu_k = 0$. Therefore $\{Pv_1, Pv_2, \ldots, Pv_k\}$ is a linearly independent set.

Now let $v \in \mathcal{N}(A)$. Then, $Av = 0$. As $A = PBP^{-1}$, this gives $PBP^{-1}v = 0$ and P is invertible implies that $B(P^{-1}v) = 0$. Hence, $P^{-1}v \in \mathcal{N}(B)$. Then there exist scalars $\xi_1, \xi_2, \ldots, \xi_k \in \mathbb{K}$ such that

$$P^{-1}v = \xi_1 v_1 + \xi_2 v_2 + \cdots + \xi_k v_k$$

which implies that

$$v = \xi_1 P v_1 + \xi_2 P v_2 + \cdots + \xi_k P v_k$$

Thus, $\{Pv_1, Pv_2, \ldots, Pv_k\}$ spans $\mathcal{N}(A)$. Therefore $Nullity(A) = Nullity(B)$. Now by *Rank-Nullity theorem*, we get $Rank(A) = Rank(B)$.

Definition 4.5 (*Diagonalizability*) If $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is similar to a diagonal matrix, then A is said to be diagonalizable. That is, A is diagonalizable if there exists a diagonal matrix D such that $P^{-1}AP = D$ for some invertible matrix $P \in \mathbb{M}_{n\times n}(\mathbb{K})$.

Example 4.12 Consider the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

from Example 4.1. For $P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$, we get

$$P^{-1}AP = -\frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = D$$

Therefore A is diagonalizable. But how to find P such that $P^{-1}AP$ is a diagonal matrix? Notice that the columns of P are the eigenvectors of A. Also observe that the diagonal entries of D are not just any scalars but the eigenvalues of A. Interesting!!!

Example 4.13 Consider the matrix

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

from Example 4.7. For $P = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$, we get

$$P^{-1}AP = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = D$$

Therefore A is diagonalizable. Here also, observe that the diagonal entries of D the eigenvalues of A. What about the columns of P? Verify for yourself that the columns of P are the eigenvectors of A.

What about the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$ from Example 4.3? We will later see that there does not exist a matrix P such that $P^{-1}AP$ is a diagonal matrix. What could be the reason for non-diagonalizability? In Example 4.3, we have seen that A has only two linearly independent eigenvectors. Could this be the reason? The next theorem will provide us with an answer. The theorem establishes a necessary and sufficient condition for the diagonalizability of a square matrix.

Theorem 4.12 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Proof Suppose that $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is diagonalizable. Then there exist a diagonal matrix D, such that $P^{-1}AP = D$ for some nonsingular matrix $P \in \mathbb{M}_{n\times n}(\mathbb{K})$. Now

$$P^{-1}AP = D \Rightarrow AP = PD$$

We know that e_i is an eigenvector of D with eigenvalue d_{ii}, where d_{ii} denotes the ith diagonal entry of D and e_i is the ith element in the standard ordered basis for $\mathbb{K}^n$. That is, $De_i = d_{ii}e_i$. We will show that Pe_i is an eigenvector of A with eigenvalue d_{ii} for each $i = 1, 2, \ldots, n$. We have

$$A(Pe_i) = (AP)e_i = (PD)e_i = P(De_i) = d_{ii}(Pe_i), \ \forall\, i = 1, 2, \ldots, n$$

which implies that $\{Pe_1, Pe_2, \ldots, Pe_n\}$ are eigenvectors of A. Now will show that this set is linearly independent. Suppose that there exist $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{K}$ such that

$$\mu_1 Pe_1 + \mu_2 Pe_2 + \cdots + \mu_n Pe_n = 0$$

This implies that,

$$P(\mu_1 e_1 + \mu_2 e_2 + \cdots + \mu_n e_n) = 0$$

Since P is invertible, multiplying by P^{-1} on both sides, we get

$$\mu_1 e_1 + \mu_2 e_2 + \cdots + \mu_n e_n = 0$$

As $\{e_1, e_2, \ldots, e_n\}$ is linearly independent, we get $\mu_1 = \mu_2 = \cdots = \mu_n = 0$. Therefore $\{Pe_1, Pe_2, \ldots, Pe_n\}$ is linearly independent and hence A has n linearly independent eigenvectors.

Conversely, suppose that A has n linearly independent eigenvectors, say $v_1, v_2, \ldots, v_n$. Then take P as the $n \times n$ matrix with $v_1, v_2, \ldots, v_n$ as its columns. That is,

$$P = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}$$

Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be eigenvalues of A. Take D as the diagonal matrix with $d_{ii} = \lambda_i$. Now

$$AP = \begin{bmatrix} Av_1 & Av_2 & \dots & Av_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_n v_n \end{bmatrix} = PD$$

As the columns of P are linearly independent, P has $Rank\ n$ and hence is invertible. This implies that $P^{-1}AP = D$. Hence, A is diagonalizable.

Corollary 4.5 *If $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ has n distinct eigenvalues, then A is diagonalizable.*

Proof Suppose that $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ has n distinct eigenvalues. Then by Theorem 4.2, A has n linearly independent eigenvectors. Therefore A is diagonalizable.

Corollary 4.6 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then A is diagonalizable if and only if there exists a basis of $\mathbb{K}^n$ consisting of eigenvectors of A.*

Proof Suppose that $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is diagonalizable. Then A has n distinct eigenvectors. Then the collection of these linearly independent eigenvectors of A is a maximal linearly independent set in $\mathbb{K}^n$ and hence is a basis of $\mathbb{K}^n$.

Conversely, suppose that $\mathbb{K}^n$ has a basis consisting of eigenvectors of A. Then clearly, A has n linearly independent eigenvectors. Therefore A is diagonalizable.

Observe that the converse of Corollary 4.5 is not true. That is, A matrix is diagonalizable need not imply that it has n distinct eigenvalues. For example, consider the matrix $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$. We have seen Example 4.13 that the matrix A is diagonalizable. But A has only two distinct eigenvalues. Another example is the $n \times n$ identity matrix, I. As I is a diagonal matrix, it is clearly diagonalizable (Why?). But it does not have n distinct eigenvalues.

Example 4.14 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The characteristic polynomial of A is $(1-\lambda)^2(-1-\lambda)$. Therefore the eigenvalues of A are 1 and -1. Let us find the eigenspace corresponding to $\lambda_1 = 1$.

$$(A - I)v = 0 \Rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 3 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow v_1 = v_2$$

Thus

$$\mathcal{N}(A - I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = v_2\} = span\{(1, 1, 0), (0, 0, 1)\}$$

So we can pick two linearly independent eigenvectors of A corresponding to $\lambda_1 = 1$, say $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Now for $\lambda_2 = -1$,

$$(A + I)v = 0 \Rightarrow \begin{bmatrix} 3 & -1 & 0 \\ 3 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow 3v_1 - v_2 = 0, v_3 = 0$$

Thus

$$\mathcal{N}(A + I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_2 = 3v_1, v_3 = 0\} = span\{(1, 3, 0)\}$$

As $\mathcal{N}(A + I)$ is one dimensional space, pick one eigenvector corresponding to $\lambda_2 = -1$, say $\begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$. Thus A has three linearly independent eigenvectors. Therefore, by Theorem 4.12, A is diagonalizable. Note that, for $P = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}$, we get

$$P^{-1}AP = \frac{1}{2} \begin{bmatrix} 3 & -1 & 0 \\ 0 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Also, notice A does not have three distinct eigenvalues.

Although diagonalizability is an important property for matrices, every matrix need not be diagonalizable. The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 7 \end{bmatrix}$ from Example 4.3 is not diagonalizable as it has only two linearly independent eigenvectors. Here is another example for a non-diagonalizable matrix.

Example 4.15 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

The characteristic polynomial of A is $(2 - \lambda)^2(3 - \lambda)$. Therefore the eigenvalues of A are 2 and 3. Let us find the eigenspace corresponding to $\lambda_1 = 2$.

$$(A - 2I)v = 0 \Rightarrow \begin{bmatrix} 0 & -1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow v_2 = v_3 = 0$$

Thus

$$\mathcal{N}(A-2I)=\{(v_1,v_2,v_3)\in\mathbb{R}^3 \mid v_2=v_3=0\}=span\{(1,0,0)\}$$

Therefore we can pick one eigenvector corresponding to $\lambda_1=2$ is $\begin{bmatrix}1\\0\\0\end{bmatrix}$ from $\mathcal{N}(A-I)$. Now for $\lambda_2=3$

$$(A-3I)v=0 \Rightarrow \begin{bmatrix}-1&-1&2\\0&-1&-2\\0&0&0\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \Rightarrow v_1=-2v_2,\ v_3=-\frac{1}{2}v_2$$

Thus

$$\mathcal{N}(A-3I)=\left\{(v_1,v_2,v_3)\in\mathbb{R}^3 \mid v_1=-2v_2,\ v_3=-\frac{1}{2}v_2\right\}=span\{(4,-2,1)\}$$

Thus we can pick one eigenvector corresponding to $\lambda_2=3$, say $\begin{bmatrix}4\\-2\\1\end{bmatrix}$. The given matrix does not have three linear independent eigenvectors and hence by Theorem 4.12, A is not diagonalizable.

The next theorem gives a necessary and sufficient condition for diagonalizability of a square matrix A in terms of algebraic multiplicity and geometric multiplicity of its eigenvalues.

Theorem 4.13 *Let $A\in\mathbb{M}_{n\times n}(\mathbb{K})$. A is diagonalizable if and only if for every eigenvalue λ of A, the geometric multiplicity equals the algebraic multiplicity.*

Proof Suppose that A is diagonalizable. Then there exists a diagonal matrix D with $P^{-1}AP=D$, for some nonsingular matrix $P\in M_n(\mathbb{K})$. Let $\lambda_1,\lambda_2,\ldots,\lambda_k$ be distinct eigenvalues of A with algebraic multiplicities $m_1,m_2,\ldots,m_k$, respectively. The geometric multiplicity of an eigenvalue λ_i is equal to $Nullity\,(A-\lambda_i I)$. Since

$$D-\lambda_i I=P^{-1}AP-\lambda_i I=P^{-1}AP-\lambda_i P^{-1}P=P^{-1}(A-\lambda_i I)P$$

$A-\lambda_i I$ and $D-\lambda_i I$ are similar. Then by Theorem 4.11, they have same nullity. Therefore the geometric multiplicity of an eigenvalue λ_i is equal to $Nullity\,(D-\lambda_i I)=m_i=$ algebraic multiplicity of λ_i as D is a diagonal matrix with $\lambda_1,\lambda_2,\ldots,\lambda_k$ as diagonal entries and each λ_i repeats m_i times. That is, for every eigenvalue λ of A, the geometric multiplicity equals the algebraic multiplicity.

Conversely, suppose that for every eigenvalue λ of A, the geometric multiplicity equals the algebraic multiplicity. Let $\lambda_1,\lambda_2,\ldots,\lambda_k$ be distinct eigenvalues of A with algebraic multiplicities $m_1, m_2, \ldots, m_k$ respectively. Since the geometric multiplicity equals the algebraic multiplicity for every eigenvalue of A, $Nullity(A - \lambda_i I) = m_i$ for each i. Now corresponding to each λ_i, consider the basis $\{v_i^1, v_i^2, \ldots v_i^{m_i}\}$ of $\mathcal{N}(A - \lambda_i I)$, where m_i is the geometric multiplicity of λ_i for $i = 1, 2, \ldots, k$. Now consider the set $B = \{v_1^1, v_1^2, \ldots v_1^{m_1}, \ldots, v_k^1, v_k^2, \ldots v_k^{m_k}\}$. We will show that B is a basis for $\mathbb{K}^n$. Since $m_1 + m_2 + \cdots + m_k = n$, it is enough to prove that B is linearly independent. Now let $\mu_1^1, \mu_1^2, \ldots \mu_1^{m_1}, \ldots, \mu_k^1, \mu_k^2, \ldots \mu_k^{m_k}$ be scalars such that

$$\mu_1^1 v_1^1 + \mu_1^2 v_1^2 + \cdots + \mu_1^{m_1} v_1^{m_1} + \cdots + \mu_k^1 v_k^1 + \mu_k^2 v_k^2 + \cdots + \mu_k^{m_k} v_k^{m_k} = 0$$

Consider the collection $\chi = \left\{\sum_{i=1}^{m_1} \mu_i^i v_i^i, \ldots, \sum_{i=1}^{m_k} \mu_i^i v_i^i\right\}$. Since $\{v_i^1, v_i^2, \ldots v_i^{m_i}\}$ is a basis of $\mathcal{N}(A - \lambda_i I)$, a linear combination of $v_i^1, v_i^2, \ldots v_i^{m_i}$ is either zero vector or an eigenvector corresponding to λ_i. This is true for every $i = 1, 2, \ldots, k$. For an element in χ to be the zero vector, the coefficients must be zero since each element χ is a linear combination of a linearly independent set. Now consider the remaining non-zero elements in χ. Since eigenvectors corresponding to distinct eigenvalues are linearly independent, a linear combination of non-zero vectors in χ implies that all the coefficients are zero. Therefore B is linearly independent and hence is a basis of $\mathbb{R}^n$. Therefore A is diagonalizable.

Example 4.16 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

from Example 4.14. We have seen that A has two eigenvalues 1 and -1. From Example 4.14, $AM(1) = GM(1) = 2$ and $AM(-1) = GM(-1) = 1$. As the algebraic multiplicity equals geometric multiplicity for every eigenvalue of A, the matrix A is diagonalizable.

Example 4.17 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

in Example 4.15. A has two eigenvalues 2 and 3. From Example 4.15, we have $AM(2) = 2$, but $GM(2) = 1$. Therefore A is not diagonalizable.

The next theorem shows that an $n \times n$ matrix A is diagonalizable if and only if $\mathbb{K}^n$ can be written as a direct sum of eigenspaces of A.

Theorem 4.14 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then A is diagonalizable if and only if*

$$\mathbb{K}^n = \mathcal{N}(A-\lambda_1 I) \oplus \mathcal{N}(A-\lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A-\lambda_k I)$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A.

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A, with algebraic multiplicities $m_1, m_2, \ldots, m_k$, respectively. Suppose that A is diagonalizable. Then by Theorem 4.13, $m_1 + m_2 + \cdots + m_k = n$ and the union of basis vectors of $\mathcal{N}(A-\lambda_i I)$, where $i = 1, 2, \ldots, k$ forms a basis for $\mathbb{K}^n$. Then by Theorem 2.22, $\mathbb{K}^n = \mathcal{N}(A-\lambda_1 I) \oplus \mathcal{N}(A-\lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A-\lambda_k I)$.

Now suppose that $\mathbb{K}^n = \mathcal{N}(A-\lambda_1 I) \oplus \mathcal{N}(A-\lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A-\lambda_k I)$. Again the union of basis vectors of $\mathcal{N}(A-\lambda_i I)$, where $i = 1, 2, \ldots, k$ forms a basis for $\mathbb{K}^n$. That is, $\mathbb{K}^n$ has a basis consisting of eigenvectors. Then by Corollary 4.6, A is diagonalizable.

Thus if A has n distinct eigenvalues, we can write $\mathbb{K}^n$ as a direct sum of n one dimensional spaces. Let us consider some examples to verify Theorem 4.14.

Example 4.18 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

from Example 4.14 which is diagonalizable. We have seen that the eigenvalues of A are 1 and -1. Also,

$$\mathcal{N}(A-I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = v_2\}$$

and

$$\mathcal{N}(A+I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_2 = 3v_1, v_3 = 0\}$$

Clearly, $\mathcal{N}(A-I) + \mathcal{N}(A+I) = \mathbb{R}^3$ and $\mathcal{N}(A-I) \cap \mathcal{N}(A+I) = \{0\}$(Verify). Then by Theorem 2.20,

$$\mathcal{N}(A-I) \oplus \mathcal{N}(A+I) = \mathbb{R}^3$$

which verifies Theorem 4.14.

Example 4.19 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

from Example 4.15, which is not diagonalizable. We have seen that the eigenvalues of A are 2 and 3. Also

$$\mathcal{N}(A-2I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_2 = v_3 = 0\} = span\{(1, 0, 0)\}$$

and

$$\mathcal{N}(A-3I) = \left\{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = -2v_2, v_3 = -\frac{1}{2}v_2\right\} = span\{(4, -2, 1)\}$$

Both $\mathcal{N}(A-2I)$ and $\mathcal{N}(A-3I)$ are of dimension 1 each. Clearly

$$\mathcal{N}(A-2I) + \mathcal{N}(A-3I) \neq \mathbb{R}^3$$

We can also check whether a matrix is diagonalizable or not by finding its minimal polynomial. The next theorem states that a matrix A is diagonalizable if and only if its minimal polynomial does not have any repeated roots.

Theorem 4.15 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then A is diagonalizable if and only if the minimal polynomial of A has no repeated roots.*

Proof Suppose that A is diagonalizable. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be distinct eigenvalues of A. Then by the above theorem,

$$\mathbb{K}^n = \mathcal{N}(A-\lambda_1 I) \oplus \mathcal{N}(A-\lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A-\lambda_k I)$$

Therefore $\mathbb{K}^n$ have a basis consisting of eigenvectors of A. We will show that

$$q(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$$

is the minimal polynomial of A.Since characteristic polynomial and minimal polynomial have same roots it is enough to show that

$$q(A) = (A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_k I) = 0$$

Let v be an eigenvector of A, then $(A-\lambda_i I)v = 0$ for some i. Since the matrices $(A-\lambda_i I)$ commutes with each other, $q(A)v = 0$ for every eigenvector of A and as collection of all eigenvectors forms a basis, $q(A) = 0$. Therefore $q(\lambda)$ is the minimal polynomial of A and it has no repeated roots.

Conversely, suppose that the minimal polynomial of A has no repeated roots. Let

$$q(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$$

be the minimal polynomial of A. Then $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the distinct eigenvalues of A. We will show that

$$\mathbb{K}^n = \mathcal{N}(A - \lambda_1 I) \oplus \mathcal{N}(A - \lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A - \lambda_k I)$$

We have already shown that the union of basis vectors of $\mathcal{N}(A - \lambda_i I)$, where $i = 1, 2, \ldots, k$ is a linearly independent set. Now it is enough to show that

$$\mathbb{K}^n = \mathcal{N}(A - \lambda_1 I) + \mathcal{N}(A - \lambda_2 I) + \cdots + \mathcal{N}(A - \lambda_k I)$$

Now consider the polynomials

$$f_i(\lambda) = q(\lambda_i)/(\lambda - \lambda_i) = \Pi_{j \neq i}(\lambda - \lambda_i)$$

Since $\lambda_i' s$ are distinct, by Theorem 1.16, there exists polynomials $g_1(\lambda), g_2(\lambda), \ldots, g_k(\lambda) \in \mathbb{K}[\lambda]$ such that

$$\sum_{i=1}^{k} g_i(\lambda) f_i(\lambda) = 1$$

Then

$$g_1(A) f_1(A)v + g_2(A) f_2(A)v + \cdots + g_2(A) f_2(A)v = v$$

for any $v \in \mathbb{K}^n$. Also

$$(A - \lambda_i I) f_i(A)v = q(A)v = 0$$

for each i. This implies that $f_i(A)v \in \mathcal{N}(A - \lambda_i I)$ for each i. That is, any $v \in \mathbb{K}^n$ can be written as a linear combination of elements in $\mathcal{N}(A - \lambda_i I)$. Therefore

$$\mathbb{K}^n = \mathcal{N}(A - \lambda_1 I) + \mathcal{N}(A - \lambda_2 I) + \cdots + \mathcal{N}(A - \lambda_k I)$$

and hence A is diagonalizable.

Example 4.20 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

from Example 4.14. The characteristic polynomial of A is $(1 - \lambda)^2(-1 - \lambda)$. The minimal polynomial has possibilities, $(1 - \lambda)(-1 - \lambda)$ and $(1 - \lambda)^2(-1 - \lambda)$. As

$$(I - A)(-I - A) = \begin{bmatrix} 1 & -1 & 0 \\ 3 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & -1 & 0 \\ 3 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$(1 - \lambda)(-1 - \lambda)$ is the minimal polynomial of A. Clearly, minimal polynomial has no repeated roots. We have already seen that A is diagonalizable.

Example 4.21 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

from Example 4.15. The characteristic polynomial of A is $(2-\lambda)^2(3-\lambda)$. The minimal polynomial of A has two possibilities, $(2-\lambda)(3-\lambda)$ and $(2-\lambda)^2(3-\lambda)$. As

$$(2I - A)(3I - A) = \begin{bmatrix} 0 & -1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$(2-\lambda)^2(3-\lambda)$ is the minimal polynomial. Clearly, minimal polynomial has repeated roots. We have seen that A is not diagonalizable.

4.3 Schur Triangularization Theorem

We have seen that for a matrix A, there may not exist n linearly independent eigenvectors. In this case, A is not diagonalizable. But we may obtain an *almost diagonal representation* for A in such cases. Next, we will be discussing the *almost diagonal representation* of a non-diagonalizable matrix A.

Definition 4.6 (*Triangular Matrix*) A square matrix is called upper (lower) triangular if all the entries below(above) the main diagonal are zero. A square matrix which is either upper triangular or lower triangular is called a triangular matrix.

Remark 4.5 The collection of all upper (lower) triangular matrices forms a vector space under matrix addition and scalar multiplication. Under matrix multiplication, it does not form a group, since the eigenvalues of an upper triangular matrices are its diagonal entries, an upper triangular matrix is invertible if and only if the diagonal entries are non-zero.

Definition 4.7 (*Triangularizable Matrix*) A square matrix is called triangularizable if it is similar to an upper(lower) triangular matrix.

Definition 4.8 (*Unitary matrix*) A matrix $A \in \mathbb{M}_{n\times n}(\mathbb{C})$ is called a unitary matrix if $AA^* = I = A^*A$. That is, if $A^* = A^{-1}$. If $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ and $A^TA = I = AA^T$, then A is called an orthogonal matrix.

Theorem 4.16 (Schur Triangularization Theorem) *Let $A \in \mathbb{M}_{n\times n}(\mathbb{C})$, then there exists a unitary matrix $U \in \mathbb{M}_{n\times n}(\mathbb{C})$ such that U^*AU is upper triangular.*

Proof The proof is by induction on n. If $n = 1$, then clearly A is an upper triangular matrix. Now suppose that the result is true for all $(n-1) \times (n-1)$ matrices. Now let $A \in \mathbb{M}_{n\times n}(\mathbb{C})$. Let $\lambda_1 \in \mathbb{C}$ be an eigenvalue of A (Such an eigenvalue exists, by *Fundamental theorem of Algebra*) and $v \in \mathbb{C}^n$ be the corresponding eigenvector. Take $u = \frac{v}{\|v\|}$. Then $\| u \| = 1$. Let $U_1 \in \mathbb{M}_{n\times n}(\mathbb{C})$ be a unitary matrix with u as its first column. The existence of such a matrix is guaranteed by *Gram–Schmidt Orthonormalization* (which will be discussed later) process. Then consider the matrix $U_1^* A U_1$. Its first column is given by,

$$U_1^* A U_1 e_1 = U_1^* A u = \lambda_1 U_1^* u = \lambda_1 U_1^* U_1 e_1 = \lambda_1 e_1$$

Therefore $U_1^* A U_1$ is of the form $\begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & A_1 \end{bmatrix}$, where A_1 is an $(n-1) \times (n-1)$ matrix and $\mathbf{0}$ is an $(n-1)$ column vector. Now by induction hypothesis, there exists a unitary matrix $\tilde{U_1} \in M_{n\times n}(\mathbb{C})$ such that $\tilde{U_1}^* A_1 \tilde{U_1}$ is an upper triangular matrix. Take

$$U_2 = \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \tilde{U_1} \end{bmatrix}$$

Then U_2 is unitary and

$$U_2{}^*(U_1^* A U_1) U_2 = \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & \tilde{U_1}^* A_1 \tilde{U_1} \end{bmatrix}$$

which is an upper triangular matrix. Take $U = U_1 U_2$. Then

$$UU^* = U_1 U_2 (U_1 U_2)^* = U_1 U_2 U_2^* U_1^* = I$$

and

$$U^* U = (U_1 U_2)^* U_1 U_2 = U_2^* U_1^* U_1 U_2 = I$$

Therefore U is unitary and $U^* A U$ is upper triangular.

Corollary 4.7 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ be with all its eigenvalues are real. Then there exists an orthogonal matrix $Q \in \mathbb{M}_{n\times n}(\mathbb{R})$ such that $Q^T A Q$ is an upper triangular matrix.*

4.4 Generalized Eigenvectors

In Sect. 4.2, we have seen some necessary and sufficient conditions for the diagonalizability of an $n \times n$ matrix A. We can summarize these conditions as follows. $A \in \mathbb{K}^n$ is diagonalizable if and only if

(1) A has n linearly independent eigenvectors.
(2) for every eigenvalue λ of A, the geometric multiplicity equals the algebraic multiplicity.
(3) $\mathbb{K}^n = \mathcal{N}(A - \lambda_1 I) \oplus \mathcal{N}(A - \lambda_2 I) \oplus \cdots \oplus \mathcal{N}(A - \lambda_k I)$ where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A.
(4) the minimal polynomial of A has no repeated roots.

We can say that each eigenspace W is *invariant* under the action of A. By the term *invariant*, we mean that if A acts on any element in W, we get the image in W itself. Thus the diagonalizability of a matrix $A \in \mathbb{K}^n$, could mean that $\mathbb{K}^n$ can be represented as the direct sum of its subspaces that are invariant under the action of A. Or it could mean that $\mathbb{K}^n$ has a basis consisting of eigenvectors of A. Remember that, we have also seen examples for matrices that are not diagonalizable. If A is not diagonalizable, is it possible to represent $\mathbb{K}^n$ as a direct sum of subspaces that are invariant under the action of the matrix A? In this section, for a given non-diagonalizable matrix A, we will find subspaces of $\mathbb{K}^n$ that invariant under the action of A and we will represent $\mathbb{K}^n$ as a direct sum of these invariant subspaces. In order to achieve this goal, we will introduce the concept of generalized eigenvectors for a matrix. Let us start by an example.

Example 4.22 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

from Example 4.15. We have seen that 3 and 2 are the eigenvalues of A with corresponding eigenvectors $u = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$ and $v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ respectively. In Example 4.17, we have noted that $AM(2)$ is 2, but $GM(2)$ is 1. That is, $GM(2)$ is one less than $AM(2)$. Now let us find a vector $w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \in \mathbb{R}^3$ such that $(A - 2I)w = v$. Now

$$(A - 2I)w = v \Rightarrow \begin{bmatrix} 0 & -1 & 2 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \Rightarrow w_2 = -1, w_3 = 0 \qquad (4.9)$$

Let us denote the set of all $w \in \mathbb{R}^3$ which satisfy $(A - 2I)w = v$ by W. Then by Eq. 4.9, we have

$$W = \{(w_1, w_2, w_3) \in \mathbb{R}^3 \mid w_2 = -1, w_3 = 0\}$$

Observe that, for any $w \in W$, $(A - 2I)w = v$ and $(A - 2I)^2 w = (A - 2I)v = 0$. Take the vector $\tilde{w} \in W$ and consider the span of $\{v, w\}$, which is the subspace

$U = \{(u_1, u_2, 0) \in \mathbb{R}^3 \mid u_1, u_2 \in \mathbb{R}\}$ (Verify). Then for any vector $\tilde{u} \in U$, there exists scalars $\mu_1, \mu_2 \in \mathbb{R}$ such that

$$\tilde{u} = \mu_1 v + \mu_2 \tilde{w}$$

Then

$$A\tilde{u} = A(\mu_1 v + \mu_2 \tilde{w}) = \mu_1 Av + \mu_2 A\tilde{w} \tag{4.10}$$

As $(A - 2I)\tilde{w} = v$, we have $A\tilde{w} = v + 2\tilde{w}$. Also $Av = 2v$, as v is an eigenvector of A corresponding to the eigenvalue 2. Substituting this in Eq. (4.10), we get

$$A\tilde{u} = 2\mu_1 v + \mu_2 v + 2\mu_2 \tilde{w} = (2\mu_1 + \mu_2)v + 2\mu_2 \tilde{w} \in span\{v, \tilde{w}\} = U$$

Therefore we can say that U is invariant under A.

Now let $\tilde{U}$ denote the span of $u = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$. Then we can observe that $U + \tilde{U} = \mathbb{R}^3$ and $U \cap \tilde{U} = \{0\}$. Hence $U \oplus \tilde{U} = \mathbb{R}^3$ and $\{u, v, \tilde{w}\}$ is a basis for $\mathbb{R}^3$. Thus U is a subspace of $\mathbb{R}^3$ that can be associated with A and satisfying all our requirements. This idea motivates the definition of generalized eigenvector associated with a matrix.

Definition 4.9 (*Invariant Subspace*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and let W be a subspace of $\mathbb{K}^n$, then W is called $A-$ invariant if $Aw \in W$ for each vector $w \in W$. Clearly the one dimensional invariant subspaces correspond to eigenvectors.

Example 4.23 Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then, clearly $\{0\}$, $\mathbb{K}^n$, $\mathcal{N}(A)$ and $\mathcal{R}(A)$ are invariant subspaces of A.(Verify)

The geometrical significance of an eigenvector v of a matrix A corresponding to a real eigenvalue λ was that the vector under action by A will remain on its span. We have observed this fact for the eigenvectors of the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ in Fig. 4.1. Thus we can say that the eigenspaces of a matrix A are invariant under the action of A. Consider the following example.

Example 4.24 Let A be an $n \times n$ matrix with an eigenvalue λ. To prove $\mathcal{N}(A - \lambda I)$ is invariant under A, it is enough to show that for any $v \in \mathcal{N}(A - \lambda I)$, $Av = \lambda v$ also belongs to $\mathcal{N}(A - \lambda I)$. We already know that for any scalar $\mu \in \mathbb{K}$,

$$A(\mu v) = \mu(Av) = \mu(\lambda v) = \lambda(\mu v)$$

Clearly $Av \in \mathcal{N}(A - \lambda I)$ for any $v \in \mathcal{N}(A - \lambda I)$. Therefore $\mathcal{N}(A - \lambda I)$ is invariant under A.

Definition 4.10 (*Generalized Eigenvector*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. A vector $v \in \mathbb{K}^n$ is said to be a generalized eigenvector of rank m of the matrix A and the corresponding eigenvalue $\lambda \in \mathbb{K}$, if $(A - \lambda I)^m v = 0$ but $(A - \lambda I)^{m-1} v \neq 0$. Clearly, an ordinary eigenvector is a generalized eigenvector of rank 1.

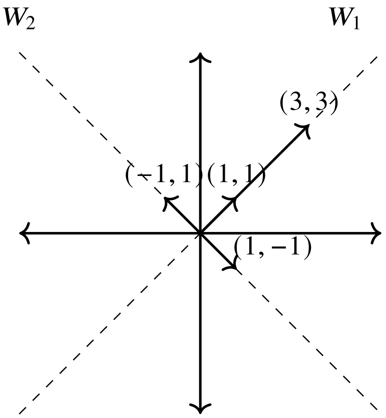


Fig. 4.1 As we can see, the vector $(-1, 1)$ is mapped to $(1, -1) = (-1)(-1, 1)$ and $(1, 1)$ is mapped to $(3, 3) = 3(1, 1)$. It is interesting to observe that the eigenvalue -1 changes the direction of any vector that lies on the line $y = -x$ and the eigenvalue 3 scales the magnitude of any vector that lies on the line $y = x$, three times

The generalized eigenspace of A corresponding to λ is denoted by E_λ and is given by $E_\lambda = \cup_{i=1}^{\infty} \mathcal{N}(A - \lambda I)^i$.

Definition 4.11 (*Chain*) Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and let $\lambda \in \mathbb{K}$ be an eigenvalue of A, then a set of non-zero vectors $v_1, v_2, \ldots, v_m \in \mathbb{K}^n$ is called a chain of generalized eigenvectors of length m corresponding to λ if

$$(A - \lambda I)v_i = \begin{cases} v_{i+1}, & when\ i < m \\ 0, & when\ i = m \end{cases}$$

Remark 4.6 Let $v_1, v_2, \ldots, v_m \in \mathbb{K}$ be a chain of generalized eigenvectors of length m of the matrix $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ corresponding to the eigenvalue λ, then by definition $(A - \lambda I)^i v_1 = v_{i+1}$ if $i < m$ and $(A - \lambda I)^m v_1 = 0$. Therefore the chain of generalized eigenvectors of length m corresponding to λ can also be written as $v_1, (A - \lambda I)v_1, \ldots (A - \lambda I)^{m-1} v_1$. Also $(A - \lambda I)^k v_i = 0$ for all $k \geq m$ for all $i = 1, 2, \ldots, m$.

Example 4.25 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

From Example 4.22, the chain of generalized eigenvectors associated with the eigenvalue 2 is $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and the chain associated with the eigenvalue 3 is $\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$.

Observe that the chain of generalized eigenvectors corresponding to 2 is linearly independent. This will be true for every eigenvalue. Consider the following theorem.

Theorem 4.17 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and let $\lambda \in \mathbb{K}$ be an eigenvalue of A, then the chain of generalized eigenvectors corresponding to λ is linearly independent.*

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $\lambda \in \mathbb{K}$ is an eigenvalue of A. Let $v_1, v_2, \ldots, v_m \in \mathbb{K}^n$ be the chain of generalized eigenvectors of length m corresponding to λ. Let $\mu_1, \mu_2, \ldots, \mu_m \in \mathbb{K}$ be scalars such that

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_m v_m = 0$$

Then we have

$$\mu_1 v_1 + \mu_2 (A - \lambda I) v_1 + \cdots + \mu_m (A - \lambda I)^{m-1} v_1 = 0 \tag{4.11}$$

Now multiplying both sides with $(A - \lambda I)^{m-1}$, we get

$$\mu_1 (A - \lambda I)^{m-1} v_1 = 0$$

as $(A - \lambda I)^k v_i = 0$ for all $k \geq m$. Since $(A - \lambda I)^{m-1} v_1 \neq 0$, this implies that $\mu_1 = 0$. Now multiply Eq. (4.11) by $(A - \lambda I)^{m-2}$. Then we will get

$$\mu_2 (A - \lambda I)^{m-1} v_1 = 0$$

which implies that $\mu_2 = 0$. Proceeding like this, we get $\mu_i = 0$ for all $i = 1, 2, \ldots, m$. Therefore $\{v_1, v_2, \ldots, v_m\}$ is linearly independent.

The following theorem gives the relation between the null spaces of the powers of a matrix A.

Theorem 4.18 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then*

$$\{0\} = \mathcal{N}(A^0) \subseteq \mathcal{N}(A^1) \subseteq \mathcal{N}(A^2) \subseteq \cdots$$

If $\mathcal{N}(A^m) = \mathcal{N}(A^{m+1})$, for some integer m, then

$$\mathcal{N}(A^m) = \mathcal{N}(A^{m+1}) = \mathcal{N}(A^{m+2}) = \mathcal{N}(A^{m+3}) = \cdots$$

Proof Suppose that $v \in \mathcal{N}(A^k)$ for some positive integer k, then

$$A^{k+1} v = A(A^k v) = 0 \Rightarrow v \in \mathcal{N}(A^{k+1})$$

Therefore $\mathcal{N}(A^k) \subseteq \mathcal{N}(A^{k+1})$ for any positive integer k. Now let $\mathcal{N}(A^m) = \mathcal{N}(A^{m+1})$, for some integer m. Let $v \in \mathcal{N}(A^{m+k+1})$ for some positive integer k, then

$$A^{m+k+1}v = A^{m+1}(A^k v) = 0 \Rightarrow A^k v \in \mathcal{N}(A^{m+1}) = \mathcal{N}(A^m) \Rightarrow A^{m+k}(v) = 0$$

Therefore $\mathcal{N}(A^{m+k+1}) \subseteq \mathcal{N}(A^{m+k})$ for every positive integer k. This implies that $\mathcal{N}(A^m) = \mathcal{N}(A^{m+1}) = \mathcal{N}(A^{m+2}) = \mathcal{N}(A^{m+3}) = \cdots$.

Remark 4.7 Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then

$$V = \mathcal{R}(A^0) \supseteq \mathcal{R}(A^1) \supseteq \mathcal{R}(A^2) \supseteq \cdots$$

Also, by *Rank-Nullity Theorem*, $Rank(A^i) = n - Nullity(A^i)$ for all i. Since $\mathcal{N}(A^i)$ form an increasing sequence of subspaces of a finite-dimensional space V, it will eventually stop increasing at some point j. Then we have proven that $\mathcal{N}(A^i) = \mathcal{N}(A^j)$ for all $i \geq j$. At this point, the sequence $\mathcal{R}(A^i)$ also stops decreasing and $Rank(A^i) = n - Nullity(A^j)$ for all $i \geq j$.

In Example 4.24, we have shown that the eigenspaces of a matrix A are invariant under A. Now we will show that the generalized eigenspace is also an invariant subspace of A.

Theorem 4.19 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $\lambda \in \mathbb{K}$ be an eigenvalue of A, then*

(a) $E_\lambda = \mathcal{N}\left((A - \lambda I)^n\right)$.
(b) E_λ *is invariant under A.*

Proof (a) Assume that $\mathcal{N}\left((A - \lambda I)^n\right) \neq \mathcal{N}\left((A - \lambda I)^{n+1}\right)$, for all n, then by the previous theorem

$$\{0\} = \mathcal{N}(A - \lambda I)^0 \subset \mathcal{N}(A - \lambda I)^1 \subset \mathcal{N}(A - \lambda I)^2 \subset \cdots \subset \mathcal{N}(A - \lambda I)^{n+1}$$

Then

$$0 < Nullity(A - \lambda I) < Nullity(A - \lambda I)^2 < \cdots < Nullity(A - \lambda I)^{n+1} \leq dim(V) = n$$

which is not possible, since it implies that $Nullity(A - \lambda I), \ldots, Nullity(A - \lambda I)^{n+1}$ are $n + 1$ distinct integers in $\{1, 2, \ldots, n\}$. Therefore

$$\mathcal{N}\left((A - \lambda I)^n\right) = \mathcal{N}\left((A - \lambda I)^{n+1}\right)$$

and hence $E_\lambda = \mathcal{N}\left((A - \lambda I)^n\right)$.

(b) Let $v \in E_\lambda = \mathcal{N}\left((A - \lambda I)^n\right)$, then

$$(A - \lambda I)^n (Av) = A\left((A - \lambda I)^n v\right) = A(0) = 0$$

Therefore $Av \in E_\lambda = \mathcal{N}\left((A - \lambda I)^n\right)$. That is, E_λ is invariant under A.

Example 4.26 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

from Example 4.15. We have already seen that

$$\mathcal{N}(A - 2I) = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_2 = v_3 = 0\}$$

Now $(A-2I)^2 = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ and hence

$$(A - 2I)^2 v = 0 \Rightarrow \begin{bmatrix} 0 & 0 & 4 \\ 0 & 0 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow v_3 = 0$$

Therefore $\mathcal{N}(A-2I)^2 = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_3 = 0\}$. Observe that $(A-2I)^2 = (A-2I)^3$. Thus we have $\mathcal{N}(A-2I)^2 = \mathcal{N}(A-2I)^3$. Therefore the generalized eigenspace associated with the eigenvalue 2 is

$$E_2 = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_3 = 0\}$$

Also, as $\mathcal{N}(A-3I) = \mathcal{N}(A-3I)^2 = \left\{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = -2v_2, v_3 = -\frac{1}{2}v_2\right\}$, we get

$$E_3 = \left\{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = -2v_2, v_3 = -\frac{1}{2}v_2\right\}$$

as we have seen in Example 4.22.

Theorem 4.20 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Let $f(\lambda), g(\lambda) \in \mathbb{K}[\lambda]$ be polynomials such that $(f, g) = 1$ and $f(A)g(A) = 0$. Then $\mathbb{K}^n = \mathcal{N}(f(A)) \oplus \mathcal{N}(g(A))$.*

Proof Let $f(\lambda)$ and $g(\lambda)$ be polynomials such that $(f, g) = 1$ and $f(A)g(A) = 0$. Then by Theorem 1.16, there exists $r(\lambda), s(\lambda) \in \mathbb{K}[\lambda]$, such that $rf + sg = 1$. Now, for $v \in \mathbb{K}^n$, define

$$v_1 = s(A)g(A)v \text{ and } v_2 = r(A)f(A)v$$

Then

$$v = Iv = (s(A)g(A) + r(A)f(A))v = v_1 + v_2$$

Also,

$$f(A)v_1 = f(A)s(A)g(A)v = 0 \text{ and } g(A)v_2 = g(A)r(A)f(A)v = 0$$

Thus, for $v \in \mathbb{K}^n$ there exists $v_1 \in \mathcal{N}(f(A))$ and $v_2 \in \mathcal{N}(g(A))$ such that $v = v_1 + v_2$.

Now we have to prove that, this expression is unique. Suppose that $v = w_1 + w_2$ where $w_1 \in \mathcal{N}(f(A))$ and $w_2 \in \mathcal{N}(g(A))$. Then, since $f(A)(v_1 - w_1) = 0$ and $g(A)(v_2 - w_2) = 0$, we have $v_1 - w_1 \in \mathcal{N}(f(A))$ and $v_2 - w_2 \in \mathcal{N}(g(A))$. Now

$$v_1 - w_1 = I(v_1 - w_1) = (s(A)g(A) + r(A)f(A))(v_1 - w_1) = 0$$

and

$$v_2 - w_2 = I(v_2 - w_2) = (s(A)g(A) + r(A)f(A))(v_2 - w_2) = 0$$

Hence, $v_1 = w_1$ and $v_2 = w_2$. Therefore for every $v \in \mathbb{K}^n$, there exists a unique expression as a sum of elements from $\mathcal{N}(f(A))$ and $\mathcal{N}(g(A))$ and hence $\mathbb{K}^n = \mathcal{N}(f(A)) \oplus \mathcal{N}(g(A))$.

Corollary 4.8 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $f(\lambda) \in \mathbb{K}[\lambda]$. Let $f_1(\lambda), f_2(\lambda), \ldots, f_k(\lambda) \in \mathbb{K}[\lambda]$ be such that $(f_i, f_j) = 1$ for all $i \neq j$, $f = f_1 f_2 \cdots f_k$ and $f(A) = 0$. Then $\mathbb{K}^n = \mathcal{N}(f_1(A)) \oplus \mathcal{N}(f_2(A)) \oplus \cdots \oplus \mathcal{N}(f_k(A))$.*

The following theorem shows that for any matrix A, we can write $\mathbb{K}^n$ as the direct sum of generalized eigenspaces associated with A.

Theorem 4.21 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $p(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$ be the characteristic polynomial of A, then*

$$\mathbb{K}^n = \mathcal{N}((A - \lambda_1 I)^{n_1}) \oplus \mathcal{N}((A - \lambda_2 I)^{n_2}) \oplus \cdots \oplus \mathcal{N}((A - \lambda_k I)^{n_k})$$

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $p(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_k)^{n_k}$ be the characteristic polynomial of A. Now, take $f = p$ and $f_i = (\lambda - \lambda_i)^{n_i}$ where $i = 1, 2, \ldots, k$. Then $(f_i, f_j) = 1$ for all $i \neq j$ and $f = f_1 f_2 \cdots f_k$. Also $f(A) = p(A) = 0$. Therefore $\mathbb{K}^n = \mathcal{N}((A - \lambda_1 I)^{n_1}) \oplus \mathcal{N}((A - \lambda_2 I)^{n_2}) \oplus \cdots \oplus \mathcal{N}((A - \lambda_k I)^{n_k})$ by the above corollary.

Corollary 4.9 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Let $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ be distinct eigenvalues of A with algebraic multiplicities $n_1, n_2, \ldots, n_k$, respectively, then $dim\ (\mathcal{N}(A - \lambda_i I)^n) = n_i$.*

Proof By Theorem 4.19, as λ_i is the only eigenvalue of A when $\mathcal{N}(A - \lambda_i I)^n$ is considered as domain. Therefore $dim\ (\mathcal{N}(A - \lambda_i I)^n) \leq n_i$ and by the above theorem, $\sum_{i=1}^{k} dim\ (\mathcal{N}(A - \lambda_i I)^n) = n = \sum_{i=1}^{k} n_i$. Hence $dim\ (\mathcal{N}(A - \lambda_i I)^n) = n_i$.

In Theorem 4.2, we have seen that the eigenvectors associated with distinct eigenvalues are linearly independent. This result is squarely applicable for generalized eigenvectors also. Consider the following corollary.

Corollary 4.10 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Let $v_1, v_2, \ldots, v_k$ be generalized eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$. Then $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.*

Example 4.27 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

In Example 4.21, we have seen that the minimal polynomial of A is $(2-\lambda)^2(3-\lambda)$. Also, from Example 4.26,

$$\mathcal{N}(A-2I)^2 = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_3 = 0\}$$

and

$$\mathcal{N}(A-3I) = \left\{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = -2v_2, v_3 = -\frac{1}{2}v_2\right\}$$

Clearly, $\mathbb{R}^3 = \mathcal{N}(A-2I)^2 \oplus \mathcal{N}(A-3I)$.

Theorem 4.22 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A. Let r_i be the least positive integer such that $(A-\lambda_i I)^{r_i} v = 0$ for every $v \in \mathcal{N}((A-\lambda_i I)^n)$. Let $m(\lambda) = (\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}$, then $m(\lambda)$ is the minimal polynomial of A.*

Proof Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ and $q(\lambda) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$ be the minimal polynomial of A. Then by the previous theorem, $\mathbb{K}^n = \mathcal{N}((A-\lambda_1 I)^{m_1}) \oplus \mathcal{N}((A-\lambda_2 I)^{m_2}) \oplus \cdots \oplus \mathcal{N}((A-\lambda_k I)^{m_k})$. Therefore for every $v \in \mathbb{K}^n$, we get $m(A)v = 0$. But as $q(\lambda)$ is the least degree polynomial such that $q(A) = 0$, $m = q$.

4.5 Jordan Canonical Form

In this section, using the idea of similar matrices, we will reduce a square matrix to a block diagonal matrix form, where each diagonal blocks are upper triangular matrices.

Definition 4.12 (*Jordan Block*) A Jordan block corresponding to λ of size m is an $m \times m$ matrix of the form $J_\lambda^m = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}$, where λ lies on the diagonal entries, 1 lies on the super-diagonal (some authors prefer sub-diagonal) and the missing entries are all zero.

Definition 4.13 (*Jordan Form*) A square matrix is said to be in Jordan form, if it is a block diagonal matrix with each block as a Jordan block.

Theorem 4.23 *Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ with characteristic polynomial $p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_k)^{n_k}$ and minimal polynomial $q(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$. Then A is similar to a matrix J with Jordan blocks along the main diagonal. All other elements of J are zero. Corresponding to each eigenvalue λ_i of A, there exists atleast one Jordan block of size m_i and other Jordan blocks corresponding to λ_i have size less than or equal to m_i. Also, the sum of size of each Jordan blocks corresponding to λ_i is n_i. The number of Jordan blocks corresponding to an eigenvalue λ_i is the geometric multiplicity of λ_i.*

Proof Let $A \in M_n(\mathbb{K})$ with characteristic polynomial

$$p(\lambda) = (\lambda - \lambda_1)^{n_1} \cdots (\lambda - \lambda_k)^{n_k}$$

and minimal polynomial

$$q(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_k)^{m_k}$$

Then by Theorem 4.21,

$$\mathbb{K}^n = \mathcal{N}((A - \lambda_1 I)^{n_1}) \oplus \mathcal{N}((A - \lambda_2 I)^{n_2}) \oplus \cdots \oplus \mathcal{N}((A - \lambda_k I)^{n_k})$$

Let $l_1, l_2, \ldots, l_k$ be the geometric multiplicities the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ respectively. Now we can construct a basis for $\mathcal{N}((A - \lambda_i I)^{n_i})$ for each $i = 1, 2, \ldots, k$. Let $v_{i1}^1, v_{i1}^2, \ldots, v_{i1}^{l_i}$ be the linearly independent generalized eigenvectors of A with rank $r_{i1}, r_{i2}, \ldots, r_{il_i}$ corresponding to λ_i. Then $r_{ij} \le m_i$ for all $j = 1, 2, \ldots, l_i$ and at least one of $r_{i1}, r_{i2}, \ldots, r_{il_i}$ must be equal to m_i, since m_i is the least positive integer such that $(A - \lambda_i I)^{m_i} v = 0$ for all $v \in \mathcal{N}((A - \lambda_i I)^{n_i})$. Now consider the Jordan chain associated with each of these generalized eigenvectors and let

$$B_i = \{v_{i1}^1, v_{i2}^1, \ldots v_{ir_{i1}}^1, v_{i1}^2, v_{i2}^2, \ldots v_{ir_{i2}}^2, \ldots, v_{i1}^{l_i}, v_{i2}^{l_i}, \ldots v_{ir_{il_i}}^{l_i}\}$$

denote the union of these Jordan chains. By Theorem 4.17 and Corollary 4.9, B_i forms a basis for $\mathcal{N}((A - \lambda_i I)^{n_i})$. Now let $B = \cup_{i=1}^k B_i$. Then B forms a basis for $\mathbb{K}^n$. Now consider the matrix P with elements from B as its columns. Since B is a basis for $\mathbb{K}^n$, P is invertible. i.e., Let

$$P = \left[v_{11}^1 \cdots v_{1r_{11}}^1 \; v_{11}^2 \cdots v_{1r_{12}}^2 \cdots v_{k1}^{l_k} \cdots v_{kr_{kl_k}}^{l_k} \right]$$

Then

$$AP = \left[Av_{11}^1 \cdots Av_{1r_{11}}^1 \; Av_{11}^2 \cdots Av_{1r_{12}}^2 \cdots Av_{k1}^{l_k} \cdots Av_{kr_{kl_k}}^{l_k}\right]$$
$$= \left[\lambda_1 v_{11}^1 + v_{12}^1 \cdots \lambda_1 v_{1r_{11}}^2 \; \lambda_1 v_{11}^2 + v_{12}^2 \cdots \lambda_1 v_{1r_{12}}^2 \cdots \lambda_k v_{k1}^{l_k} + v_{k2}^{l_k} \cdots \lambda_k v_{kn_k}^{l_k}\right]$$
$$= P \begin{bmatrix} J_{\lambda_1}^{r_{11}} & & & & & \\ & J_{\lambda_1}^{r_{12}} & & & & \\ & & \ddots & & & \\ & & & J_{\lambda_1}^{r_{1l_1}} & & \\ & & & & \ddots & \\ & & & & & J_{\lambda_k}^{r_{kl_k}} \end{bmatrix}$$

i.e., $P^{-1}AP = \begin{bmatrix} J_{\lambda_1}^{r_{11}} & & & & & \\ & J_{\lambda_1}^{r_{12}} & & & & \\ & & \ddots & & & \\ & & & J_{\lambda_1}^{r_{1l_1}} & & \\ & & & & \ddots & \\ & & & & & J_{\lambda_k}^{r_{kl_k}} \end{bmatrix} = J$. Then J satisfies all the requirements mentioned in the theorem.

Example 4.28 Consider the matrix

$$A = \begin{bmatrix} 2 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{bmatrix}$$

From Example 4.17, the eigenvalues of A are 2 and 3 with $AM(2) = 2$, $GM(2) = 1$ and $AM(3) = GM(3) = 1$. From Example 4.15, $v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $v_3 = \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$ are the linearly independent eigenvectors of A with respect to 2 and 3, respectively. From Example 4.25, the Jordan chain associated with the eigenvalue 2 is $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Take $P = \begin{bmatrix} 1 & 0 & 4 \\ 0 & -1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$. Then $P^{-1}BP = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

Example 4.29 Suppose that the characteristic and minimal polynomial of a matrix A is given by

$$p(A) = (\lambda - 3)^4(\lambda - 1)^3 \quad \text{and} \quad m(A) = (\lambda - 3)^2(\lambda - 1)^3$$

As the eigenvalues 3 and 1 have multiplicity 2 and 3 in the minimal polynomial, the Jordan canonical form must have a Jordan block of order 2 corresponding to the eigenvalue 3 and a Jordan block of order 3 corresponding to the eigenvalue 1. As the multiplicity of 1 in both minimal and characteristic polynomial of A is the same, there exists only one Jordan block corresponding to 1. The number of Jordan blocks corresponding to the eigenvalue 3 depends on the number of linearly independent eigenvectors corresponding to 3. Then there are two possible Jordan forms for A.

(1) If A has two linearly independent eigenvectors belonging to the eigenvalue 3. Then the Jordan form of A is given by

$$\begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(2) If A has three independent eigenvectors belonging to the eigenvalue 3. Then the Jordan form of A is given by

$$\begin{bmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

4.6 Exercises

1. Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$. If μ is an eigenvalue of A and ν is an eigenvalue of B, give an example such that

   (a) $\mu + \nu$ need not be an eigenvalue of $A + B$.
   (b) $\mu\nu$ need not be an eigenvalue of AB.

2. Find the characteristic polynomial of the following matrices

$$\text{a)}\begin{bmatrix} 1 & 5 \\ 7 & 3 \end{bmatrix} \quad \text{b)}\begin{bmatrix} 4 & -1 \\ 11 & 7 \end{bmatrix} \quad \text{c)}\begin{bmatrix} 2 & 3 & 1 \\ 5 & 0 & 4 \\ 6 & -1 & 7 \end{bmatrix}$$

   Also, verify *Cayley–Hamilton theorem.*

3. Find a matrix $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ with 2 and 3 as eigenvalues and $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}$ as corresponding eigenvectors.
4. Let V be a finite-dimensional vector space over $\mathbb{K}$ with a basis B and $T : V \to V$ be a linear transformation. Show that $\lambda \in \mathbb{K}$ is an eigenvalue of T if and only if λ is an eigenvalue of $[T]_B$.
5. (**Gerschgorin's Theorem**) Let λ be an eigenvalue of a matrix $A = [a_{ij}] \in \mathbb{M}_{n\times n}(\mathbb{K})$. Then for some integer j, $(1 \leq j \leq n)$, show that

$$|a_{jj} - \lambda| \leq |a_{j1}| + |a_{j2}| + \cdots + |a_{j(j-1)}| + |a_{j(j+1)}| + \cdots + |a_{jn}|$$

6. Let $A, B \in \mathbb{M}_{9\times 9}(\mathbb{K})$ be such that $Rank(A) = 3$ and $Rank(B) = 5$. Show that there exists $v \in \mathbb{R}^9$ such that $Av = Bv = 0$.
7. Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that the sum of elements in each row are the same, say λ. Show that λ is an eigenvalue of A. Will this be true if sum of elements in each column are the same?
8. Let $A \in \mathbb{M}_{2\times 2}(\mathbb{K})$ be such that $tr(A) = 8$ and $det(A) = 15$. Find the eigenvalues of A.
9. Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Let λ be an eigenvalue of A with eigenvectors $v_1, v_2, \ldots, v_k$. Show that $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_k v_k$ is an eigenvector of A corresponding to λ for any scalars $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$.
10. Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$. Show that AB and BA have the same characteristic polynomial. Does they have the same minimal polynomial?
11. Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$. Check which of the following statements are true.

    (a) A and A^T have the same eigenvalues.
    (b) If $Rank(A) = k$, then A has at most $k + 1$ distinct eigenvalues.
    (c) If A is singular, then $A + I$ is nonsingular, where I is the identity matrix.
    (d) If every $v \in \mathbb{R}^n$ is an eigenvector of A, then $A = \lambda I$ for some $\lambda \in \mathbb{R}$.
    (e) Let $f(x) \in \mathbb{R}[x]$. Then $\tilde{\lambda}$ is an eigenvalue of $f(A)$ if and only if $\tilde{\lambda} = f(\lambda)$ for some eigenvalue λ of A.
    (f) If A satisfies the equation $A^3 = A$, then the characteristic equation of A is $x^3 - x$.
    (g) If A and B have the same eigenvalues, then they have the same characteristic and minimal polynomials.
    (h) $\{0\}$ and $\mathbb{R}^n$ itself are the only subspaces of $\mathbb{R}^n$ which are invariant under every A.
    (i) $\mathcal{N}(A - \lambda I)$ is an invariant subspace of A for any $\lambda \in \mathbb{R}$.
    (j) If A is diagonalizable, then there exists one dimensional invariant subspaces $V_1, V_2, \ldots, V_n$ of A such that $\mathbb{R}^n = V_1 \oplus V_2 \oplus \cdots \oplus V_n$.

12. Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that $A^k = 0$ for some k (A is called *nilpotent matrix*). Then show that 0 is the only eigenvalue of A. Can A be diagonalizable?
13. Let A be a square matrix with real entries such that $A^{2022} = 0$. Then what are the possible values of $Tr(A^2)$?

14. Check whether the following matrices are diagonalizable over $\mathbb{C}$ or not.

$$\text{a)} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{b)} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{c)} \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{d)} \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 6 \\ 1 & 0 & 5 \end{bmatrix} \quad \text{e)} \begin{bmatrix} 3 & 2 & 0 \\ 0 & -1 & 6 \\ 0 & 0 & 5 \end{bmatrix}$$

15. Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that $A^2 = I$ (A is called *involutory matrix*). Then show that
    (a) 1 and -1 are the only possible eigenvalues of A.
    (b) A is diagonalizable.
16. Let $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that $A^2 = A$(A is called *idempotent matrix*). Then show that
    (a) 0 and 1 are the only possible eigenvalues of A.
    (b) $\mathbb{K}^n = Im(A) \oplus Ker(A)$.
    (c) A is diagonalizable.
17. Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that $B = P^{-1}AP$. Then show that if v is an eigenvector of A corresponding to the eigenvalue λ, $P^{-1}v$ is an eigenvector of B corresponding to the eigenvalue λ.
18. Find the characteristic and minimal polynomial of the matrix $\begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}_{n\times n}$.
19. Show that if A is a block diagonal matrix, then the minimal polynomial of A is the least common multiple of minimal polynomials of the diagonal blocks.
20. Let $A \in \mathbb{M}_{7\times 7}(\mathbb{R})$ have three distinct eigenvalues λ_1, λ_2 and λ_3. The eigenspace corresponding to λ_1 is two dimensional and the eigenspace corresponding to one of the other two eigenvalue is three dimensional. Is A diagonalizable?
21. Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{K})$ be such that $AB = BA$. Show that if A is diagonalizable, then B is also diagonalizable.
22. Show that if $A \in \mathbb{M}_{3\times 3}(\mathbb{R})$ is not triangularizable over $\mathbb{R}$, then it is diagonalizable over $\mathbb{C}$.
23. Let $V_1, V_2, \dots, V_k$ be invariant subspaces under $A \in \mathbb{M}_{n\times n}(\mathbb{K})$. Show that $V_1 + V_2 + \dots + V_k$ is also an invariant subspace under A.
24. How many invariant subspaces does the zero matrix and identity matrix of order n have?
25. Consider a matrix $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ given by

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

    (a) Find the eigenvalues and eigenvectors of A.

(b) Find matrix P such that $PAP^{-1} = D$, where D is a diagonal matrix, if it exists.
(c) Find all invariant subspaces of A.

26. Find the possible Jordan Canonical forms of $A \in \mathbb{M}_{8\times 8}(\mathbb{R})$ if the characteristic and minimal polynomials are given by

$$p(A) = (\lambda - 3)^2(\lambda - 2)^3(\lambda + 1)^3$$

and

$$m(A) = (\lambda - 3)^2(\lambda - 2)(\lambda + 1)^2$$

27. Let A be a real $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ repeated according to their multiplicities. Then show that there exist a basis of generalized eigenvectors $v_1, v_2, \ldots, v_n$ with $P = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}$ invertible and $A = S + N$, where $P^{-1}SP = diag\{\lambda_1, \ldots, \lambda_n\}$ ($diag\{\lambda_1, \ldots, \lambda_n\}$ denotes the diagonal matrix with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$). Further, the matrix $N = A - S$ is nilpotent of order $k \leq n$ and S and N commute.

Solved Questions related to this chapter are provided in Chap. 10.

Chapter 5
Normed Spaces and Inner Product Spaces

This chapter delves into the fundamental mathematical structures of normed linear spaces and inner product spaces, providing a solid comprehension of these essential mathematical structures. Normed spaces are defined as vector spaces that have been reinforced with a norm function that quantifies the magnitude or *length* of a vector from the origin. Several examples, such as Euclidean space with the well-known Euclidean norm, demonstrate the use of normed spaces. Building on this, inner product spaces are investigated, with the goal of broadening the concept of normed spaces by integrating an inner product that generalizes the dot product. Euclidean space is one example, where the inner product can characterize orthogonality and angle measurements. The chapter expands on the importance of orthogonality in inner product spaces, providing insights into geometric relationships and applications in a variety of domains. Gram–Schmidt orthogonalization technique is introduced, which provides a mechanism for constructing orthogonal bases from any bases of an inner product space. The concept of orthogonal complement and projection onto subspaces broadens our understanding by demonstrating the geometrical interpretation and practical application of these fundamental mathematical constructs. Proficiency in these topics is essential for advanced mathematical study and a variety of real-world applications in a variety of areas.

5.1 Normed Linear Spaces

In this section, we will introduce a metric structure called a *norm* on a vector space and then study in detail the resultant space. A vector space with a norm defined on it is called normed linear space. A norm, which intuitively measures the magnitude or size of a vector in a normed space, enables the definition of distance and convergence. Normed spaces provide an adaptive environment for various mathematical and scientific applications, providing a deeper understanding of vector spaces and

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_5

accommodating numerous norm functions to meet various needs. Let us start with the following definition.

Definition 5.1 (*Normed linear space*) Let V be a vector space over the field $\mathbb{K}$, where $\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$. *Norm* is a real-valued function on V ($\|.\| : V \to \mathbb{R}$) satisfying the following three conditions for all $u, v \in V$ and $\lambda \in \mathbb{K}$:

(N1) $\|v\| \geq 0$, and $\|v\| = 0$ if and only if $v = 0$
(N2) $\|\lambda v\| = |\lambda|\, \|v\|$
(N3) $\|u + v\| \leq \|u\| + \|v\|$.(Triangle Inequality)

Then V together with a norm defined on it, denoted by $(V, \|.\|)$, is called a *Normed linear space*.

Example 5.1 Consider the vector space $\mathbb{R}$ over $\mathbb{R}$. Define $\|v\|_0 = |v|$ for $v \in \mathbb{R}$. Then by the properties of modulus function, $\|.\|_0$ is a norm on $\mathbb{R}$.

Example 5.2 Consider the vector space $\mathbb{R}^n$ over $\mathbb{R}$. For $v = (v_1, v_2, \dots, v_n)$ in $\mathbb{R}^n$, define $\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}}$. This norm is called the *2-norm*.

(N1) Clearly $\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} \geq 0$ and $\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} = 0 \Leftrightarrow |v_i|^2 = 0$ for all $i = 1, 2, \dots, n \Leftrightarrow v = 0$.
(N2) For $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^n$,

$$\begin{aligned}\|\lambda v\|_2 = \left(\sum_{i=1}^{n} |\lambda v_i|^2\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} |\lambda|^2 |v_i|^2\right)^{\frac{1}{2}} &= \left(|\lambda|^2 \sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} \\ &= |\lambda| \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} = |\lambda|\, \|v\|_2\end{aligned}$$

(N3) For $u, v \in \mathbb{R}^n$,

$$\begin{aligned}\sum_{i=1}^{n} (|u_i| + |v_i|)^2 &= \sum_{i=1}^{n} (|u_i| + |v_i|)\,(|u_i| + |v_i|) \\ &= \sum_{i=1}^{n} |u_i|\,(|u_i| + |v_i|) + \sum_{i=1}^{n} |v_i|\,(|u_i| + |v_i|) \\ &\leq \left(\sum_{i=1}^{n} |u_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} (|u_i| + |v_i|)^2\right)^{\frac{1}{2}} \\ &+ \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} (|u_i| + |v_i|)^2\right)^{\frac{1}{2}} \\ &= \left(\sum_{i=1}^{n} (|u_i| + |v_i|)^2\right)^{\frac{1}{2}} \left[\left(\sum_{i=1}^{n} |u_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}}\right]\end{aligned}$$

which implies

$$\left(\sum_{i=1}^{n}(|u_i|+|v_i|)^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n}|u_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n}|v_i|^2\right)^{\frac{1}{2}}$$

Since $|u_i + v_i| \leq |u_i| + |v_i|$, we have

$$\left(\sum_{i=1}^{n}|u_i+v_i|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n}|u_i|^2\right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n}|v_i|^2\right)^{\frac{1}{2}}$$

Therefore $\mathbb{R}^n$ is a normed linear space with respect to $2-norm$. In general, $\mathbb{R}^n$ is a normed linear space with respect to the $p-norm$ defined by $\|v\|_p = \left(\sum_{i=1}^{n}|v_i|^p\right)^{\frac{1}{p}}$, $p \geq 1$.(Verify)

Example 5.3 Consider the vector space $\mathbb{R}^n$ over $\mathbb{R}$. For $v = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$, define $\|v\|_\infty = \max\{|v_1|, |v_2|, \ldots, |v_n|\} = \max\limits_{i\in\{1,\ldots,n\}}\{|v_i|\}$. This norm is called the *infinity norm*.

Example 5.4 Let $V = C[a, b]$, the space of continuous real-valued functions on $[a, b]$. For $f \in V$, define $\|f\| = \max\limits_{x\in[a,b]}|f(x)|$. This norm is called *supremum* norm.

(N1) Clearly $\|f\| = \max\limits_{x\in[a,b]}|f(x)| \geq 0$. Also, $\|f\| = \max\limits_{x\in[a,b]}|f(x)| = 0 \Leftrightarrow |f(x)| = 0$ for all $x \in [a, b] \Leftrightarrow f(x) = 0$ for all $x \in [a, b]$.

(N2) For $\lambda \in \mathbb{R}$ and $f \in C[a, b]$,

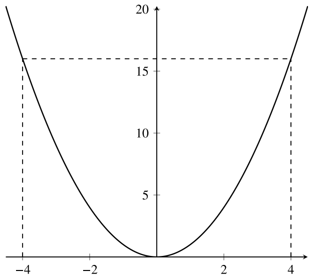
$$\|\lambda f\| = \max_{x\in[a,b]}|(\lambda f)(x)| = \max_{x\in[a,b]}|\lambda\,(f(x))| = \max_{x\in[a,b]}|\lambda||f(x)| = |\lambda| \max_{x\in[a,b]}|f(x)| = |\lambda|\,\|f\|$$

(N3) Since $|a + b| \leq |a| + |b|$, for $f, g \in C[a, b]$ we have

$$\|f+g\| = \max_{x\in[a,b]}|(f+g)(x)| = \max_{x\in[a,b]}|f(x)+g(x)| \leq \max_{x\in[a,b]}|f(x)| + \max_{x\in[a,b]}|g(x)| = \|f\| + \|g\|$$

Then $C[a, b]$ is a normed linear space with the *supremum* norm (Fig. 5.1).

We have shown that $\|f\| = \max\limits_{x\in[a,b]}|f(x)|$ defines a norm in $C[a, b]$. Now let us define $\|f\| = \min\limits_{x\in[a,b]}|f(x)|$. Does that function defines a norm on $C[a, b]$? No, it doesn't! Clearly, we can observe that $\|f\| = 0$ does not imply that $f = 0$. For example, consider the function $f(x) = x^2$ in $C[-4, 4]$. Then $\|f\| = \min\limits_{x\in[-4,4]}|f(x)| = 0$, but $f \neq 0$. As $(N1)$ is violated, $\|f\| = \min\limits_{x\in[-4,4]}|f(x)|$ does not defines a norm on $C[-4, 4]$.



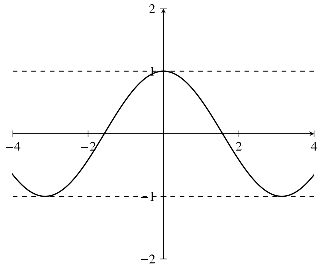


Fig. 5.1 Consider the functions $f(x) = x^2$ and $g(x) = \cos x$ in $C[-4, 4]$. Then $\|f\| = \max_{x\in[-4,4]} |x^2| = 16$ and $\|g\| = \max_{x\in[-4,4]} |\cos x| = 1$

Definition 5.2 (*Subspace*) Let $(V, \|.\|)$ be a normed linear space. A subspace of V is a vector subspace W of V with the same norm as that of V. The norm on W is said to be induced by the norm on V.

Example 5.5 Consider $C[a, b]$ with the supremum norm, then $\mathbb{P}[a, b]$ is a subspace of $C[a, b]$ with supremum norm as the induced norm.

We will now show that every normed linear space is a metric space. Consider the following theorem.

Theorem 5.1 *Let $(V, \|.\|)$ be a normed linear space. Then $d(v_1, v_2) = \|v_1 - v_2\|$ is a metric on V.*

Proof Let $v_1, v_2, v_3 \in V$. Then

(M1) By $(N1)$, we have

$$d(v_1, v_2) = \|v_1 - v_2\| \geq 0$$

and

$$d(v_1, v_2) = \|v_1 - v_2\| = 0 \Leftrightarrow v_1 - v_2 = 0 \Leftrightarrow v_1 = v_2$$

(M2) By $(N2)$, we have

$$d(v_1, v_2) = \|v_1 - v_2\| = \|v_2 - v_1\| = d(v_2, v_1)$$

(M3) Now we have to prove the triangle inequality.

$$\begin{aligned} d(v_1, v_2) &= \|v_1 - v_2\| \\ &= \|v_1 - v_3 + v_3 - v_2\| \\ &\leq \|v_1 - v_3\| + \|v_3 - v_2\| \quad (By (N3)) \\ &= d(v_1, v_3) + d(v_3, v_2) \end{aligned}$$

The metric defined in the above theorem is called *metric induced by the norm*. The above theorem implies that *every normed linear space is a metric space with respect to the induced metric*. Is the converse true? Consider the following example.

Example 5.6 In Example 1.25, we have seen that for any non-empty set X, the function d defined by

$$d(x, y) = \begin{cases} 1\,, & x \neq y \\ 0\,, & x = y \end{cases}$$

defines a metric on X. Let V be a vector space over the field $\mathbb{K}$. Clearly (V, d) is a metric space. If V is a normed linear space, by Theorem 5.1, we have

$$\|v\| = d(v, 0) = \begin{cases} 1\,, & v \neq 0 \\ 0\,, & v = 0 \end{cases}$$

As you can observe that, for any $\lambda \neq 0 \in \mathbb{K}$,

$$\|\lambda v\| = \begin{cases} 1\,, & v \neq 0 \\ 0\,, & v = 0 \end{cases} \neq |\lambda|\,\|v\| = \begin{cases} |\lambda|\,, & v \neq 0 \\ 0\,, & v = 0 \end{cases}$$

the discrete metric cannot be obtained from any norm. Therefore, *every metric space need not be a normed linear space.*

Now that you have understood the link between normed spaces and metric spaces, let us discuss a bit more in detail about defining a distance notion on vector spaces. In Example 5.2, we have defined a number of norms on $\mathbb{R}^n$. What is the significance of defining several norms on a vector space? Consider a simple example as depicted in Fig. 5.2.

In real life, we can justify the significance of defining various notions of distances on vector spaces with many practical applications. Therefore, while dealing with a normed linear space we choose the norm which meets our need accordingly (Fig. 5.3).

Now we understand that different norms on a vector space can give rise to different geometrical and analytical structures. Now we will discuss whether these structures are related or not. As a prerequisite for the discussion, let us define the "fundamental sets" on a normed linear space

Definition 5.3 (*Open ball*) Let $(V, \|.\|)$ be a normed linear space. For any point $v_0 \in V$ and $\epsilon \in \mathbb{R}^+$,

$$B_\epsilon(v_0) = \{v \in V \mid \|v - v_0\| < \epsilon\}$$

is called an open ball centered at v_0 with radius ϵ. The set $\{v \in V \mid \|v\| = 1\}$ is called the unit sphere in V

We can see that this definition follows from the Definition 1.23 of an open ball in a metric space.

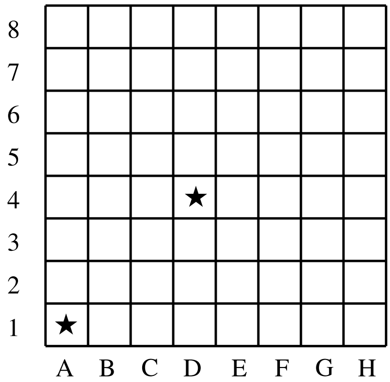
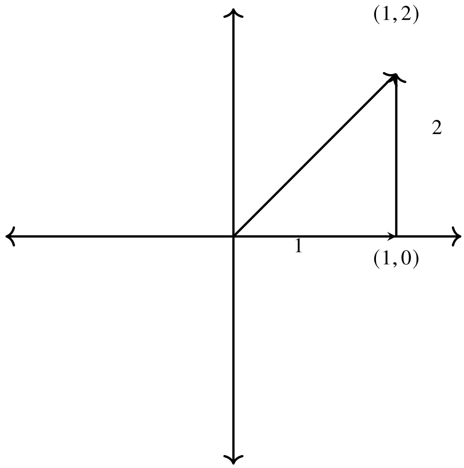


Fig. 5.2 Suppose that you have to move a chess piece from $A1$ to $D4$ in least number of moves. If the piece is a bishop we can move the piece directly from A_1 to D_4. If the piece is a rook, first we will have to move the piece either to $A4$ or $D1$ and then to $D4$. Now, if the piece is king, the least number of moves would be $3(A_1 \to B_2 \to C_3 \to D_4)$. Observe that the path chosen by different pieces to move from A_1 to D_4 in least number of moves are different. Now try to calculate the distance traveled by the piece in each of these cases. Are they the same? We need different notions of distances, right? Interestingly, the metric induced from the infinity norm, $d(u, v) = max_i\{|u_i - v_i|\}$ is known as the *chess distance* or *Chebyshev distance* (In honor of the Russian mathematician, *Pafnuty Chebyshev (1821–1894)*) as the Chebyshev distance between two spaces on a chess board gives the minimum number of moves required by the king to move between them

Fig. 5.3 Consider $\mathbb{R}^2$ with different norms defined on it. If we are using the 2-norm, the distance from the origin to the point $(1, 2)$ is $\sqrt{|1|^2 + |2|^2} = \sqrt{5}$ as it is length of the hypotenuse of a triangle with base 1 and height 2. If we are using 1-norm the distance will be 3 as it is the sum of the absolute values of the coordinates and if we are using infinity norm, the distance will be 2 as it is the maximum of absolute values of the coordinates



Example 5.7 Consider $(\mathbb{R}, \|.\|_0)$. In Example 1.26, we have seen that the open balls in $(\mathbb{R}, \|.\|_0)$ are open intervals in the real line. Now, consider the set $S = \{(v_1, 0) \mid v_1 \in \mathbb{R}, 1 < v_1 < 4\}$ in $(\mathbb{R}^2, \|.\|_2)$. Is S an open ball in $(\mathbb{R}^2, \|.\|_2)$? Is there any way to generalize the open balls in $(\mathbb{R}^2, \|.\|_2)$? Yes, we can!! Take an arbitrary point $w = (w_1, w_2) \in \mathbb{R}^2$, and $\epsilon \in \mathbb{R}^+$. Then

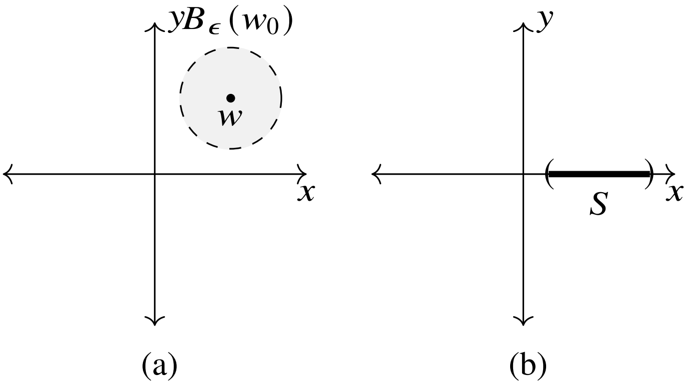
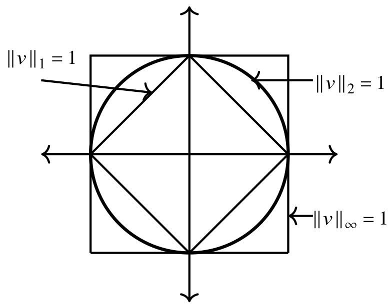


Fig. 5.4 The open balls in $(\mathbb{R}^2, \|.\|_2)$ are open circles as given in (a). Clearly, S is not an open ball in $(\mathbb{R}^2, \|.\|_2)$

Fig. 5.5 Unit spheres in $\mathbb{R}^2$ with respect to 1− norm, 2-norm and infinity norm. Observe that the interior portion of the unit spheres represents the open unit ball, $B_1(0) = \{v \in V \mid \|v\| < 1\}$ in each of the norms



$$
\begin{aligned}
B_\epsilon(w) &= \{v = (v_1, v_2) \in \mathbb{R}^2 \mid \|v - w\| < \epsilon\} \\
&= \{v = (v_1, v_2) \in \mathbb{R}^2 \mid (v_1 - w_1)^2 + (v_2 - w_2)^2 < \epsilon^2\}
\end{aligned}
$$

That is, open balls in $(\mathbb{R}^2, \|.\|_2)$ are "open circles" (Fig. 5.4).

Example 5.8 Let us compute the open unit balls centered at the origin in $\mathbb{R}^2$ with respect to *1-norm*, *2-norm* and *infinity norm*. Let B_ϵ^p denote the open ball in $\left(\mathbb{R}^2, \|.\|_p\right)$. Then

$$
\begin{aligned}
B_1^1 &= \left\{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1| + |v_2| < 1\right\} \\
B_1^2 &= \left\{(v_1, v_2) \in \mathbb{R}^2 \mid |v_1|^2 + |v_2|^2 < 1\right\}
\end{aligned}
$$

and (Fig. 5.5)

$$
B_1^\infty = \left\{(v_1, v_2) \in \mathbb{R}^2 \mid max\{|v_1|, |v_2|\} < 1\right\}
$$

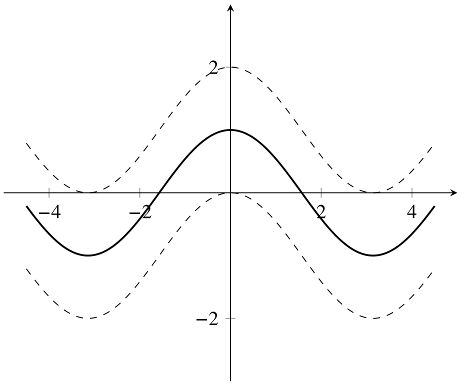


Fig. 5.6 Consider a function f in $C[-4, 4]$ with supremum norm. Continuous functions that lie between the dotted lines constitute $B_1(f) = \{g \in C[-4, 4] \mid \|f - g\| < 1\}$

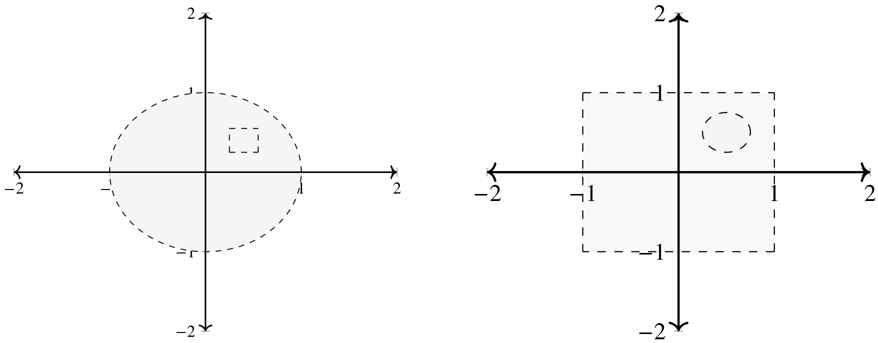


Fig. 5.7 Clearly, we can observe that every point in an open ball generated by the infinity norm is inside an open ball generated by 2-norm and vice versa

Observe that the open balls in $\mathbb{R}^2$ corresponding to different norms may not have the same shape even if the center and radius are the same. Now, let us give you an example of open ball in $C[-4, 4]$ with *supremum norm* (Fig. 5.6).

Earlier, we have posed a question, does there exist any link between the topology generated by the different norms defined on a vector space? It is interesting to note that the topology generated by any norms on a finite-dimensional space is the same. That is, the open sets defined by these norms are topologically same. The following figure illustrates this idea by taking the open balls in $\mathbb{R}^2$ generated by the infinity norm and 2-norm as an example (Fig. 5.7).

Now we will prove algebraically that, in a finite-dimensional space the open sets generated by any norms are topologically the same. For that, we will have the following definition.

Definition 5.4 (*Equivalence of norms*) A norm $\|.\|$ on a vector space V is equivalent to $\|.\|_0$ on V if there exists positive scalars λ and μ such that for all $v \in V$, we have

$$\lambda \|v\|_0 \leq \|v\| \leq \mu \|v\|_0$$

Example 5.9 Let us consider the 1-norm, 2-norm and *infinity norm* in $\mathbb{R}^n$. For any element $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, we have

$$\|v\|_\infty = \max_{i \in \{1,2,\ldots,n\}} \{|v_i|\} \leq |v_1| + |v_2| + \cdots + |v_n| = \|v\|_1$$

Also by *Holder's inequality* (Exercise 5, Chap. 1), we have

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| = \sum_{i=1}^{n} |v_i|.1 \leq \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} 1^2\right)^{\frac{1}{2}} = \sqrt{n}\,\|v\|_2$$

and finally,

$$\|v\|_2 = \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} \left(\max_{i \in \{1,2,\ldots,n\}} \{|v_i|\} |v_i|\right)^2\right)^{\frac{1}{2}} = \left(n\,\|v\|_\infty^2\right)^{\frac{1}{2}} = \sqrt{n}\,\|v\|_\infty$$

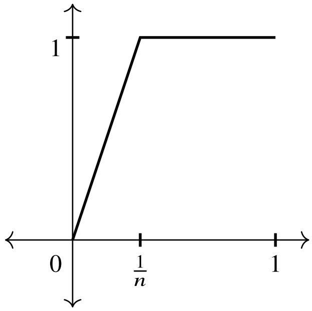
Thus 1$-$ norm, 2-norm and *infinity norm* in $\mathbb{R}^n$ are equivalent.

In fact, we can prove that every norm in a finite-dimensional space is equivalent. But this is not the case if the space is infinite- dimensional. Consider the following example.

Example 5.10 Consider the linear space $C[0, 1]$ over the field $\mathbb{R}$. In Example 5.4, we have seen that $\|f\| = \max_{x \in [0,1]} |f(x)|$ defines a norm on $C[0, 1]$, called the *supremum norm*. Also, we can show that $\|f\|_1 = \int_0^1 |f(x)| dx$ defines a norm on $C[0, 1]$(Verify!). We will show that there doesn't exist any scalar λ such that $\|f\| \leq \|f\|_1$ for all $f \in C[0, 1]$. For example, consider a function defined as in Fig. 5.8. Then we can observe that $\|f_n\| = 1$ and $\|f_n\|_1 = \frac{1}{2n}$(How?). Clearly, we can say that there doesn't exists any scalar λ such that $1 \leq \frac{\lambda}{2n}$ for all n.

We have discussed the equivalence of norms in terms of defining topologically identical open sets. This can also be discussed in terms of sequences. In Chap. 1, we have seen that the addition of metric structure to an arbitrary set enables us to discuss the convergence or divergence of sequences, limit and continuity of functions, etc., in detail. The same happens with normed linear spaces also. The difference is that we are adding the metric structure not just to any set, but a vector space. All these notions can be discussed in terms of induced metric as well as norm. We will start by defining a Cauchy sequence in a normed linear space.

Fig. 5.8 Define $f_n(x) = \begin{cases} nx, & x \in \left[0, \frac{1}{n}\right] \\ 1, & x \in \left[\frac{1}{n}, 1\right] \end{cases}$ as shown in the figure. Clearly $f_n(x)$ belongs to $C[0, 1]$ for all n



Definition 5.5 (*Cauchy Sequence*) A sequence $\{v_n\}$ in a normed linear space $(V, \|.\|)$ is said to be Cauchy if for every $\epsilon > 0$ there exists an $N_\epsilon \in \mathbb{N}$ such that $\|v_n - v_m\| < \epsilon$ for all $m, n > N$.

Definition 5.6 (*Convergence*) Let $\{v_n\}$ be a sequence in $(V, \|.\|)$, then $v_n \to v$ in V if and only if $\|v_n - v\| \to 0$ as $n \to \infty$.

In Chap. 1, we have seen that in a metric space every *Cauchy sequence* need not necessarily be convergent. Now the important question of whether a *Cauchy sequence* is convergent or not in a normed linear space pops up. The following example gives us an answer.

Example 5.11 Consider the normed linear space $\mathbb{P}[0, 1]$ over $\mathbb{R}$ with the supremum norm. Consider the sequence, $\{p_n(x)\}$, where

$$p_n(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$$

Is the sequence convergent? If so, is the limit function a polynomial? Clearly, not! We know that $p_n(x) \to e^x$, $x \in [0, 1]$(Verify!). Is it the only sequence in $\mathbb{P}[0, 1]$ over $\mathbb{R}$ that converge to a function which is not a polynomial? Let us consider another sequence $\{q_n(x)\}$, where

$$q_n(x) = 1 + \frac{x}{2} + \frac{x^2}{4} + \cdots + \frac{x^n}{2^n}$$

First we will prove that $\{q_n\}$ is a Cauchy sequence. For $n > m$,

$$\begin{aligned} \|q_n(x) - q_m(x)\| &= \max_{x \in [0,1]} \left| \sum_{i=0}^{n} \frac{x^i}{2^i} - \sum_{i=0}^{m} \frac{x^m}{2^m} \right| \\ &= \max_{x \in [0,1]} \left| \sum_{i=m+1}^{n} \frac{x^i}{2^i} \right| \\ &\leq \frac{1}{2^m} \end{aligned}$$

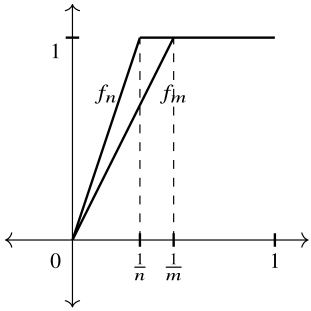


Fig. 5.9 As $\|f_n - f_m\|$ is the area of the triangle depicted in the figure, it is easy to observe that $\{f_n\}$ is Cauchy

which shows that $\{q_n(x)\}$ is a *Cauchy sequence.* Now for any $x \in [0, 1]$, we have $q_n(x) \to q(x)$ as $n \to \infty$ where $q(x) = \dfrac{1}{1 - \frac{x}{2}}$ (How?) and clearly $q(x) \notin \mathbb{P}[0, 1]$ as it is not a polynomial function. Hence $\{q_n(x)\}$ is not convergent in $\mathbb{P}[0, 1]$. What about $\mathbb{P}_n[0, 1]$? Is it complete?

Here is another example of an incomplete normed linear space.

Example 5.12 Consider $C[0, 1]$ with $\|f\| = \int_0^1 |f(x)|dx$ for $f \in C[0, 1]$. Consider the sequence of functions $f_n \in C[0, 1]$ where

$$f_n(x) = \begin{cases} nx, & x \in \left[0, \frac{1}{n}\right] \\ 1, & x \in \left[\frac{1}{n}, 1\right] \end{cases}$$

We will show that $\{f_n\}$ is Cauchy but not convergent (Fig. 5.9).

For $n > m$,

$$|f_n(x) - f_m(x)| = \begin{cases} nx - mx, & x \in \left[0, \frac{1}{n}\right] \\ 1 - mx, & x \in \left[\frac{1}{n}, \frac{1}{m}\right] \\ 0, & x \in \left[\frac{1}{m}, 1\right] \end{cases}$$

Then

$$\begin{aligned} \int_0^1 |f_n(x) - f_m(x)|dx &= \int_0^{\frac{1}{n}} (n-m)\, x.dx + \int_{\frac{1}{n}}^{\frac{1}{m}} (1 - mx)\, dx \\ &= (n-m)\frac{1}{2n^2} + \frac{1}{m} - \frac{1}{n} - \frac{1}{2m} + \frac{m}{2n^2} \\ &= \frac{1}{2}\left[\frac{1}{m} - \frac{1}{n}\right] \end{aligned}$$

Now for any $\epsilon > 0$, take $N > \frac{2}{\epsilon}$. Then for $m, n > N$,

$$\int_0^1 |f_n(x) - f_m(x)|dx = \frac{1}{2}\left[\frac{1}{m} - \frac{1}{n}\right] < \frac{1}{m} + \frac{1}{n} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Therefore the sequence is *Cauchy*. Now consider

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \in (0, 1] \end{cases}$$

Then $\|f_n - f\| = \frac{1}{n} \to 0$ as $n \to \infty$. That is, f_n converges to f but $f \notin C[0, 1]$.

Normed linear spaces where every *Cauchy sequence* is convergent are of greater importance in Mathematics. Such spaces are named after the famous Polish mathematician *Stefan Banach (1892–1945)* who started a systematic study in this area.

Definition 5.7 (*Banach Space*) A complete normed linear space is called a Banach space.

Example 5.13 Consider the normed linear space $\mathbb{R}^n$ over $\mathbb{R}$ with 2-norm. We will show that this space is a *Banach space*. Let $\{v_k\}$ be a Cauchy sequence in $\mathbb{R}^n$. As $v_k \in \mathbb{R}^n$, we can take $v_k = \left(v_1^k, v_2^k, \ldots, v_n^k\right)$ for each k. Since $\{v_k\}$ is a Cauchy sequence, for every $\epsilon > 0$ there exists an N such that

$$\|v_k - v_m\|^2 = \sum_{i=1}^{n} \left(v_i^k - v_i^m\right)^2 < \epsilon^2$$

for all $k, m \geq N$. This implies that $\left(v_i^k - v_i^m\right)^2 < \epsilon^2$ for each $i = 1, 2, \ldots, n$ and $k, m \geq N$ and hence $|v_i^k - v_i^m| < \epsilon$ for each $i = 1, 2, \ldots, n$ and $k, m \geq N$. Thus for a fixed i, the sequence $v_i^1, v_i^2, \ldots$ forms a Cauchy sequence of real numbers. Since $\mathbb{R}$ is complete, $v_i^k \to v_i$ as $k \to \infty$ for each i. Take $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$. Then

$$\|v_k - v\|^2 = \sum_{i=1}^{n} \left(v_i^k - v_i\right)^2 \to 0 \; as \; n \to \infty$$

Hence, $\|v_k - v\| \to 0 \; as \; n \to \infty$. Therefore $\mathbb{R}^n$ over $\mathbb{R}$ with 2-norm is a Banach space. What about $\mathbb{C}^n$ over $\mathbb{C}$ with 2-norm?

In fact, we can prove that every finite-dimensional normed linear space is complete. We have seen that this is not true when the normed linear space is infinite-dimensional. Here is an example of infinite-dimensional Banach space.

Example 5.14 Consider $C[a, b]$ with supremum norm. Let $\{f_n\}$ be a Cauchy sequence in $C[a, b]$. Then for every $\epsilon > 0$ there exists an N such that

$$\|f_n - f_m\| = \max_{x \in [a,b]} |f_n(x) - f_m(x)| < \epsilon \tag{5.1}$$

Hence for any fixed $x_0 \in [a, b]$, we have

$$|f_n(x_0) - f_m(x_0)| < \epsilon$$

for all $m, n > N$. This implies that $f_1(x_0), f_2(x_0), f_3(x_0), \ldots$ is a Cauchy sequence of real numbers. Since $\mathbb{R}$ is complete (by Theorem 1.2), this sequence converges, say $f_n(x_0) \to f(x_0)$ as $n \to \infty$. Proceeding like this for each point in $[a, b]$, we can define a function $f(x)$ on $[a, b]$. Now we have to prove that $f_n \to f$ and $f \in C[a, b]$. Then from, Equation 5.1, as $m \to \infty$, we have

$$\max_{x \in [a,b]} |f_m(x) - f(x)| \leq \epsilon$$

for all $m > N$. Hence for every $x \in [a, b]$,

$$|f_m(x) - f(x)| \leq \epsilon$$

for all $m > N$. This implies that $\{f_m(x)\}$ converges to $f(x)$ uniformly on $[a, b]$. Since f_m's are continuous on $[a, b]$ and the convergence is uniform, the limit function is continuous on $[a, b]$(See Exercise 12, Chap. 1). Thus $f \in C[a, b]$ and $f_n \to f$. Therefore $C[a, b]$ is complete.

5.2 Inner Product Spaces

In the previous section, we have added a metric structure to vector spaces which enabled as to find the distance between any two vectors. Now we want to study the geometry of vector spaces which will be useful in many practical applications. In this section, we will give another abstract structure that will help us to study the orthogonality of vectors, projection of one vector over another vector, etc.

$\mathbb{R}^2$ *and Dot product*

$\mathbb{R}^2$ **and Dot product** First we will discuss the properties of the dot product in the space $\mathbb{R}^2$ and then generalize these ideas to abstract vector spaces.

Definition 5.8 (*Dot Product*) Let $v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2$. The dot product of v and w is denoted by $'\ v.w\ '$ and is given by

$$v.w = v_1 w_1 + v_2 w_2$$

Theorem 5.2 *For $u, v, w \in \mathbb{R}^2$ and $\lambda \in \mathbb{K}$,*

(a) $v.v \geq 0$ *and* $v.v = 0$ *if and only if* $v = 0$.
(b) $u.(v + w) = u.v + u.w$ *(distributivity of dot product over addition)*
(c) $(\lambda u).v = \lambda(u.v)$
(d) $u.v = v.u$ *(commutative)*

Proof (a) Let $v = (v_1, v_2) \in \mathbb{R}^2$. Clearly, $v.v = v_1^2 + v_2^2 \geq 0$ and

$$v.v = v_1^2 + v_2^2 = 0 \Leftrightarrow v_1 = v_2 = 0 \Leftrightarrow v = 0$$

(b) For $u = (u_1, u_2), v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2$,

$$\begin{aligned} u.(v + w) &= u_1(v_1 + w_1) + u_2(v_2 + w_2) \\ &= u_1v_1 + u_2v_2 + u_1w_1 + u_2w_2 \\ &= u.v + u.w \end{aligned}$$

(c) For $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$ and $\lambda \in \mathbb{K}$,

$$\begin{aligned} (\lambda u).v &= (\lambda u_1, \lambda u_2).(v_1, v_2) \\ &= \lambda u_1v_1 + \lambda u_2v_2 \\ &= \lambda(u_1v_1 + u_2v_2) = \lambda(u.v) \end{aligned}$$

(d) For $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$,

$$u.v = u_1v_1 + u_2v_2 = v_1u_1 + v_2u_2 = v.u$$

Definition 5.9 (*Length of a vector*) Let $v = (v_1, v_2) \in \mathbb{R}^2$. The length of v is denoted by $|v|$ and is defined by $|v| = \sqrt{v.v} = \sqrt{v_1^2 + v_2^2}$.

Theorem 5.3 *Let $u, v \in \mathbb{R}^2$, then $u.v = |u||v|\cos\theta$ where $0 \leq \theta \leq \pi$ is the angle between u and v.*

Proof Let $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$. If either u or v is the zero vector, say $u = 0$, then

$$u.v = 0v_1 + 0v_2 = 0$$

Then as $|u| = 0$, $|u||v|\cos\theta = 0$. Therefore, the theorem holds. Now suppose that, both $u, v \neq 0$. Consider the triangle with sides u, v and w. Then $w = v - u$ and by the law of cosines of triangle,

$$|w|^2 = |u|^2 + |v|^2 - 2|u||v|\cos\theta \tag{5.2}$$

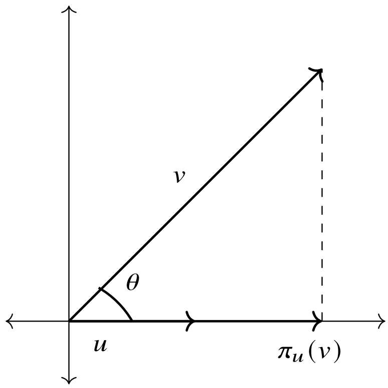


Fig. 5.10 Orthogonal projection of v on u

where $0 \leq \theta \leq \pi$ is the angle between u and v. Also,

$$|w|^2 = w.w = (v-u).(v-u) = (v-u).v - (v-u).u = v.v + u.u - 2u.v \tag{5.3}$$

Then equating (5.2) and (5.3), we get, $u.v = |u||v| \cos\theta$.

Remark 5.1 Let u and v be two vectors in $\mathbb{R}^2$ and let θ be the angle between u and v. Then

1. $\theta = cos^{-1}\left(\dfrac{u.v}{|u||v|}\right)$.
2. If $\theta = \frac{\pi}{2}$, then $u.v = 0$. Then we say that u is orthogonal to v and is denoted by $u \perp v$.

Let $v \in \mathbb{R}^2$ be any vector and $u \in \mathbb{R}^2$ be a vector of unit length. We want to find a vector in $span\ (\{u\})$ such that it is near to v than any other vector in $span\ (\{u\})$ (Fig. 5.10). We know that the shortest distance from a point to a line is the segment perpendicular to the line from the point. We will proceed using this intuition. From the above figure, we get

$$\pi_u(v) = (|v| \cos\theta)\, u$$

From Theorem 5.3, $\cos\theta = \frac{u.v}{|u||v|}$. Substituting this in the above equation, we get $\pi_u(v) = (u.v)u$. The vector $\pi_u(v)$ is called the *orthogonal projection* of v on u as $v - \pi_u(v)$ is perpendicular to $span\ (\{u\})$.

Definition 5.10 (*Projection*) Let $v \in \mathbb{R}^2$ be any vector and $u \in \mathbb{R}^2$ be a vector of unit length. Then the projection of v onto $span\ (\{u\})$ (a line passing through origin) is defined by $\pi_u(v) = (u.v)u$.

Inner Product Spaces

Norm defined on a vector space generalizes the idea of the length of a vector in $\mathbb{R}^2$. Likewise, we will generalize the idea of the dot product in $\mathbb{R}^2$ to arbitrary vector spaces to obtain a more useful structure, where we can discuss the idea of orthogonality, projection, etc.

Definition 5.11 (*Inner product space*) Let V be a vector space over a field $\mathbb{K}$. An inner product on V is a function that assigns, to every ordered pair of vectors $u, v \in V$, a scalar in $\mathbb{K}$, denoted by $\langle u, v\rangle$, such that for all u, v and w in V and all $\lambda \in \mathbb{K}$, the following hold:

(IP1) $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0 \Leftrightarrow v = 0$
(IP2) $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$
(IP3) $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$
(IP4) $\overline{\langle u, v\rangle} = \langle v, u\rangle$, where the bar denotes complex conjugation.

Then V together with an inner product defined on it is called an *Inner product space.* If $\mathbb{K} = \mathbb{R}$, then $(IP4)$ changes to $\langle u, v\rangle = \langle v, u\rangle$.

Remark 5.2 1. If $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ and $w, v_1, v_2, \ldots, v_n \in V$, then

$$\left\langle \sum_{i=1}^{n} \lambda_i v_i, w \right\rangle = \sum_{i=1}^{n} \lambda_i \langle v_i, w\rangle$$

2. By $(IP2)$ and $(IP3)$, for a fixed $v \in V$, $\langle u, v\rangle$ is a linear transformation on V.
3. Dot product is an inner product on the vector space $\mathbb{R}^2$ over $\mathbb{R}$.

Example 5.15 Consider the vector space $\mathbb{K}^n$ over $\mathbb{K}$. For $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$ in $\mathbb{K}^n$, define $\langle u, v\rangle = \sum_{i=1}^{n} u_i\overline{v_i}$, here $\overline{v}$ denote the conjugate of v. This inner product is called *standard inner product* in $\mathbb{K}^n$.

(IP1) We have

$$\langle u, u\rangle = \sum_{i=1}^{n} u_i\overline{u_i} = \sum_{i=1}^{n} |u_i|^2 \geq 0$$

and

$$\langle u, u\rangle = \sum_{i=1}^{n} |u_i|^2 = 0 \Leftrightarrow |u_i|^2 = 0, \forall i = 1, 2, \ldots, n \Leftrightarrow u_i = 0, \forall i = 1, 2, \ldots, n \Leftrightarrow u = 0$$

(IP2) For, $w = (w_1, w_2, \ldots, w_n) \in \mathbb{K}^n$

$$\begin{aligned}\langle u + w, v\rangle &= \sum_{i=1}^{n}(u_i + w_i)\overline{v_i} \\ &= \sum_{i=1}^{n} u_i\overline{v_i} + \sum_{i=1}^{n} w_i\overline{v_i} = \langle u, v\rangle + \langle u, w\rangle\end{aligned}$$

(IP3) $\langle \lambda u, v\rangle = \sum_{i=1}^{n} \lambda u_i\overline{v_i} = \lambda \sum_{i=1}^{n} u_i\overline{v_i} = \lambda\langle u, v\rangle$, where $\lambda \in \mathbb{K}$.
(IP4) $\overline{\langle u, v\rangle} = \overline{\sum_{i=1}^{n} u_i\overline{v_i}} = \sum_{i=1}^{n} \overline{u_i\overline{v_i}} = \sum_{i=1}^{n} v_i\overline{u_i} = \langle v, u\rangle$

Therefore $\mathbb{K}^n$ is an inner product space with respect to the standard inner product. Observe that if $\mathbb{K} = \mathbb{R}$, the inner product, $\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i$ is the usual dot product in $\mathbb{R}^n$.

Example 5.16 Let $V = C[a, b]$, the space of real-valued functions on $[a, b]$. For $f, g \in V$, define $\langle f, g\rangle = \int_a^b f(x)g(x)dx$. Then V is an inner product space with the defined inner product.

(IP1) We have

$$\langle f, f\rangle = \int_a^b f(x)f(x)dx = \int_a^b [f(x)]^2\, dx \geq 0$$

and

$$\langle f, f\rangle = \int_a^b [f(x)]^2\, dx = 0 \Leftrightarrow f(x) = 0, \forall\, x \in [a, b]$$

(IP2) For, $h \in C[a, b]$

$$\begin{aligned}\langle f + h, g\rangle &= \int_a^b [f(x) + h(x)]\, g(x)dx \\ &= \int_a^b f(x)g(x)dx + \int_a^b h(x)g(x)dx = \langle f, g\rangle + \langle h, g\rangle\end{aligned}$$

(IP3) $\langle \lambda f, g\rangle = \int_a^b \lambda f(x)g(x)dx = \lambda \int_a^b f(x)g(x)dx = \lambda\langle f, g\rangle$ where $\lambda \in \mathbb{R}$.
(IP4) $\langle f, g\rangle = \int_a^b f(x)g(x)dx = \int_a^b g(x)f(x)dx = \langle g, f\rangle$.

Thus $C[a, b]$ is an inner product space with respect to the inner product $\langle f, g\rangle = \int_a^b f(x)g(x)dx$. Let us consider a numerical example here for better understanding. Consider $f(x) = x^2 - 1,\ g(x) = x + 1 \in C[0, 1]$. Then

$$\langle f, g\rangle = \int_0^1 (x^3 + x^2 - x - 1)dx = \left[\frac{x^4}{4} + \frac{x^3}{3} - \frac{x^2}{2} - x\right]_0^1 = \frac{-11}{12}$$

$$\langle f, f\rangle = \int_0^1 (x^4 - 2x^2 + 1)dx = \left[\frac{x^5}{5} - 2\frac{x^3}{3} + x\right]_0^1 = \frac{8}{15}$$

and

$$\langle g, g\rangle = \int_0^1 (x^2 + 2x + 1)dx = \left[\frac{x^3}{3} + x^2 + x\right]_0^1 = \frac{7}{3}$$

What if we define, $\langle f, g\rangle = \int_0^1 f(x)g(x)dx - 1$ for $f, g \in C[0, 1]$? Does it define an inner product on $C[0, 1]$? No, it doesn't! Observe that, for $f(x) = x^2 - 1$, we get $\langle f, f\rangle = \frac{8}{15} - 1 = \frac{-7}{15} < 0$. This is not possible for an inner product as it violates $(IP1)$. Now, let us discuss some of the basic properties of inner product spaces.

Theorem 5.4 *Let V be an inner product space. Then for $u, v, w \in V$ and $\lambda \in \mathbb{K}$, the following statements are true.*

(a) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$
(b) $\langle u, \lambda v\rangle = \overline{\lambda}\langle u, v\rangle$
(c) $\langle u, 0\rangle = \langle 0, u\rangle = 0$
(d) If $\langle u, v\rangle = \langle u, w\rangle$ for all $u \in V$, then $v = w$.

Proof For $u, v, w \in V$ and $\lambda \in \mathbb{K}$,

(a) $\langle u, v + w\rangle = \overline{\langle v + w, u\rangle} = \overline{\langle v, u\rangle + \langle w, u\rangle} = \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle} = \langle u, v\rangle + \langle u, w\rangle$
(b) $\langle u, \lambda v\rangle = \overline{\langle \lambda v, u\rangle} = \overline{\lambda\langle v, u\rangle} = \overline{\lambda}\ \overline{\langle v, u\rangle} = \overline{\lambda}\langle u, v\rangle$
(c) $\langle u, 0\rangle = \langle u, 0 + 0\rangle = \langle u, 0\rangle + \langle u, 0\rangle \Rightarrow \langle u, 0\rangle = 0$. Similarly $\langle 0, u\rangle = \langle 0 + 0, u\rangle = \langle 0, u\rangle + \langle 0, u\rangle = 0$.
(d) Suppose that $\langle u, v\rangle = \langle u, w\rangle$ for all $u \in V$.

$$\langle u, v\rangle = \langle u, w\rangle \Rightarrow \langle u, v\rangle - \langle u, w\rangle = 0 \Rightarrow \langle u, v - w\rangle = 0$$

That is, $\langle u, v\rangle = \langle u, w\rangle$ for all $u \in V$ implies that $\langle u, v - w\rangle = 0\ \forall\ u \in V$. In particular, $\langle v - w, v - w\rangle = 0$. This implies $v - w = 0$. That is, $v = w$.

The following theorem gives one of the most important and widely used inequalities in mathematics, called the Cauchy-Schwarz Inequality, named after the French mathematician *Augustin-Louis Cauchy (1789–1857)* and the German mathematician *Hermann Schwarz (1843–1921).*

Theorem 5.5 (Cauchy-Schwarz Inequality) *Let V be an inner product space. For $v, w \in V$,*

$$|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$$

where equality holds if and only if $\{v, w\}$ is linearly dependent.

Proof Let $v, w \in V$. Consider

$$u = \langle w, w\rangle v - \langle v, w\rangle w$$

Then

$$\begin{aligned} 0 \le \langle u, u\rangle &= \langle \langle w, w\rangle v - \langle v, w\rangle w, \langle w, w\rangle v - \langle v, w\rangle w\rangle \\ &= |\langle w, w\rangle|^2\langle v, v\rangle - \langle w, w\rangle|\langle v, w\rangle|^2 - \langle w, w\rangle|\langle v, w\rangle|^2 + \langle w, w\rangle|\langle v, w\rangle|^2 \\ &= \langle w, w\rangle\left[\langle v, v\rangle\langle w, w\rangle - |\langle v, w\rangle|^2\right] \end{aligned}$$

Now suppose that $\langle w, w\rangle > 0$, then $\langle v, v\rangle\langle w, w\rangle - |\langle v, w\rangle|^2 \ge 0$, which implies that $|\langle v, w\rangle|^2 \le \langle v, v\rangle\langle w, w\rangle$. If $\langle w, w\rangle = 0$, then by $(IP4)$, $w = 0$. Therefore by Theorem 5.4(c), $\langle v, w\rangle = 0$ and hence $\langle v, v\rangle\langle w, w\rangle = 0 = |\langle v, w\rangle|^2$.

Now suppose that equality holds. That is, $|\langle v, w\rangle|^2 = \langle v, v\rangle\langle w, w\rangle$. Then $\langle u, u\rangle = 0$. Then $\langle w, w\rangle v = \langle v, w\rangle w$ and hence $\{v, w\}$ is linearly dependent. Conversely, suppose that $\{v, w\}$ is linearly dependent. Then by Corollary 2.1, one is a scalar multiple of the other. That is, there exists $\lambda \in \mathbb{K}$ such that $v = \lambda w$ or $w = \lambda v$. Then

$$\langle v, v\rangle\langle w, w\rangle = \langle \lambda w, \lambda w\rangle\langle w, w\rangle = |\lambda|^2|\langle w, w\rangle|^2 = |\langle v, w\rangle|^2$$

Hence the proof.

Example 5.17 Consider $\mathbb{R}^n$ with standard inner product. For $(u_1, \ldots, u_n)$, $(v_1, \ldots, v_n) \in \mathbb{R}^n$, by Cauchy-Schwarz inequality, we have

$$(u_1v_1 + u_2v_2 + \cdots + u_nv_n)^2 \le (u_1 + u_2 + \cdots + u_n)^2(v_1 + v_2 + \cdots + v_n)^2$$

That is, $\left(\sum_{i=1}^n u_iv_i\right)^2 \le \left(\sum_{i=1}^n u_i\right)^2\left(\sum_{i=1}^n v_i\right)^2$. If we consider, $C[a, b]$ with the inner product, $\langle f, g\rangle = \int_a^b f(x)g(x)dx$, then by Cauchy-Schwarz inequality, we have

$$\left[\int_a^b f(x)g(x)dx\right]^2 \le \int_a^b f^2(x)dx \int_a^b g^2(x)dx$$

That is, $|\langle f, g\rangle|^2 \le \langle f, f\rangle\langle g, g\rangle$. Consider $f, g \in C[0, 1]$ as defined in Example 5.16. We have seen that $\langle f, g\rangle = \frac{-11}{12}$, $\langle f, f\rangle = \frac{8}{15}$ and $\langle g, g\rangle = \frac{7}{3}$. Clearly,

$$|\langle f, g\rangle|^2 = \frac{121}{144} \le \frac{56}{45} = \langle f, f\rangle\langle g, g\rangle$$

In the previous section, we have seen that every normed linear space is a metric space. Now, we will show that every inner product space is a normed linear space. The following theorem gives a method to define a norm on an inner product space using the inner product.

Theorem 5.6 *Let V be an inner product space. For $v \in V$, $\|v\| = \sqrt{\langle v, v\rangle}$ is a norm on V.*

Proof(N1) Let $v \in V$. Since $\langle v, v\rangle \ge 0$, we have $\|v\| = \sqrt{\langle v, v\rangle} \ge 0$. Also $\langle v, v\rangle = 0 \Leftrightarrow v = 0$, implies that $\|v\| = \sqrt{\langle v, v\rangle} = 0 \Leftrightarrow v = 0$.

(N2) $\|\lambda v\| = \sqrt{\langle \lambda v, \lambda v\rangle} = \sqrt{\lambda\bar{\lambda}\langle v, v\rangle} = \sqrt{|\lambda|^2\,\|v\|^2} = |\lambda|\,\|v\|$, where $\lambda \in \mathbb{K}$.
(N3) For $u, v \in V$,

$$\begin{aligned}\|u+v\|^2 &= \langle u+v, u+v\rangle \\ &= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\ &= \|u\|^2 + \|v\|^2 + 2Re(\langle u, v\rangle) \\ &\le \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle| \\ &\le \|u\|^2 + \|v\|^2 + 2\,\|u\|\,\|v\| \;\; (Cauchy - Schwarz) \\ &= (\|u\| + \|v\|)^2\end{aligned}$$

Hence $\|u+v\| \le \|u\| + \|v\|$.
Therefore $\|v\| = \sqrt{\langle v, v\rangle}$ is a norm on V.

Remark 5.3 The norm defined in the above theorem is called the norm induced by the inner product. Every inner product space is a normed linear space with respect to the induced norm.

Example 5.18 Consider $\mathbb{R}^n$ with standard inner product. Observe that for $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, we get

$$\|v\| = \sqrt{\langle v, v\rangle} = \left(\sum_{i=1}^{n} v_i^2\right)^{\frac{1}{2}} = \|v\|_2$$

Thus the standard inner product on $\mathbb{R}^n$ induces 2-norm. Similarly, the inner product $\langle f, g\rangle = \int_a^b f(x)g(x)dx$ on $C[a, b]$ induces the norm,

$$\|f\| = \sqrt{\langle f, f\rangle} = \left(\int_a^b f^2(x)dx\right)^{\frac{1}{2}}$$

This norm is called, *energy norm.*

The following inclusion can be derived between the collections of these abstract spaces.

$$\{\text{Inner product spaces}\} \subset \{\text{Normed spaces}\} \subset \{\text{Metric spaces}\}$$

Now we have to check whether the reverse inclusion is true or not. The following theorem gives a necessary condition for an inner product space.

Theorem 5.7 (Parallelogram Law) *Let V be an inner product space. Then for all $u, v \in V$,*

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$$

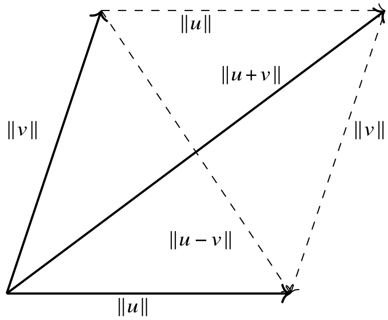


Fig. 5.11 Parallelogram law

Proof For all $u, v \in V$,

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle$$
$$\|u-v\|^2 = \langle u-v, u-v\rangle = \langle u,u\rangle - \langle u,v\rangle - \langle v,u\rangle + \langle v,v\rangle$$

Therefore $\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$ (Fig. 5.11).

Example 5.19 In Example 5.4, we have seen that $C[a,b]$, the space of continuous real-valued functions on $[a,b]$ is a normed linear space with the supremum norm given by, $\|f\| = \max_{x\in[a,b]} |f(x)|$ where $f \in C[a,b]$. This space gives an example of a normed linear space which is not an inner product space. Consider the elements $f_1(x) = 1$ and $f_2(x) = \dfrac{(x-a)}{(b-a)}$ in $C[a,b]$. Then $\|f_1\| = 1$ and $\|f_2\| = 1$. We have

$$(f_1 + f_2)(x) = 1 + \frac{(x-a)}{(b-a)} \text{ and } (f_1 - f_2)(x) = 1 - \frac{(x-a)}{(b-a)}$$

Hence $\|f_1 + f_2\| = 2$ and $\|f_1 - f_2\| = 1$. Now

$$\|f_1 + f_2\|^2 + \|f_1 - f_2\|^2 = 5 \text{ but } 2\left(\|f_1\|^2 + \|f_2\|^2\right) = 4$$

Clearly, parallelogram law is not satisfied. Thus supremum norm on $C[a,b]$ cannot be obtained from an inner product.

From the above example, we can conclude that not all normed linear spaces are inner product spaces. Now, we will prove that *a normed linear space is an inner product space if and only if the norm satisfies parallelogram law*.

Theorem 5.8 *Let* $(V, \|.\|)$ *be a normed linear space. Then there exists an inner product* $\langle , \rangle$ *on* V *such that* $\langle v, v\rangle = \|v\|^2$ *for all* $v \in V$ *if and only if the norm satisfies the parallelogram law.*

Proof Suppose that we have an inner product on V with $\langle v, v\rangle = \|v\|^2$ for all $v \in V$. Then by Theorem 5.7, parallelogram law is satisfied.

Conversely, suppose that the norm on V satisfies parallelogram law. For any $u, v \in V$, define

$$4\langle u, v\rangle = \|u+v\|^2 - \|u-v\|^2 + i\,\|u+iv\|^2 - i\,\|u-iv\|^2$$

Now we will prove that the inner product defined above will satisfy the conditions $(IP1) - (IP4)$.

(IP1) For any $v \in V$, we have

$$\begin{aligned} 4\langle v, v\rangle &= \|v+v\|^2 - \|v-v\|^2 + i\,\|v(1+i)\|^2 - i\,\|v(1-i)\|^2 \\ &= 4\,\|v\|^2 + i|1+i|^2\,\|v\|^2 - i|1-i|^2\,\|v\|^2 \\ &= 4\,\|v\|^2 + 2i\,\|v\|^2 - 2i\,\|v\|^2 \\ &= 4\,\|v\|^2 \end{aligned}$$

This implies that $\langle v, v\rangle = \|v\|^2$ for all $v \in V$. Hence $\langle v, v\rangle \geq 0$ for all $v \in V$ and $\langle v, v\rangle = 0$ if and only if $v = 0$.

(IP2) For any $u, v, w \in V$, we have

$$4\langle u+w, v\rangle = \|(u+w)+v\|^2 - \|(u+w)-v\|^2 + i\,\|(u+w)+iv\|^2 - i\,\|(u+w)-iv\|^2$$

rewriting $u+w+v$ as $\left(u+\frac{v}{2}\right)+\left(w+\frac{v}{2}\right)$ and applying parallelogram law, we have

$$\left\|\left(u+\frac{v}{2}\right)+\left(w+\frac{v}{2}\right)\right\|^2 + \left\|\left(u+\frac{v}{2}\right)-\left(w+\frac{v}{2}\right)\right\|^2 = 2\left\|u+\frac{v}{2}\right\|^2 + 2\left\|w+\frac{v}{2}\right\|^2$$

This implies

$$\|u+w+v\|^2 = 2\left\|u+\frac{v}{2}\right\|^2 + 2\left\|w+\frac{v}{2}\right\|^2 - \|u-w\|^2$$

Similarly,

$$\|u+w-v\|^2 = 2\left\|u-\frac{v}{2}\right\|^2 + 2\left\|w-\frac{v}{2}\right\|^2 - \|u-w\|^2$$

Then

$$\|u+w+v\|^2 - \|u+w-v\|^2 = 2\left[\left\|u+\frac{v}{2}\right\|^2 - \left\|u-\frac{v}{2}\right\|^2 + \left\|w+\frac{v}{2}\right\|^2 - \left\|w-\frac{v}{2}\right\|^2\right] \tag{5.4}$$

Multiplying both sides by i and replacing v by iv in the above equation,

$$i\left[\|u+w+iv\|^2-\|u+w-iv\|^2\right]=2i\left[\left\|u+\frac{iv}{2}\right\|^2-\left\|u-\frac{iv}{2}\right\|^2+\left\|w+\frac{iv}{2}\right\|^2-\left\|w-\frac{iv}{2}\right\|^2\right] \tag{5.5}$$

adding (5.4) and (5.5),we get

$$\begin{aligned}4\langle u+w,v\rangle &= 2\left[\left\|u+\frac{v}{2}\right\|^2-\left\|u-\frac{v}{2}\right\|^2+i\left\|u+\frac{iv}{2}\right\|^2-i\left\|u-\frac{iv}{2}\right\|^2\right]\\ &\quad+2\left[\left\|w+\frac{v}{2}\right\|^2-\left\|w-\frac{v}{2}\right\|^2+i\left\|w+\frac{iv}{2}\right\|^2-i\left\|w-\frac{iv}{2}\right\|^2\right]\\ &= 8\left[\left\langle u,\frac{v}{2}\right\rangle+\left\langle w,\frac{v}{2}\right\rangle\right]\end{aligned}$$

No taking $w=0$ and then $u=0$ separately in the above equation, we get $\langle u,v\rangle = 2\left\langle u,\frac{v}{2}\right\rangle$ and $\langle w,v\rangle = 2\left\langle w,\frac{v}{2}\right\rangle$. Thus we get, $4\langle u+w,v\rangle = 4\langle u,v\rangle + 4\langle w,v\rangle$ for all $u,v,w\in V$.

(IP3) Now we will prove that $\langle \lambda u,v\rangle = \lambda\langle u,v\rangle$. We will prove this as four separate cases.

(a) λ is an integer.
For all $u,v,w\in V$, we have

$$\langle u+w,v\rangle = \langle u,v\rangle + \langle w,v\rangle$$

Replacing w by u, we get $\langle 2u,v\rangle = 2\langle u,v\rangle$. Thus the result is true for $\lambda = 2$. Suppose that the result is true for any positive integer n. That is, $\langle nu,v\rangle = n\langle u,v\rangle$ for all $u,v\in V$. Now

$$\langle (n+1)u,v\rangle = \langle nu+u,v\rangle = \langle nu,v\rangle + \langle u,v\rangle = (n+1)\langle u,v\rangle$$

hence by the principle of mathematical induction, the result is true for all positive integers n. Now, to prove this for any negative integer n, first we prove that $\langle -u,v\rangle = -\langle u,v\rangle$, for any $u,v\in V$. We have

$$4\langle u,v\rangle = \|u+v\|^2-\|u-v\|^2+i\,\|u+iv\|^2-i\,\|u-iv\|^2$$

Replacing u by $-u$, we get

$$\begin{aligned}4\langle -u,v\rangle &= \|-u+v\|^2-\|-u-v\|^2+i\,\|-u+iv\|^2-i\,\|-u-iv\|^2\\ &= \|-(u-v)\|^2-\|-(u+v)\|^2+i\,\|-(u-iv)\|^2-i\,\|-(u+iv)\|^2\\ &= \|u-v\|^2-\|u+v\|^2+i\,\|u-iv\|^2-i\,\|u+iv\|^2\\ &= -4\langle u,v\rangle\end{aligned}$$

Thus we have $\langle -u, v\rangle = -\langle u, v\rangle$ for any $u, v \in V$. Let $\lambda = -\mu$ be any negative integer, where $\mu > 0$. Then we have,

$$\langle \lambda u, v\rangle = \langle -\mu u, v\rangle = \langle -(\mu u), v\rangle = -\langle \mu u, v\rangle = -\mu\langle u, v\rangle = \lambda\langle u, v\rangle$$

Thus the result is true for any integer λ.

(b) $\lambda = \frac{p}{q}$ is a rational number, where $p, q \neq 0$ are integers.
Then we have

$$p\langle u, v\rangle = \langle pu, v\rangle = \left\langle q\left(\frac{p}{q}\right)u, v\right\rangle = q\left\langle \frac{p}{q}u, v\right\rangle$$

Thus we have $\left\langle \frac{p}{q}u, v\right\rangle = \frac{p}{q}\langle u, v\rangle$ for all $u, v \in V$. Thus the result is true for all rational numbers.

(c) λ is a real number.
Then there exists a sequence of rational numbers $\{\lambda_n\}$ such that $\lambda_n \to \lambda$ as $n \to \infty$ (See Exercise 13, Chap. 1). Observe that, as $n \to \infty$

$$|\lambda_n\langle u, v\rangle - \lambda\langle u, v\rangle| = |(\lambda_n - \lambda)\langle u, v\rangle| = |\lambda_n - \lambda||\langle u, v\rangle| \to 0$$

Hence, $\lambda_n\langle u, w\rangle \to \lambda\langle u, v\rangle$ as $n \to \infty$. Now, by (b), $\lambda_n\langle u, v\rangle = \langle \lambda_n u, v\rangle$. Also,

$$\begin{aligned}4\langle \lambda_n u, v\rangle &= \|\lambda_n u + v\|^2 - \|\lambda_n u - v\|^2 + i\,\|\lambda_n u + iv\|^2 - i\,\|\lambda_n u - iv\|^2\\ &\to \|\lambda u + v\|^2 - \|\lambda u - v\|^2 + i\,\|\lambda u + iv\|^2 - i\,\|\lambda u - iv\|^2\\ &= 4\langle \lambda u, v\rangle\end{aligned}$$

That is, $\langle \lambda_n u, v\rangle \to \langle \lambda u, v\rangle$ as $n \to \infty$. This implies that $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for any $u, v \in V$.

(d) λ is a complex number.
First we will show that $\langle iu, v\rangle = i\langle u, v\rangle$. We have

$$4\langle u, v\rangle = \|u + v\|^2 - \|u - v\|^2 + i\,\|u + iv\|^2 - i\,\|u - iv\|^2$$

Replacing u by iu, we have

$$\begin{aligned}4\langle iu, v\rangle &= \|iu + v\|^2 - \|iu - v\|^2 + i\,\|iu + iv\|^2 - i\,\|iu - iv\|^2\\ &= \|i(u - iv)\|^2 - \|i(u + iv)\|^2 + i\,\|i(u + v)\|^2 - i\,\|i(u - v)\|^2\\ &= \|u - iv\|^2 - \|u + iv\|^2 + i\,\|u + v\|^2 - i\,\|u - v\|^2\\ &= -i^2\,\|u - iv\|^2 + i^2\,\|u + iv\|^2 + i\,\|u + v\|^2 - i\,\|u - v\|^2\\ &= i\left[\|u + v\|^2 - \|u - v\|^2 + i\,\|u + iv\|^2 - i\,\|u - iv\|^2\right]\\ &= i4\langle u, v\rangle\end{aligned}$$

which implies that $\langle iu, v\rangle = i\langle u, v\rangle$. Now, for any complex number $\lambda = a + ib$, then

$$\begin{aligned}\langle \lambda u, v\rangle &= \langle (a+ib)u, v\rangle \\ &= \langle au + ibu, v\rangle \\ &= \langle au, v\rangle + \langle ibu, v\rangle \\ &= a\langle u, v\rangle + ib\langle u, v\rangle \\ &= (a+ib)\langle u, v\rangle = \lambda\langle u, v\rangle\end{aligned}$$

Thus $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $u, v \in V$ and for all scalars λ.

(IP4) For any $u, v \in V$, we have

$$\begin{aligned}4\overline{\langle u, v\rangle} &= \|u+v\|^2 - \|u-v\|^2 + i\,\|u+iv\|^2 - i\,\|u-iv\|^2 \\ &= \|v+u\|^2 - \|v-u\|^2 + i\,\|i(v-iu)\|^2 - i\,\|(-i)(v+iu)\|^2 \\ &= \|v+u\|^2 - \|v-u\|^2 + i|i|^2\,\|v-iu\|^2 - i|-i|^2\,\|v+iu\|^2 \\ &= \|v+u\|^2 - \|v-u\|^2 - i\,\|v-iu\|^2 + i\,\|v+iu\|^2 \\ &= 4\langle v, u\rangle\end{aligned}$$

Hence, $\overline{\langle u, v\rangle} = \langle v, u\rangle$ for all $u, v \in V$.

Thus all the conditions for an inner product are satisfied and hence $(V, \langle, \rangle)$ is an inner product space.

Similar to what we have done in normed linear spaces, the concept of convergence of sequences in inner product spaces follows from the definition of convergence in metric spaces as given below.

Definition 5.12 (*Convergence*) Let $\{v_n\}$ be a sequence in an inner product space V, then $v_n \to v$ if and only if $\langle v_n, v\rangle \to 0$ as $n \to \infty$.

Again the question of completeness rises. The following example shows that every inner product space need not necessarily be complete.

Example 5.20 Consider $C[0, 1]$ with the inner product $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. We have already seen that $C[0, 1]$ is an inner product space with respect to the given inner product. Now, consider the sequence,

$$f_n = \begin{cases} 0, \ x \in \left[0, \frac{1}{2}\right] \\ n\left(x - \frac{1}{2}\right), \ x \in \left[\frac{1}{2}, \frac{1}{2} + \frac{1}{n}\right] \\ 1, \ x \in \left[\frac{1}{2} + \frac{1}{n}, 1\right] \end{cases}$$

If we proceed as in Example 5.12, we can show that $\{f_n\}$ is Cauchy but not convergent.

Complete inner product spaces are named after the famous German mathematician *David Hilbert (1862–1943)* who started a systematic study in this area.

Definition 5.13 (*Hilbert Space*) A complete inner product space is called a Hilbert space.

Example 5.21 Consider $\mathbb{K}^n$ over $\mathbb{K}$ with standard inner product. Then $\|v\| = \sqrt{\langle v, v\rangle} = \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}}$ for $v = (v_1, v_2, \ldots, v_n) \in \mathbb{K}^n$. Then from Example 5.13, $\mathbb{K}^n$ over $\mathbb{K}$ with standard inner product is a Hilbert space. In fact, we can prove that every finite-dimensional space over the fields $\mathbb{R}$ or $\mathbb{C}$ is complete(Prove). Is $\mathbb{Q}$ over the field $\mathbb{Q}$ complete?

5.3 Orthogonality of Vectors and Orthonormal Sets

Orthogonality of vectors in vector spaces is one of the important basic concepts in mathematics which is generalized from the idea that the dot product of two vectors is zero implies that the vectors are perpendicular in $\mathbb{R}^2$ (Fig. 5.12).

Orthogonal/orthonormal bases are of great importance in functional analysis, which we will be discussing in the coming sections. We will start with the definition of an orthogonal set.

Definition 5.14 (*Orthogonal set*) Let V be an inner product space. Vectors $v, w \in V$ are *orthogonal* if $\langle v, w\rangle = 0$. A subset S of V is orthogonal if any two distinct vectors in S are orthogonal.

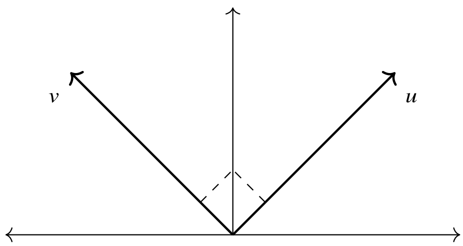
We are all familiar with the fundamental relation from Euclidean geometry that, "in a right-angled triangle, the square of the hypotenuse is equal to the sum of squares of the other two sides", named after the famous Greek mathematician, *Pythagoras(570-495 BC)* (Fig. 5.13).

This relation can be generalized to higher-dimensional spaces, to spaces that are not Euclidean, to objects that are not right triangles, and to objects that are not even triangles. Consider the following theorem.

Theorem 5.9 (Pythagoras Theorem) *Let V be an inner product space and $\{v_1, v_2, \ldots, v_n\}$ be an orthogonal set in V. Then*

$$\|v_1 + v_2 + \cdots + v_n\|^2 = \|v_1\|^2 + \|v_2\|^2 + \cdots + \|v_2\|^2$$

Fig. 5.12 Example for orthogonal vectors in $\mathbb{R}^2$



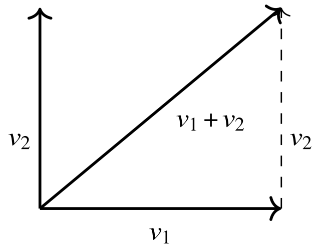


Fig. 5.13 Pythagoras theorem illustrated in $\mathbb{R}^2$

Proof As $\{v_1, v_2, \ldots, v_n\}$ is an orthogonal set in V, we have $\langle v_i, v_j \rangle = 0,\ \forall i \neq j$. Then

$$\begin{aligned}\|v_1 + v_2 + \cdots + v_n\|^2 &= \langle v_1 + v_2 + \cdots + v_n, v_1 + v_2 + \cdots + v_n \rangle \\ &= \sum_{i,j=1}^{n} \langle v_i, v_j \rangle \\ &= \sum_{i=1}^{n} \langle v_i, v_i \rangle \\ &= \|v_1\|^2 + \|v_2\|^2 + \cdots + \|v_2\|^2\end{aligned}$$

Definition 5.15 (*Orthonormal set*) A vector $v \in V$ is a *unit vector* if $\|v\| = 1$. A subset S of V is *orthonormal* if S is orthogonal and consists entirely of unit vectors. A subset of V is an *orthonormal basis* for V if it is an ordered basis that is orthonormal.

Example 5.22 Consider the set $S = \{v_1, v_2, v_3\}$ in $C[-1, 1]$, where

$$v_1 = \frac{1}{\sqrt{2}}, v_2 = \sqrt{\frac{3}{2}}x \text{ and } v_3 = \sqrt{\frac{5}{8}}(3x^2 - 1)$$

Then

$$\langle v_1, v_1 \rangle = \int_{-1}^{1} \frac{1}{2} dx = 1, \langle v_2, v_2 \rangle = \frac{3}{2} \int_{-1}^{1} x^2 dx = 1, \langle v_3, v_3 \rangle = \frac{5}{8} \int_{-1}^{1} (9x^4 - 6x^2 + 1) dx = 1$$

and

$$\langle v_1, v_2 \rangle = \frac{\sqrt{3}}{2} \int_{-1}^{1} x dx = 0, \langle v_1, v_3 \rangle = \frac{\sqrt{5}}{4} \int_{-1}^{1} (3x^2 - 1) dx = 0,$$

$$\langle v_2, v_3 \rangle = \frac{\sqrt{15}}{4} \int_{-1}^{1} (3x^3 - x) dx = 0$$

Thus S is an orthonormal set in $C[-1, 1]$. As $\mathbb{P}_2[-1, 1]$ is a subspace of $C[-1, 1]$ with dimension 3, S can be considered as an orthonormal basis for $\mathbb{P}_2[-1, 1]$.

Example 5.23 Consider the standard ordered basis $S = \{e_1, e_2, \dots, e_n\}$ in $\mathbb{R}^n$ with standard inner product. Clearly $\langle e_i, e_j \rangle = 0$ for $i \neq j$ and $\|e_i\| = \sqrt{\langle e_i, e_i \rangle} = 1$ for all $i = 1, 2, \dots, n$. Therefore the standard ordered basis of $\mathbb{R}^n$ is an orthonormal basis.

In the previous chapters, we have seen that bases are the building blocks of a vector space. Now, suppose that this basis is orthogonal. Do we have any advantage? Consider the following example.

Example 5.24 Consider the vectors $v_1 = (2, 1, 2)$, $v_2 = (-2, 2, 1)$ and $v_3 = (1, 2, -2)$ in $\mathbb{R}^3$. Clearly, we can see that $\{v_1, v_2, v_3\}$ is an orthogonal basis for $\mathbb{R}^3$(verify). Then we know that any non-zero vector in $\mathbb{R}^3$ can be written as a linear combination of $\{v_1, v_2, v_3\}$ in a unique way. That is, any $v \in \mathbb{R}^3$ can be expressed as $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$ for some scalars $\lambda_1, \lambda_2, \lambda_3$. Because of the orthogonality of basis vectors, here we can observe that,

$$\langle v, v_1 \rangle = \langle \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3, v_1 \rangle = \lambda_1 \langle v_1, v_1 \rangle = \lambda_1 \|v_1\|^2$$

Hence, $\lambda_1 = \frac{\langle v, v_1 \rangle}{\|v_1\|^2}$. Similarly, we can compute λ_2 and λ_3 as $\frac{\langle v, v_2 \rangle}{\|v_1\|^2}$ and $\frac{\langle v, v_3 \rangle}{\|v_1\|^2}$, respectively. This is interesting! right? Let us consider a numerical example. Take $v = (6, 12, -3) \in \mathbb{R}^3$. We have

$$(6, 12, -3) = 2(2, 1, 2) + 1(-2, 2, 1) + 4(1, 2, -2)$$

Observe that $\frac{\langle v, v_1 \rangle}{\|v_1\|^2} = 2$, $\frac{\langle v, v_2 \rangle}{\|v_2\|^2} = 1$ and $\frac{\langle v, v_3 \rangle}{\|v_3\|^2} = 4$. Is this possible in any arbitrary inner product space? Yes, it is possible!! That is, if we have an orthogonal basis for an inner product space V, it is easy to represent any vector $v \in V$ as a linear combination of the basis vectors. For, if $\{v_1, v_2, \dots, v_n\}$ is an orthogonal basis for an inner product space V, then for any $v \in V$, we have

$$v = \frac{\langle v, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle v, v_2 \rangle}{\|v_2\|^2} v_2 + \cdots + \frac{\langle v, v_n \rangle}{\|v_n\|^2} v_n$$

and if $\{v_1, v_2, \dots, v_n\}$ is an orthonormal basis for V, we have

$$v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \cdots + \langle v, v_n \rangle v_n$$

This fact is formulated as the following theorem.

Theorem 5.10 *Let V be an inner product space and $S = \{v_1, v_2, \dots, v_n\}$ be an orthogonal subset of V consisting of non-zero vectors. If $w \in span(S)$, then*

$$w = \sum_{i=1}^{n} \frac{\langle w, v_i \rangle}{\|v_i\|^2} v_i$$

Further if S is an orthonormal set,

$$w = \sum_{i=1}^{n} \langle w, v_i \rangle v_i$$

Proof Since $w \in span(S)$, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$ such that $w = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$. Now for $i = 1, 2, \ldots, n$, we have

$$\begin{aligned} \langle w, v_i \rangle &= \langle \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n, v_i \rangle \\ &= \lambda_1 \langle v_1, v_i \rangle + \lambda_2 \langle v_2, v_i \rangle + \cdots + \lambda_n \langle v_n, v_i \rangle \end{aligned}$$

Since $S = \{v_1, v_2, \ldots, v_n\}$ is an orthogonal set, $\langle v_i, v_j \rangle = 0$ for all $i \neq j$ and $\langle v_i, v_i \rangle = \|v_i\|^2 \neq 0$. Therefore

$$\langle w, v_i \rangle = \lambda_i \, \|v_i\|^2$$

and hence $\lambda_i = \dfrac{\langle w, v_i \rangle}{\|v_i\|^2}$ for $i = 1, 2, \ldots, n$. This implies that $w = \sum_{i=1}^{n} \dfrac{\langle w, v_i \rangle}{\|v_i\|^2} v_i$. If S is orthonormal, $v_1, v_2, \ldots, v_n$ are unit vectors and hence $\|v_i\| = 1$ for $i = 1, 2, \ldots, n$. Therefore $w = \sum_{i=1}^{n} \langle w, v_i \rangle v_i$.

Remark 5.4 The coefficients $\dfrac{\langle w, v_i \rangle}{\|v_i\|^2}$ is called the *Fourier coefficients* of v with respect to the basis $\{v_1, v_2, \ldots, v_n\}$, named after the French mathematician *Jean-Baptiste Joseph Fourier* (1768–1830).

The following corollary shows that the matrix representation of a linear operator defined on a finite-dimensional vector space with orthonormal basis can be easily computed using the idea of an inner product.

Corollary 5.1 *Let V be an inner product space, and let $B = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis of V. If T is a linear operator on V, and $A = [T]_B$. Then $A_{ij} = \langle T(v_j), v_i \rangle$, where $1 \leq i, j \leq n$.*

Proof Since B is a basis of V and as T is from V to V, from the above theorem

$$T(v_j) = \sum_{i=1}^{n} \langle T(v_j), v_i \rangle v_i$$

which clearly implies that $A_{ij} = \langle T(v_j), v_i \rangle$, where $1 \leq i, j \leq n$.

Example 5.25 Consider $\mathbb{P}_2[-1, 1]$ with the basis defined in Example 5.22. Take an arbitrary element, say $w = x^2 + 2x + 3 \in \mathbb{P}_2[-1, 1]$. Then we have,

$$\langle w, v_1 \rangle = \frac{1}{\sqrt{2}} \int_{-1}^{1} (x^2 + 2x + 3) dx = \frac{10\sqrt{2}}{3}$$

$$\langle w, v_2 \rangle = \sqrt{\frac{3}{2}} \int_{-1}^{1} (x^3 + 2x^2 + 3x)dx = \frac{2\sqrt{6}}{3}$$

and

$$\langle w, v_3 \rangle = \frac{1}{\sqrt{2}} \int_{-1}^{1} (3x^2 - 1)(x^2 + 2x + 3)dx = \frac{\sqrt{40}}{15}$$

Observe that $w = \frac{10\sqrt{2}}{3} v_1 + \frac{2\sqrt{2}}{\sqrt{3}} v_2 + \frac{\sqrt{40}}{15} v_3$.

Define $T : V \to V$ by

$$(Tp)(x) = p'(x)$$

Then

$$T(v_1) = 0, T(v_2) = \sqrt{\frac{3}{2}} \text{ and } T(v_3) = \frac{\sqrt{15}}{2} x$$

Clearly $\langle T(v_1), v_i \rangle = 0$ where $i = 1, 2, 3$. Also

$$\langle T(v_2), v_1 \rangle = \frac{\sqrt{3}}{2} \int_{-1}^{1} dx = \sqrt{3}, \langle T(v_2), v_2 \rangle = \frac{3}{2} \int_{-1}^{1} x dx = 0,$$

$$\langle T(v_2), v_3 \rangle = \frac{\sqrt{15}}{4} \int_{-1}^{1} (3x^2 - 1)dx = 0$$

And

$$\langle T(v_3), v_1 \rangle = \frac{\sqrt{15}}{2\sqrt{2}} \int_{-1}^{1} x dx = 0, \langle T(v_3), v_2 \rangle = \frac{3\sqrt{5}}{2\sqrt{2}} \int_{-1}^{1} x^2 dx = \sqrt{\frac{5}{2}},$$

$$\langle T(v_3), v_3 \rangle = \frac{5\sqrt{3}}{4\sqrt{2}} \int_{-1}^{1} (3x^3 - x)dx = 0$$

Therefore

$$[T]_B = \begin{bmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{bmatrix}$$

Corollary 5.2 *Let V be an inner product space, and $S = \{v_1, v_2, \ldots, v_k\}$ be an orthogonal subset of V consisting of non-zero vectors. Then S is linearly independent.*

Proof Let $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{K}$ be such that $\sum_{i=1}^{k} \lambda_i v_i = 0$. Then for $v_j \in S$,

$$0 = \left\langle \sum_{i=1}^{k} \lambda_i v_i, v_j \right\rangle = \lambda_j \left\| v_j \right\|^2$$

Since S is a collection of non-zero vectors, this implies that $\lambda_j = 0$ for all $j = 1, 2, \ldots, k$. Therefore S is linearly independent.

Gram–Schmidt Orthonormalization

Corollary 5.2 shows that any orthogonal set of non-zero vectors is linearly independent. In this section, we will show that from a linearly independent set, we can construct an orthogonal set. In fact, we can construct an orthonormal set from a linearly independent set, with the same span using *Gram–Schmidt Orthonormalization* process. The process is named after the Danish mathematician *Jørgen Pedersen Gram (1850–1916)* and Baltic-German mathematician *Erhard Schmidt (1876–1959).*

Theorem 5.11 (Gram–Schmidt Orthonormalization) *Let* $\{v_1, v_2, \ldots v_n\}$ *be a linearly independent subset of an inner product space* V*. Define*

$$w_1 = v_1, \ u_1 = \frac{w_1}{\|w_1\|}$$

$$w_2 = v_2 - \langle v_2, u_1\rangle u_1, \ u_2 = \frac{w_2}{\|w_2\|}$$

$$w_3 = v_3 - \langle v_3, u_1\rangle u_1 - \langle v_3, u_2\rangle u_2, u_3 = \frac{w_3}{\|w_3\|}$$

$$\vdots$$

$$w_n = v_n - \langle v_n, u_1\rangle u_1 - \cdots - \langle v_n, u_{n-1}\rangle u_{n-1}, \quad u_n = \frac{w_n}{\|w_n\|}$$

Then $\{u_1, u_2, \ldots u_n\}$ *is an orthonormal set in* V *and*

$$span\{u_1, u_2, \ldots, u_n\} = span\{v_1, v_2, \ldots, v_n\}$$

Proof Since $\{v_1, v_2, \ldots v_n\}$ is linearly independent, $v_i \neq 0$ for all $i = 1, 2, \ldots, n$. We prove by induction on i. Consider $\{v_1\}$. Clearly $\{v_1\}$ is linearly independent. Take $w_1 = v_1$ and $u_1 = \frac{w_1}{\|w_1\|}$. Then $\|u_1\| = \frac{\|w_1\|}{\|w_1\|} = 1$ and $span\{u_1\} = span\{v_1\}$ (Fig. 5.14).

For $0 \leq i \leq n-1$, define

$$w_i = v_i - \langle v_i, u_1\rangle u_1 - \cdots - \langle v_i, u_{i-1}\rangle u_{i-1}, \quad u_i = \frac{w_i}{\|w_i\|}$$

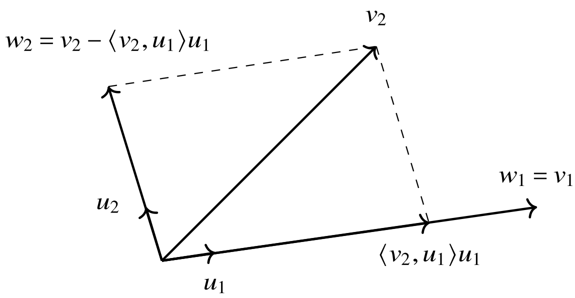


Fig. 5.14 Geometrical representation of first two steps of Gram–Schmidt process

and suppose that $\{u_1, u_2, \dots u_{n-1}\}$ is an orthonormal set with

$$span\{u_1, u_2, \dots, u_{n-1}\} = span\{v_1, v_2, \dots, v_{n-1}\}$$

Now define,

$$w_n = v_n - \langle v_n, u_1\rangle u_1 - \cdots - \langle v_n, u_{n-1}\rangle u_{n-1}$$

Since $\{v_1, v_2, \dots v_n\}$ is a linearly independent set $v_n \notin span\{v_1, v_2, \dots, v_{n-1}\} = span\{u_1, u_2, \dots, u_{n-1}\}$. Since $w_n \neq 0$, take $u_n = \dfrac{w_n}{\|w_n\|}$. Then clearly $\|u_n\| = 1$. Now for $i \leq n-1$, we have

$$\begin{aligned}\langle w_n, u_i\rangle &= \langle v_n - \langle v_n, u_1\rangle u_1 - \cdots - \langle v_n, u_{n-1}\rangle u_{n-1}, u_i\rangle \\ &= \langle v_n, u_i\rangle - \langle v_n, u_1\rangle\langle u_1, u_i\rangle - \cdots - \langle v_n, u_{n-1}\rangle\langle u_{n-1}, u_i\rangle \\ &= \langle v_n, u_i\rangle - \langle v_n, u_i\rangle \\ &= 0\end{aligned}$$

as $\{u_1, u_2, \dots u_{n-1}\}$ is an orthonormal set. Therefore $\langle w_n, w_i\rangle = 0$ for $0 \leq i \leq n-1$ and hence $\{u_1, u_2, \dots u_n\}$ is an orthonormal set. Also

$$\begin{aligned}span\{u_1, u_2, \dots u_n\} &= span\{v_1, v_2, \dots, v_{n-1}, u_n\} \\ &= span\left\{v_1, v_2, \dots, v_{n-1}, \frac{w_n}{\|w_n\|}\right\} \\ &= span\{v_1, v_2, \dots, v_n\}\end{aligned}$$

Hence the proof.

Example 5.26 Let $V = \mathbb{R}^4$ and

$$S = \{v_1 = (0, 1, 1, 0), v_2 = (1, 2, 1, 0), v_3 = (1, 0, 0, 1)\}$$

Since $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ is of rank 3, S is linearly independent. Also as $\langle v_1, v_2 \rangle = 2 + 1 = 3$, S is not orthogonal. Now we may apply, *Gram–Schmidt process* to obtain an orthonormal set. Take $w_1 = v_1 = (1, 0, 1, 0)$. Then $u_1 = \dfrac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}}(0, 1, 1, 0)$.

Now

$$\begin{aligned} w_2 &= v_2 - \langle v_2, u_1 \rangle u_1 \\ &= (1, 2, 1, 0) - \langle (1, 2, 1, 0), \frac{1}{\sqrt{2}}(0, 1, 1, 0) \rangle \frac{1}{\sqrt{2}}(0, 1, 1, 0) \\ &= \frac{1}{2}\,(2, 1, -1, 0) \end{aligned}$$

and hence $u_2 = \dfrac{w_2}{\|w_2\|} = \dfrac{1}{\sqrt{6}}(2, 1, -1, 0)$. Finally,

$$\begin{aligned} w_3 &= v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2 \\ &= (1, 0, 0, 1) - \langle (1, 0, 0, 1), \frac{1}{\sqrt{2}}(0, 1, 1, 0) \rangle \frac{1}{\sqrt{2}}(0, 1, 1, 0) \\ &\qquad\qquad - \langle (1, 0, 0, 1), \frac{1}{3}(2, 1, -1, 0) \rangle \frac{1}{3}(2, 1, -1, 0) \\ &= \frac{1}{3}(1, -1, 1, 3) \end{aligned}$$

and hence $u_3 = \dfrac{w_3}{\|w_3\|} = \frac{1}{2\sqrt{3}}(1, -1, 1, 3)$. The set $\{u_1, u_2, u_3\}$ is an orthonormal set and $span\{u_1, u_2, u_3\} = span\{v_1, v_2, v_3\}$.

Remark 5.5 Consider a matrix A with columns v_1, v_2, v_3 from the above example. That is, $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then

$$A = \begin{bmatrix} 0 & \sqrt{\frac{2}{3}} & \frac{\sqrt{3}}{6} \\ \frac{\sqrt{2}}{2} & \frac{1}{\sqrt{6}} & -\frac{\sqrt{3}}{6} \\ \frac{\sqrt{2}}{2} & -\frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{6} \\ 0 & 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} \sqrt{2} & \frac{3\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \sqrt{\frac{3}{2}} & \sqrt{\frac{2}{3}} \\ 0 & 0 & \frac{2\sqrt{3}}{3} \end{bmatrix} = QR$$

Clearly, the columns of the matrix Q forms an orthonormal set and R is an upper triangular matrix with entries $R_{ii} = \|w_i\| \; \forall \; i = 1, 2, 3$ and $R_{ij} = \langle v_j, u_i \rangle \; \forall \; j > i (i, j = 1, 2, 3)$. This decomposition of a matrix with linearly independent columns into the product of an upper triangular matrix and a matrix whose columns form an orthonormal set is called the QR- decomposition.

Example 5.27 Consider $V = \mathbb{P}_2[-1, 1]$ and $S = \{1, x, x^2\}$. We have already seen that S is a basis of V and hence is linearly independent. Also as $\int_{-1}^{1} 1.x^2 dx = \frac{2}{3}$, S is not orthogonal. Therefore take $w_1 = 1$. As $\|w_1\|^2 = \int_{-1}^{1} 1dx = 2$, we get $u_1 = \frac{1}{\sqrt{2}}$. Now

$$w_2 = v_2 - \langle v_2, u_1 \rangle u_1 = x - \frac{1}{4}\int_{-1}^{1} xdx = x, u_2 = \sqrt{\frac{3}{2}}x$$

and

$$w_3 = v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2 = x^2 - \int_{-1}^{1} x^2 dx, u_3 = \sqrt{\frac{5}{8}}(3x^2 - 1)$$

Thus $\left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2 - 1)\right\}$ is an orthonormal basis for $\mathbb{P}_2[-1, 1]$.

The above example makes it clear that given a basis, one could construct an orthonormal basis from it. Hence, we could assure that *"Every finite-dimensional vector space has an orthonormal basis"*.

5.4 Orthogonal Complement and Projection

In Sect. 5.2, we have discussed about orthogonal projection on $\mathbb{R}^2$. We will extend this idea to the general inner product space structure here. Representing an inner product space as the direct sum of a closed subspace and its orthogonal complement has many useful applications in mathematics.

Definition 5.16 Let S be a non-empty subset of an inner product space V, then the set $\{v \in V \mid \langle v, s \rangle = 0, \forall s \in S\}$,i.e., the set of all vectors of V that are orthogonal to every vector in S is called the orthogonal complement of S and is denoted by $S^\perp$. Clearly $\{0\}^\perp = V$ and $V^\perp = \{0\}$. Also $S \cap S^\perp = \{0\}$.

Remark 5.6 $S^\perp$ is a subspace of V for any subset of V. For

$$\langle \lambda s_1 + s_2, s \rangle = \lambda \langle s_1, s \rangle + \langle s_2, s \rangle = 0$$

for all $s_1, s_2 \in S^\perp$ and $\lambda \in \mathbb{K}$ (Fig. 5.15).

Example 5.28 Consider $V = \mathbb{R}^3$ and let $S_1 = \{(1, 2, 3)\}$. Then

$$\begin{aligned} S_1^\perp &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid \langle (v_1, v_2, v_3), (1, 2, 3) \rangle = 0\} \\ &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + 2v_2 + 3v_3 = 0\} \\ &= \text{plane passing through origin and perpendicular to the point } (1, 2, 3) \end{aligned}$$

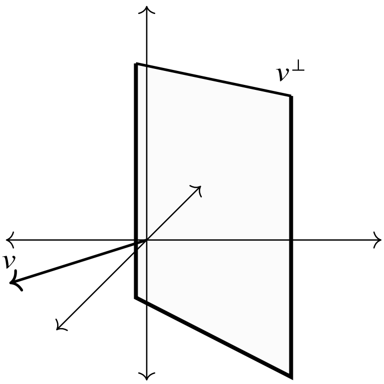


Fig. 5.15 Suppose v is a non-zero vector in $\mathbb{R}^3$. Then $v^\perp$ is the plane passing through origin O and perpendicular to the vector v

Take $S_2 = \{(1, 0, 1), (1, 2, 3)\}$. Then

$$\begin{aligned}
S_2^\perp &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid \langle (v_1, v_2, v_3), (1, 0, 1)\rangle = 0, \langle (v_1, v_2, v_3), (1, 2, 3)\rangle = 0\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + v_3 = 0, v_1 + 2v_2 + 3v_3 = 0\} \\
&= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = v_2 = -v_3\} \\
&= \text{line passing through origin and passing through the point } (1, 1, -1)
\end{aligned}$$

Observe that if S is a singleton set (with non-zero element), $S^\perp$ will be a plane passing through the origin as we will have to solve a homogeneous equation of three variables to find $S^\perp$. Similarly, if S is a set with two linearly independent elements, $S^\perp$ will be a line passing through the origin.

Example 5.29 Consider $V = \mathbb{P}_2[0, 1]$ and let $S = \{x\}$. Then

$$\begin{aligned}
S^\perp &= \{ax^2 + bx + c \in \mathbb{P}_2[0, 1] \mid \langle x, ax^2 + bx + c\rangle = 0\} \\
&= \{ax^2 + bx + c \in \mathbb{P}_2[0, 1] \mid \int_0^1 (ax^3 + bx^2 + cx)dx = 0\} \\
&= \{ax^2 + bx + c \in \mathbb{P}_2[0, 1] \mid 3a + 4b + 6c = 0\}
\end{aligned}$$

Given a subspace of an inner product space V, it is not always easy to find the orthogonal complement. The following theorem simplifies our effort in finding the orthogonal complement of a subspace.

Theorem 5.12 *Let V be an inner product space and W be a finite-dimensional subspace of V. Then for any $v \in V$, $v \in W^\perp$ if and only if $\langle v, w_i\rangle = 0$ for all $w_i \in B$, where B is a basis for W.*

Proof Let $B = \{w_1, w_2, \ldots, w_k\}$ be a basis for W. Then for $w \in W$, there exists scalars $\lambda_1, \lambda_2, \ldots, \lambda_k$ such that $w = \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k$. Then for any $v \in V$,

$$\begin{aligned}\langle v, w\rangle &= \langle v, \lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_k w_k\rangle \\ &= \sum_{i=1}^{k} \lambda_i \langle v, w_i\rangle\end{aligned}$$

Therefore $\langle v, w_i\rangle = 0$ for all $w_i \in B$ implies that $\langle v, w\rangle = 0$. Hence, $v \in W^{\perp}$. Conversely, suppose that $v \in W^{\perp}$. Then by the definition of orthogonal complement $\langle v, w_i\rangle = 0$ for all $w_i \in B$.

In Sect. 5.2, we have introduced the concept of projection of a vector to a one-dimensional subspace of $\mathbb{R}^2$. We have seen that a vector $v \in \mathbb{R}^2$ can be written as a sum of vectors, $(u.v)u \in span\{u\}$ where u is a unit vector and $v - (u.v)u$ which is orthogonal to $(u.v)u$. That is, $v - (u.v)u$ is an element of $span\{u\}^{\perp}$. The vector $(u.v)u$ is called the projection of v on $span\{u\}$. We will extend this result to any finite-dimensional subspace W of an inner product space V. We will proceed by considering an orthonormal basis $\{w_1, w_2, \ldots, w_k\}$ for W, projecting $v \in V$ on each one-dimensional subspace $span\{w_i\}$ of W and taking the sum. That is, the projection of $v \in V$ on W will be $w = \sum_{i=1}^{k} \langle v, w_i\rangle w_i$.

Theorem 5.13 *Let V be an inner product space and W be a finite-dimensional subspace of V. Then for any $v \in V$, there exist unique vectors $w \in W$ and $\tilde{w} \in W^{\perp}$ such that $v = w + \tilde{w}$. Furthermore, $w \in W$ is the unique vector that has the shortest distance from v.*

Proof Let $B = \{w_1, w_2, \ldots, w_k\}$ be an orthonormal basis for W and consider $w = \sum_{i=0}^{k} \langle v, w_i\rangle w_i \in W$. Take $\tilde{w} = v - w$. Then for any $w_j \in B$,

$$\begin{aligned}\langle \tilde{w}, w_j\rangle &= \left\langle v - \sum_{i=0}^{k} \langle v, w_i\rangle w_i, w_j \right\rangle \\ &= \langle v, w_j\rangle - \sum_{i=1}^{k} \langle v, w_i\rangle \langle w_i, w_j\rangle \\ &= \langle v, w_j\rangle - \langle v, w_j\rangle = 0\end{aligned}$$

That is, $\langle \tilde{w}, w_j\rangle = 0$ for all $w_j \in B$. Then by Theorem 5.12, $\tilde{w} \in W^{\perp}$. Also, $v = w + \tilde{w}$. To prove the uniqueness of w and $\tilde{w}$ suppose that $v = w + \tilde{w} = u + \tilde{u}$ where $u \in W$ and $\tilde{u} \in W^{\perp}$. This implies that $v = w - u = \tilde{u} - \tilde{w}$. Then as $w - u \in W$ and $\tilde{u} - \tilde{w} \in W^{\perp}$, $v \in W \cap W^{\perp} = \{0\}$. Hence, $w = u$ and $\tilde{w} = \tilde{u}$.

Now we have to prove that $w = \sum_{i=1}^{k} \langle v, w_i\rangle w_i$ in W is the unique vector that has the shortest distance from v. Now for any $w' \in W$,

$$\left\|v - w'\right\|^2 = \left\|w + \tilde{w} - w'\right\|^2 = \left\|(w - w') + \tilde{w}\right\|^2$$

As $w - w' \in W$ and $\tilde{w} \in W^{\perp}$, by *Pythagoras theorem*,

$$\left\|v - w'\right\|^2 = \left\|(w - w')\right\|^2 + \|\tilde{w}\|^2 \geq \|\tilde{w}\|^2 = \|v - w\|^2$$

Thus for any $w' \in W$, we get $\left\|v - w'\right\| \geq \|\tilde{w}\|^2 = \|v - w\|$.

Corollary 5.3 *Let V be an inner product space and W be a finite-dimensional subspace of V. Then $V = W \oplus W^{\perp}$.*

Proof From the above theorem, clearly $V = W + W^{\perp}$. Also, $W \cap W^{\perp} = \{0\}$. Then by Theorem 2.20, $V = W \oplus W^{\perp}$.

The above decomposition is called the *orthogonal decomposition* of V with respect to the subspace W. In general, W can be any closed subspace of V.

Definition 5.17 (*Orthogonal Projection*) Let V be an inner product space and W be a finite-dimensional subspace of V. Then the orthogonal projection π_W of V onto W is the function $\pi_W(v) = w$, where $v = w + \tilde{w}$ is the orthogonal decomposition of v with respect to W.

Example 5.30 Consider $\mathbb{R}^3$ over $\mathbb{R}$ with standard inner product. Let

$$W = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 = 0\}$$

That is, the *yz-plane*. Consider the vector $v_1 = (2, 4, 5) \in \mathbb{R}^3$. Now we will find the projection of v on W. Clearly $\{(0, 1, 0), (0, 0, 1)\}$ is an orthonormal basis for W. Then the projection of v_1 on W is given by

$$\pi_W(v_1) = \langle(2, 4, 5), (0, 1, 0)\rangle(0, 1, 0) + \langle(2, 4, 5), (0, 0, 1)\rangle(0, 0, 1) = (0, 4, 5)$$

For an arbitrary vector $v = (a, b, c) \in \mathbb{R}^3$

$$\pi_W(v) = \langle(a, b, c), (0, 1, 0)\rangle(0, 1, 0) + \langle(a, b, c), (0, 0, 1)\rangle(0, 0, 1) = (0, b, c)$$

Also observe that $W^{\perp} = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_2 = v_3 = 0\}$, i.e., the *x-axis* and hence $(a, b, c) = (0, b, c) + (a, 0, 0)$ is the orthogonal decomposition of v with respect to W.

Example 5.31 Consider $\mathbb{P}_2[-1, 1]$. Let $W = \{a + bx \mid a, b \in \mathbb{R}\}$. Clearly W is a subspace of $\mathbb{P}_2[0, 1]$ and we have already seen that $\left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x\right\}$ is an orthonormal basis for W. Consider the element $v = x^2 + 2x + 3 \in \mathbb{P}_2[-1, 1]$. Then from Example 5.25,

$$\left\langle\frac{1}{\sqrt{2}}, x^2 + 2x + 3\right\rangle = \frac{10\sqrt{2}}{3} \text{ and } \left\langle\sqrt{\frac{3}{2}}x, x^2 + 2x + 3\right\rangle = \frac{2\sqrt{6}}{3}$$

Therefore the projection of v on W is $\pi_W(v) = \dfrac{10}{3} + 2x$.

Now we will discuss some of the important properties of projection map in the following theorem.

Theorem 5.14 *Let W be a finite-dimensional subspace of an inner product space V and let π_W be the orthogonal projection of V onto W. Then*

(a) π_W *is linear.*
(b) $\mathcal{R}(\pi_W) = W$ *and* $\mathcal{N}(\pi_W) = W^\perp$
(c) $\pi_W^2 = \pi_W$

Proof (a) Let $v_1, v_2 \in V$. Then by Theorem 5.13, there exists unique vectors $w_1, w_2 \in W$ and $\tilde{w_1}, \tilde{w_2} \in W^\perp$ such that $v_1 = w_1 + \tilde{w_1}$ and $v_2 = w_2 + \tilde{w_2}$. Then $\pi_W(v_1) = w_1$ and $\pi_W(v_2) = w_2$. Now for $\lambda \in \mathbb{K}$,

$$\lambda v_1 + v_2 = \lambda(w_1 + \tilde{w_1}) + (w_2 + \tilde{w_2}) = (\lambda w_1 + w_2) + (\lambda \tilde{w_1} + \tilde{w_2})$$

where $\lambda w_1 + w_2 \in W$ and $\lambda \tilde{w_1} + \tilde{w_2} \in W^\perp$ as W and $W^\perp$ are subspaces of V. Therefore

$$\pi_W(\lambda v_1 + v_2) = \lambda w_1 + w_2 = \lambda \pi_W(v_1) + \pi_W(v_2)$$

therefore, π_W is linear.

(b) From Theorem 5.13, we have $V = W \oplus W^\perp$ and any vector $v \in V$ can be written as $v = \pi_W(v) + (v - \pi_W(v))$. Clearly $\mathcal{R}(\pi_W) \subseteq W$. Now we have prove the converse part. Let $w \in W$, then $\pi_W(w) = w$ as $w = w + 0 \in W + W^\perp$. Therefore $\mathcal{R}(\pi_W) = W$.
Similarly, it is clear that $\mathcal{N}(\pi_W) \subseteq W^\perp$. Now let $\tilde{w} \in W^\perp$. As $\tilde{w} = 0 + \tilde{w}$, we have $\pi_W(\tilde{w}) = 0$ and hence $\mathcal{N}(\pi_W) = W^\perp$.

(c) Take any $v \in V$. By Theorem 5.13, there exists unique vectors $w \in W$ and $\tilde{w} \in W^\perp$ such that $v = w + \tilde{w}$. Then

$$\pi_W^2(v) = \pi_W(\pi_W(v)) = \pi_W(w) = w = \pi_W(v)$$

Therefore $\pi_W^2 = \pi_W$.

In Theorem 5.13, we decomposed V as the direct sum of two subspaces where one is the orthogonal complement of the other. There may exist decompositions of V as the direct sum of two subspaces where one subspace is not the orthogonal complement of the other. For example, consider $\mathbb{R}^3$. Let $W_1 = span\{(1, 0, 0), (0, 1, 0)\}$ and $W_2 = span\{(1, 1, 1)\}$. Observe that $V = W_1 \oplus W_2$ and $W_1 \not\perp W_2$. In such cases also we can define a linear map.

Theorem 5.15 *Let V be an inner product space and W_1, W_2 be subspaces of V with $V = W_1 \oplus W_2$. Then the map P defined by $P(v) = w_1$, where $v = w_1 + w_2$ is the unique representation of $v \in V$ is linear.*

Proof Similar to the proof of Theorem 5.14(a).

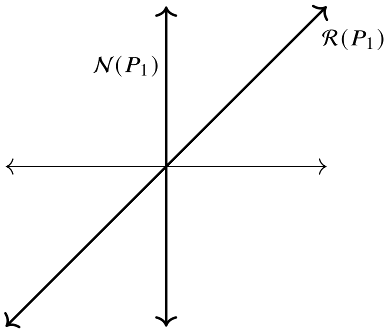


Fig. 5.16 Observe that $\mathcal{R}(P_1) \not\perp \mathcal{N}(P_1)$. Therefore P_1 is not an orthogonal projection

The above defined map P is called projection map. Observe that an orthogonal projection map is a projection map P with $[\mathcal{R}(P)]^{\perp} = \mathcal{N}(P)$.

Example 5.32 Consider $\mathbb{R}^2$ over $\mathbb{R}$ with standard inner product. Let $P_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear map defined by

$$P_1(v_1, v_2) = (v_1, 0)$$

Observe that $\mathcal{R}(P_1)$ is the straight line $y = x$ and $\mathcal{N}(P_1)$ is the $y-$ axis. Clearly, $\mathbb{R}^2 = \mathcal{R}(P_1) \oplus \mathcal{N}(P_1)$. Thus P_1 is a projection but not an orthogonal projection (Fig. 5.16).

Example 5.33 Consider $\mathbb{P}_2[0, 1]$ with the inner product $\langle p, q\rangle = \int_0^1 p(x)q(x)dx$. Let $P_2 : \mathbb{P}_2[0, 1] \to \mathbb{P}_2[0, 1]$ be a linear map defined by

$$P_2(a_0 + a_1x + a_2x^2) = a_1x$$

We have $\mathcal{R}(P_2) = span\{x\}$ and $\mathcal{N}(P_2) = span\{1, x^2\}$. Observe that $\mathbb{P}_2[0, 1] = \mathcal{R}(P_2) \oplus \mathcal{N}(P_2)$, but $\mathcal{R}(P_2) \not\perp \mathcal{N}(P_2)$. Therefore P_2 is a projection but not an orthogonal projection.

The following theorem gives an algebraic method to check whether a linear operator is a projection map or not.

Theorem 5.16 *Let V be a finite-dimensional inner product space and T be a linear operator on V. Then T is a projection of V if and only if $T^2 = T$.*

Proof Suppose that T is a projection on V, then clearly $T^2 = T$ by definition. Now suppose that T is a linear operator on V such that $T^2 = T$. We will show that $V = \mathcal{R}(T) \oplus \mathcal{N}(T)$. Let $v \in \mathcal{R}(T) \oplus \mathcal{N}(T)$. Then there exists $\tilde{v} \in V$ such that $T(\tilde{v}) = v$. Also $T(v) = 0$. Now $T^2(\tilde{v}) = T(v) = 0 = T(\tilde{v}) = v$ as $T^2 = T$. Thus T is a projection on V.

Example 5.34 Consider the linear operators P_1 and P_2 from Examples 5.32 and 5.33 respectively. Clearly, we can see that $P_1^2 = P_1$ and $P_2^2 = P_2$.

5.5 Exercises

1. Show that $(\mathbb{R}, d)$ is a metric space, where $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by
   (a) $d(x, y) = |e^x - e^y|$ for $x, y \in \mathbb{R}$.
   (b) $d(x, y) = \frac{|x-y|}{1+|x-y|}$ for $x, y \in \mathbb{R}$.

   Check whether d is induced by any norm on $\mathbb{R}$?
2. Let $v = (v_1, v_2, \ldots v_n) \in \mathbb{R}^n$. Show that
   (a) $\|v\|_\infty = max\{|v_1|, \ldots, |v_n|\}$ defines a norm on $\mathbb{R}^n$ called *infinity norm.*
   (b) for $p \geq 1$, $\|v\|_p = \left(\sum_{i=1}^n |v_i|^p\right)^{\frac{1}{p}}$ defines a norm on $\mathbb{R}^n$ called *p-norm.*
3. Show that the following functions define a norm on $\mathbb{M}_{m\times n}(\mathbb{R})$. Let $A = \left[a_{ij}\right] \in \mathbb{M}_{m\times n}(\mathbb{R})$.
   (a) $\|A\|_1 = \max\limits_{1\leq j\leq n} \sum_{i=1}^m |a_{ij}|$
   (b) $\|A\|_\infty = \max\limits_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|$
   (c) $\|A\|_2 = \sqrt{\lambda_{max}(A^T A)}$, where λ_{max} denotes the highest eigenvalue of A.
4. Show that in a finite-dimensional space V every norm defined on it are equivalent.
5. Show that every finite-dimensional normed linear space is complete.
6. Show that
   (a) $\|v\|_p = \left(\sum_{i=1}^\infty |v_i|^p\right)^{\frac{1}{p}}$ defines a norm on l^p.
   (b) $\|v\|_\infty = \sup\limits_{i\in\mathbb{N}} |v_i|$ defines a norm on l^∞.
   (c) for $1 \leq p < r < \infty$, $l^p \subset l^r$. Also $l^p \subset l^\infty$.
7. Show that the following collections

$$
\begin{aligned}
c &= \{v = (v_1, v_2, \ldots) \in l^\infty \mid v_i \to \lambda \in \mathbb{K} \text{ as } i \to \infty\} \\
c_0 &= \{v = (v_1, v_2, \ldots) \in l^\infty \mid v_i \to 0 \text{ as } i \to \infty\} \\
c_{00} &= \{v = (v_1, v_2, \ldots) \in l^\infty \mid \text{all but finitely many } v_i's \text{ are equal to } 0\}
\end{aligned}
$$

   are subspaces of l^∞.
8. Show that c, c_0 are complete, whereas c_{00} is not complete with respect to the norm defined on l^∞.
9. Let V be a vector space over a field $\mathbb{K}$. A set $B \subset V$ is a **Hamel basis** for V if $span(B) = V$ and any finite subset of B is linearly independent. Show that if $(V, \|.\|)$ is an infinite-dimensional Banach space with a Hamel basis B, then B is uncountable. (**Hint**: Use *Baire's Category theorem.*)
10. Let $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$. Check whether the following defines an inner product on $\mathbb{R}^2$ or not.
    (a) $\langle u, v\rangle = v_1(u_1 + 2u_2) + v_2(2u_1 + 5v_2)$
    (b) $\langle u, v\rangle = v_1(2u_1 + u_2) + v_2(u_1 + v_2)$

11. Show that $\langle z_1, z_2\rangle = Re(z_1\overline{z_2})$ defines an inner product on $\mathbb{C}$, where $Re(z)$ denotes the real part of the complex number $z = a + ib$.
12. Show that $\langle A, B\rangle = Tr\,(B^*A)$ defines an inner product on $\mathbb{M}_{m\times n}\,(\mathbb{K})$.
13. Prove or disprove:

    (a) The sequence spaces l^p with $p \neq 2$ are not inner product spaces.
    (b) l^2 with $\langle u, v\rangle = \sum_{i=1}^{\infty} u_i\overline{v_i}$, where $u = (u_1, u_2, \ldots)$, $v = (v_1, v_2, \ldots) \in l^2$ is a Hilbert space.

14. Let $(V, \langle , \rangle)$ be an inner product space . Then show that for all $u, v \in V$

$$\langle u, v\rangle = \frac{1}{4}\left[\langle u+v, u+v\rangle - \langle u-v, u-v\rangle\right]$$

if $\mathbb{K} = \mathbb{R}$. Also show that if $\mathbb{K} = \mathbb{C}$, we have

$$\langle u, v\rangle = \frac{1}{4}\left[\langle u+v, u+v\rangle - \langle u-v, u-v\rangle + i\langle u+iv, u+iv\rangle - i\langle u-iv, u-iv\rangle\right]$$

15. Show that in an inner product space V, $u_n \to u$ and $v_n \to v$ implies that $\langle u_n, v_n\rangle \to \langle u, v\rangle$.
16. Show that l^p with $\langle u, v\rangle = \sum_{n=1}^{\infty} u_n v_n$ is a Hilbert space.
17. Let V be an inner product space with an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$. Then for any $v \in V$, show that $\|v\|^2 = \sum_{i=1}^{n}|\langle v, v_i\rangle|^2$.
18. (**Bessel's Inequality**) Let S be a countable orthonormal set in an inner product space V. Then for every $v \in V$, show that $\sum_{u_i\in S}|\langle v, u_i\rangle|^2 \leq \|v\|^2$.
19. Let S be an orthonormal set in an inner product space V. Then for every $v \in V$, show that the set $S_v = \{u \in S \mid \langle v, u\rangle = 0\}$ is a countable set. (**Hint**: Use *Bessel's Inequality*)
20. Construct an orthonormal basis using *Gram–Schmidt orthonormalization* process

    (a) for $\mathbb{R}^3$ with standard inner product, using the basis $\left\{\begin{bmatrix}1\\2\\2\end{bmatrix}, \begin{bmatrix}-1\\0\\2\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$
    (b) for $\mathbb{P}_3[0, 1]$ with $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$, using the basis $\{1, x, x^2\}$

21. Show that, for $A \in \mathbb{M}_{n\times n}\,(\mathbb{R})$, $AA^T = I$ if and only if the rows of A form an orthonormal basis for $\mathbb{R}^n$.
22. Consider $\mathbb{R}^2$ with standard inner product. Find $S^{\perp}$, when S is

    (a) $\{u\}$, where $u = (u_1, u_2) \neq 0$
    (b) $\{u, v\}$, where u,v are two linearly independent vectors.

23. Let S_1, S_2 be two non-empty subsets of an inner product space V, with $S_1 \subset S_2$. Then show that
(a) $S_1 \subset S_1^{\perp\perp}$ (b) $S_2^{\perp} \subset S_1^{\perp}$ (c) $S_1^{\perp\perp\perp} = S_1^{\perp}$

24. Let $S = \{(3, 5, -1)\} \subset \mathbb{R}^3$.
    (a) Find an orthonormal basis B for $S^\perp$.
    (b) Find the projection of $(2, 3, -1)$ onto $S^\perp$.
    (c) Extend B to an orthonormal basis of $\mathbb{R}^3$.
25. Let V be a finite-dimensional inner product space. Let W_1, W_2 be subspaces of V. Then show that
    (a) $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$
    (b) $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$
26. Prove or disprove: Let W be any subspace of $\mathbb{R}^n$ and let $S \subset \mathbb{R}^n$ spans W. Consider a matrix A with elements of S as columns. Then $W^\perp = ker(A)$.
27. Find the orthogonal projection of the given vector v onto the given subspace W of an inner product space V.
    (a) $v = (1, 2)$, $W = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 + x_2 = 0\}$
    (b) $v = (3, 1, 2)$, $W = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_3 = 2x_1 + x_2\}$
    (c) $v = 1 + 2x + x^2$, $W = \{a_0 + a_1 x + a_2 x^2 \in \mathbb{P}_2[0, 1] \mid a_2 = 0\}$
28. Let V be an inner product space and W be a finite-dimensional subspace of V. If T is an orthogonal projection of V onto W, then $I - T$ is the orthogonal projection of V onto $W^\perp$.
29. Consider $C[-1, 1]$ with the inner product $\langle f, g \rangle = \int_{-1}^{1} f(s)g(s)ds$, for all $f, g \in C[0, 1]$. Let W be the subspace of $C[0, 1]$ spanned by $\{x + 1, x^2 + x\}$.
    (a) Find an orthonormal basis for $span\,(W)$.
    (b) What will be the projection of x^3 onto $span\,(W)$?
30. Show that a bounded linear operator on a Hilbert space V is an orthogonal projection if and only if P is self-adjoint and P is idempotent($P^2 = P$).

Solved Questions related to this chapter are provided in Chap. 11.

Chapter 6
Bounded Linear Maps

In this chapter, the exploration of advanced linear algebra and functional analysis unfolds the notion of bounded linear maps, which elegantly combine linearity and boundedness, crucial in various mathematical applications. The concept of the adjoint operator which is introduced, enabling the study of self-adjoint, normal, and unitary operators, each possessing for distinct properties and widespread utility. Singular value decomposition (SVD) emerges as a powerful factorization method, revolutionizing linear equation solving. When standard matrix inverses do not exist, generalized inverses, such as the Moore–Penrose inverse, provide a flexible structure for solving systems of linear equations, enabling least square solutions to otherwise ill-posed problems in a number of mathematical and practical situations. Banach contraction principle offers a profound insight into mappings on metric spaces, underpinning algorithms across disciplines. Lastly, iterative methods, including Gauss–Jacobi and Gauss–Seidel, are introduced for solving linear systems, catering to large-scale numerical problems.

6.1 Bounded Linear Maps

As a linear map is a function, the question of continuity arises naturally. Because every normed space is a metric space, the definition of a continuous function in a normed space follows from Definition 1.28. We have seen that a function f from a metric space (X, d_1) to a metric space (Y, d_2) is said to be continuous at a point $x_0 \in X$, if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$d_2(f(x), f(x_0)) < \epsilon \text{ whenever } d_1(x, x_0) < \delta$$

Then, by Theorem 5.1, we have the following definition for continuity of a linear operator.

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_6

Definition 6.1 (*Continuous Linear Operator*) Let V and W be normed spaces and $T \in \mathcal{L}(V, W)$ be a linear operator, then T is said to be continuous at $v_0 \in V$, if for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$\|Tv - Tv_0\| < \epsilon \text{ whenever } \|v - v_0\| < \delta$$

T is said to be continuous on V if T is continuous at every $v \in V$.

Now that we have defined the continuity of a linear map, we will give some alternate characterizations of continuity in the following theorem.

Theorem 6.1 *Let V and W be normed linear spaces and let $T \in \mathcal{L}(V, W)$. Then the following are equivalent:*

(a) T is continuous.
(b) T is continuous at 0.
(c) There exist $\lambda > 0$ such that $\|Tv\| \leq \lambda \|v\|$ for all $v \in V$.

Proof The implication $(a) \Rightarrow (b)$ is obvious. To prove $(b) \Rightarrow (c)$, suppose that T is continuous at 0. Then by Definition 6.1, for $\epsilon = 1$, there exists $\delta > 0$ such that $\|Tv\| < 1$ when $\|v\| < \delta$. Now take $u = \frac{\delta}{2}\frac{v}{\|v\|}$ for $v \neq 0$. Then $\|u\| = \frac{\delta}{2} < \delta$ and by continuity of T at 0,

$$T(u) = T\left(\frac{\delta}{2}\frac{v}{\|v\|}\right) = \frac{\delta}{2\|v\|}T(v) < 1$$

which implies that $\|T(v)\| \leq \frac{2}{\delta}\|v\|$ for all $v \in V \setminus \{0\}$. Also $\|T(0)\| = 0 \leq \frac{2}{\delta}\|0\|$. Take $\lambda = \frac{2}{\delta}$. Then $\|Tv\| \leq \lambda \|v\|$ for all $v \in V$.

Now to prove $(c) \Rightarrow (a)$, suppose that such a λ exists. Since T is a linear transformation

$$\|T(u) - T(v)\| = \|T(u - v)\| \leq \lambda \|u - v\|$$

for all $u, v \in V$. Let $\epsilon > 0$ and $\delta = \frac{\epsilon}{\lambda}$. Then for $u, v \in V$ with $\|u - v\| < \delta$, we have

$$\|T(u) - T(v)\| \leq \lambda \|u - v\| < \lambda\left(\frac{\epsilon}{\lambda}\right) = \epsilon$$

Therefore, T is continuous.

As every normed space is a metric space, sequential continuity can also be considered as an alternative criterion for the continuity of a linear map. Let T be a linear map between two normed linear spaces V and W. T is sequentially continuous if $\{v_n\}$ is a sequence in V with $v_n \to v$, then $T(v_n) \to T(v)$.

Example 6.1 Consider $C[0, 1]$ with supremum norm. Define $T : C[0, 1] \to C[0, 1]$ by $T(f) = \int_0^1 f(x)dx$. We have already seen that T is linear (See, Example 3.5). Now

$$\begin{aligned}\|T(f)\| &= \max_{x\in[0,1]}\left|\int_0^1 f(x)dx\right| \\ &\le \max_{x\in[0,1]}\int_0^1 |f(x)|dx \\ &\le \max_{x\in[0,1]}|f(x)|\int_0^1 dx \\ &= \|f\|\end{aligned}$$

for all $f \in C[0,1]$. Hence, the integral operator is a continuous linear operator.

Example 6.2 Consider $C[0,1]$ with supremum norm. Define $T : C[0,1] \to C[0,1]$ by $T(f(x)) = \frac{d}{x}(f(x))$. We have already seen that T is linear (See, Example 3.4). Consider $f_n(x) = x^n$. Then

$$\|T(f_n)\| = \max_{x\in[0,1]}|nx^{n-1}| = n \max_{x\in[0,1]}|x^{n-1}| = n$$

As $n \to \infty$, $\|T(f_n)\| \to \infty$. That is, there does not exists $\lambda > 0$ such that $\|T(f)\| \le \lambda \|f\|$ for all $f \in C[0,1]$. Therefore differential operator is an example of a linear operator which is not continuous.

Theorem 6.2 *Let V and W be normed spaces where V is finite- dimensional and $T \in \mathcal{L}(V, W)$. Then T is continuous.*

Proof Suppose that V is finite-dimensional. We will show that for any sequence $\{v_n\}$ with $v_n \to v$ in V, $T(v_n) \to T(v)$. Then by Theorem 1.6, T is continuous. As V is finite-dimensional, it has a finite basis, say $\{v_1, v_2, \ldots, v_n\}$. Thus for each $n \in \mathbb{N}$, $v_n \in V$ can be represented as

$$v_n = \lambda_1^n v_1 + \lambda_2^n v_2 + \cdots + \lambda_m^n v_m$$

where $\lambda_i^n \in \mathbb{K}, i = 1, 2, \ldots, m$. Also, as $v \in V$, $v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_m v_m$, where $\lambda_i \in \mathbb{K}, i = 1, 2, \ldots, m$. Now $v_n \to v$ implies that $\lambda_i^n \to \lambda_i$ for each $i = 1, 2, \ldots, m$. Therefore

$$\begin{aligned}T(v_n) &= T\left(\lambda_1^n v_1 + \lambda_2^n v_2 + \cdots + \lambda_m^n v_m\right) \\ &= \lambda_1^n T(v_1) + \lambda_2^n T(v_2) + \cdots + \lambda_m^n T(v_m) \\ &\to \lambda_1 T(v_1) + \lambda_2 T(v_2) + \cdots + \lambda_m T(v_m) = T(v)\end{aligned}$$

by linearity of T. Therefore, T is continuous.

Consider condition (c) in Theorem 6.1. Always keep in mind that two different norms are used. One is from V and the other from W. For simplicity, no distinction is made. Maps satisfying this condition are of great importance in the field of Mathematics.

Definition 6.2 (*Bounded Linear Operator*) Let V and W be normed spaces and $T \in \mathcal{L}(V, W)$, then T is said to be bounded, if there exists $\lambda \in \mathbb{R}$ such that $\|Tv\| \le \lambda \|v\|$.

By Theorem 6.1, for a linear transformation the term continuous and bounded can be used interchangeably. Bounded linear transformations essentially maps a bounded set to a bounded set. Note that a bounded linear map does not demand a bounded range set. For example, consider $T : \mathbb{R} \to \mathbb{R}$ defined by $T(v) = v$. As $\|Tv\| = \|v\|$ for all $v \in V$, T is a bounded linear operator, but the range set is $\mathbb{R}$, which is unbounded.

Now observe that the definition of a bounded linear operator can also be stated as follows. A linear operator between the normed spaces V and W is bounded if and only if there exists a real number λ such that $\frac{\|Tv\|}{\|v\|} \leq \lambda$ for all $v \in V \setminus \{0\}$. That is, if and only if the set $\left\{\frac{\|Tv\|}{\|v\|} \mid v \in V \setminus \{0\}\right\}$ is bounded. Now consider the supremum of this set. We will show that this gives us a norm on the vector space of all bounded linear operators from V to W, denoted by $\mathcal{B}(V, W)$.

Theorem 6.3 *Let V and W be normed spaces and $T : V \to W$ be a bounded linear operator. Then*

$$\|T\| = \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|}{\|v\|} \tag{6.1}$$

defines a norm on $\mathcal{B}(V, W)$ called as operator norm.

Proof Consider the norm defined as above.

(N1) Clearly $\|T\| \geq 0$ for all $T \in \mathcal{B}(V, W)$ as we are taking supremum of the set $\left\{\frac{\|Tv\|}{\|v\|} \mid v \in V \setminus \{0\}\right\}$, which contains non-negative elements only. Also

$$\begin{aligned} \|T\| = 0 &\Leftrightarrow \sup_{v \in V \setminus \{0\}} \frac{\|Tv\|}{\|v\|} = 0 \\ &\Leftrightarrow \|Tv\| = 0 \; \forall \, v \in V \; (\because \; T(0) = 0) \\ &\Leftrightarrow T = \mathbf{0} \end{aligned}$$

where $\mathbf{0}$ is the zero operator on V. Thus $(N1)$ is satisfied.

(N2) Now for any $\lambda \in \mathbb{K}$,

$$\begin{aligned} \|\lambda T\| &= \sup_{v \in V \setminus \{0\}} \frac{\|(\lambda T)(v)\|}{\|v\|} \\ &= \sup_{v \in V \setminus \{0\}} \frac{\|\lambda(Tv)\|}{\|v\|} \\ &= \sup_{v \in V \setminus \{0\}} \frac{|\lambda| \, \|Tv\|}{\|v\|} \\ &= |\lambda| \, \|T\| \end{aligned}$$

Thus $(N2)$ is satisfied.

(N3) Now for any $T_1, T_2 \in \mathcal{B}(V, W)$,

$$\begin{aligned}\|T_1 + T_2\| &= \sup_{v \in V\setminus\{0\}} \frac{\|(T_1 + T_2)(v)\|}{\|v\|} \\ &\leq \sup_{v \in V\setminus\{0\}} \left\{ \frac{\|T_1 v\|}{\|v\|} + \frac{\|T_2 v\|}{\|v\|} \right\} \\ &\leq \sup_{v \in V\setminus\{0\}} \frac{\|T_1 v\|}{\|v\|} + \sup_{v \in V\setminus\{0\}} \frac{\|T_2 v\|}{\|v\|} \\ &= \|T_1\| + \|T_2\|\end{aligned}$$

Thus $(N3)$ is also satisfied.

Therefore $\mathcal{B}(V, W)$ is a normed linear space with the norm defined as in (6.1)

So if $T \in \mathcal{B}(V, W)$ is a bounded linear operator, we can write,

$$\|Tv\| \leq \|T\| \, \|v\|$$

for all $v \in V$.

Remark 6.1 For $T \in \mathcal{B}(V, W)$, the norm defined as

$$\|T\| = \sup_{v \in V, \|v\| \leq 1} \|Tv\|$$

gives an alternating way of calculating operator norm.

Example 6.3 Consider $C[0, 1]$ with supremum norm. Define $T : C[0, 1] \to C[0, 1]$ as in Example 6.1. We have proved that $\|T(f)\| \leq \|f\|$ for all $f \in C[0, 1]$. This implies that $\|T\| \leq 1$. Now for the function $f(x) = 1$, $\|T(f)\| = 1$. Therefore $\|T\| = 1$.

We have seen that every linear map from a finite-dimensional space is bounded. Now we will give an example for a linear map which is not bounded, otherwise called as an *unbounded linear map*. As boundedness and continuity are synonyms for linear operators, the differential operator on $C[0, 1]$ is an example of an unbounded operator.

In Chap. 4, we have defined eigenvalues and eigenvectors for square matrices. We have already seen that a $n \times n$ matrix is nothing but a linear operator from a finite-dimensional space to itself. The same definition can be used for linear operators on infinite-dimensional spaces.

Definition 6.3 Let $T : V \to V$ be a linear operator on a vector space V over a field $\mathbb{K}$. A vector $v \neq 0$ is said to be an eigenvector of T if there exists $\lambda \in \mathbb{K}$, such that $T(v) = \lambda v$.

The set of all eigenvalues of T is called as the *eigen spectrum* of T. There are other sets of scalars that are closely related to operators and are beyond the scope of this book.

6.2 Adjoint of a Bounded Linear Map

Let T be a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. We know that there exists a corresponding matrix $A = [a_{ij}]_{m\times n} \in M_{m\times n}(\mathbb{R})$ such that

$$T(v) = \left[\sum_{j=1}^n a_{1j}v_j \ \sum_{j=1}^n a_{2j}v_j \ \dots \ \sum_{j=1}^n a_{mj}v_j\right]$$

Then

$$\begin{aligned}\langle Av, w\rangle_{\mathbb{R}^m} &= \left\langle \left[\sum_{j=1}^n a_{1j}v_j \ \sum_{j=1}^n a_{2j}v_j \ \dots \ \sum_{j=1}^n a_{mj}v_j\right], \left[w_1, w_2, \dots w_m\right]\right\rangle_{\mathbb{R}^m} \\ &= \left(\sum_{i=1}^m a_{i1}w_i\right)v_1 + \left(\sum_{i=1}^m a_{i2}w_i\right)v_2 + \dots + \left(\sum_{i=1}^m a_{in}w_i\right)v_n \\ &= \left\langle \left[v_1, v_2, \dots v_n\right], \left[\sum_{i=1}^m a_{i1}w_i \ \sum_{i=1}^m a_{i2}w_i \ \dots \ \sum_{i=1}^m a_{in}w_i\right]\right\rangle_{\mathbb{R}^n} \\ &= \langle v, A^T w\rangle_{\mathbb{R}^n}\end{aligned}$$

Therefore we can say that for every linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, there exists a linear map T^* from $\mathbb{R}^m$ to $\mathbb{R}^n$ with corresponding matrix A^T. If $\mathbb{C}$ is considered instead of $\mathbb{R}$, the matrix corresponding to T^* will be A^*. In this section, we will generalize this idea to abstract spaces.

Let us start by establishing a connection between a Hilbert space and its dual space. The following theorem asserts a one-one correspondence between an element in a Hilbert space and its corresponding dual space.

Theorem 6.4 (Riesz Theorem) *Let V be a Hilbert space and $f : V \to \mathbb{K}$ be a linear transformation, then f can be represented in terms of*

$$f(v) = \langle v, w\rangle$$

where w is uniquely determined by f and has norm $\|w\| = \|f\|$.

Proof If $f = 0$, take $w = 0$. Clearly, all conditions of the theorem are satisfied. Now if $f \neq 0$, $\mathcal{N}(f) \neq V$ and hence by Theorem 5.13, $\mathcal{N}(f)^\perp \neq \{0\}$. Then there exists at least one element in $\mathcal{N}(f)^\perp$ say $w_0 \neq 0$. Now take,

$$\tilde{w} = f(v)w_0 - f(w_0)\,v$$

where $v \in V$. Then

$$f(\tilde{w}) = f(v)f(w_0) - f(w_0)\,f(v) = 0$$

which implies that $\tilde{w} \in \mathcal{N}(f)$. Since $w_0 \in \mathcal{N}(f)^\perp$, we have

$$0 = \langle \tilde{w}, w_0\rangle = \langle f(v)w_0 - f(w_0)\,v, w_0\rangle = f(v)\langle w_0, w_0\rangle - f(w_0)\langle v, w_0\rangle$$

As $\langle w_0, w_0 \rangle = \|w_0\| \neq 0$, we get

$$f(v) = \frac{f(w_0)}{\langle w_0, w_0 \rangle} \langle v, w_0 \rangle$$

If we take $w = \frac{\overline{f(w_0)}}{\langle w_0, w_0 \rangle} w_0$, we get $f(v) = \langle v, w \rangle$ for any $v \in V$. Now to prove that such an element $w \in V$ is unique, suppose that there exists $w_1 \in V$ such that

$$f(v) = \langle v, w \rangle = \langle v, w_1 \rangle$$

for all $v \in V$. Then by Theorem 5.4(d), we have $w_1 = w$.
Now to prove that $\|w\| = \|f\|$. We have

$$\|w\|^2 = \langle w, w \rangle = f(w) \leq \|f\| \, \|w\|$$

Therefore $\|w\| \leq \|f\|$. Also

$$\|f\| = \sup_{v \in V, \|v\| \leq 1} |f(v)| = \sup_{v \in V, \|v\| \leq 1} |\langle v, w \rangle| \leq \|w\|$$

by *Schwarz* inequality. Thus $\|w\| = \|f\|$.

Example 6.4 Let $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f(v_1, v_2, v_3) = v_1 + 2v_2 + 3v_3$. Clearly f is a linear map. Now consider the standard ordered basis $B = \{e_1, e_2, e_3\}$ for $\mathbb{R}^3$ which is orthonormal. Take $\overline{v} = f(e_1)e_1 + f(e_2)e_2 + f(e_3)e_3 = (1, 2, 3)$. Then $\langle (v_1, v_2, v_3), (1, 2, 3) \rangle = v_1 + 2v_2 + 3v_3 = T(v_1, v_2, v_3)$.

Definition 6.4 (*Sesquilinear function*) Let V and W be vector spaces over the same field $\mathbb{K}$. A sesquilinear function is a mapping $f : V \times W \to \mathbb{K}$ such that for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$

$$f(\lambda v_1 + v_2, w) = \lambda f(v_1, w) + f(v_2, w)$$

and

$$f(v, \lambda w_1 + w_2) = \overline{\lambda} f(v, w_1) + f(v, w_2)$$

In other words, we can say that $f : V \times W \to \mathbb{K}$ is sesquilinear if it is linear in first variable and conjugate linear in second variable. Clearly, the inner product is a *sesquilinear function.*

Definition 6.5 (*Boundedness*) Let V and W be normed spaces over the same field $\mathbb{K}$ and let f be a sesquilinear mapping from $V \times W$. Then f is said to be bounded if there exists $\lambda \in \mathbb{K}$ such that for all $v \in V$ and $w \in W$, we have $|f(v, w)| \leq \lambda \|v\| \, \|w\|$, and the number

$$\|f\| = \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|f(v,w)|}{\|v\|\,\|w\|} = \sup_{\|v\|=\|w\|=1} |f(v,w)|$$

is called norm of f.

Example 6.5 Let V be an inner product space over the field $\mathbb{K}$. We have already seen that the inner product is an example of a sesquilinear function. Now we will show that it is bounded. By *Schwarz inequality*, we have $|\langle v, w\rangle| \leq \|v\|\,\|w\|$ for all $v, w \in V$. Therefore $f(v, w) = \langle v, w\rangle$ is a bounded sesquilinear function on $V \times V$.

Now let us discuss a more general version of Theorem 6.4. We will use the following theorem to prove the existence of an adjoint map for a bounded linear map in a Hilbert space.

Theorem 6.5 (Riesz Representation Theorem) *Let V be Hilbert space and W be an inner product space over the field $\mathbb{K}$ and $f : V \times W \to \mathbb{K}$ be a sesquilinear function. Then f has a representation $f(v, w) = \langle \mathcal{F}(v), w\rangle$, where $\mathcal{F} : V \to W$ is a bounded linear map. $\mathcal{F}$ is uniquely determined by f and $\|\mathcal{F}\| = \|f\|$.*

Proof Fix $v \in V$. Define $g : W \to \mathbb{K}$ by $g(w) = \overline{f(v, w)}$. Then for all $w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$, we have

$$\begin{aligned}
g(\lambda w_1 + w_2) &= \overline{f(v, \lambda w_1 + w_2)} \\
&= \overline{\bar{\lambda} f(v, w_1) + f(v, w_2)} \\
&= \lambda \overline{f(v, w_1)} + \overline{f(v, w_2)} \\
&= \lambda g(w_1) + g(w_2)
\end{aligned}$$

Thus, g is linear. By Theorem 6.4, there exists a unique element, $\tilde{w} \in W$ such that $g(w) = \overline{f(v, w)} = \langle w, \tilde{w}\rangle$. Therefore $f(v, w) = \langle \tilde{w}, w\rangle$. Clearly $\tilde{w}$ depends on $v \in V$. Now using this fact, define $\mathcal{F} : V \to W$ by $\mathcal{F}(v) = \tilde{w}$. Now for $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$, we have

$$\begin{aligned}
\langle \mathcal{F}(\lambda v_1 + v_2), w\rangle &= f(\lambda v_1 + v_2, w) \\
&= \lambda f(v_1, w) + f(v_2, w) \\
&= \lambda \langle \mathcal{F}(v_1), w\rangle + \langle \mathcal{F}(v_2), w\rangle \\
&= \langle \lambda \mathcal{F}(v_1) + \mathcal{F}(v_2), w\rangle
\end{aligned}$$

which is true for all $w \in W$. Then by Theorem 5.4(d), $\mathcal{F}(\lambda v_1 + v_2) = \lambda \mathcal{F}(v_1) + \mathcal{F}(v_2)$ for all $v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$. Therefore $\mathcal{F}$ is linear. Now we have to prove that $\|\mathcal{F}\| = \|f\|$. If $f = 0$, then $\mathcal{F} = 0$ and $\|\mathcal{F}\| = \|f\|$. Otherwise,

$$\begin{aligned}\|f\| &= \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|f(v,w)|}{\|v\|\,\|w\|}\\ &= \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|\langle \mathcal{F}(v), w\rangle|}{\|v\|\,\|w\|}\\ &\geq \sup_{v\neq 0\in V, \mathcal{F}(v)\neq 0\in W} \frac{|\langle \mathcal{F}(v), \mathcal{F}(v)\rangle|}{\|v\|\,\|\mathcal{F}(v)\|}\\ &= \sup_{v\neq 0\in V} \frac{\|\mathcal{F}(v)\|}{\|v\|} = \|\mathcal{F}\|\end{aligned}$$

That is, $\|\mathcal{F}\| \leq \|f\|$. Thus $\mathcal{F}$ is bounded. By Schwarz inequality, we have $|\langle \mathcal{F}(v), w\rangle| \leq \|\mathcal{F}(v)\|\,\|w\|$. Using this, we have

$$\|f\| = \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|\langle \mathcal{F}(v), w\rangle|}{\|v\|\,\|w\|} \leq \sup_{v\neq 0\in V, w\neq 0\in W} \frac{\|\mathcal{F}(v)\|\,\|w\|}{\|v\|\,\|w\|} = \|\mathcal{F}\|$$

Thus $\|\mathcal{F}\| = \|f\|$. Now to prove the uniqueness part, suppose that there exists $\mathcal{F}_1 : V \to W$ such that

$$f(v,w) = \langle \mathcal{F}(v), w\rangle = \langle \mathcal{F}_1(v), w\rangle$$

for all $v \in V$ and $w \in W$. Then by Theorem 5.4(d), $\mathcal{F}(v) = \mathcal{F}_1(v)$ for all $v \in V$. Thus $\mathcal{F} = \mathcal{F}_1$.

Now we will show that every bounded linear map on a Hilbert space will have an adjoint map and it is unique. If the domain is not a Hilbert space, a bounded linear map need not have an adjoint (See Example 6.8).

Theorem 6.6 *Let V be Hilbert space and W be an inner product space over the same field $\mathbb{K}$. Let $T : V \to W$ be a bounded linear map. Then, there exists a unique bounded linear map $T^* : W \to V$ such that $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ for all $v \in V$ and $w \in W$ with $\|T\| = \|T^*\|$.*

Proof Define $f : W \times V \to \mathbb{K}$ by $f(w,v) = \langle w, T(v)\rangle$. Now for $v \in V$, $w_1, w_2 \in W$ and $\lambda \in \mathbb{K}$, we have

$$f(\lambda w_1 + w_2, v) = \langle \lambda w_1 + w_2, T(v)\rangle = \lambda\langle w_1, T(v)\rangle + \langle w_2, T(v)\rangle = \lambda f(w_1, v) + f(w_2, v)$$

and for $v_1, v_2 \in V$, $w \in W$, $\lambda \in \mathbb{K}$, we have

$$f(w, \lambda v_1 + v_2) = \langle w, T(\lambda v_1 + v_2)\rangle = \overline{\lambda}\langle w, T(v_1)\rangle + \langle w, T(v_2)\rangle = \overline{\lambda} f(w, v_1) + f(w, v_2)$$

Thus, f is a sesquilinear function. Also, by Cauchy–Schwarz inequality, we have

$$|f(w,v)| = |\langle w, T(v)\rangle| \leq \|w\|\,\|T(v)\| \leq \|T\|\,\|v\|\,\|w\|$$

Thus

$$\|f\| = \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|f(w,v)|}{\|w\|\,\|v\|} \leq \|T\|$$

Also

$$\begin{aligned}\|f\| &= \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|f(w,v)|}{\|w\|\,\|v\|} \\ &= \sup_{v\neq 0\in V, w\neq 0\in W} \frac{|\langle w, T(v)\rangle|}{\|w\|\,\|v\|} \\ &\geq \sup_{v\neq 0\in V, T(v)\neq 0\in W} \frac{|\langle T(v), T(v)\rangle|}{\|T(v)\|\,\|v\|} = \|T\|\end{aligned}$$

That is, $\|f\| = \|T\|$. Thus f is a bounded sesquilinear function. Then by Theorem 6.5, there exists a unique bounded linear map T^* from W to V with $f(w, v) = \langle w, T(v)\rangle = \langle T^*(w), v\rangle$ and $\|f\| = \|T\| = \|T^*\|$. Taking conjugates, we get $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ for all $v \in V$ and $w \in W$.

Definition 6.6 (*Adjoint of a Bounded Linear Map*) Let V be Hilbert space and W be an inner product space over the same field $\mathbb{K}$. and $T : V \to W$ be a bounded linear map. Then the linear map $T^* : W \to V$ satisfying $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ for all $v \in V$ and $w \in W$ is called adjoint of T.

Remark 6.2 Let V be a Hilbert space and W be an inner product space over the same field $\mathbb{K}$ and $T : V \to W$ be a bounded linear map and $T^* : W \to V$ be its adjoint, then $\langle T^*(w), v\rangle = \langle w, T(v)\rangle$ for all $v \in V$ and $w \in W$. For,

$$\langle T^*(w), v\rangle = \overline{\langle v, T^*(w)\rangle} = \overline{\langle T(v), w\rangle} = \langle w, T(v)\rangle$$

is true for all $v \in V$ and $w \in W$. Also, if W is a Hilbert space, as

$$\langle v, T(w)\rangle = \langle T^*(v), w\rangle = \langle v, \left(T^*\right)^*(w)\rangle$$

by Theorem 5.4(d), we get $(T^*)^* = T$.

Example 6.6 Let V be a Hilbert space over the field $\mathbb{K}$. Define the identity operator $I : V \to V$, $I(v) = v$ for all $v \in V$ and the zero operator, $O : V \to V$, by $O(v) = 0$. Then for all $v, w \in V$,

$$\langle v, I^*(w)\rangle = \langle I(v), w\rangle = \langle v, w\rangle = \langle v, I(v)\rangle$$

and

$$\langle v, O^*(w)\rangle = \langle O(v), w\rangle = \langle 0, w\rangle = 0 = \langle v, 0\rangle = \langle v, O(w)\rangle$$

Thus $I^* = I$ and $O^* = O$.

Example 6.7 Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by

$$T(v_1, v_2, \ldots, v_n) = (0, v_1, v_2, \ldots, v_{n-1})$$

Then for $v = (v_1, v_2, \ldots, v_n)$, $w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$, we have

$$\begin{aligned}\langle T(v), w\rangle &= \langle (0, v_1, v_2, \ldots, v_{n-1}), (w_1, w_2, \ldots, w_n)\rangle \\ &= v_1w_2 + v_2w_3 + \cdots + v_{n-1}w_n \\ &= v_1w_2 + v_2w_3 + \cdots + v_{n-1}w_n + v_n.0 \\ &= \langle (v_1, v_2, \ldots, v_n), (w_2, w_3, \ldots, w_n, 0)\rangle\end{aligned}$$

Define $T^*(w) = (w_2, w_3, \ldots, w_n, 0)$. Clearly T^* is linear and $\langle T(v), w\rangle = \langle v, T^*(w)\rangle$ for all $v, w \in V$. T and T^* are called **right** and **left shift operators**, respectively. Observe that the matrices of these operators are given by $[T] = \begin{bmatrix} 0\ 0 \ldots 0\ 0 \\ 1\ 0 \ldots 0\ 0 \\ 0\ 1 \ldots 0\ 0 \\ \vdots\ \vdots \ddots \vdots\ \vdots \\ 0\ 0 \ldots 1\ 0 \end{bmatrix}$ and $[T^*] = \begin{bmatrix} 0\ 1\ 0 \ldots 0 \\ 0\ 0\ 1 \ldots 0 \\ \vdots\ \vdots\ \vdots \ddots \vdots \\ 0\ 0\ 0 \ldots 1 \\ 0\ 0\ 0 \ldots 0 \end{bmatrix}$, respectively. Clearly, $[T^*] = [T]^T$ as we have seen in the introduction.

For a bounded linear map between two general inner product spaces, the existence of an adjoint is not assured. But if we are considering the bounded linear map from a Hilbert space to an inner product space, there definitely exists an adjoint.

Example 6.8 Consider c_{00}, the linear space of all real sequences having only a finite number of non-zero terms with the inner product $\langle v, w\rangle = \sum_{n=1}^{\infty} v_n w_n$. Then c_{00} is an incomplete space with respect to the given inner product (Why?). Define $T : c_{00} \to c_{00}$ by

$$T(v) = \left(\sum_{n=1}^{\infty} \frac{v_n}{n}, 0, 0, \ldots\right)$$

Clearly, T is linear (verify!). Also,

$$\|T\| \leq \sqrt{\sum_{n=1}^{\infty} \frac{1}{n^2}} = \frac{\pi}{\sqrt{6}}$$

Thus T is a bounded linear operator on c_{00}. We will show that there does not exists an operator $\tilde{T}$ such that $\langle T(v), w\rangle = \langle v, \tilde{T}(w)\rangle$. Consider the sequence $\{e_n\}$, where $e_n = (0, 0, \ldots, 0, 1, 0)$ (1 is in the nth position and all other entries are zero). Clearly, $\{e_n\} \in c_{00}$ (is it?). If $(\tilde{T}(v))_n$ denotes the element in the nth position of $\tilde{T}(v)$, we have

$$(\tilde{T}(e_1))_n = \langle e_n, \tilde{T}(e_1)\rangle = \langle T(e_n), e_1\rangle = \frac{1}{n}$$

Thus $\tilde{T}(e_1) = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right)$. This is not possible as $\tilde{T}(e_1)$ must be an element of c_{00}. Therefore T does not have an adjoint operator on c_{00}. What if we define T on l^2 with the given inner product? Remember that l^2 is complete with respect to the given inner product (See, Exercise 13, Chap. 5).

Now, let us discuss some properties of the adjoint of a bounded linear map.

Theorem 6.7 *Let V be a Hilbert space and W be an inner product space over the same field $\mathbb{K}$. Let $T, \tilde{T} : V \to W$ be bounded linear maps. Then*

(a) $(\lambda T + \tilde{T})^* = \overline{\lambda} T^* + \tilde{T}^*$, *where* $\lambda \in \mathbb{K}$.
(b) $\left(T\tilde{T}\right)^* = \tilde{T}^* T^*$.
(c) $\|T^*T\| = \|TT^*\| = \|T\|^2$.
(d) $T^*T = 0$ *if and only if* $T = 0$.

Proof (a) For all $v, w \in V$ and $\lambda \in \mathbb{K}$, we have

$$\begin{aligned}\langle(\lambda T + \tilde{T})(v), w\rangle &= \langle \lambda T(v) + \tilde{T}(v), w\rangle \\ &= \lambda\langle T(v), w\rangle + \langle \tilde{T}(v), w\rangle \\ &= \lambda\langle v, T^*(w)\rangle + \langle v, \tilde{T}^*(w)\rangle \\ &= \langle v, \overline{\lambda} T^*(w) + \tilde{T}^*(w)\rangle\end{aligned}$$

As the adjoint of a linear operator is unique, we get $(\lambda T + \tilde{T})^* = \overline{\lambda} T^* + \tilde{T}^*$, where $\lambda \in \mathbb{K}$.

(b) For all $v, w \in V$

$$\langle (T\tilde{T})(v), w\rangle = \langle T\left(\tilde{T}(v)\right), w\rangle = \langle \tilde{T}(v), T^*(w)\rangle = \langle v, \tilde{T}^* T^*(w)\rangle$$

Therefore $\left(T\tilde{T}\right)^* = \tilde{T}^* T^*$.

(c) We have $T^*T : V \to V$. Now,

$$\left\|T^*T(v)\right\| \leq \left\|T^*\right\| \|Tv\| \leq \left\|T^*\right\| \|T\| \|v\| = \|T\|^2 \|v\|$$

for all $v \in V$. Thus $\|T^*T\| \leq \|T\|^2$. Also

$$\|T(v)\|^2 = \langle Tv, Tv\rangle = \langle\left(T^*T\right)(v), v\rangle \leq \left\|\left(T^*T\right)(v)\right\| \|v\| \leq \left\|T^*T\right\| \|v\|^2$$

Taking supremum over all $v \in V$ with $\|v\| = 1$, we get $\|T\|^2 \leq \|T^*T\|$. Thus $\|T^*T\| = \|T\|^2$. Similarly, we can prove that $\|TT^*\| = \|T\|^2$.

(d) Suppose that $T = 0$. That is, $T(v) = 0$ for all $v \in V$. Then $(T^*T)(v) = T^*(T(v)) = 0$ for all $v \in V$ and hence $T^*T = 0$. Conversely, suppose that $T^*T = 0$. Then from (c), we have $\|T^*T\| = \|T\|^2 = 0$. Therefore $T = 0$.

Now we will use the concept of adjoint operators to define some special class of bounded linear operators.

6.3 Self-adjoint Operators

In this section, we will study a special class of bounded linear maps which are of great importance in applications of linear operator theory and are defined using the adjoint of a linear operator.

Definition 6.7 (*Self-adjoint Operators*) Let H be a Hilbert space and $T : H \to H$ be a bounded linear operator. Then T is self-adjoint if $T = T^*$.

Theorem 6.8 *Let $T_1, T_2 : H \to H$ be self-adjoint operators on a Hilbert space H. Then $T_1 + T_2$ is self-adjoint and T_1T_2 is self-adjoint if and only if T_1 and T_2 commute.*

Proof Let $T_1, T_2 : H \to H$ be self-adjoint operators. Then for all $v, w \in H$, we have

$$\begin{aligned}\langle (T_1 + T_2)(v), w\rangle &= \langle T_1(v), w\rangle + \langle T_2(v), w\rangle \\ &= \langle v, T_1(w)\rangle + \langle v, T_2(w)\rangle \\ &= \langle v, (T_1 + T_2)(w)\rangle\end{aligned}$$

Thus $T_1 + T_2$ is self-adjoint. Also

$$\langle (T_1T_2)(v), w\rangle = \langle T_2(v), T_1^*(w)\rangle = \langle T_2(v), T_1(w)\rangle = \langle v, (T_2T_1)(w)\rangle$$

Thus $(T_1T_2)^* = T_2T_1 \Leftrightarrow T_2T_1 = T_1T_2$.

Example 6.9 Let V be a Hilbert space. From Example 6.6, the identity operator and the zero operator on V are self-adjoint operators.

Example 6.10 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T(v_1, v_2) = (2v_1 + v_2, v_1 + 3v_2)$$

Then for all $(v_1, v_2), (u_1, u_2) \in \mathbb{R}^2$, we have

$$\begin{aligned}\langle T(v_1, v_2), (u_1, u_2)\rangle &= \langle (2v_1 + v_2, v_1 + 3v_2), (u_1, u_2)\rangle \\ &= (2v_1 + v_2)u_1 + (v_1 + 3v_2)u_2 \\ &= v_1(2u_1 + u_2) + v_2(u_1 + 3u_2) \\ &= \langle (v_1, v_2), T^*(u_1, u_2)\rangle\end{aligned}$$

That is, $T^*(v_1, v_2) = (2v_1 + v_2, v_1 + 3v_2) = T(v_1, v_2)$ for all $(v_1, v_2) \in \mathbb{R}^2$. Thus T is self-adjoint.

Example 6.11 Let $T : C[0, 1] \to C[0, 1]$ be defined by $(Tf)(x) = x(f(x))$. Then for all $f, g \in C[0, 1]$, we have

$$\langle T(f), g\rangle = \int_0^1 (xf(x))\, g(x)dx = \int_0^1 f(x)\,(xg(x))\, dx = \langle f, T^*(g)\rangle$$

That is, $(T^*f)(x) = xf(x) = (Tf)(x)$ for all $f \in C[0, 1]$. Therefore T is self-adjoint.

Observe that the above linear operator does not have an eigenvalue. For, let λ be an eigenvalue of T with eigenvector $f \in C[0, 1]$. That is, $T(f(x)) = xf(x) = \lambda f(x)$ for all $x \in [0, 1]$. Then $(x - \lambda)f(x) = 0$ which implies that $f(x) = 0$ for all $x \in [0, 1]$. This is a contradiction as, an eigenvector must be a non-zero element. Thus the existence of eigenvalues, is not guaranteed even for bounded linear self-adjoint operators. Now in the following theorem, we will prove that if H is a complex Hilbert space, the eigenvalues of a bounded linear self-adjoint operator are real numbers if they exist.

Theorem 6.9 *Let $T : H \to H$ be a self-adjoint operator on a complex Hilbert space H. Then*

(a) the eigenvalues of T, if they exist, are real numbers.
(b) the eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof Let $T : H \to H$ be a self-adjoint operator on a complex Hilbert space H. That is, $\langle T(v), v\rangle = \langle v, T(v)\rangle$ for all $v \in H$.

(a) Let λ be an eigenvalue of T with eigenvector v. Then,

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle T(v), v\rangle = \langle v, T(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle$$

gives $\left(\lambda - \overline{\lambda}\right)\langle v, v\rangle = 0$. As $\langle v, v\rangle \neq 0$, we get $\lambda = \overline{\lambda}$. Hence, λ is real.

(b) Let λ_1 and λ_2 be two distinct eigenvalues of T with eigenvectors $v_1, v_2 \in H$ respectively. From (a), both λ_1 and λ_2 are real. Then

$$\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \langle T(v_1), v_2\rangle = \langle v_1, T(v_2)\rangle = \langle v_1, \lambda_2 v_2\rangle = \lambda_2\langle v_1, v_2\rangle$$

gives $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$. As $\lambda_1 \neq \lambda_2$, we get $\langle v_1, v_2\rangle = 0$. Thus, the eigenvectors corresponding to distinct eigenvalues are orthogonal.

The existence of an eigenvalue is not guaranteed in the above theorem. But the following lemma guarantees that a linear operator T on a finite-dimensional complex vector space V has at least one eigenvalue.

Lemma 6.1 *Let $T : V \to V$ be a linear operator on a finite-dimensional complex vector space V. Then T has at least one eigenvalue.*

Proof Consider a basis B for V. Let A be the matrix representation of T with respect to B. Let $p(\lambda)$ be the characteristic polynomial of A. Then $p(\lambda)$ is a polynomial of finite degree, say n. Then by *Fundamental theorem of algebra*, A has at least one eigenvalue. That is, T has at least one eigenvalue.

Now we will discuss one of the important theorem in the theory of self-adjoint operators.

Theorem 6.10 *Let $T : V \to V$ be a self-adjoint linear operator on a finite- dimensional complex vector space V. Then, there exists an orthonormal basis for V that consists only of eigenvectors of T and the matrix representation of T with respect to B is diagonal.*

Proof Let $T : V \to V$ be a self-adjoint linear operator on a finite- dimensional complex vector space V. Let S denote the span of eigenvectors of T. First, we will prove that $S^{\perp} = \{0\}$. Let $v \in S^{\perp}$ and $v_i \in S$ be an eigenvector of T corresponding to the eigenvalue λ_i. Then

$$\langle T(v), v_i \rangle = \langle v, T(v_i) \rangle = \langle v, \lambda_i v_i \rangle = \lambda_i \langle v, v_i \rangle = 0$$

Thus, for every $v \in S^{\perp}$, we have $T(v) \in S^{\perp}$. In other words $S^{\perp}$ is invariant under T. If $S^{\perp} \neq \{0\}$, so that its dimension is one or more, then, by Lemma 6.1, T has an eigenvalue λ and with $v \in S^{\perp}$ being its corresponding eigenvector. But then v, being an eigenvector, is also in S. As $S \cap S^{\perp} = \{0\}$, we get $v = 0$. Then by Corollary 5.3, $S = V$. That is, we have V as the span of eigenvectors of T. Now using, Gram–Schmidt orthonormalization process, we will get an orthonormal basis for V consisting of eigenvectors of T. Clearly, the matrix representation of T with respect to B is diagonal.

In matrix terms, the above theorem states that for a self-adjoint matrix A, there exists a matrix P, with the eigenvectors (orthonormal) of A as columns such that $PAP^{-1} = D$, where D is a diagonal matrix with the eigenvalues of A as diagonal entries. Now we will characterize self-adjoint operators in complex Hilbert spaces.

Theorem 6.11 *Let $T : H \to H$ be a bounded linear operator on a Hilbert space H. If T is self-adjoint, then $\langle T(v), v \rangle$ is real for all $v \in H$. In particular, if a complex Hilbert space is considered, $\langle T(v), v \rangle$ is real for all $v \in H$ implies T is self-adjoint.*

Proof Since T is self-adjoint, for all $v \in H$, we have

$$\overline{\langle T(v), v \rangle} = \langle v, T(v) \rangle = \langle T(v), v \rangle$$

That is, $\langle T(v), v \rangle$ is equal to its complex conjugate for all $v \in H$. Thus $\langle T(v), v \rangle$ is real for all $v \in H$.

Now, suppose that H is a complex Hilbert space. Then, for all $v, w \in H$, we have $v + iw \in H$. As $\langle T(v), v \rangle$ is real for all $v \in H$, we get $\langle v, T(v) \rangle = \langle T(v), v \rangle$ for all $v \in H$. In particular,

$$\langle T(v+iw), v+iw\rangle = \langle v+iw, T(v+iw)\rangle$$

This implies that

$$-i\langle T(v), w\rangle + i\langle T(w), v\rangle = -i\langle v, T(w)\rangle + i\langle w, T(v)\rangle$$

and hence

$$\langle T(v), w\rangle + \langle w, T(v)\rangle = \langle T(w), v\rangle + \langle v, T(w)\rangle$$

Then by using $(IP4)$, we get $Re\,(\langle T(w), v\rangle) = Re\,(\langle v, T(w)\rangle)$ for all $v, w \in V$. Now, to prove that $Im\,(\langle T(w), v\rangle) = Im\,(\langle v, T(w)\rangle)$, it is enough to consider the element $v + w \in V$ instead of $v + iw$, where $v, w \in V$. Thus we have $\langle T(w), v\rangle = \langle v, T(w)\rangle$ for all $v, w \in V$.

We have defined the norm of an operator in Sect. 6.1. Let $T : V \to W$ be a bounded linear map between two inner product spaces V and W. For $v \in V$, $|\langle T(v), w\rangle| \leq \|T(v)\|\,\|w\|$ by *Schwarz inequality* for all $w \in W$. Also, for $T(v) \neq 0$, if we take $w = \frac{T(v)}{\|T(v)\|}$, we get $\langle T(v), w\rangle = \|T(v)\|$. Thus

$$\|T(v)\| = Sup\{|\langle T(v), w\rangle| : \; w \in W, \; \|w\| \leq 1\}$$

Then $\|T\|$ can also be defined as

$$\|T\| = Sup\{|\langle T(v), w\rangle| : v \in V, \; w \in W, \|v\| \leq 1, \; \|w\| \leq 1\}$$

The following theorem shows that if we are considering a self-adjoint operator, we need not have to take the supremum of such a big set.

Theorem 6.12 *Let $T : H \to H$ be a bounded self-adjoint operator on a Hilbert space H. Then $\|T\| = Sup\{|\langle T(v), v\rangle| : v \in V, \|v\| \leq 1\}$.*

Proof Let $T : H \to H$ be a bounded self-adjoint operator on a Hilbert space H. Take

$$\alpha = Sup\{|\langle T(v), w\rangle| : v, w \in H, \|v\| \leq 1, \; \|w\| \leq 1\}$$

and

$$\beta = Sup\{|\langle T(v), v\rangle| : v \in H, \|v\| \leq 1\}$$

Clearly, $\beta \leq \alpha$. Now, observe that for every $v, w \in V$, we have

$$\langle T(v+w), v+w\rangle = \langle T(v), v\rangle + \langle T(v), w\rangle + \langle T(w), v\rangle + \langle T(w), w\rangle$$

$$\langle T(v-w), v-w\rangle = \langle T(v), v\rangle - \langle T(v), w\rangle - \langle T(w), v\rangle + \langle T(w), w\rangle$$

Then

$$\begin{aligned}\langle T(v+w), v+w\rangle - \langle T(v-w), v-w\rangle &= 2\langle T(v), w\rangle + 2\langle T(w), v\rangle \\ &= 2\langle T(v), w\rangle + 2\langle w, T(v)\rangle \\ &= 2\langle T(v), w\rangle + 2\overline{\langle T(v), w\rangle} \\ &= 4Re\,(\langle T(v), w\rangle)\end{aligned}$$

As $\frac{v+w}{\|v+w\|}$ has norm 1, we have

$$|\langle T(v+w), v+w\rangle| \le \alpha \|v+w\|^2$$

Similarly,

$$|\langle T(v-w), v-w\rangle| \le \alpha \|v-w\|^2$$

Thus we have

$$\begin{aligned}4Re\,(\langle T(v), w\rangle) &\le |\langle T(v+w), v+w\rangle| + |\langle T(v-w), v-w\rangle| \\ &\le \alpha \left(\|v+w\|^2 + \|v-w\|^2\right) \\ &\le 2\alpha \left(\|v\|^2 + \|w\|^2\right)\end{aligned}$$

by using Parallelogram law. Now for $v, w \in H$ with $\|v\|, \|w\| \le 1$, the above inequality implies that $4Re\,(\langle T(v), w\rangle) \le 4\alpha$. Now, if $T(v) \ne 0$, take $w = \frac{T(v)}{\|T(v)\|}$, then $\|w\| = 1$ and $\|T(v)\| = \langle T(v), w\rangle = Re\,(\langle T(v), w\rangle)$ so that $\beta \le \alpha$. Therefore, we have $\|T\| = Sup\{|\langle T(v), v\rangle| : v \in V, \|v\| \le 1\}$.

Corollary 6.1 *Let $T : H \to H$ be a bounded self-adjoint operator on a Hilbert space H. Then $T = 0$ if and only if $\langle T(v), v\rangle = 0$ for all $v \in H$.*

Proof Let $T : H \to H$ be a bounded self-adjoint operator on a HIlbert space H. Then, we have $T = 0$ if and only if $\|T\| = Sup\{|\langle T(v), v\rangle| : v \in V, \|v\| \le 1\} = 0$. That is, if and only if $\langle T(v), v\rangle = 0$ for all $v \in H$.

This is an important consequence of Theorem 6.12, which will be used later to characterize normal operators.

Positive Operators

Theorem 6.11 establishes an important property of self-adjoint operators, which will be used for further classification of self- adjoint operators. In this section, we will be classifying self-adjoint operators on finite-dimensional inner product spaces based on the first part of Theorem 6.11 which will give as an important class of self-adjoint operators.

Definition 6.8 (*Positive Definite*) Let $T : V \to V$ be a self-adjoint linear operator on a finite- dimensional inner product space V, then T is called positive definite operator if $\langle T(v), v\rangle > 0$ for all non-zero $v \in V$. T is called positive semi-definite operator if $\langle T(v), v\rangle \geq 0$ for all non-zero $v \in V$.

Example 6.12 Let V be a Hilbert space and I be the identity operator on V. In Example 6.9, we have seen that I is self-adjoint. Also,

$$\langle I(v), v\rangle = \langle v, v\rangle = \|v\|^2 > 0$$

for all non-zero $v \in V$. Therefore I is positive definite.

Example 6.13 Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T_1(v_1, v_2) = (v_1, 2v_2)$. Clearly, T_1 is a self- adjoint linear operator on $\mathbb{R}^2$(Verify). Also, $\langle T_1(v_1, v_2), (v_1, v_2)\rangle = v_1^2 + 2v_2^2 > 0$ for all non-zero $(v_1, v_2) \in \mathbb{R}^2$. Thus, T_1 is positive definite.

Now define $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T_2(v_1, v_2) = (0, 2v_2)$. Clearly, T_2 is a self-adjoint linear operator on $\mathbb{R}^2$(Verify). Observe that, $\langle T_2(v_1, 0), (v_1, 0)\rangle = 0$ for all $v_1 \in \mathbb{R}$ and $\langle T_2(v_1, v_2), (v_1, v_2)\rangle = 2v_2^2 \geq 0$ otherwise. Hence, T_2 is positive semi-definite.

In Sect. 5.4, we have defined orthogonal projection operator on an inner product space projecting a vector to a finite- dimensional subspace. In the next example, we will show that orthogonal projection operators are positive operators.

Example 6.14 Let V be a finite-dimensional vector space and W be a subspace of V. By Theorem 5.13, for any $v \in V$, there exists $w \in W$ and $\tilde{w} \in W^{\perp}$ such that $v = w + \tilde{w}$. We define $T : V \to V$ by $\pi_W(v) = w$. Let $v_1 = w_1 + \tilde{w_1}, v_2 = w_2 + \tilde{w_2} \in V$. Then,

$$\langle \pi_W(v_1), v_2\rangle = \langle w_1, w_2 + \tilde{w_2}\rangle = \langle w_1, w_2\rangle = \langle w_1 + \tilde{w_1}, w_2\rangle = \langle v_1, \pi_W(v_2)\rangle$$

Thus, π_W is self-adjoint. Now for every non-zero $v = w + \tilde{w} \in V$, we have

$$\langle \pi_W(v), v\rangle = \langle w, w + \tilde{w}\rangle = \langle w, w\rangle > 0$$

Therefore π_W is positive definite.

Now, we will characterize positive operators on finite-dimensional inner product spaces.

Theorem 6.13 *Let $T : V \to V$ be a self-adjoint linear operator on a finite- dimensional inner product space V. Then T is positive definite if and only if all the eigenvalues of T are positive. Similarly, T is positive semi-definite if and only if all the eigenvalues of T are non-negative.*

Proof Suppose that $T : V \to V$ is positive definite. Let λ be an eigenvalue of T with eigenvector v. As T is positive definite, we have

$$\langle T(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle > 0$$

Clearly, $\lambda > 0$. Now let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of T without counting multiplicity. Suppose that they are all positive. As T is self-adjoint, by Theorem 6.10, we know that V has a basis consisting of eigenvectors of T, say $B = \{v_1, v_2, \ldots, v_n\}$, where v_i is an eigenvector of λ_i. Then, for any $v \in V$, there exists scalars $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ such that $v = \sum_{i=1}^{n} \alpha_i v_i$. Therefore, for all $v \in V$, we have

$$\langle T(v), v\rangle = \left\langle T\left(\sum_{i=1}^{n} \alpha_i v_i\right), \sum_{i=1}^{n} \alpha_i v_i \right\rangle = \sum_{i=1}^{n} \lambda_i \langle \alpha_i v_i, \alpha_i v_i\rangle > 0$$

Thus, T is positive definite. Proof is similar for positive semi-definite operators.

In matrix terms, the above theorem means that, an $n \times n$ matrix A is positive definite (or positive semi-definite) if and only if all the eigenvalues are positive(or non-negative).

Theorem 6.14 *Let V and W finite-dimensional inner product spaces over the same field $\mathbb{K}$. Let $T : V \to W$ be a linear map. Then*

(a) T^*T and TT^* are positive semi-definite.
(b) $Rank\ (T^*T) = Rank\ (TT^*) = Rank\ (T)$

Proof We have, $T^*T : V \to V$ and $TT^* : W \to W$. As

$$\langle \left(T^*T\right)(v), v\rangle = \langle T(v), T(v)\rangle = \langle v, T^*T(v)\rangle$$

for all $v \in V$, T^*T is self-adjoint. Similarly, TT^* is also self-adjoint.

(a) Now for $v \in V$, $\langle (T^*T)(v), v\rangle = \langle T(v), T(v)\rangle \geq 0$ and for $w \in W$, $\langle (TT^*)(v), v\rangle = \langle T^*(v), T^*(v)\rangle \geq 0$. Thus both T^*T and TT^* are positive semi-definite.
(b) By Theorem 6.7, $\mathcal{N}(T) = \mathcal{N}(T^*T) = \mathcal{N}(TT^*)$. Then by Rank–Nullity theorem, $Rank\ (T^*T) = Rank\ (TT^*) = Rank\ (T)$.

We will be using the positive semi-definiteness of the operator T^*T later in this chapter.

6.4 Normal, Unitary Operators

In the previous section, we have discussed one of the important class of linear operators, called self-adjoint operators. Although the scope of this book covers linear operators on finite-dimensional spaces mostly, we will discuss two more important classes of linear operators defined using adjoint operator, namely, unitary operators and normal operators.

Definition 6.9 (*Normal, Unitary Operators*) Let H be a Hilbert space and $T : H \to H$ be a bounded linear operator. Then T is normal if $TT^* = T^*T$ and unitary if $TT^* = T^*T = I$.

Example 6.15 Consider the operator defined in Example 6.10, Observe that $TT^* = TT^* \neq I$. Thus T is normal, but not unitary. Similarly, the operator defined in Example 6.11 is normal, but not unitary.

Example 6.16 Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(v_1, v_2) = (v_2, v_1)$. Then for all (v_1, v_2), $(u_1, u_2) \in \mathbb{R}^2$, we have

$$\langle T(v_1, v_2), (u_1, u_2)\rangle = \langle (v_2, v_1), (u_1, u_2)\rangle = v_2u_1 + v_1u_2 = \langle (v_1, v_2), (u_2, u_1)\rangle$$

Thus $T^* = T$. Also $TT^* = T^*T = I$. Thus T is normal and unitary.

Remark 6.3 Observe that both self-adjoint operators and unitary operators are normal operators. But a normal operator need not be either of them. For, consider the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(v_1, v_2) = (v_1 + v_2, v_2 - v_1)$. Then for all $(v_1, v_2), (u_1, u_2) \in \mathbb{R}^2$, we have

$$\begin{aligned} \langle T(v_1, v_2), (u_1, u_2)\rangle &= \langle (v_1 + v_2, v_2 - v_1), (u_1, u_2)\rangle \\ &= (v_1 + v_2)u_1 + (v_2 - v_1)u_2 \\ &= v_1(u_1 - u_2) + v_2(u_1 + u_2) \\ &= \langle (v_1, v_2), (u_1 - u_2, u_1 + u_2)\rangle \\ &= \langle (v_1, v_2), T^*(u_1, u_2)\rangle \end{aligned}$$

Therefore $T^* : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T^*(v_1, v_2) = (v_1 - v_2, v_1 + v_2)$. Here $TT^* = 2(v_1, v_2) = T^*T$. Therefore T is normal. As $T \neq T^*$, T is not self-adjoint and as $TT^* = T^*T = 2I \neq I$, T is not unitary.

We have characterized self-adjoint operators on complex Hilbert spaces in Theorem 6.11. In the following theorem, we will characterize normal and unitary operators.

Theorem 6.15 *Let $T : H \to H$ be a bounded linear operator, where H is a Hilbert space. Then,*

(a) T is normal if and only if $\|T(v)\| = \|T^*(v)\|$ for all $v \in H$.
(b) T is unitary if and only if T is onto and $\|T(v)\| = \|v\|$ for all $v \in H$.

Proof Let T be a bounded linear operator on a Hilbert space H.

(a) Suppose that T is normal. Then for all $v \in H$, we have

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle = \langle T^*(v), T^*(v)\rangle = \left\|T^*(v)\right\|^2$$

Now, suppose that $\|T(v)\| = \|T^*(v)\|$ for all $v \in H$. Then

$$\langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle\ \forall\, v \in H$$

Thus, $\langle (T^*T - TT^*)(v), v\rangle = 0$ for all $v \in H$. Then by Corollary 6.1, we have $T^*T = TT^*$. Therefore T is normal.

(b) Suppose that T is unitary. Then as T is invertible, T is onto. Also

$$\|T(v)\|^2 = \langle T(v), T(v)\rangle = \langle T^*T(v), v\rangle = \langle v, v\rangle = \|v\|^2$$

for all $v \in H$. Now, suppose that T is onto and $\|T(v)\| = \|v\|$ for all $v \in H$. Then $\langle (T^*T)(v), v\rangle = \langle I(v), v\rangle$ for all $v \in H$. This implies that

$$\langle \left(T^*T - I\right)(v), v\rangle = 0\ \forall\, v \in H$$

As both, T^*T and I are self-adjoint operators, by Corollary 6.1, $T^*T = I$. Now, for all $v, w \in H$, as

$$\|T(v) - T(w)\| = \|T(v - w)\| = \|v - w\|$$

T is one-one. Thus T is an invertible operator and hence by the uniqueness of the inverse operator $T^{-1} = T^*$. Therefore T is normal.

Now we will check whether the sum and product of normal and unitary operators are respectively normal and unitary.

Theorem 6.16 *Let H is a Hilbert space.*

(a) Let $T_1, T_2 : H \to H$ be normal operators, then $T_1 + T_2$ and T_1T_2 is normal if T_1 commutes with T_2^* and T_1^* commutes with T_2.

(b) Let $T_1, T_2 : H \to H$ be unitary operators, then T_1T_2 is unitary and $T_1 + T_2$ is unitary if it is onto and Re $(\langle T_1(v), T_2(v)\rangle) = -\frac{1}{2}$ for every $v \in H$ with $\|v\| = 1$.

Proof (a) Let $T_1, T_2 : H \to H$ be normal operators, where H is a Hilbert space. Suppose that $T_1T_2^* = T_2^*T_1$ and $T_1^*T_2 = T_2T_1^*$. Then

$$\begin{aligned}(T_1 + T_2)(T_1 + T_2)^* &= (T_1 + T_2)\left(T_1^* + T_2^*\right)\\ &= T_1T_1^* + T_1T_2^* + T_2T_1^* + T_2T_2^*\\ &= T_1^*T_1 + T_1^*T_2 + T_2^*T_1 + T_2^*T_2\\ &= (T_1 + T_2)^*(T_1 + T_2)\end{aligned}$$

Thus $T_1 + T_2$ is normal. As,

$$(T_1T_2)\left(T_2^*T_1^*\right) = T_1\left(T_2T_2^*\right)T_1^* = \left(T_1T_2^*\right)\left(T_2T_1^*\right) = \left(T_2^*T_1\right)\left(T_1^*T_2\right) = \left(T_2^*T_1^*\right)(T_1T_2)$$

we get $(T_1T_2)(T_1T_2)^* = (T_1T_2)^*(T_1T_2)$. Thus, T_1T_2 is normal.

(b) Let $T_1, T_2 : H \to H$ be unitary operators, where H is a Hilbert space. Now for all $v \in H$, we have

$$\langle (T_1 + T_2)(v), (T_1 + T_2)(v) \rangle = \langle T_1(v), T_1(v) \rangle + \langle T_1(v), T_2(v) \rangle + \langle T_2(v), T_1(v) \rangle + \langle T_2(v), T_2(v) \rangle$$

By Theorem 6.15(b), $\langle T_1(v), T_1(v) \rangle = \langle T_2(v), T_2(v) \rangle = \langle v, v \rangle$ and $\langle T_2(v), T_1(v) \rangle = \overline{\langle T_1(v), T_2(v) \rangle}$. Therefore,

$$\langle (T_1 + T_2)(v), (T_1 + T_2)(v) \rangle = 2\langle v, v \rangle + 2Re\left(\langle T_1(v), T_2(v) \rangle\right)$$

Again by Theorem 6.15(b), we have $T_1 + T_2$ is unitary if and only if it is onto and $\langle (T_1 + T_2)(v), (T_1 + T_2)(v) \rangle = \langle v, v \rangle$. This happens only if $\langle v, v \rangle + 2Re\left(\langle T_1(v), T_2(v) \rangle\right) = 0$. Hence, $Re\left(\langle T_1(v), T_2(v) \rangle\right) = -\frac{1}{2}$ for every $v \in H$ with $\|v\| = 1$. As,

$$(T_1T_2)(T_1T_2)^* = T_1\left(T_2T_2^*\right)T_1^* = T_1IT_1^* = I$$

$$(T_1T_2)^*(T_1T_2) = T_2^*\left(T_1^*T_1\right)T_2 = T_2^*IT_2 = I$$

T_1T_2 is unitary.

Now let us discuss some of the properties regarding the eigenvalues and eigenvectors of a normal operator. Keep in mind that as both self-adjoint operators and unitary operators are normal operators, they will also possess these properties.

Theorem 6.17 *Let $T : H \to H$ be a normal operator, where H is a Hilbert space. Then,*

(a) If λ is an eigenvalue of T with eigenvector v, then $\overline{\lambda}$ is an eigenvalue of T^* with v as eigenvector.
(b) Eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof Let $T : H \to H$ be a normal operator.

(a) Let λ be an eigenvalue of T with eigenvector v. Then $(T - \lambda I)(v) = 0$. Now, by Theorem 6.15(a),

$$\|(T - \lambda I)(v)\| = \left\|(T - \lambda I)^*(v)\right\| = \left\|\left(T^* - \overline{\lambda}I\right)(v)\right\| = 0$$

Therefore, $T^*(v) = \overline{\lambda}v$.

(b) Let $v_1, v_2 \in H$ be eigenvectors corresponding to eigenvalues λ_1, λ_2 respectively, where $\lambda_1 \neq \lambda_2$. Now,

$$\lambda_1\langle v_1, v_2 \rangle = \langle \lambda_1 v_1, v_2 \rangle = \langle T(v_1), v_2 \rangle = \langle v_1, T^*(v_2) \rangle$$

Now, from (a), we have $\overline{\lambda_2}$ is an eigenvalue of T, with v_2 as an eigenvector. Therefore $\lambda_1\langle v_1, v_2 \rangle = \langle v_1, \overline{\lambda_2}v_2 \rangle = \lambda_2\langle v_1, v_2 \rangle$. Then $(\lambda_1 - \lambda_2)\langle v_1, v_2 \rangle = 0$. As $\lambda_1 \neq \lambda_2$, we get $\langle v_1, v_2 \rangle = 0$.

For general linear operators, the above results need not be true. Consider the right and left shift operators defined as in Example 6.7, observe that 0 is an eigenvalue of left shift operator with eigenvector $(1, 0, \ldots, 0) \in \mathbb{R}^n$. But 0 is not an eigenvalue of right shift operator.

6.5 Singular Value Decomposition

In earlier chapters, we have seen some decompositions of matrices which exist only for square matrices. In this section, we will generalize the concept of decomposition of matrices to general $m \times n$ matrices.

Theorem 6.18 *Let $T : V \to W$ be a linear transformation of rank r , where V and W are finite-dimensional inner product spaces. Then there exists orthonormal bases $\{v_1, v_2, \ldots, v_n\}, \{w_1, w_2, \ldots, w_m\}$ of V, W respectively and unique scalars $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ such that*

$$T(v_i) = \begin{cases} \sigma_i w_i, & \textit{if } 1 \leq i \leq r \\ 0, & \textit{if } i > r \end{cases}$$

Proof For any linear map $T : V \to W$, by Theorem 6.14, T^*T is a positive definite operator of rank r on V. Now by Theorem 6.10, there is an orthonormal basis $\{v_1, v_2, \ldots, v_n\}$ of V consisting of eigenvectors of T^*T with corresponding eigenvalues λ_i where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0$ and $\lambda_i = 0$ for $i > r$. Now, define $\sigma_i = \sqrt{\lambda_i}$ and $w_i = \frac{1}{\sigma_i} T(v_i)$ where $1 \leq i \leq r$. First, we will prove that $\{w_1, w_2, \ldots, w_r\}$ is an orthonormal subset of W. For $i \leq i, j \leq r$, we have

$$\langle w_i, w_j \rangle = \left\langle \frac{1}{\sigma_i} T(v_i), \frac{1}{\sigma_i} T(v_i) \right\rangle = \frac{1}{\sigma_i \sigma_j} \langle T^*T(v_i), v_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle \lambda_i v_i, v_j \rangle$$

As $\sigma_i = \sqrt{\lambda_i}$ and $\{v_1, v_2, \ldots, v_n\}$ is an orthonormal basis for V this implies that $\langle w_i, w_j \rangle = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$. Thus $\{w_1, w_2, \ldots, w_r\}$ is an orthonormal set in W. By Theorem 2.14 and Gram–Schmidt orthonormalization process, we can extend this set to an orthonormal basis, $\{w_1, w_2, \ldots, w_m\}$, of W. Then, $T(v_i) = \sigma_i w_i$ for $1 \leq i \leq r$. For $i > r$, we have $(T^*T)(v_i) = 0$. Then by Theorem 6.7(d), we have $T(v_i) = 0$, where $i > r$. Thus, we have

$$T(v_i) = \begin{cases} \sigma_i w_i, & \text{if } 1 \leq i \leq r \\ 0, & \text{if } i > r \end{cases}$$

Now we have to prove that the scalars $\sigma_1, \sigma_2, \ldots, \sigma_r$ are unique. For $1 \leq i \leq m$ and $1 \leq j \leq n$, we have

$$\langle T^*(w_i), v_j\rangle = \langle w_i, T(v_j)\rangle = \begin{cases} \sigma_i, & \text{if } i = j \le r \\ 0, & \text{otherwise} \end{cases}$$

And by Theorem 5.10, for any $1 \le i \le m$,

$$T^*(w_i) = \sum_{j=1}^{n} \langle T^*(w_i), v_j\rangle v_j = \begin{cases} \sigma_i v_i, & \text{if } i = j \le r \\ 0, & \text{otherwise} \end{cases}$$

Now, for $i \le r$, $(T^*T)(v_i) = T^*(\sigma_i w_i) = \sigma_i T^*(w_i) = \sigma_i^2 w_i$ and for $i > r$, $(T^*T)(v_i) = T^*(0) = 0$. Thus each v_i is an eigenvector of T^*T with corresponding eigenvalue σ_i^2 if $i \le r$ and 0 if $i > r$.

Definition 6.10 (*Singular Value Decomposition*) The unique scalars $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$ defined in Theorem 6.18 are called singular values of a linear transformation and the above decomposition is called Singular Value Decomposition. If $r < min\{m, n\}$, then $\sigma_{r+1} = \cdots = \sigma_l = 0$ are also considered as singular values, where $l = min\{m, n\}$.

Example 6.17 Consider a linear transformation $T : \mathbb{P}_2[-1, 1] \to \mathbb{P}_1[-1, 1]$ defined by $T(f) = f''$. We have already seen that $\tilde{\Phi} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2 - 1)\right\}$ is an orthonormal basis for $\mathbb{P}_2[-1, 1]$. Hence $\tilde{\Psi} = \left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x\right\}$ is an orthonormal basis for $\mathbb{P}_1[-1, 1]$. Then $A = [T]_{\tilde{\Phi}}^{\tilde{\Psi}} = \begin{bmatrix} 0 & 0 & 3\sqrt{5} \\ 0 & 0 & 0 \end{bmatrix}$ and hence $A^*A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 45 \end{bmatrix}$. Here $\lambda_1 = 45$, $\lambda_2 = 0$ and $\lambda_3 = 0$. Corresponding eigenvectors are $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$. Observe here that $r = 1 < min\{2, 3\}$. Therefore the singular values are $\sigma_1 = 3\sqrt{5}$ and $\sigma_2 = 0$. Translating this to the given context, we get $v_1 = \sqrt{\frac{5}{8}}(3x^2 - 1)$, $v_2 = \sqrt{\frac{3}{2}}x$ and $v_3 = \frac{1}{\sqrt{2}}$. Then $w_1 = \frac{1}{\sigma_1}T(v_1) = \frac{1}{\sqrt{2}}$. Take $w_2 = \sqrt{\frac{3}{2}}x$. $\{w_1, w_2\}$ forms an orthonormal basis for $\mathbb{P}_1[-1, 1]$.

From Theorem 6.18, it is clear that the scalars $\sigma_1, \sigma_2, \ldots, \sigma_r$ are determined uniquely by T. Now, we know that, for every linear transformation $T : V \to W$, there exists a corresponding matrix $A = [T]_{\tilde{\Phi}}^{\tilde{\Psi}}$, where $\tilde{\Phi}$ is a basis of V and $\tilde{\Psi}$ is a basis for W. Take Φ as the matrix with the above orthonormal basis elements $v_1, v_2, \ldots, v_n$ as columns. That is, the matrix with eigenvectors of A^*A as columns in the decreasing order of their corresponding eigenvalues. Take Ψ as the matrix with the orthonormal basis elements $w_1, w_2, \ldots, w_m$ of as columns, where $w_i = \frac{1}{\sigma_i}Av_i$. Then in matrix terms, the above theorem can be stated as follows.

Theorem 6.19 *Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times m$ unitary matrix Ψ and an $n \times n$ unitary matrix Φ such that $A = \Psi\Sigma\Phi^*$, where*

$$\Sigma_{ij} = \begin{cases} \sigma_i, \ if\ i = j \leq r \\ 0, \ otherwise \end{cases}$$

and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ *are scalars such that* $\sigma_i = \sqrt{\lambda_i}$ *where* $\lambda_i, i = 1, 2, \ldots, r$ *are the non-zero eigenvalues of* A^*A.

Example 6.18 Consider a matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Then, $A^*A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, which has eigenvalues $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0$. Observe here that the corresponding eigenvectors are $v_1 = (1, 0, 0)$, $v_2 = (0, 1, 0)$ and $v_3 = (0, 0, 1)$ respectively. Thus $\Phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The non-zero singular values are $\sigma_1 = \sigma_2 = 1$. Also, $u_1 = \frac{1}{\sigma_1} Av_1 = (1, 0)$ and $u_2 = \frac{1}{\sigma_2} Av_2 = (0, 1)$. Therefore $\Psi = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Now, observe that

$$\Psi \Sigma \Phi^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = A$$

Example 6.19 Consider the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 2 & 2 \end{bmatrix}$. We have $A^*A = \begin{bmatrix} 14 & 10 \\ 10 & 14 \end{bmatrix}$, which has eigenvalues $\lambda_1 = 24$ and $\lambda_2 = 4$ with eigenvectors $v_1 = \frac{1}{\sqrt{2}}(1, 1)$ and $v_2 = \frac{1}{\sqrt{2}}(1, -1)$ respectively. Then $u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{\sqrt{3}}(1, 1, 1)$ and $u_2 = \frac{1}{\sigma_2} Av_2 = \frac{1}{\sqrt{2}}$ $(-1, 1, 0)$. Choose $u_3 = \frac{1}{\sqrt{6}}(1, 1, -2)$. Observe that,

$$\Psi \Sigma \Phi^* = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} \sqrt{24} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = A$$

6.6 Generalized Inverse of a Matrix

We know that every $n \times n$ non-singular matrix A has a unique inverse denoted by A^{-1}. What about singular matrices and rectangular matrices? There is a need for partial inverse or generalized inverse for such matrices having some properties of the usual inverse in numerous mathematical problems. One of the vital applications is in solving a system of linear equations. For example, consider the system $Ax = b$. If A is non-singular, the given system has a unique solution for x given by $x = A^{-1}b$. When A is singular or rectangular, the system can have no solution or infinitely many solutions. Suppose that there exists a matrix X having the property that $AXA = A$. Take $x = Xb$. Then, we have

$$Ax = A(Xb) = AX(Ax) = (AXA)x = Ax = b \tag{6.2}$$

That is, $x = Xb$ solves the system $Ax = b$. This implies that to solve a system of linear equations, $Ax = b$, a matrix X having the property that $AXA = A$ is really useful. For a given matrix A, we characterize all matrices X having the property $AXA = A$ as the generalized inverses of A, which helps to solve the system $Ax = b$. Likewise, we may need other "relationships" of A and X to solve different problems. This will give us more restricted classes of generalized inverses. When A is non-singular, A^{-1} trivially satisfies these "relationships". One of the interesting facts to observe here is that the generalized inverse for a matrix need not be unique. Consider the following example.

Example 6.20 Let $A = \begin{bmatrix} 2 & 3 \end{bmatrix}$. A generalized inverse for A is a matrix $X = \begin{bmatrix} a \\ b \end{bmatrix}$ with $AXA = A$. We have,

$$AXA = \begin{bmatrix} 2 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \begin{bmatrix} 2 & 3 \end{bmatrix} = (2a + 3b) \begin{bmatrix} 2 & 3 \end{bmatrix} = \begin{bmatrix} 4a + 6b & 6a + 9b \end{bmatrix}$$

Now $AXA = A$ implies that any matrix $X = \begin{bmatrix} a \\ b \end{bmatrix}$ with $2a + 3b = 1$ is a generalized inverse of A.

Now, we will show that if X is a generalized inverse of A, then both AX and XA are projection matrices onto the column space of A and A^T respectively. Consider the following theorem.

Theorem 6.20 *Let $A \in \mathbb{M}_{m\times n}(\mathbb{K})$ with generalized inverse $X \in \mathbb{M}_{n\times m}(\mathbb{K})$, then $AX \in \mathbb{M}_{m\times m}(\mathbb{K})$ is an orthogonal projection onto the column space of A*

Proof First, we will prove that AX and A have the same column space. Let $w \in Im(A)$. Then there exists $v \in \mathbb{K}^n$ such that $Av = w$. We have

$$w = Av = (AXA)v = (AX)(Av)$$

Therefore $w \in Im(AX)$.

Conversely, let $w \in Im(AX)$. Then there exists $v \in \mathbb{K}^n$ such that $(AX)v = w$. Clearly, $w = A(Xv) \in Im(A)$. Thus, both A and AX have the same column space. Now to prove that AX is an orthogonal projection, it is enough to show that AX is both self-adjoint and idempotent (See, Excercise 5.30). We have

$$(AX)(AX) = (AXA)X = AX$$

and by condition 6.5, $(AX)^* = AX$. Thus, AX is an orthogonal projection onto the column space of A.

Similarly, we can prove that XA is also a projection matrix onto the column space of A^*. We have already seen that the generalized inverse of a matrix need not be unique. In 1955, Roger Penrose[1] proved that for every matrix A (square or rectangular), there exists a unique matrix X satisfying the following conditions.

$$AXA = A \tag{6.3}$$

$$XAX = X \tag{6.4}$$

$$(AX)^* = AX \tag{6.5}$$

$$(XA)^* = XA \tag{6.6}$$

For, suppose X and Y satisfy conditions (6.3)–(6.6). Then,

$$\begin{aligned} X &= XAX = X(AX)^* = XX^*A^* = XX^*(AYA)^* = X(AX)^*(AY)^* \\ &= X(AXA)Y = XAY = XA(YAY) = (XA)^*(YA)^*Y = A^*X^*A^*Y^*Y \\ &= (AXA)^*Y^*Y = A^*Y^*Y = (YA)^*Y = YAY = Y \end{aligned} \tag{6.7}$$

The conditions (6.3)–(6.6) are collectively known as *Penrose conditions*. Based on these conditions, some classifications are made for generalized inverses. Generalized inverse of a matrix A satisfying conditions (6.3)–(6.6) are named after the American mathematician *Eliakim Hastings Moore (1862–1932)* and the English mathematician *Roger Penrose (1931–)*.

Definition 6.11 (*Moore–Penrose Inverse*) Let A be a $m \times n$ matrix. If X is a matrix such that it satisfies condition (6.3), then it is a generalized inverse of A. If X satisfies both conditions (6.3) and (6.4), then it is a reflexive generalized inverse of A. If X satisfies all the four conditions, then it is the *Moore–Penrose inverse* of A, denoted by $A^\dagger$.

Example 6.21 Consider the matrix $A = \begin{bmatrix} 2 & 3 \end{bmatrix}$ as given in Example 6.20. We have seen that any element of the set

$$X = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \mid 2a + 3b = 1 \right\}$$

is a generalized inverse of A. Consider $X_1 = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix} \in X$. Then, we can observe that

[1] Penrose, R. (1955, July). A generalized inverse for matrices. In Mathematical proceedings of the Cambridge philosophical society (Vol. 51, No. 3, pp. 406–413). Cambridge University Press.

$$AX_1A = \begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}\frac{1}{2}\\0\end{bmatrix}\begin{bmatrix}2 & 3\end{bmatrix} = \begin{bmatrix}2 & 3\end{bmatrix} = A$$

$$X_1AX_1 = \begin{bmatrix}\frac{1}{2}\\0\end{bmatrix}\begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}\frac{1}{2}\\0\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\0\end{bmatrix} = X_1$$

$$(AX_1)^* = \begin{bmatrix}1\end{bmatrix} = AX_1$$

and

$$(X_1A)^* = \begin{bmatrix}1 & 0\\ \frac{3}{2} & 0\end{bmatrix} \neq \begin{bmatrix}1 & \frac{3}{2}\\ 0 & 0\end{bmatrix} = X_1A$$

That is, X_1 satisfies conditions (6.3)–(6.5) but not condition (6.6). Thus X_1 is a reflexive generalized inverse but not Moore–Penrose inverse. So, how do we find the Moore–Penrose inverse of A? Clearly, conditions (6.3) and (6.4) are trivially satisfied for any element in X. Observe here that, for a matrix $\tilde{X} = \begin{bmatrix}a\\b\end{bmatrix} \in X$ to satisfy

$$(\tilde{X}A)^* = \begin{bmatrix}2a & 2b\\ 3a & 3b\end{bmatrix} = \begin{bmatrix}2a & 3a\\ 2b & 3b\end{bmatrix}\tilde{X}A$$

a and b must be $\frac{2}{13}$ and $\frac{3}{13}$ respectively. Consider, $X_2 = \begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix} \in X$. Then,

$$X_2AX_2 = \begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix}\begin{bmatrix}2 & 3\end{bmatrix}\begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix} = \begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix} = X_2 \text{ and } (X_2A)^* = \begin{bmatrix}\frac{4}{13} & \frac{6}{13}\\ \frac{6}{13} & \frac{9}{13}\end{bmatrix} = X_2A$$

Thus $X_2 = \begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix}$ satisfies conditions (6.3)–(6.6) and is the Moore–Penrose inverse for A. That is, $A^\dagger = \begin{bmatrix}\frac{2}{13}\\ \frac{3}{13}\end{bmatrix}$.

Now, we will define a generalized inverse for an $m \times n$ matrix using singular value decomposition. Consider the following definition.

Theorem 6.21 *Let A be an $m \times n$ matrix with rank r. Let $A = \Psi\Sigma\Phi^*$ be the singular decomposition of A, where Ψ, Σ and Φ are as described in Theorem 6.19. Then the matrix, $X = \Phi\Sigma^+\Psi^*$ satisfies conditions (6.3)–(6.6), where Σ^+ is an $n \times m$ matrix with*

$$\Sigma^+_{ij} = \begin{cases}\frac{1}{\sigma_i}, & if\ i = j \leq r\\ 0, & otherwise\end{cases}$$

and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are singular values of A.

Proof We have,

$$AXA = \left(\Psi\Sigma\Phi^*\right)\left(\Phi\Sigma^+\Psi^*\right)\left(\Psi\Sigma\Phi^*\right) = \Psi\left(\Sigma\Sigma^+\right)\left(\Psi^*\Psi\right)\Sigma\Phi^* = \Psi\Sigma\Phi^* = A$$

Also,

$$XAX = \left(\Phi\Sigma^+\Psi^*\right)\left(\Psi\Sigma\Phi^*\right)\left(\Phi\Sigma^+\Psi^*\right) = \Phi\left(\Sigma^+\Sigma\right)\left(\Phi^*\Phi\right)\Sigma^+\Psi^* = \Phi\Sigma^+\Psi^* = X$$

Observe that

$$AX = \Psi\Sigma\left(\Phi^*\Phi\right)\Sigma^+\Psi^* = \Psi\left(\Sigma\Sigma^+\right)\Psi^*$$

and

$$XA = \Phi\Sigma^+\left(\Psi^*\Psi\right)\Sigma\Phi^* = \Phi\left(\Sigma^+\Sigma\right)\Phi^*$$

where $\Sigma\Sigma^+$ and $\Sigma^+\Sigma$ are diagonal matrices with 1 as first r diagonal entries and the remaining entries zero. Therefore, $(AX)^* = AX$ and $(XA)^* = XA$. This implies that $X = \Phi\Sigma^+\Psi^*$ satisfies conditions the Penreose conditions, (6.3)–(6.6).

The matrix $X = \Phi\Sigma^+\Psi^*$ is known as the *pseudo inverse* of A, and is denoted by $A^\dagger$. By (6.7), we can observe that $A^\dagger$ is unique. Obviously, if $A \in \mathbb{M}_{n\times n}(\mathbb{K})$ is an invertible matrix, then $A^\dagger = A^{-1}$.

Example 6.22 Consider the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ from Example 6.18. We have seen that the singular value decomposition of A is

$$\Psi\Sigma\Phi^* = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore, the pseudo inverse for A is

$$A^\dagger = \Phi\Sigma^+\Psi^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Example 6.23 In Example 6.19, we have seen that the singular value decomposition of the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 2 & 2 \end{bmatrix}$ is

$$\Psi\Sigma\Phi^* = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \end{bmatrix}\begin{bmatrix} \sqrt{24} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Then

$$A^{\dagger} = \Phi\Sigma^{+}\Psi^{*} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{24}} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} \end{bmatrix} = \frac{1}{12} \begin{bmatrix} -2 & 4 & 1 \\ 4 & -2 & 1 \end{bmatrix}$$

Consider a system $Ax = b$, where A is a $m \times n$ matrix and $b \in \mathbb{R}^m$.

Pseudo Inverse and System of Linear Equations

Consider the system of linear equations $Ax = b$, where $A \in \mathbb{M}_{m\times n}(\mathbb{K})$ and $b \in \mathbb{K}^n$. We have seen that three possibilities can occur while solving the system. It can have a unique solution, an infinite number of solutions, or no solutions. The system has a unique solution when A is invertible, and the solution is given by $x = A^{-1}b$. We know that, if A is invertible $A^{-1} = A^{\dagger}$, and hence the solution can also be represented as $x = A^{\dagger}b$. From (6.2), we know that $x = A^{\dagger}b$ is a solution to the given system if the system is consistent (either the system has a unique solution or has an infinite number of solutions). Now the question arises: What is the meaning of $A^{\dagger}b$ when the system does not have any solution?

Suppose that the system is inconsistent with A having rank n. That is, b lies outside the column space of A. So we will try to get as "close" as to b. As $Im(A)$ is a subspace of $\mathbb{R}^m$, by Theorem 5.13, we can consider the orthogonal projection of b to $Im(A)$ (say $\tilde{b}$) for this purpose. Since $\tilde{b}$ lies in the column space of A, we can solve $Ax = \tilde{b}$. Let $\tilde{x}$ be the solution for this equation. Then,

$$(b - A\tilde{x}) \perp Im(A)$$

This implies that

$$A^T(b - A\tilde{x}) = 0$$

and hence

$$A^T b = A^T A\tilde{x}$$

If A has rank n, then $A^T A$ is invertible. Then we get

$$\tilde{x} = (A^T A)^{-1} A^T b$$

That is, if the system $Ax = b$ is inconsistent, $\tilde{x} = (A^T A)^{-1}A^T b$ gives us a best approximate to the solution. Interesting! Now, is there any relation between A and $(A^T A)^{-1}A^T$? or between $A^{\dagger}$ and $(A^T A)^{-1}A^T$? The following theorem provides an answer.

Theorem 6.22 *Let $A \in \mathbb{M}_{m\times n}(\mathbb{K})$, $m \geq n$ be a matrix with $Rank(A) = n$. Then*

$$A^{\dagger} = \left(A^*A\right)^{-1} A^*$$

Proof As $A \in \mathbb{M}_{m\times n}(\mathbb{K})$, by *Sylvester inequality* A^*A has rank n. Therefore A^*A is invertible and $X = (A^*A)^{-1} A^*$ is well defined. To show that X is the pseudo inverse of A, it is enough to prove that X satisfies (6.3)–(6.6). We have

$$AXA = A\left(A^*A\right)^{-1} A^*A = \left(AA^{-1}\right)\left((A^*)^{-1}A^*\right) A = A$$

and

$$\begin{aligned} XAX &= \left((A^*A)^{-1}A^*\right) A(A^*A)^{-1}A^* \\ &= \left((A^*A)^{-1}A^*\right)\left(AA^{-1}\right)\left((A^*)^{-1}A^*\right) \\ &= \left(A^*A\right)^{-1} A^* = X \end{aligned}$$

Also

$$\begin{aligned} (AX)^* &= \left(A(A^*A)^{-1}A^*\right)^* = A(A^*A)^{-1}A^* = AX \\ (XA)^* &= \left((A^*A)^{-1}A^*A\right)^* = I = (A^*A)^{-1}A^*A = XA \end{aligned}$$

Thus $A^{\dagger} = (A^*A)^{-1} A^*$.

This result has significant role in solving least square problems, which will be discussed later in this chapter. Let us consider an example first.

Example 6.24 Consider the matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 1 \\ 2 & 2 \end{bmatrix}$ from Example 6.19. Clearly A has rank 2. Then by Theorem 6.22, we have

$$A^{\dagger} = \left(A^*A\right)^{-1} A^* = \frac{1}{96}\begin{bmatrix} 14 & -10 \\ -10 & 14 \end{bmatrix}\begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 2 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} -2 & 4 & 1 \\ 4 & -2 & 1 \end{bmatrix}$$

which we have obtained in Example 6.23. Thus, if the matrix A has full column rank, we can compute the psuedo inverse easily using Theorem 6.22.

Remark 6.4 If A is an $m \times n$ matrix with $m \geq n$ and $Rank(A) = n$, then the pseudo inverse of A is

$$A^{\dagger} = (A^*A)^{-1}A^*$$

and if A is an $m \times n$ matrix with $m \leq n$ and $Rank(A) = m$, then the pseudo inverse of A is

$$A^{\dagger} = A^*(AA^*)^{-1}$$

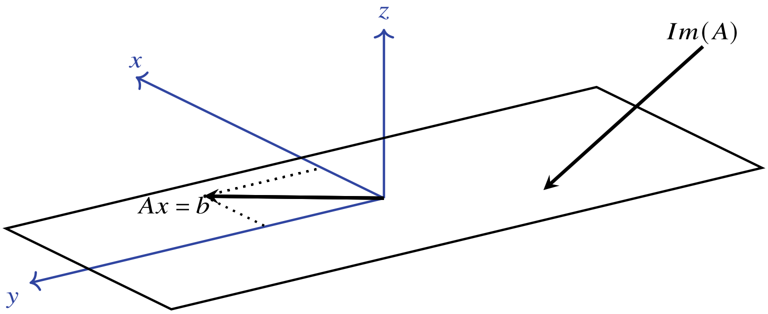


Fig. 6.1 Consistent system

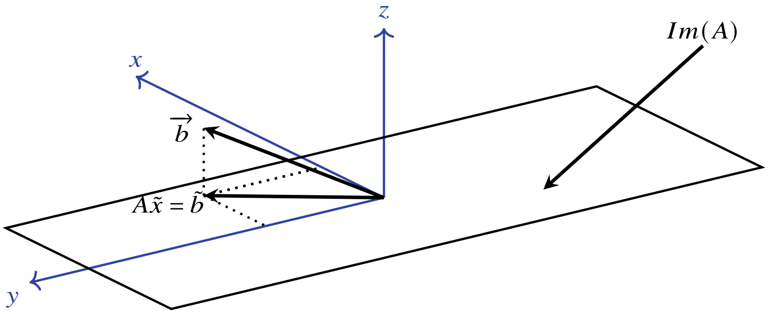


Fig. 6.2 Inconsistent system

Theorem 6.23 *Consider the system of linear equations $Ax = b$, where $A \in \mathbb{M}_{m\times n}(\mathbb{K})$ and $b \in \mathbb{K}^n$. If $\tilde{x} = A^\dagger b$, then the following statements are true.*

(i) If the system is consistent, then $\tilde{x}$ is the unique solution of the system having minimum norm (Fig. 6.1).
(ii) If the system is inconsistent, then $\tilde{x}$ is the unique best approximation of a solution having minimum norm (Fig. 6.2).

Proof (i) If the system is consistent b lies in the column space of A.
We have seen in (6.2) that, if the system is consistent then $x = Xb$ is a solution for $Ax = b$, where X is any generalized inverse of A. Thus $\tilde{x} = A^\dagger b$ is a solution of $Ax = b$. We have to show that $\|\tilde{x}\| \leq \|\hat{x}\|$ for any solution of $Ax = b$. Now suppose that $\hat{x}$ is a solution of the given system. Then

$$A^\dagger A\hat{x} = A^\dagger b = \tilde{x}$$

By Theorem 6.20, we can say that $\tilde{x}$ is the orthogonal projection of $\hat{x}$ to the column space of A. Then by Theorem 5.13, we have $\|\tilde{x}\| \leq \|\hat{x}\|$.

(ii) Suppose that the system is inconsistent. By Theorem 6.20, $A\tilde{x} = AA^\dagger b$ is the orthogonal projection of b to the column space of A. Then by Theorem 5.13, then the image of $\tilde{x}$ is the vector in $Im(A)$ "closest" to b. That is, $\tilde{x}$ is the best approximation of a solution of $Ax = b$.

To prove the uniqueness part, suppose that there exists a vector $\hat{x} \in \mathbb{K}^n$ with $A\tilde{x} = A\hat{x} = \tilde{b}$. Then,

$$A^\dagger \tilde{b} = A^\dagger A\tilde{x} = A^\dagger A A^\dagger b = A^\dagger b = \tilde{x}$$

That is, both $\tilde{x}$ and $\hat{x}$ are two independent solutions of the system $Ax = \tilde{b}$ with $\tilde{x} = A^\dagger \tilde{b}$. Then by part (i), $\|\tilde{x}\| \leq \|\hat{x}\|$.

Example 6.25 Consider the system of equations,

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 1 \\ 4x_1 + 2x_2 - 2x - 3 &= 2 \end{aligned} \tag{6.8}$$

The system can be converted into the form $Ax = b$, where $A = \begin{bmatrix} 2 & 1 & -1 \\ 4 & 2 & -2 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. We can observe the system has an infinite number of solutions. Suppose we want to find the solution of the system (6.8) with minimum norm. Here,

$$A^\dagger = \frac{1}{30}\begin{bmatrix} 2 & 4 \\ 1 & 2 \\ -1 & -2 \end{bmatrix}$$

By Theorem 6.23, we can say that

$$x = A^\dagger b = \frac{1}{30}\begin{bmatrix} 2 & 4 \\ 1 & 2 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

is the required solution.

Now, suppose that b is changed to $\tilde{b} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Then the system $Ax = \tilde{b}$ is inconsistent. Here,

$$x = A^\dagger \tilde{b} = \frac{1}{30}\begin{bmatrix} 2 & 4 \\ 1 & 2 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{1}{30}\begin{bmatrix} 14 \\ 7 \\ -7 \end{bmatrix}$$

is not a solution of the system, but it is the best approximation to a solution and having minimum norm.

Least Square Problems

A researcher is collecting data on marine food exports from India to Europe over a particular period. Given n points, suppose that the data is of the form $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ and he/she wants to represent the data using a curve. Suppose the data are plotted as points on a plane, as shown in the figure below.

From this plot, suppose that the researcher feels a linear relation exists between x_i's and y_i's. That is, there may exist a line, say $y = ax + b$, that fits the data appropriately, or we may be able to find a line that represents the data with less *error*. As we can observe from Fig. 6.3, the distance from (x_i, y_i) to $(x_i, ax_i + b)$ is $|y_i - ax_i - b|$, which is the error between the actual output and the computed output. The sum of squares of the errors for the entire data is

$$\epsilon = \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

As ϵ depends on a and b, a necessary condition for ϵ to be minimum is

$$\begin{cases} \frac{\partial \epsilon}{\partial a} = -2 \sum_{i=1}^{m} x_i (y_i - ax_i - b) = 0 \\ \frac{\partial \epsilon}{\partial b} = -2 \sum_{i=1}^{m} (y_i - ax_i - b) = 0 \end{cases} \tag{6.9}$$

This implies that

$$\begin{cases} a \sum_{i=1}^{m} x_i^2 + b \sum_{i=1}^{m} x_i = \sum_{i=1}^{m} x_i y_i \\ a \sum_{i=1}^{m} x_i + bm = \sum_{i=1}^{m} y_i \end{cases} \tag{6.10}$$

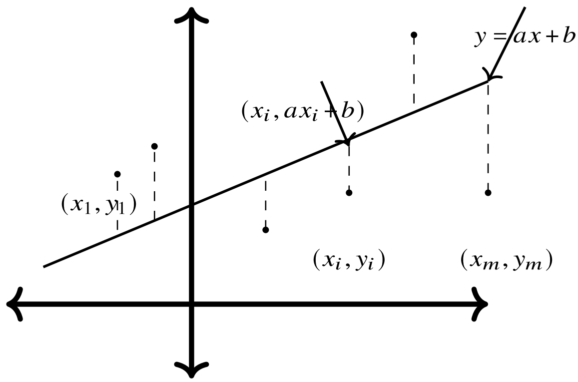


Fig. 6.3 fitting the data points on a straight line

These equations are called *normal equations* corresponding to our problem. Observe that (6.10) can be written in the compact form

$$\begin{bmatrix} \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i \\ \sum_{i=1}^{m} x_i & m \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} x_i y_i \\ \sum_{i=1}^{m} y_i \end{bmatrix} \tag{6.11}$$

We solve this system for a and b which will provide the straight line which has least deviation from the given data.

Example 6.26 Let the data collected be $(1, 4)$, $(2, 4)$, $(3, 5)$, $(4, 6)$ and $(5, 7)$. To obtain the normal equations consider the following table.

Tabular representation of data to obtain normal equations

x_i	y_i	$x_i y_i$	x_i^2
1	4	4	1
2	4	8	4
3	5	15	9
4	6	24	16
5	7	35	25
$\sum_{i=1}^{5} x_i = 15$	$\sum_{i=1}^{5} y_i = 15$	$\sum_{i=1}^{5} x_i y_i = 86$	$\sum_{i=1}^{5} x_i^2 = 55$

Then the normal equations are

$$\begin{cases} 55a + 15b = 86 \\ 15a + 5b = 26 \end{cases} \tag{6.12}$$

Solving (6.12), we get $a = 0.8$ and $b = 2.8$. Thus the line $y = 0.8x + 2.8$ best fits the given data. Also, we can compute the error, $\epsilon = 0.4$.

We can also get the least square solution by using the Moore–Penrose inverse method. Our aim was to fit the data $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$ to a line $y = ax + b$. Then, the data must satisfy the following equations:

$$\begin{aligned} y_1 &= ax_1 + b \\ y_2 &= ax_2 + b \\ &\vdots \\ y_n &= ax_n + b \end{aligned}$$

This can be written in the form $Ax = y$, where

$$A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad x = \begin{bmatrix} a \\ b \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

We know that if the number of unknowns is less than the number of equations, the system is *overdetermined*. Thus we will be finding an approximate solution in such cases. First, let us multiply both sides of the equation $Ax = y$ by A^T. Then, we have

$$(A^T A)x = A^T y$$

As x_i's are all different $Rank(A^T A) = 2$ by *Sylvester's Inequality* and hence $A^T A$ is invertible. Hence we have the approximate solution $\tilde{x} = (A^T A)^{-1} A^T y$. From the previous section, we know that $(A^T A)^{-1} A$ is the Moore–Penrose inverse of A. Consider the following example.

Example 6.27 Consider the data given in Example 6.26. Then

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} \text{ and } y = \begin{bmatrix} 4 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix}$$

Now,

$$A^T A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} = \begin{bmatrix} 55 & 15 \\ 15 & 5 \end{bmatrix}$$

and

$$\left(A^T A\right)^{-1} = \frac{1}{50} \begin{bmatrix} 5 & -15 \\ -15 & 55 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 1 & -3 \\ -3 & 11 \end{bmatrix}$$

Therefore

$$\tilde{x} = \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 1 & -3 \\ -3 & 11 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 5 \\ 6 \\ 7 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 8 \\ 28 \end{bmatrix}$$

It follows that the line $y = 0.8x + 2.8$ best fits the data as we have seen in Example 6.26. This line is called as the *least square line*.

Observe that $(A^T A)x = A^T y$ is nothing but the compact form of normal equations as given in Eq. (6.11). Also, observe that $\epsilon = \sum_{i=1}^{n} (y_i - ax_i - b)^2 = \|y - Ax\|_2^2$.

This method also can be used to fit the data to a polynomial of degree $k < n - 1$, i.e., $p(x) = a_0 + a_1 x + \cdots + a_k x^k$. Suppose we need to fit the data to a quadratic polynomial $p(x) = a_0 + a_1 x + a_2 x^2$ (i.e., $k = 2$). Then, the sum of the squared error ϵ is given by

$$\epsilon = \sum_{i=1}^{n} (y_i - p(x_i))^2$$

and the coefficients a_0, a_1 and a_2 that minimize the sum of the squared error can be obtained by solving the time equations

$$\frac{\partial \epsilon}{\partial a_0} = 0, \ \frac{\partial \epsilon}{\partial a_1} = 0, \ \frac{\partial \epsilon}{\partial a_2} = 0$$

This gives rise to the normal equations:

$$\begin{cases} a_0 m + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + a_2 \sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} x_i y_i \\ a_0 \sum_{i=1}^{n} x_i^2 + a_1 \sum_{i=1}^{n} x_i^3 + a_2 \sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} x_i^2 y_i \end{cases} \tag{6.13}$$

and in the matrix form

$$\begin{bmatrix} m & \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 \\ \sum_{i=1}^{m} x_i & \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i^3 \\ \sum_{i=1}^{m} x_i^2 & \sum_{i=1}^{m} x_i^3 & \sum_{i=1}^{m} x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} y_i \\ \sum_{i=1}^{m} x_i y_i \\ \sum_{i=1}^{m} x_i^2 y_i \end{bmatrix} \tag{6.14}$$

which can be solved for a_0, a_1 and a_2. Let $A = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{bmatrix}$. Again, observe that $(A^T A)x = A^T y$ is the normal equations given by (6.14) and the least square solution is given by the Moore–Penrose inverse as

$$x = (A^T A)^{-1} A^T y = A^{\dagger} y$$

Let us generalize the ideas we have discussed so far with the following theorem. Given an $m \times n$ matrix A, we can find $\tilde{x} \in \mathbb{R}^n$ such that $\|y - A\tilde{x}\| \leq \|y - Ax\|$ for all $x \in \mathbb{R}^n$. We can use this method to find a polynomial of degree at most $k < n - 1$, for any positive integer k that best fits the data.

Theorem 6.24 *Let $A \in \mathbb{M}_{m\times n}(\mathbb{R})$, $m \geq n$ and $y \in \mathbb{R}^m$. Then there exists $\tilde{x} \in \mathbb{R}^n$ such that $\left(A^T A\right)\tilde{x} = A^T y$ and $\|y - A\tilde{x}\| \leq \|y - Ax\|$ for all $x \in \mathbb{R}^n$. Furthermore, if $Rank(A) = n$, then $\tilde{x} = \left(A^T A\right)^{-1} A^T y$.*

Proof Let $W = \{Ax \mid x \in \mathbb{R}^n\}$. Clearly, W is a subspace of $\mathbb{R}^m$. Then by Theorem 5.13, there exists a unique vector $\tilde{w} \in W$, with $\|y - \tilde{w}\| \leq \|y - w\|$ for all $w \in W$. As $\tilde{w} \in W$, $\tilde{w} = Ax$ for some $x \in \mathbb{R}^n$. Call this vector x as our $\tilde{x}$. Then $\|y - A\tilde{x}\| \leq \|y - Ax\|$ for all $x \in \mathbb{R}^n$.

To find $\tilde{x}$, we have $y - A\tilde{x} \in W^\perp$ so that $\langle Ax, y - A\tilde{x}\rangle = \langle x, A^T(y - A\tilde{x})\rangle = 0$ for all $x \in \mathbb{R}^n$. This implies that, $A^T(y - A\tilde{x}) = 0$ and hence $A^T y = A^T A\tilde{x}$. Then, if $Rank(A) = n$, we have $\tilde{x} = \left(A^T A\right)^{-1} A^T y$.

Remark 6.5 Consider the system $Ax = y$, where $y \in \mathbb{R}^m$, A is a real $m \times n$ matrix with $m \geq n$ and $Rank(A) < n$. Then the given system has infintely many least square solutions.

Now, suppose that we need to fit the data

$$(\theta_1, y_1), (\theta_2, y_2), \ldots, (\theta_n, y_n)$$

to a trigonometric curve, say $y = asin\ \theta + bcos\ \theta$. Can we convert this to a least square problem? Yes, we can!! Observe that, if we take

$$A = \begin{bmatrix} sin\ \theta_1 & cos\ \theta_1 \\ sin\ \theta_2 & cos\ \theta_2 \\ \vdots & \\ sin\ \theta_n & cos\ \theta_n \end{bmatrix}, x = \begin{bmatrix} a \\ b \end{bmatrix}, \text{ and } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

we can convert the problem to the form $Ax = y$ and solve it for x to obtain the desired trigonometric curve that best fit the given data. Another situation where the least square problem comes handy is when we have to fit a data of the form $(x_1^1, x_2^1, x_3^1), (x_1^2, x_2^2, x_3^2), \ldots, (x_1^n, x_2^n, x_3^n)$ to a plane $x_3 = a + bx_1 + cx_2$. Here also, if we take

$$A = \begin{bmatrix} 1 & x_1^1 & x_2^1 \\ 1 & x_1^2 & x_2^2 \\ \vdots & & \\ 1 & x_1^n & x_2^n \end{bmatrix}, x = \begin{bmatrix} a \\ b \\ c \end{bmatrix}, \text{ and } y = \begin{bmatrix} x_3^1 \\ x_3^2 \\ \vdots \\ x_3^n \end{bmatrix}$$

we can convert the problem to the form $Ax = y$ and solve it for x to obtain the plane that best fits the given data. Due to their effectiveness in handling challenging data fitting and optimization problems, these adaptable approaches have applications across a wide range of fields.

6.7 Iterative Methods for System of Linear Equations

There are two approaches to solving systems of linear equations: direct methods and iterative methods. Direct approaches like *Gauss elimination method*, *LU-decomposition method*, etc. seek an exact solution in a finite number of steps, with guaranteed convergence. They are computationally demanding, especially for large systems, and necessitate a huge amount of memory. Iterative approaches, on the other hand, begin with an initial guess and repeatedly refine the solution, providing computational efficiency and lower memory requirements for big and sparse systems. Iterative approaches, on the other hand, provide approximate solutions that may need various numbers of iterations to achieve a desired degree of precision, making them ideal in situations when exact solutions are not required and computational resources are limited. Consider an equation of the form:

$$Ax = b \tag{6.15}$$

where $A = [a_{ij}]$ is $n \times n$ matrix and b is an $n \times 1$ matrix. First we will write $A = M - N$, where M is an invertible matrix. Then, (6.15) becomes of the form;

$$Mx = Nx + b$$

As M is invertible, we can write it as

$$x = M^{-1}(Nx + b) = M^{-1}Nx + M^{-1}b \tag{6.16}$$

Let us define a function $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(x) = M^{-1}Nx + M^{-1}b \tag{6.17}$$

If there exists an element $x_0 \in \mathbb{R}^n$ with $T(x_0) = x_0$, then we can observe that x_0 is a solution for the system (6.15). For,

$$\begin{aligned} T(x_0) = x_0 &\Rightarrow M^{-1}Nx_0 + M^{-1}b = x_0 \\ &\Rightarrow Nx_0 + b = Mx_0 \\ &\Rightarrow (M - N)x_0 = b \Rightarrow Ax_0 = b \end{aligned}$$

Thus we can conclude that any element in $\mathbb{R}^n$ which is mapped onto itself by T is a solution to the system (6.15). Such elements are called *fixed points* of T in $\mathbb{R}^n$. Now, let us define fixed points of functions defined on arbitrary sets. Let X be a given set and f be any function defined from X to itself. Then fixed points of f are points in X that remains unchanged under the action of f. We have the following formal definition.

Definition 6.12 (*Fixed Point*) A fixed point of a function $f : X \to X$ is a point $x \in X$ which is mapped onto itself. That is, $f(x) = x$.

Example 6.28 Consider $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ defined by $f_1(x) = -x$ and $f_2(x) = x + 1$. Then f_1 has exactly one fixed point, which is 0. For,

$$f_1(x) = x \Rightarrow x = -x \Rightarrow 2x = 0 \Rightarrow x = 0$$

Clearly, f_2 has no fixed points as it translates any number on the real line to one unit right of it.

Example 6.29 Let V be any vector space and $T : V \to V$ be any linear transformation. Then the zero element in V is a fixed point for T.

Example 6.30 Let $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T_1(x_1, x_2) = (x_1, 0)$ and $T_2(x_1, x_2) = (x_2, x_1)$. Then all the points on the x-axis are fixed points for T_1 and all the points on the line $y = x$ are fixed points for T_2.

Now, consider the system (6.15) and function T as defined in (6.17). We have seen that a fixed point of T is a solution for the system (6.15). The essential question now is whether we can guarantee the existence of such an element for the function T. Will it be unique if it exists? We can guarantee the existence of fixed points for functions having certain characteristics. In this section, we will be focusing on the existence of one particular class of functions called "contractions".

Definition 6.13 (*Contraction*) Let $(X, \|.\|)$ be a normed space. A function $f : X \to X$ is said to be a contraction on X if there exists a positive real number $\alpha < 1$ such that

$$\|f(x) - f(y)\| \leq \alpha \|x - y\|$$

In other words, the distance between any two points in the domain will always be greater than the distance between their respective images, that is, a contraction brings points closer together. Consider the following example.

Example 6.31 Consider the normed space $(\mathbb{R}, \|.\|_0)$. Define $f_1 : \mathbb{R} \to \mathbb{R}$ by $f_1(x) = \frac{1}{2}x$. Then f is a contraction with $\alpha = \frac{1}{2}$. For,

$$\|f_1(x) - f_1(y)\| = \left|\frac{1}{2}x - \frac{1}{2}y\right| = \frac{1}{2}|x - y| = \frac{1}{2}\|x - y\|$$

Now, define $f_2 : \mathbb{R} \to \mathbb{R}$ by $f_2(x) = x$. Is f_2 a contraction? Does there exists a positive real number α with $\alpha < 1$ such that $|x - y| \leq \alpha|x - y|$? Clearly, such an α with $0 < \alpha < 1$ does not exists. Thus f_2 is not a contraction.

Example 6.32 Consider the metric space $([-1, 1], \|.\|_0)$. Define $f : [-1, 1] \to [-1, 1]$ defined by $f(x) = cos\ x$. By *Mean-Value Theorem*[2], there exists a point $p \in (-1, 1)$ with

$$|cos\ x - cos\ y| = |sin\ p||x - y|,\ \forall x, y \in [-1, 1]$$

As the sine function does not achieve its maximum value(which is 1) in this interval, we have $|cos\ x - cos\ y| < |x - y|,\ \forall x, y \in [-1, 1]$. Thus f is a contraction on $[-1, 1]$. What if we change the domain to $\mathbb{R}$?

We will now prove one of the fundamental results in mathematics that provides the conditions under which a contraction mapping from a normed space to itself has a unique fixed point.

Banach Contraction Principle

Banach contraction principle is a crucial mathematical result that guarantees the existence of a fixed point for contraction mapping defined on a Banach space. Essentially, it provides a powerful mathematical instrument for establishing the existence and uniqueness of solutions in a variety of contexts, including optimization problems, differential equations, and iterative numerical approaches. The importance of this theorem stems from its broad applicability across mathematics as well as its function in illustrating the convergence of iterative algorithms in solving real-world issues.

Theorem 6.25 (Banach Contraction Principle) *Let $(X, \|.\|)$ be a Banach space and $f : X \to X$ be a contraction on X. Then f has exactly one fixed point.*

Proof We will start the proof by defining a Cauchy sequence $\{x_n\}$ in X using the function f. As X is complete $\{x_n\}$ will converge to a point $x \in X$. We will show that x is the unique fixed point for f in X. Choose an $x_0 \in X$ and define

$$x_{n+1} = f(x_n),\ n = 1, 2, 3, \ldots$$

First we will show that this sequence is Cauchy. As f is a contraction, we have

[2] **Mean-Value Theorem**: Suppose $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$ and that f has a derivative in the open interval (a, b). Then there exists atleast one point $c \in (a, b)$ such that $f(b) - f(a) = f'(c)(b - a)$, where f' denotes the derivative of f.

$$\begin{aligned}\|x_{m+1} - x_m\| = \|f(x_m) - f(x_{m-1})\| &\leq \alpha \|x_m - x_{m-1}\| \\ &\leq \alpha \|f(x_{m-1}) - f(x_{m-2})\| \\ &\leq \alpha^2 \|x_{m-1} - x_{m-2}\| \\ &\vdots \\ &\leq \alpha^m \|x_1 - x_0\|\end{aligned}$$

By Triangle inequality and summation formula for geometric series, for $n > m$ we have

$$\begin{aligned}\|x_m - x_n\| &\leq \|x_m - x_{m+1}\| + \|x_{m+1} - x_{m+2}\| + \cdots + \|x_{n-1} - x_n\| \\ &\leq \left(\alpha^m + \alpha^{m+1} + \cdots + \alpha^{n-1}\right) \|x_1 - x_0\| \\ &= \alpha^m \frac{1 - \alpha^{n-m}}{1 - \alpha} \|x_1 - x_0\|\end{aligned}$$

As α is a positive real number with $\alpha < 1$, we have $1 - \alpha^{n-m} < 1$ and hence for all $n > m$

$$\|x_m - x_n\| \leq \frac{\alpha^m}{1 - \alpha} \|x_1 - x_0\|$$

Since $0 < \alpha < 1$ and $\|x_1 - x_0\|$ is fixed if we choose m sufficiently large, we can make $\|x_m - x_n\|$ as small as possible. Thus $\{x_n\}$ is Cauchy and as X is complete $x_n \to x \in X$.

Now, we will show that x is a fixed point of f. By Triangle Inequality, we have

$$\begin{aligned}\|x - f(x)\| &\leq \|x - x_n\| + \|x_n - f(x)\| \\ &= \|x - x_n\| + \|f(x_{n-1}) - f(x)\| \\ &\leq \|x - x_n\| + \alpha \|x_{n-1} - x\|\end{aligned}$$

As $x_n \to x$ we can make this distance as small as possible by choosing m as sufficiently large. Thus $\|x - f(x)\| = 0$ and hence we can conclude that $f(x) = x$. That is, x is a fixed point of f.

Now we will prove that x is the only fixed point of x. Suppose that there exists another fixed point for f in X, say $\tilde{x}$. That is, we have $f(\tilde{x}) = \tilde{x}$. Then

$$\|x - \tilde{x}\| = \|f(x) - f(\tilde{x})\| \leq \alpha \|x - \tilde{x}\|$$

and this implies $\|x - \tilde{x}\| = 0$. Thus $x = \tilde{x}$ and hence f has exactly one fixed point.

Graphically identifying fixed points of a function requires identifying the points on the graph where the function intersects the line $y = x$. We have seen examples for contractions on complete metric spaces in Example 6.31 and Example 6.32. Then by *Banach Fixed Point Theorem*, we can say that both these functions have exactly one fixed point and it can be visualized as follows (Fig. 6.4).

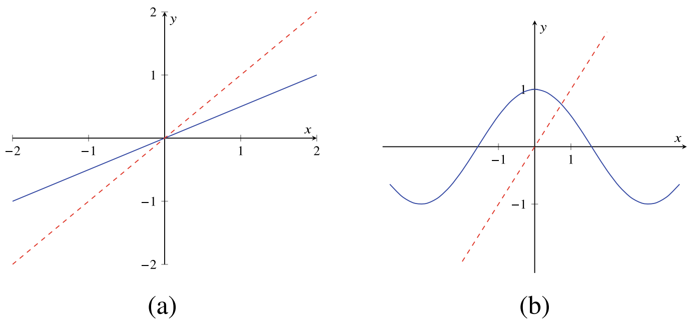


Fig. 6.4 Observe that the line $y = x$(represented by the dotted line) touches the graph of $(a)\, f(x) = \frac{1}{2}x$ and $f(x) = cos\ x$ exactly once in their respective domains as mentioned in Examples 6.31 and 6.32

Table 6.1 Fixed-point iteration for $f(x) = \cos(x)$ with $x_0 = 0.7$

Iteration (n)	Approximation (x_n)
0	0.7
1	0.7648
2	0.7215
3	0.7508
⋮	⋮
16	0.7392
17	0.7390
18	0.7391
19	0.7391

We can clearly identify $x = 0$ is the unique fixed point for $f(x) = \frac{1}{2}x$. Now let us find the fixed point of $f(x) = cos\ x$ using *fixed point iteration method*. Consider the following example.

Example 6.33 Consider $f : [-1, 1] \to [-1, 1]$ defined by $f(x) = cos\ x$ as given in Example 6.32. Choose $x_0 = 0.7$(an initial approximation can be identified from the Figure 6.4(b)). Now, define $x_{n+1} = cos\ x_n$ (Table 6.1).
Proceeding like this, we can approximate the fixed point of $f(x) = cos\ x$ as $x \approx 0.739$ after a certain number of iterations. The more the number of iterations, the less will be the error associated with it. Keep in mind that any initial point will give us the fixed point, however, the number of iterations required may vary.

Remark 6.6 To prove the *Banach contraction principle*, we use only the properties of distance notion on X provided by the *infinity norm*, but we do not use any properties of vector space structure of X. We can prove that this result is valid in a complete metric space also. That is, if (X, d) is a complete metric space and $f : X \to X$ is a contraction on X, then f has exactly one fixed point.

Gauss–Jacobi Method and Gauss–Siedel Method

The *Banach contraction principle* plays a vital role in iterative algorithms for solving systems of linear equations, such as the Gauss–Jacobi and Gauss–Seidel methods. These methods reduce a complex problem to a series of simpler ones, with each iteration attempting to get the solution closer to the actual solution. The *Banach contraction principle* establishes a theoretical basis for their convergence by ensuring that, under specific conditions, the iterations converge to a single fixed point that corresponds to the solution of the original problem. This theorem guarantees that, when used correctly, iterative approaches will produce accurate solutions for linear systems.

Theorem 6.26 *Consider the system*

$$x = Cx + \tilde{b} \tag{6.18}$$

of n *linear equations in* n *unknowns, say* $x_1, x_2, \ldots, x_n$. *Here* $C = \left[c_{ij}\right]_{n\times n}$, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ *and* $\tilde{b} = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_n \end{bmatrix}$ *is a fixed vector. If* C *satisfies*

$$\sum_{j=1}^{n} |c_{ij}| < 1, \ \forall i = 1, 2, \ldots, n \tag{6.19}$$

the system (6.18) *has exactly one solution and this can be obtained by the iterative scheme*

$$x^{m+1} = Cx^m + \tilde{b} \tag{6.20}$$

where x^0 *is arbitrary.*

Proof Consider $\mathbb{R}^n$ with the *infinity norm*, $\|x\|_\infty = \max\limits_{i\in\{1,\ldots,n\}} |x_i|$, where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ is an elements in $\mathbb{R}^n$. Then $(\mathbb{R}^n, \|.\|_\infty)$ is a complete metric space(Verify!). Define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$y = T(x) = Cx + \tilde{b} \tag{6.21}$$

where $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n$. Then for $i = 1, 2, \ldots, n$, we have $y_i = \sum\limits_{j=1}^{n} c_{ij}x_j + \tilde{b}_i$. Take $w = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, $z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{R}^n$ with $Tw = z$. Now,

$$\begin{aligned}\|y - z\|_\infty = \|T(x) - T(w)\|_\infty &= \max_{i\in\{1,\ldots,n\}} |y_i - z_i| \\ &= \max_{i\in\{1,\ldots,n\}} \left|\sum_{j=1}^{n} c_{ij}(x_j - w_j)\right| \\ &\le \max_{i\in\{1,\ldots,n\}} \sum_{j=1}^{n} |c_{ij}| \max_{j\in\{1,\ldots,n\}} |x_j - w_j| \\ &= \max_{i\in\{1,\ldots,n\}} \sum_{j=1}^{n} |c_{ij}| \, \|x - w\|_\infty\end{aligned}$$

Then by (6.19), we have

$$\|T(x) - T(w)\|_\infty < \|x - w\|_\infty$$

Thus T is a contraction on $\mathbb{R}^n$. Then by *Banach contraction principle*, T has exactly one fixed point.

Observe that the condition (6.19) which helped us in proving that T is a contraction is useful only when we are using the *infinity norm* on $\mathbb{R}^n$. What if we use another metric? (Think!) Now we will discuss *Gauss–Jacobi method* and *Gauss–Siedel method* which are popular iterative techniques for solving systems of linear equations. First, we will discuss the *Gauss–Jacobi method* named after the German mathematicians *Johann Carl Friedrich Gauss (1777–1855)* and *Carl Gustav Jacob Jacobi (1804–1851).*

Theorem 6.27 (Gauss–Jacobi Iteration) *Let* $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ *denote an approximate solution for* (6.15). *If*

$$\sum_{j=1, j\ne i}^{n} |a_{ij}| < |a_{ii}| \tag{6.22}$$

for all $i = 1, 2, \ldots, n$*, then the iteration method defined by*

$$x_i^{m+1} = \frac{1}{a_{ii}}\left(b_i - \sum_{j=1, j\ne i}^{n} a_{ij} x_j^m\right) \tag{6.23}$$

converges to x.

Proof Consider the system (6.15). Write $A = L + D + U$, where L and U are the strict lower and upper triangular part of A and D is the diagonal part of A. Then, we can write (6.15) in the form (6.18), where

$$C = -D^{-1}(L+U) = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \frac{1}{a_{ii}} \end{bmatrix} \begin{bmatrix} 0 & a_{12} & \dots & a_{1n} \\ a_{22} & 0 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & 0 \end{bmatrix}$$

and

$$\tilde{b} = D^{-1}b = \begin{bmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \vdots & \frac{1}{a_{ii}} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Then $c_{ii} = 0$ for all $i = 1, 2, \dots, n$ and $c_{ij} = \frac{a_{ij}}{a_{ii}}$ otherwise. As $\sum\limits_{j=1, j\neq i}^{n} |a_{ij}| < |a_{ii}|$ for all $i = 1, 2, \dots, n$ we have $\sum\limits_{j=1}^{n} |c_{ij}| < 1\, \forall i = 1, 2, \dots, n$. Then, by Theorem 6.26, the iteration scheme converges.

A matrix $A = [a_{ij}]$ satisfying condition (6.24) is called *strictly diagonally dominant matrix*. Thus, if the matrix A is strictly diagonally dominant, then we can say that the *Gauss–Jacobi iteration* scheme converges.

Example 6.34 Consider the system,

$$\begin{bmatrix} 4 & 1 & 2 \\ 1 & 5 & 1 \\ 2 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 5 \\ 3 \end{bmatrix}$$

Let us solve this system using *Gauss–Jacobi iteration* method with initial condition $x_1 = x_2 = x_3 = 0$. Clearly, the matrix $A = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 5 & 1 \\ 2 & 1 & 4 \end{bmatrix}$ is strictly diagonally dominant.

Now,

$$C = -D^{-1}(L+U) = \begin{bmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{5} & 0 & -\frac{1}{5} \\ -\frac{1}{2} & -\frac{1}{4} & 0 \end{bmatrix}$$

and

$$\tilde{b} = D^{-1}b = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{5} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} -1 \\ 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{4} \\ 1 \\ \frac{3}{4} \end{bmatrix}$$

Table 6.2 The table contains the first 10 iterations of the *Gauss–Jacobi iteration method*. Clearly, at each step the solution approaches the actual solution $(-1, 1, 1)$

m	x_m
0	$(0, 0, 0)$
1	$(-0.25, 1, 0.75)$
2	$(-0.875, 0.9, 0.625)$
3	$(-0.7875, 1.05, 0.9625)$
4	$(-0.9938, 0.965, 0.8812)$
5	$(-0.9319, 1.0225, 1.0056)$
6	$(-1.0084, 0.9852, 0.9603)$
7	$(-0.9765, 1.0096, 1.0079)$
8	$(-1.0064, 0.9937, 0.9858)$
9	$(-0.9913, 1.0041, 1.0048)$
10	$(-1.0034, 0.9973, 0.9946)$

Table 6.3 The table contains the first 7 iterations of the *Gauss–Siedel iteration method*. Observe that these iterates approaches to $(-1, 1, 1)$ faster than the *Gauss–Jacobi* iterates

m	x_m
0	$(0, 0, 0)$
1	$(-0.25, 1.05, 0.6125)$
2	$(-0.8188, 1.0412, 0.8991)$
3	$(-0.9598, 1.0122, 0.9769)$
4	$(-0.9915, 1.0029, 0.995)$
5	$(-0.9982, 1.0006, 0.999)$
6	$(-0.9996, 1.0001, 0.9998)$
7	$(-0.9999, 1, 1)$

Then the *Gauss–Jacobi iteration scheme* is given by,

$$x^{m+1} = \begin{bmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ -\frac{1}{5} & 0 & -\frac{1}{5} \\ -\frac{1}{2} & -\frac{1}{4} & 0 \end{bmatrix} x^m + \begin{bmatrix} -\frac{1}{4} \\ 1 \\ \frac{3}{4} \end{bmatrix}$$

Observe that in *Gauss–Jacobi iteration* method the values of the variables are updated simultaneously using the values of the previous iterations. However, in *Gauss–Siedel iteration* method, the values of the variables are updated one at a time using updated values within the same iteration. This method is named after the German mathematicians *Johann Carl Friedrich Gauss (1777–1855)* and *Philipp Ludwig von Seidel (1821–1896)* (Tables 6.2 and 6.3).

Theorem 6.28 (Gauss–Siedel Iteration) *Let* $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ *denote an approximate solution for* (6.15). *If*

$$\sum_{j=1, j\neq i}^{n} |a_{ij}| < |a_{ii}| \tag{6.24}$$

for all $i = 1, 2, \ldots, n$, *then the iteration method defined by*

$$x_i^{m+1} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{m+1} - \sum_{j=i+1}^{n} a_{ij} x_j^{m} \right) \tag{6.25}$$

converges to x.

The proof this theorem is similar to the proof of Theorem 6.27. Here, we take $C = -(L + D)^{-1}U$ and $\tilde{b} = (L + D)^{-1}b$.

Example 6.35 Consider the system $Ax = b$ as given in Example 6.34. Then, Now,

$$C = -(L + D)^{-1}U = \begin{bmatrix} 0 & -\frac{1}{4} & -\frac{1}{2} \\ 0 & \frac{1}{20} & -\frac{1}{10} \\ 0 & \frac{1}{8} & \frac{11}{40} \end{bmatrix}$$

and

$$\tilde{b} = (L + D)^{-1}U = \begin{bmatrix} -\frac{1}{4} \\ \frac{21}{20} \\ \frac{49}{80} \end{bmatrix}.$$

6.8 Exercises

1. Let V, W be normed spaces and $T : V \to W$ be a linear operator. Show that if T is continuous at one point, then it is continuous.
2. Fix $A = \left[a_{ij}\right] \in \mathbb{M}_{m\times n}$. Define $T : \mathbb{R}^n \to \mathbb{R}^m$ by $T(v) = Av$. Show that $\|Tv\| \leq \lambda \|v\|$, where $\lambda = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$.
3. Check whether the following statements are true or false.
   (a) Let V, W be normed spaces and $T : V \to W$ be a bounded linear operator. Then $v_n \to v$ in V implies that $T(v_n) \to T(v)$.
   (b) Let V be an inner product space over $\mathbb{C}$, then the set of all self-adjoint operators on V forms a subspace of $\mathcal{B}(V)$
   (c) Let T be a linear operator on $\mathbb{R}^3$ such that there is a basis of $\mathbb{R}^3$ consisting of eigenvectors of T, then T is self-adjoint.
   (d) Let T be a self-adjoint operator on a finite-dimensional inner product space over $\mathbb{R}$. Then the matrix representation of T with respect to any basis is symmetric.
   (e) Let T_1, T_2 be two positive operators on a Hilbert space H. Then T_1T_2 is positive.

(f) Let T be a linear map on a Hilbert space H with $T^2 = T$. Then T is self-adjoint.

4. Let V, W be normed spaces and $T : V \to W$ be a bounded linear operator. Then

   (a) Show that $\mathcal{N}(T)$ is closed.
   (b) Give an example to show that $\mathcal{R}(T)$ need not necessarily be closed.

5. If W is a Banach space, show that $\mathcal{B}(V, W)$ with operator norm is also a Banach space.
6. Let U, V, W be normed linear spaces and $T_1 : U \to V$, $T_2 : V \to W$ be bounded linear maps. Then show that $T_2T_1 : U \to W$ is a bounded linear map and $\|T_2T_1\| \leq \|T_2\| \, \|T_1\|$.
7. Let H be a Hilbert space and T be a bounded linear map on H. Then show that $\mathcal{N}(T) = [\mathcal{R}(T^*)]^\perp$ and $\mathcal{N}(T^*) = [\mathcal{R}(T)]^\perp$
8. Let V and W be normed linear spaces and $T : V \to W$ be a linear operator. Then show that there exists $\alpha > 0$ such that $\|Tv\| \geq \alpha \|v\|$, $\forall\, v \in V$ if and only if T is injective. Also show that $T^{-1} : \mathcal{R}(T) \to V$ is continuous and $\left\|T^{-1}w\right\| \leq \frac{1}{\alpha} \|w\| \;\forall\, w \in \mathcal{R}(T)$.
9. Let H be a Hilbert space and $T : H \to H$ be a bounded linear map whose inverse is bounded. Show that $(T^*)^{-1} = \left(T^{-1}\right)^*$.
10. Does there exists a self-adjoint linear operator T on $\mathbb{R}^3$ with $T(1, 0, 1) = (0, 0, 0)$ and $T(1, 2, 0) = (2, 4, 0)$?
11. If T is a bounded self-adjoint linear operator on a complex Hilbert space H, then show that the T^2 cannot have a negative eigenvalue. Which theorem on matrices does this generalize to?
12. Let H be a Hilbert space and $P : H \to H$ be a bounded linear map. Then P is an orthogonal projection if and only if $P^2 = P$ and P is self-adjoint.
13. Let $T : l^2 \to l^2$ be defined by $(Tx)_n = x_{n-1}$, where, $x = \{x_n\} \in l^2$. Then show that T is unitary but not self-adjoint.
14. Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be defined by $T(z_1, z_2) = (\lambda z_1, \mu z_2)$, where $\lambda, \mu \in \mathbb{C}$. Then show that

    (a) T is normal.
    (b) T is self-adjoint if and only if λ and μ are real numbers.
    (c) T is unitary and only if $|\lambda| = |\mu| = 1$.

15. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(v) = \sum_{i=1}^{n} \lambda_i \langle v, e_i \rangle e_i$, where $\{e_1, e_2, \ldots, e_n\}$ is an orthonormal basis for $\mathbb{R}^n$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ are some fixed scalars. Then show that

    (a) T is normal.
    (b) T is self-adjoint.
    (c) T is unitary and only if $\lambda_i = 1$ or -1.

16. Does there exists a linear map T with $TT^* = I$ but $T^*T \neq I$?
17. Let T be a linear map on a Hilbert space H which is normal. If $T^2 = T$, then show that T is self-adjoint.

18. Let T be a normal operator on a finite-dimensional inner product space V. Show that if B is an orthonormal basis for V, $T(B)$ is also an orthonormal basis for V.
19. Let H be a complex Hilbert space and T be a bounded linear operator on H. Then show that there exists unique operators T_1, T_2 on H such that $T = T_1 + iT_2$. Also, show that:

    (a) T is normal if and only if $T_1T_2 = T_2T_1$.
    (b) T is self-adjoint if and only if $T_2 = 0$.
    (c) T is unitary and only if $T_1T_2 = T_2T_1$ and $T_1^2 + T_2^2 = I$.

20. Find the singular value decomposition of the following matrices
    a) $\begin{bmatrix} 4 & 0 \\ 3 & -5 \end{bmatrix}$ b) $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ c) $\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$ d) $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$
21. Show that $A \in \mathbb{M}_{m\times n}(\mathbb{R}$ maps the 2-norm unit sphere in $\mathbb{R}^n$ to an ellipsoid in $\mathbb{R}^r$, where $r \leq min\{m, n\}$.
22. Show that if A is a positive semi-definite matrix, then the singular values of A are the eigenvalues of A.
23. Let A be a positive definite matrix with singular value decomposition $A = \Psi\Sigma\Phi^*$. Then show that $\Psi = \Phi$.
24. Let $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ be a fixed matrix. Show that $\langle u, v\rangle = u^TAv$ defines an inner product on $\mathbb{R}^2$ if and only if A is positive definite.
25. Let $A \in \mathbb{M}_{m\times n}(\mathbb{R})$ be a matrix with rank r. Then show that

    (a) (**Full-Rank factorization**) There exists an $m \times r$ matrix P with full column rank and $r \times n$ matrix Q with full row rank such that $A = PQ$.
    (b) $A^\dagger = Q^*(P^*AQ^*)P^*$, where P and Q are as described in part (a).

26. Find the Psuedo inverse of each of the following matrices.
    (a) $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ -1 & 2 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 2 & 3 \\ 3 & -2 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ -1 & 0 & 1 & 1 \end{bmatrix}$
27. Show that for every matrix A:
    (a) $(A^\dagger)^\dagger = A$ (b) $(A^*)^\dagger = (A^\dagger)^*$ (c) $(A^T)^\dagger = (A^\dagger)^T$
    Also, give example to show that for matrices A and B with AB defined, $(AB)^\dagger$ need not be equal to $B^\dagger A^\dagger$.
28. For each of the following system of equations; if the system is consistent, find the solution with minimum norm, otherwise, find the best approximation to a solution having minimum norm

    (a)

$$2x_1 + 3x_2 - x_3 + 2x_4 = 5$$

$$x_1 + x_2 + 2x_3 - 2x_4 = 7$$

$$4x_1 + 5x_2 + 3x_3 - 2x_4 = 12$$

(b)

$$x_1 + 3x_2 = 2$$

$$2x_1 + x_2 = 5$$

$$x_1 - 2x_2 = 3$$

29. Suppose that a manufacturing company produces circular-shaped products. The company wants to ensure that the manufactured items match industrial requirements, therefore it employs sensors to capture the coordinates of numerous points on the perimeter of the manufacturing goods. Suppose we have the measured data

$$((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

where $n \geq 4$, with respect to any manufactured product of the company. We would like to fit a circle of the form

$$(x - a)^2 + (y - b)^2 = r^2$$

to the observed data (x_i, y_i), $i = 1, 2, \ldots, n$. Model this situation to a linear least square problem of the form $Ax = y$.

30. Fit a straight line to the given points by the method of least squares.

    (a) $(1, 2), (2, 5), (3, 3), (4, 8), (5, 7)$
    (b) $(-1, 0), (0, 2), (1, 4), (2, 5)$
    (c) $(0, 12), (1, 19), (2, 29), (3, 37), (4, 45)$

31. Find a plane $x_3 = a + bx_1 + cx_2$, that best fits the following data;

$$(1, -1, 3), (2, -4, 5), (3, 8, 10), (2, 8, 12), (1, 6, 10)$$

32. Find a trigonometric curve $y = asin\ \theta + bcos\ \theta$, which best fits the points (θ_i, y_i), $i = 1, 2, 3, 4$ as given below.

$$\left(\frac{\pi}{6}, 1\right), \left(\frac{\pi}{4}, -1\right), \left(\frac{\pi}{3}, 1\right), \left(\frac{3\pi}{4}, -1\right)$$

33. *Ohm's law* states that the voltage across a conductor is directly proportional to the current flowing through it, provided all physical conditions and temperature, remain constant. Mathematically, this relation is represented by $V = IR$, where V is the voltage across the conductor, I is the current flowing through the con-

ductor and R is the resistance provided by the conductor to the flow of current. Estimate R from the given data using least square approximation method.

I	V
3	162
5	255
7	360
10	495

34. Let (X, d) be a metric space and $f : X \to X$ be a contraction. Show that f is a continuous function.
35. Find the fixed points for the following functions.

    (a) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$.
    (b) $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x, y) = (x, x^2)$.
    (c) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (x, -y)$.

36. Using *Banach fixed point theorem*, show that the equation $x^3 + x^2 - 6x + 1 = 0$ has a unique solution in the interval $[-1, 1]$.
37. (**Newton-Raphson Method**) Let $f : [a, b] \to \mathbb{R}$ be a twice continuously differentiable function and let $\tilde{x}$ be a simple zero of f in (a, b). Show that the iteration defined by
$$x_{n+1} = g(x_n), \;\; g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$
is a contraction on some neighborhood of $\tilde{x}$ and it converges to $\tilde{x}$.
38. Show that if in Theorem 6.26 if we use d_1 metric or d_2 metric instead of d_∞ metric, then we obtain the sufficient conditions given by
$$\sum_{i=1}^{n} |c_{ij}| < 1, \;\; \forall j = 1, 2, \dots n$$
and
$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2 < 1, \;\; \forall j = 1, 2, \dots n$$
respectively, for convergence of the iterative scheme (6.20).
39. Consider the integral equations of the form
$$f(x) = \lambda \int_a^b k(x, y) f(y) dy + g(x) \tag{6.26}$$
where $f : [a, b] \to \mathbb{R}$ is an unknown function, $k : [a, b] \times [a, b] \to \mathbb{R}$ is a given function with $|k(x, y)| \le \alpha$ for all $(x, y) \in [a, b] \times [a, b]$ and λ is a parameter. Consider $C[a, b]$ with supremum norm. Define $T : C[a, b] \to C[a, b]$ by

$$(Tf)(x) = \lambda \int_a^b k(x, y) f(y) dy + g(x)$$

Now, solvability of (6.26) follows from the existence of fixed point of the operator T. Show that T is a contraction when $|\lambda| < \frac{1}{\alpha(b-a)}$.

40. Set up (i) Gauss–Jacobi (ii) Gauss–Siedel iterative schemes for the following system of equations and compute the first four iterations.

(a) $\begin{bmatrix} 6 & 1 & 2 \\ 1 & 4 & 3 \\ 2 & 1 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 8 \end{bmatrix}$ (b) $\begin{bmatrix} 4 & 1 & 2 \\ 0 & 4 & 2 \\ 4 & 5 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 6 \\ 11 \end{bmatrix}$

Solved Questions related to this chapter are provided in Chap. 12.

Chapter 7
Applications

This chapter explores the numerous applications of linear algebra in diverse domains, demonstrating its tremendous impact on real-world problem-solving. Linear algebra is used in economics to support models of supply and demand, optimize resource allocation, analyze economic systems, etc. It facilitates the investigation of chemical reactions in chemistry, which is critical for drug development and materials science. Linear algebra is essential in Markov processes, assisting in the modeling of stochastic systems, predicting future states, and studying phenomena such as population dynamics and financial markets. It is the foundation of circuit analysis and signal processing in electrical engineering, making it easier to build electronic systems. Furthermore, linear algebra is used in control theory to understand dynamical systems and design control algorithms. These broad applications demonstrate linear algebra's prevalence and versatility as a foundational tool for addressing complicated challenges across a wide range of scientific and engineering disciplines, linking theory and practice. It has been shown that the general regression models can be implemented as a problem of finding the learning weights of an artificial neural network (ANN) with linear transfer functions.

7.1 Applications Involving System of Equations

A system of linear equations serves as a fundamental mathematical framework applicable across numerous domains. These systems, consisting of multiple linear equations with common variables, are employed to model a wide array of real-world problems. Whether in economics, engineering, physics, or social sciences, solving systems of linear equations helps us make informed decisions, optimize processes, and understand complex relationships. Linear algebra provides powerful techniques for solving these systems, revealing unique solutions, infinite solutions, or inconsistency, depending on the underlying equations. This mathematical tool is indispensable for problem-solving and decision-making in various fields, making it a cornerstone of applied mathematics.

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_7

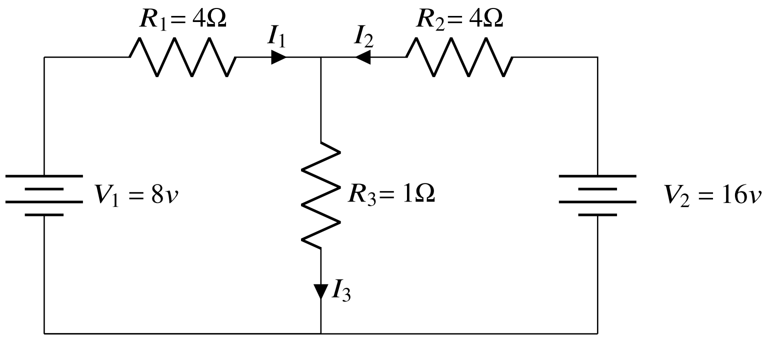


Fig. 7.1 Electrical circuit

Electrical Circuit Problem

Consider an electrical circuit as given in the following figure (Fig. 7.1).

We have the following fundamental principles in electrical circuit theory to represent the given circuit:

- *Kirchhoff's Voltage Law*—The voltage around a loop equals the sum of every voltage drop in the same loop for any closed network and equals zero.
- *Kirchhoff's Current Law*—The total current entering a junction or a node is equal to the charge leaving the node as no charge is lost.

Using the above laws, we have the following system of equations:

$$\begin{aligned} I_1 + I_2 - I_3 &= 0 \\ 4I_1 + I_3 &= 8 \\ 4I_2 + I_3 &= 16 \end{aligned}$$

This can be represented in the general form, $AX = B$ as

$$\begin{bmatrix} 1 & 1 & -1 \\ 4 & 0 & 1 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 16 \end{bmatrix}$$

Solving this, we get $I_1 = 1A$, $I_2 = 3A$ and $I_3 = 4A$.

Leontief Input–Output Models in Economics

Soviet-American economist, *Wassily Wassilyevich Leontief (1905–1999)* employed matrices to simulate economic systems. His models, also known as input-output models, segment the economy into different sectors, each of which generates commodities and services both for itself and for other sectors. Because of their interdependence, the total input and total output are always equal. For his contributions in this area, he received the Nobel Prize in Economics in 1973.

Closed Leontief Model: Leontief closed model is especially useful for understanding the relationships and dependencies within a self-contained economy. However, it does not take into account the external factors that can have an impact on demand. For example, consider an economy consisting of 3 industries, namely A, B and C. Suppose that each of the industries produces for internal consumption among themselves only. Suppose that,

- A itself consumes 40% of its product and gives 40% to B, and 20% to C.
- B itself consumes 30% of its product and gives 40% to A, and 30% to C.
- C itself consumes 50% of its product and gives 30% to A, and 20% to B.

The above data can be represented by the following table.

Tabular representation of the data

	Proportion produced by A	Proportion produced by B	Proportion produced by C
Proportion used by A	0.4	0.4	0.3
Proportion used by B	0.4	0.3	0.2
Proportion used by C	0.2	0.3	0.5

We can observe that a matrix representation would be more convenient to represent this data. In the matrix form, it can be written as

$$N = \begin{bmatrix} 0.4 & 0.4 & 0.3 \\ 0.4 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}$$

This matrix is called the input coefficient matrix. Now, suppose that A, B and C gets paid x, y and z dollars, respectively. Let us now look at A's expenses. A uses up 40% of its own production, that is, of the x dollars he gets paid, A pays itself $0.40x$ dollars, pays $0.40y$ dollars to the B, and $0.30z$ to C. As this economy is self contained, it can be modeled as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.4 & 0.4 & 0.3 \\ 0.4 & 0.3 & 0.2 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{7.1}$$

Observe that, if we denote $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, we can write (7.1) in the form of a homogeneous system of linear equations as follows:

$$X - NX = (I - N)X = 0$$

Solving this, we get $x = \frac{29}{26}\alpha$, $y = \frac{12}{13}\alpha$ and $z = \alpha$, where $\alpha \in \mathbb{R}$.

Open Leontief Model: In an open economy, interactions occur not only between different sectors within the economy, but also with the rest of the world. This means that imports and exports, as well as internal sector interactions, are taken into account. For example, consider a simple economy, where there are only three sectors: Agriculture (A), Manufacturing (M), and Services (S). Suppose that

- Agriculture (A) requires 20% of its own output, 30% of Manufacturing's output, and 10% of Services' output as inputs.
- Manufacturing (M) requires 40% of Agriculture's output, 20% of its own output, and 20% of Services' output as inputs.
- Services (S) requires 10% of Agriculture's output, 30% of Manufacturing's output, and 40% of its own output as inputs.

Then the input coefficient matrix is,

$$N = \begin{bmatrix} 0.2 & 0.3 & 0.1 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}$$

Let the final demand values be; Agriculture(A)-\$30 million, Manufacturing (M)-\$50 million and Services (S)-\$70 million. Denote the final demand values by the vector $Y = \begin{bmatrix} 30 \\ 50 \\ 70 \end{bmatrix}$. "Final demand" denotes the external demand for the output of each sector, which originates from sources outside the modeled economy. It indicates the entire amount of products and services that are used up outside of the sectors that are being analyzed through consumption, investment, or other means. For example, spending on agricultural products by consumers, businesses, and government bodies that are not part of the Agriculture sector itself could be included in the final demand for Agriculture.

Let $X = \begin{bmatrix} X_A \\ X_M \\ X_S \end{bmatrix}$ represent the equilibrium outputs of the sectors in the Leontief model. These variables represent optimum output levels for each sector that meet both internal input-output relationships and external final demand. Then we can set up the equations in the form as follows:

$$\begin{bmatrix} X_A \\ X_M \\ X_S \end{bmatrix} = \begin{bmatrix} 0.2 & 0.3 & 0.1 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} X_A \\ X_M \\ X_S \end{bmatrix} + \begin{bmatrix} 30 \\ 50 \\ 70 \end{bmatrix} \tag{7.2}$$

We can see that, (7.2) can be re-written in the form of a system of linear equations as follows:

$$(I - N)X = Y$$

Then

$$X = (I - N)^{-1}Y$$

Solving this system of equations gives us the equilibrium outputs for each sector:

$$X = \begin{bmatrix} 138.2353 \\ 190.3361 \\ 234.8739 \end{bmatrix}.$$

Some Problems in Chemistry

Now, let us give you some problems in Chemistry involving a system of linear equations.

Example 7.1 A chemical substance is created by combining three separate constituents, A, B, and C. Before they may interact to form the chemical, A, B, and C must be dissolved in water separately. Suppose that a 2.6 g/cm^3 solution of A coupled with a 2.7 g/cm^3 solution of B combined with a 3.7 g/cm^3 solution of C yields 21.2 g/cm^3 of the chemical. If the percentages of A, B, and C in these solutions are altered to 2.4 g/cm^3, 3.75 g/cm^3, and 4 g/cm^3 respectively (keeping the volumes same), 22.7 g/cm^3 of chemical is produced. Finally, 23.6 g/cm^3 of chemical is produced if the proportions are 2.75 g/cm^3, 3.4 g/cm^3, and 3.85 g/cm^3, respectively. Suppose that we have to find the volumes of the solutions containing A, B, and C. How will you proceed? Again, the techniques of linear algebra come in handy. The above scenario can be represented as,

$$\begin{aligned} 2.6A + 2.7B + 3.7C &= 21.2 \\ 2.4A + 3.75B + 4C &= 22.7 \\ 2.75A + 3.4B + 3.85C &= 23.6 \end{aligned}$$

which can be converted into the form $AX = B$, where $A = \begin{bmatrix} 2.6 & 2.7 & 3.7 \\ 2.4 & 3.75 & 4 \\ 2.75 & 3.4 & 3.85 \end{bmatrix}$ and $B = \begin{bmatrix} 21.2 \\ 22.7 \\ 23.6 \end{bmatrix}$. Solving, we get A = 5.0403 cm^3, B = 2.2281 cm^3 and C = 0.5620 cm^3.

Another instance of employing a system of linear equations can be found in the balancing of chemical equations. When considering chemical reactions, we want to look at how much of each element was there at the start and how much of each element is present in the end result.

Example 7.2 Consider the combustion reaction of Isooctane(C_8H_{18}) given by

$$C_8H_{18} + O_2 \rightarrow CO_2 + H_2O \tag{7.3}$$

Carbon dioxide and water are produced as a result of the combustion of Isooctane. Now, the question is, What exactly is a balanced chemical equation? A balanced chemical equation is a description of a chemical reaction that indicates the relative quantities of reactants and products involved in the reaction using chemical formulas and symbols. The phrase "balanced" denotes that the equation follows the law of conservation of mass, which stipulates that matter in a chemical reaction cannot be generated or destroyed, just rearranged. For an equation to be balanced, the following conditions must be met:

- The number of atoms of each element on the left side of the equation must be equal to the number of atoms of the same element on the right side.
- The total mass of the reactants must be equal to the total mass of the products.

We can see that (7.3) is not balanced, for the number of carbon, hydrogen, and oxygen atoms on the right side of the equation is not the same as the number of carbon, hydrogen, and oxygen atoms on the left side of the equation. To balance the equation, re-write (7.3) as

$$xC_8H_{18} + yO_2 \rightarrow zCO_2 + wH_2O \tag{7.4}$$

Then, we have

$$\begin{aligned} 8x &= z \\ 18x &= 2w \\ 2y &= 2z + w \end{aligned}$$

which implies

$$\begin{aligned} 8x - z &= 0 \\ 18x - 2w &= 0 \\ 2y - 2z - w &= 0 \end{aligned}$$

This can be written in the form of $AX = B$ as

$$\begin{bmatrix} 8 & 0 & -1 & 0 \\ 8 & 0 & 0 & -2 \\ 0 & 2 & -2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving, we get $x = \alpha$, $y = \frac{25}{2}\alpha$, $z = 8\alpha$ and $w = 9\alpha$. If we take $\alpha = 2$, we get $y = 25$, $z = 16$ and $w = 18$. Clearly,

$$2C_8H_{18} + 25O_2 \rightarrow 16CO_2 + 18H_2O$$

is a balanced chemical equation for (7.4).

Traffic Flow

Consider a traffic network as given in the following figure. The streets are all one-way, with arrows showing the traffic flow direction. The traffic flow in and out is measured in units of vehicles per hour(v/h). Let us construct a mathematical model to analyze this network (Fig. 7.2).

The traffic at junction A can be represented by $x_1 + x_2 = 525$. Similarly,
at junction B, $x_1 + x_4 = 375$.
at junction C, $x_3 + x_4 = 700$.
at junction D, $x_2 + x_3 = 850$.

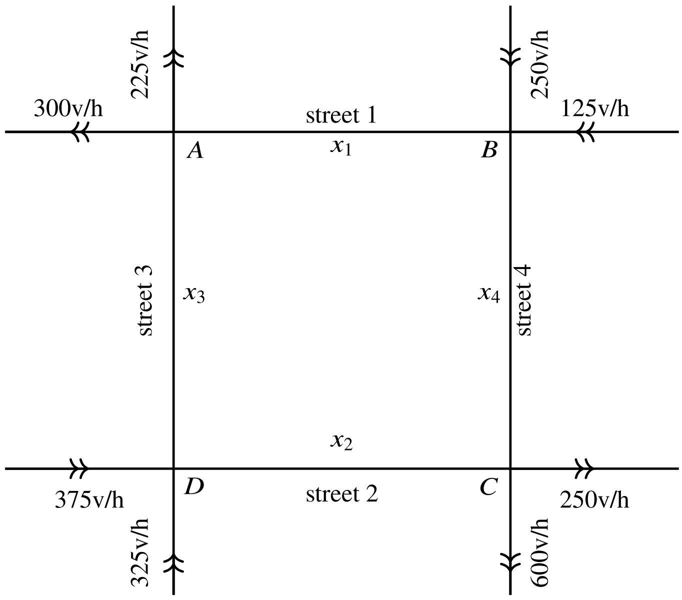


Fig. 7.2 Traffic network

This can be modeled into the form $AX = B$ as

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 525 \\ 375 \\ 700 \\ 850 \end{bmatrix}$$

and it can be reduced into

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 375 \\ 150 \\ 700 \\ 0 \end{bmatrix}$$

Clearly, the system has an infinite number of solutions. This implies here that there are an infinite number of possible traffic flows. There are certain options available to drivers at intersections. As you can see, a driver has two options at junction A. Same at other intersections as well.

This model can be used to analyze and obtain more information about the traffic flow. Suppose Street 3 needs to undergo a mandatory road maintenance. Then, it would be preferable if street 3 had the least amount of traffic feasible. That is, x_3 must be as small as possible. The question that arises then becomes, what is the smallest value of x_3 that will not cause traffic congestion? As the streets are all one-way, all the traffic flows should be non-negative. Now, from the reduced form, we can see that

$$x_3 + x_4 = 700$$

which implies that x_3 is minimum when x_4 is maximum. From the first and second equations in the reduced form, we can see that the maximum value of x_4, without x_1 and x_2 being negative, is 375. Therefore, the minimum value of x_3 is 325. That is, any road repair work on Street 3 should be done only after the appropriate arrangements for a traffic flow of $325v/h$ have been made.

7.2 Cryptography

In an increasingly interconnected digital world, providing secure communication and data security through the use of matrices in cryptography is essential. The translation of plain text into ciphertext and vice verse is made possible by the use of matrices, which provide the mathematical foundation for a wide variety of encryption and decryption procedures. Matrix-based encryption techniques, such as the Hill Cipher, which uses matrices to describe the encryption process, and more recent algorithms like the Advanced Encryption Standard (AES), which uses matrices to manipulate data blocks, all significantly influence cryptographic strategies. Matrix-based codes

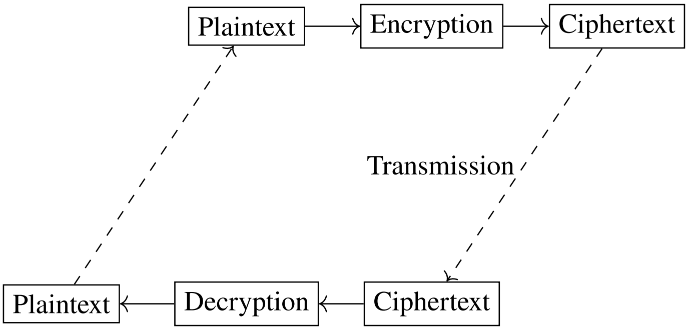


Fig. 7.3 Encryption and decryption process

Table 7.1 Numbers assigned to English alphabets

A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
1	2	3	4	5	6	7	8	9	10	11	12	13
N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
14	15	16	17	18	19	20	21	22	23	24	25	26

are used in public-key cryptography to create secure key pairs, and by error detection and correction methods to protect data integrity (Fig. 7.3).

In this section, we look at a matrix multiplication and matrix inverse-based encryption technique. *Lester S. Hill (1896–1961)*, a math professor who worked on military encryption and lectured at various US institutions, developed this technique, which is referred to as the Hill Algorithm. Modern mathematical theory and techniques entered the realm of cryptography with the development of the Hill algorithm. The Hill Algorithm is no longer regarded as a safe encryption technique because it is extremely simple to defeat using current technology. However, contemporary computing technology did not exist in 1929 when it was created. With hand computations, this procedure was too time-consuming to employ but is simple to use with today's technology. The secret message is first encoded by randomly assigning a number to each letter, creating an integer string. Let's encrypt the phrase "LINEAR ALGEBRA" by allocating a position number to each letter of the alphabet. A space is represented by the number 27, and punctuation is ignored (Table 7.1).

We divide the message's letters into two-letter groups as follows:

$$LI \quad NE \quad AR \quad \_A \quad LG \quad BE \quad RA$$

We assign the numbers from the above table to these letters and turn each pair of numbers into 2×1 matrices.

$$\begin{bmatrix} L \\ I \end{bmatrix} = \begin{bmatrix} 12 \\ 9 \end{bmatrix}, \begin{bmatrix} N \\ E \end{bmatrix} = \begin{bmatrix} 14 \\ 5 \end{bmatrix}, \begin{bmatrix} A \\ R \end{bmatrix} = \begin{bmatrix} 1 \\ 18 \end{bmatrix}, \begin{bmatrix} \_ \\ A \end{bmatrix} = \begin{bmatrix} 27 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} L \\ G \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}, \begin{bmatrix} E \\ B \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} R \\ A \end{bmatrix} = \begin{bmatrix} 18 \\ 1 \end{bmatrix}$$

So, at this point, our message is expressed using 2×1 matrices as follows:

$$\begin{bmatrix} 12 \\ 9 \end{bmatrix}, \begin{bmatrix} 14 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 18 \end{bmatrix}, \begin{bmatrix} 27 \\ 1 \end{bmatrix}, \begin{bmatrix} 12 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 18 \\ 1 \end{bmatrix} \tag{7.5}$$

The next step is to multiply this string of numbers by an inverse square matrix of our choice in order to create a new set of numbers. The coded message is represented by this new set of digits. In this, let us consider the matrix $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$. Then, our first pair $\begin{bmatrix} L \\ I \end{bmatrix}$ will be represented by

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 12 \\ 9 \end{bmatrix} = \begin{bmatrix} 69 \\ 39 \end{bmatrix}$$

Multiplying each 2×1 message in (7.5), our message can be encrypted as follows,

$$\begin{bmatrix} 69 \\ 39 \end{bmatrix}, \begin{bmatrix} 53 \\ 29 \end{bmatrix}, \begin{bmatrix} 92 \\ 55 \end{bmatrix}, \begin{bmatrix} 59 \\ 30 \end{bmatrix}, \begin{bmatrix} 59 \\ 33 \end{bmatrix}, \begin{bmatrix} 20 \\ 11 \end{bmatrix}, \begin{bmatrix} 41 \\ 21 \end{bmatrix} \tag{7.6}$$

To convert them into the alphabet form, we have to use the $mod(27)$ arithmetic. Then, the matrices in (7.6), will be of the form

$$\begin{bmatrix} 15 \\ 12 \end{bmatrix}, \begin{bmatrix} 26 \\ 2 \end{bmatrix}, \begin{bmatrix} 11 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 20 \\ 11 \end{bmatrix}, \begin{bmatrix} 14 \\ 21 \end{bmatrix} \tag{7.7}$$

and the encoded message is "QMACNCGDGGTKOU". The matrix $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ is the *key* to this encoded message. Let us decode this message. For decoding, we will use the inverse of the *key matrix*, given by $\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$. Let us multiply each 2×1 matrix in (7.7) by the inverse of the *key matrix*. Then, we obtain the following matrices,

$$\begin{bmatrix} -15 \\ 9 \end{bmatrix}, \begin{bmatrix} 68 \\ -22 \end{bmatrix}, \begin{bmatrix} 28 \\ -9 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -15 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} -63 \\ 28 \end{bmatrix} \tag{7.8}$$

Applying $mod(27)$ arithmetic on (7.8), we will get back our original message(verify). Also, try to encode this message using a 3×3 matrix as key and then decode it.

7.3 Markov Process

A Markov chain or Markov process is a stochastic model that illustrates a series of potential occurrences where the likelihood of each event is solely determined by the state it reached in the preceding event. Numerous real-world processes can be statistically modeled using Markov chains, including the dynamics of animal populations, queues or lines of passengers at airports, and cruise control systems in automobiles. In computer science, physics, biology, economics, and finance Markov chains are crucial tools for comprehending, describing, and forecasting occurrences. It is named after the Russian mathematician *Andrey Andreyevich Markov (1856–1922).*

Markov processes are concerned with the fixed probabilities of transitioning between a finite number of states. We start by defining probability vector and then probability transition matrix/stochastic matrix.

Definition 7.1 A probability vector $p = \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}$ is a vector with each component $p_i \geq 0$ and the sum of components equal to one. That is, $\sum_{i=1}^{n} p_i = 1$.

Definition 7.2 A matrix $M = \left[m_{ij}\right]_{n\times n}$ with real entries is called a stochastic matrix or probability transition matrix provided that each column of M is a probability vector, where m_{ij} denote the probability of transition from the jth state to the ith state. As the total probability of transition from the state j to any other state is 1, $0 \leq m_{ij} \leq 1$ and $\sum_{i=1}^{n} m_{ij} = 1$. That is, each column sum of M is 1.

Let $M = \left[m_{ij}\right]_{n\times n}$ be a stochastic matrix and $x^q = \begin{bmatrix} x_1^q \\ \vdots \\ x_n^q \end{bmatrix}$ denote the state of the system at time q. Assume that x^q denote the amount of some materials spread among n states. Then, x_i^q denote the amount of material in ith state at time q. The transition of material from the jth state at time 0 to the ith state at time 1 is given by $m_{ij}x_j^0$. Then the total amount of material at the ith state would be the sum of the material from all the states to state i. That is,

$$x_i^1 = \sum_{j=1}^{n} m_{ij} x_j^0$$

Therefore, we have

$$x^1 = Mx^0$$

and in general, we can write

$$x^q = Mx^{q-1}$$

As the amount of material at each state i at time q depends on the amount of material in i at the time $q-1$, this can also be represented as,

$$x^q = M^q x^0$$

Clearly, we can observe that the total amount of material at time q is the same as at time $q-1$. For,

$$\begin{aligned}\sum_{i=1}^{n} x_i^q &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} m_{ij} x_j^{q-1}\right)\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{n} m_{ij}\right) x_j^{q-1}\end{aligned}$$

and as $\sum_{i=1}^{n} m_{ij} = 1$, we get $\sum_{i=1}^{n} x_i^q = \sum_{j=1}^{n} x_i^{q-1}$. Let this amount be denoted by α. Then the proportion of the material in the ith state at time q is given by,

$$p_i^q = \frac{x_i^q}{\alpha}$$

Then the probability vector at time q is,

$$p^q = \begin{bmatrix} p_1^q \\ \vdots \\ p_n^q \end{bmatrix} = \frac{1}{\alpha}\begin{bmatrix} x_1^q \\ \vdots \\ x_n^q \end{bmatrix} = \frac{1}{\alpha} x^q$$

Also,

$$p_i^q = \frac{1}{\alpha} x^q == \frac{1}{\alpha} M x^{q-1} = M\left(\frac{1}{\alpha} x^{q-1}\right) = M p^{q-1}$$

That is, the probability vectors also transform through multiplication by the matrix M. For a stochastic matrix M, the transformation $p^q = Mp^{q-1}$ on probability vectors is called a *Markov process*. A *Markov Chain* is the sequence of iterates $p^q = M^q p^0$ obtained for a given initial probability vector, p^0.

Example 7.3 Let's consider an example of a Markov chain representing the weather states, sunny,cloudy and rainy. The transition probabilities are defined as follows;

- If it's sunny today, there's a 70% chance it will be sunny tomorrow, a 20% chance it will be cloudy, and a 10% chance it will be rainy.
- If it's cloudy today, there's a 50% chance it will be cloudy tomorrow, a 30% chance it will be sunny, and a 20% chance it will be rainy.
- If it's rainy today, there's a 60% chance it will be rainy tomorrow, a 20% chance it will be cloudy, and a 20% chance it will be sunny.

We can represent this Markov chain with a transition probability matrix,

$$M = \begin{bmatrix} 0.7 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.6 \end{bmatrix}$$

Now, consider the initial condition $x^0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, which means that initially the weather state is sunny with 100% certainty, and the probabilities of rainy and cloudy are both zero. Then, for day 1, we have

$$x^1 = Mx^0 = \begin{bmatrix} 0.7 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.7 \\ 0.2 \\ 0.1 \end{bmatrix}$$

So the likelihood of a sunny day is 0.7, a cloudy day is 0.2, and a rainy day is 0.1 the following day. For day 2, we have

$$x^2 = Mx^1 = \begin{bmatrix} 0.7 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 0.7 \\ 0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.57 \\ 0.26 \\ 0.17 \end{bmatrix}$$

and for day 3,

$$x^3 = Mx^2 = \begin{bmatrix} 0.7 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 0.57 \\ 0.26 \\ 0.17 \end{bmatrix} = \begin{bmatrix} 0.511 \\ 0.278 \\ 0.211 \end{bmatrix}$$

proceeding like this, for day 14, we have

$$x^{14} = Mx^{13} = \begin{bmatrix} 0.7 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.2 \\ 0.1 & 0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 0.4572 \\ 0.2857 \\ 0.2571 \end{bmatrix} = \begin{bmatrix} 0.4572 \\ 0.2857 \\ 0.2571 \end{bmatrix}$$

That is, the sequence of vectors have converged to the vector $\begin{bmatrix} 0.4572 \\ 0.2857 \\ 0.2571 \end{bmatrix}$. This vector is called the *steady state vector*. The steady state vector in this example indicates the equilibrium distribution of the weather. It informs us how much of the time we can expect the weather to be sunny, cloudy, or rainy in the long run, providing the transition probabilities remain constant.

From the above example, we can observe that the steady state vector is the eigenvector corresponding to the eigenvalue 1 (Clearly, 1 is an eigenvalue of M as the column sum of M is 1).

Page Rank Algorithm

Page Rank is an algorithm created by Google founders Larry Page and Sergey Brin to rank web pages in search engine results. It transformed the way search engines give relevant and authoritative results by utilizing the structure of the web graph and linear algebra concepts.

The internet is modeled as a directed graph in the context of Page Rank, which is commonly referred to as the "web graph". In the graph, web sites are represented as nodes, while connections between pages are represented as directed edges. Each link from page A to page B represents a vote of confidence or recommendation from page A to page B. Page Rank simulates the behavior of a random online surfer who navigates the web graph by clicking on hyperlinks using a Markov chain model. The web surfer begins on a random website and, at each step, either follows a hyperlink on the current page or, with a given chance, teleports to a random page. This random surfer model is based on the idea that users frequently navigate across online sites by clicking on links.A transition probability matrix is created to quantitatively depict the unpredictable movement of the surfer. Based on the hyperlinks, this matrix depicts the odds of navigating from one page to another. Each row of the matrix represents a source page, while each column represents a destination page. The entries of this matrix represent the likelihoods of transiting from the source page to each destination page. Finding the stationary distribution of the Markov chain, which depicts the long-term probability of the arbitrary surfer being on each web page, is the key to Page Rank Algorithm. This stationary distribution corresponds to the Page Rank scores of the pages. The Page Rank vector is the dominant eigenvector of the transition probability matrix, with an eigenvalue 1.

In Summary, the Page Rank Algorithm uses linear algebra principles, namely eigenvectors, eigenvalues, and matrix operations, to rank online sites based on their relevance and authority within the linked web graph. This sophisticated method illustrates how linear algebra may be used to solve real-world challenges in information retrieval and ranking.

7.4 Coupled Harmonic Oscillators

Let's now give you an illustration of how eigenvalues and eigenvectors can be employed in real-life applications. Consider an oscillator as shown in the figure (Fig. 7.4).

Two bodies of mass m are coupled by springs with spring constants k and slide on a smooth plane. The x_1 and x_2 displacements are measured from their equilibrium positions and are positive when to the right. *Hooke's law* and *Newton's laws of motion*, when combined with these conventions, yield the differential equations given below.

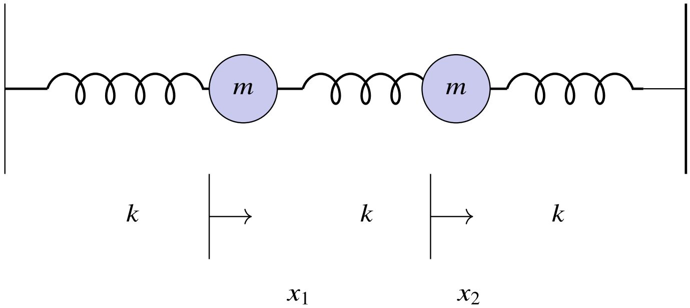


Fig. 7.4 Coupled Harmonic oscillators

$$\begin{aligned} m\ddot{x}_1 &= -kx_1 + k(x_2 - x_1) = k(x_2 - 2x_1) \\ m\ddot{x}_2 &= -k(x_2 - x_1) - kx_2 = k(x_1 - 2x_2) \end{aligned} \tag{7.9}$$

This can be written as

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \frac{k}{m} \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{7.10}$$

In this case, since there is no damping, we choose a purely oscillatory solution as we have seen in Sect. 4.1. Let

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \mu_1 e^{i\alpha t} \\ \mu_2 e^{i\alpha t} \end{bmatrix}$$

be a solution of (7.10). Then

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = -\alpha^2 \begin{bmatrix} \mu_1 e^{i\alpha t} \\ \mu_2 e^{i\alpha t} \end{bmatrix} = -\alpha^2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{7.11}$$

From (7.10) and (7.11), we have

$$\frac{k}{m} \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -\alpha^2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Clearly, this is an eigenvalue problem $Ax = \lambda x$, where $A = \frac{k}{m} \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$ and $\lambda = -\alpha^2$.

Now, let's find the value of α when $m = k = 1$. Then, we have $A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$. Clearly, the eigenvalues of A are -1 and -3 with respective eigenvectors $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Therefore, $\alpha^2 = 1$ or $\alpha^2 = 3$ and the general solution is

$$x(t) = c_1 v_1 e^{it} + c_2 v_1 e^{-it} + c_3 v_2 e^{\sqrt{3}t} + c_4 v_2 e^{-\sqrt{3}t}$$
$$= \begin{bmatrix} c_1 e^{it} + c_2 e^{-it} + c_3 e^{\sqrt{3}t} + c_4 e^{-\sqrt{3}t} \\ c_1 e^{it} + c_2 e^{-it} - c_3 e^{\sqrt{3}t} - c_4 e^{-\sqrt{3}t} \end{bmatrix}.$$

7.5 Satellite Control Problem

Linear algebra is an essential building block in control theory, supporting the analysis and design of dynamic systems for a variety of applications. It provides a strong mathematical framework for modeling, interpreting, and managing systems that evolve over time. Linear algebra allows us to formulate complicated systems as matrices and vectors using ideas such as state-space representation, matrix transformations, and eigenvalue analysis. This, in turn, makes it easier to investigate system stability, controllability, observability, and performance. The impact of linear algebra is profound, whether designing control algorithms for aerospace, robotics, industrial processes, or economic systems, as it enables control theorists to harness the elegance of matrices and eigenvectors to unravel the intricate behavior of dynamic systems, paving the way for sophisticated control strategies that improve stability, optimize performance, and achieve desired outcomes. Let us give you an example. Consider a satellite revolving around the earth. Due to various forces, a satellite injected into orbit may slightly deviate from the predicted orbit or alter orientation. To correct the deviation, the satellites include built-in control mechanisms in the form of thrusters in the radial and tangential directions.

Consider a satellite of unit mass orbiting around the earth under inverse square law field. It is convenient to choose polar coordinates, with $r(t)$ the radius from the origin to the mass, and $\theta(t)$ the angle from x- axis. We can assume that the satellite has thrusting capacity with radial thrust $u_1(t)$ and tangential thrust $u_2(t)$ (Fig. 7.5).

Then by *Newton's law*, the equations of motion have the form

$$\begin{aligned} \frac{d^2 r}{dt^2} &= r(t)\left(\frac{d\theta}{dt}\right)^2 - \frac{\beta}{r^2(t)} + u_1(t) \\ \frac{d^2\theta}{dt^2} &= -\frac{2}{r(t)}\frac{d\theta}{dt}\frac{dr}{dt} + \frac{u_2(t)}{r(t)} \end{aligned} \tag{7.12}$$

With $u_1 = u_2 = 0$, and the initial conditions $r(0) = \sigma, \dot{r}(0) = 0, \theta(0) = 0$ and $\dot{\theta}(0) = \omega$, where $\omega = \left(\frac{\beta}{\sigma^3}\right)^{\frac{1}{2}}$, the coupled Eq. (7.12) have solutions given by

$$\begin{aligned} r(t) &= \sigma \\ \theta(t) &= \omega t \end{aligned} \tag{7.13}$$

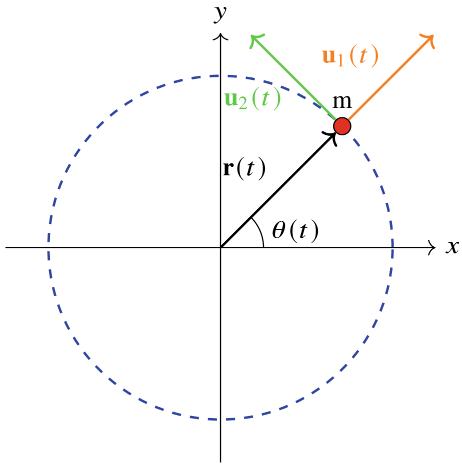


Fig. 7.5 A unit mass m in gravitational orbit

If we make the following change of variables,

$$x_1 = r - \theta, x_2 = \dot{r}, x_3 = \sigma(\theta - \omega t), x_4 = \sigma(\dot{\theta} - \omega)$$

Equation (7.12) will reduce to the form;

$$\begin{bmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \\ \frac{dx_3}{dt} \\ \frac{dx_4}{dt} \end{bmatrix} = \begin{bmatrix} x_2 \\ (x_1+\sigma)\left(\frac{x_4}{\sigma}+\omega\right)^2 - \frac{\beta}{(x_1+\sigma)^2} + u_1 \\ x_4 \\ -2\sigma\left(\frac{x_4}{\sigma}+\omega\right)\frac{x_2}{(x_1+\sigma)} + \frac{u_2\sigma}{(x_1+\sigma)} \end{bmatrix} \tag{7.14}$$

This is a system of non-linear ordinary differential equations involving the forcing functions (controls) u_1 and u_2, which can be represented in compact vector notation as follows:

$$\frac{dx}{dt} = f(x, u), \quad x(t) \in \mathbb{R}^4, u(t) \in \mathbb{R}^2 \tag{7.15}$$

Here f is a vector with components f_1, f_2, f_3, f_4 given by

$$\begin{bmatrix} f_1(x_1, x_2, x_3, x_4, u_1, u_2) \\ f_2(x_1, x_2, x_3, x_4, u_1, u_2) \\ f_3(x_1, x_2, x_3, x_4, u_1, u_2) \\ f_4(x_1, x_2, x_3, x_4, u_1, u_2) \end{bmatrix} = \begin{bmatrix} x_2 \\ (x_1+\sigma)\left(\frac{x_4}{\sigma}+\omega\right)^2 - \frac{\beta}{(x_1+\sigma)^2} + u_1 \\ x_4 \\ -2\sigma\left(\frac{x_4}{\sigma}+\omega\right)\frac{x_2}{(x_1+\sigma)} + \frac{u_2\sigma}{(x_1+\sigma)} \end{bmatrix}$$

We now linearize the non-linear system about the zero equilibrium solution to obtain the system in the form $\dot{x}(t) = Ax(t) + Bu(t)$. By linearizing the function $f(x, u)$, about $x = 0$, $u = 0$, we have

$$\hat{f}(x,u) = f'_x(0,0)x + f'_u(0,0)u = Ax + Bu$$

where,

$$A = f'_x(0,0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & \frac{\partial f_1}{\partial x_4} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & \frac{\partial f_2}{\partial x_4} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4} \\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} \end{bmatrix}_{(0,0)} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \tag{7.16}$$

and

$$B = f'_u(0,0) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} \\ \frac{\partial f_4}{\partial u_1} & \frac{\partial f_4}{\partial u_2} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{7.17}$$

Here, σ is normalized to 1. The representation of the system in the form

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{7.18}$$

is called the *state-space representation*. Dynamical systems can often be seen in state-space form, with the behavior of the systems described by a set of linear differential equations. The state vector indicates the internal variables of the system, and matrices describe how these variables change over time. This is a time-invariant linear system and the matrix e^{At} forms the *state transition matrix* of the linear homogeneous system. Now, by using the variation of parameter method the solution of (7.18) with initial condition $x(t_0) = x_0$ can be written in the form;

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau. \tag{7.19}$$

Computation of e^{At}

Computation of state transition matrix in the form of the exponential of a square matrix is a key concept in linear algebra which have wide applications in solving differential equations, analyzing dynamic systems, and understanding transformations induced by matrices on vectors in diverse fields. It involves extending the notion of exponential function from numbers to matrices through a power series expansion:

$$e^{At} := I + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{n!}A^n + \cdots \tag{7.20}$$

This expansion parallels the Taylor series for the regular exponential function. In practice, the matrix exponential calculation can be challenging, especially for bigger matrices. It can be computed quickly and precisely with the aid of methods like matrix diagonalization and Jordan form.

We know that if the matrix A is diagonalizable, then there exists a matrix P such that,

$$A = PDP^{-1}$$

where, D is a diagonal matrix containing eigenvalues of A. Then,

$$\begin{aligned} e^{At} &= I + At + \frac{t^2}{2!}A^2 + \cdots + \frac{t^n}{n!}A^n + \cdots \\ &= I + PDP^{-1}t + \frac{t^2}{2!}PD^2P^{-1} + \cdots + \frac{t^n}{n!}PD^nP^{-1} + \cdots \\ &= P\left[I + Dt + \frac{t^2}{2!}D^2 + \cdots + \frac{t^n}{n!}D^n + \cdots\right]P^{-1} \\ &= Pe^{Dt}P^{-1} \end{aligned}$$

That is, if A is diagonalizble, we can compute e^A in a simple manner. This can also be generalized as follows. From Exercise 27, Chap. 4 we know that for any real $n \times n$ matrix A, there exists a matrix P consisting of generalized eigenvectors of A, such that $A = S + N$ where $S = P^{-1}diag\{\lambda_1, \ldots, \lambda_n\}P$ and N is a nilpotent matrix. By using this representation, we can write

$$e^{At} = \left[Pdiag\left\{e^{\lambda_1 t}, \ldots, e^{\lambda_n t}\right\}P^{-1}\right]\left[I + Nt + \frac{N^2t^2}{2!} + \cdots + \frac{N^{k-1}t^{k-1}}{(k-1)!}\right] \tag{7.21}$$

Using this formula, let us compute the state transition matrix for the satellite system discussed above. We can observe that the eigenvalues of A are $0, 0, \pm i\omega$ and A is not diagonalizable. The eigenvectors corresponding to the eigenvalue $\lambda = 0$ are $\left(1, 0, 0, -\frac{3\omega}{2}\right)$ and $(0, 0, 1, 0)$ and the eigenvectors corresponding to the complex eigenvalues $\lambda = \pm i\omega$ are $(1, 0, 0, 2\omega)$ and $(0, \omega, 2, 0)$. This gives,

$$P = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & \omega \\ 0 & 1 & 0 & 2 \\ -\frac{3\omega}{2} & 0 & -2\omega & 0 \end{bmatrix} \; and \; P^{-1} = \begin{bmatrix} 4 & 0 & 0 & \frac{2}{\omega} \\ 0 & -\frac{2}{\omega} & 1 & 0 \\ -3 & 0 & 0 & -\frac{2}{\omega} \\ 0 & \frac{1}{\omega} & 0 & 0 \end{bmatrix}$$

and hence,

$$N = A - S = A - P^{-1}diag\{0, 0, i\omega, -i\omega\}P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -6\omega & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then we can compute $e^{A(t-t_0)}$ as

$$\begin{bmatrix} 4-3cos\,\omega(t-t_0) & \frac{1}{\omega}sin\,\omega(t-t_0) & 0 & \frac{2}{\omega}\left(1-cos\,\omega(t-t_0)\right) \\ 3\omega sin\,\omega(t-t_0) & cos\,\omega(t-t_0) & 0 & 2sin\,\omega(t-t_0) \\ -6\omega(t-t_0)+6sin\,\omega(t-t_0) & \frac{2}{\omega}\left[cos\,\omega(t-t_0)-1\right] & 1 & \frac{4}{\omega}sin\,\omega(t-t_0)-3(t-t_0) \\ 6\omega\left[cos\,\omega(t-t_0)-1\right] & -2sin\,\omega(t-t_0) & 0 & 4cos\,\omega(t-t_0)-3 \end{bmatrix}$$

Now, let us define the controllability of a dynamical system.

Controllability of Linear Systems

The concept of controllability analysis is crucial in the study of dynamical systems. It considers whether a system can be steered from an arbitrary initial state to any desired final state using appropriate control inputs within a given time frame. Controllability in linear systems is governed by the qualities of the system's state matrix and control matrix. A system is considered controllable if the reachable states can span its state space under the effect of control inputs. This feature is critical in engineering and control theory because it underpins the design and execution of successful control techniques for a wide range of applications, including robotics and aerospace, as well as economics and chemical processes. Let us give a formal definition for controllability first and then we will discuss on the conditions that are used to verify the controllability of a dynamical system.

Definition 7.3 (*Controllability*) The system $\dot{x}(t) = Ax(t) + Bu(t)$ with initial condition $x(t_0) = x_0$ is said to be controllable in the interval $[t_0, t_1]$ if for every $x_0, x_1 \in \mathbb{R}^n$ there exists a control input $u \in \mathcal{L}^2([t_0, t_1]; \mathbb{R}^m)$ such that the corresponding solution starting from $x(t_0) = x_0$ also satisfies $x(t_1) = x_1$.

From Eq. 7.19 it follows that the system (7.18) is controllable if and only if there exists a control function $u \in \mathcal{L}^2([t_0, t_1]; \mathbb{R}^m)$ such that

$$x(t_1) = x_1 = e^{A(t_1-t_0)}x_0 + \int_{t_0}^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau$$

That is,

$$x_1 - e^{A(t_1-t_0)}x_0 = \int_{t_0}^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau$$

If we take $x_1 - e^{A(t_1-t_0)}x_0 = w$, system (7.18) is controllable if and only if $\forall w \in \mathbb{R}^n$ there exists a control function $u \in \mathcal{L}^2([t_0, t_1]; \mathbb{R}^m)$ such that

$$w = \int_{t_0}^{t_1} e^{A(t_1-t_0)}Bu(\tau)d\tau$$

Define an operator $C : \mathcal{L}^2([t_0, t_1] : \mathbb{R}^m) \to \mathbb{R}^n$ by

$$Cu = \int_{t_0}^{t_1} e^{A(t_1-\tau)} Bu(\tau) d\tau \tag{7.22}$$

Observe that C is a bounded linear operator(Verify!) and system (7.18) is controllable if and only if $Cu = w$ has a solution for every $w \in \mathbb{R}^n$. That is, controllability of system (7.18) is equivalent to the ontoness of the operator C. The operator C defines its adjoint $C^* : \mathbb{R}^n \to \mathcal{L}^2([t_0, t_1] : \mathbb{R}^m)$ in the following way:

$$\begin{aligned} \langle Cu, v\rangle_{\mathbb{R}^n} &= \langle \int_{t_0}^{t_1} e^{A(t_1-\tau)} Bu(\tau) d\tau, v\rangle_{\mathbb{R}^n} \\ &= \int_{t_0}^{t_1} \langle e^{A(t_1-\tau)} Bu(\tau), v\rangle_{\mathbb{R}^n} d\tau \\ &= \int_{t_0}^{t_1} \langle u(\tau), B^T e^{A^T(t_1-\tau)} v\rangle_{\mathbb{R}^n} d\tau \\ &= \langle u(\tau), B^T e^{A^T(t_1-\tau)} v\rangle_{\mathcal{L}^2} \\ &= \langle u, C^* v\rangle_{\mathcal{L}^2} \end{aligned}$$

That is,

$$(C^* v)(\tau) = B^T e^{A^T(t_1-\tau)} v \tag{7.23}$$

The following theorem explains the relation between controllability of the system (7.18) with the operators C and C^*.

Theorem 7.1 *System (7.18) is controllable if and only if one of the following is satisfied.*

(a) the operator C is onto.
(b) the operator C^* is onto.
(c) the Gramian matrix

$$W(t_0, t_1) = CC^* = \int_{t_0}^{t_1} e^{A(t_1-\tau)} BB^T e^{A^T(t_1-\tau)} d\tau$$

is non-singular.

Thus, we have observed that system (7.18) is controllable if and only if there exists a control function $u \in \mathcal{L}^2([t_0, t_1]; \mathbb{R}^m)$ such that

$$w = Cu = \int_{t_0}^{t_1} e^{A(t_1-\tau)} Bu(\tau) d\tau$$

By *Cayley–Hamilton Theorem*, we can write the above equation as

$$\begin{aligned} w &= \int_{t_0}^{t_1} \left[\mathcal{P}_0(\tau)I_n + \mathcal{P}_1(\tau)A + \cdots + \mathcal{P}_{n-1}(\tau)A^{n-1}\right] Bu(\tau)d\tau, \forall w \in \mathbb{R}^n \\ &= \left[B \ AB \ \ldots \ A^{n-1}B\right] \begin{bmatrix} \int_{t_0}^{t_1} \mathcal{P}_0(\tau)u(\tau)d\tau \\ \int_{t_0}^{t_1} \mathcal{P}_1(\tau)u(\tau)d\tau \\ \vdots \\ \int_{t_0}^{t_1} \mathcal{P}_{n-1}(\tau)u(\tau)d\tau \end{bmatrix} \end{aligned}$$

where $\mathcal{P}_i(\tau), i = 1, 2, \ldots, n-1$ are polynomial functions that appears in the expansion of $e^{A(t_1-\tau)}$. Thus, we can say that system (7.18) is controllable if and only if $\mathcal{R}\left(\left[B \ AB \ \ldots \ A^{n-1}B\right]\right) = \mathbb{R}^n$. that is, if and only if $Rank\left(\left[B \ AB \ \ldots \ A^{n-1}B\right]\right) = n$. This result is proposed by the Hungarian-American electrical engineer and mathematician *Rudolf Emil Kalman (1930–2016)* and is known as *Kalman's Rank Condition* for controllability.

Theorem 7.2 (Kalman's Rank Condition) *System* (7.18) *is controllable if and only if the controllability matrix*

$$Q = [B \mid AB \mid A^2B \mid \cdots \mid A^{n-1}B]$$

has full rank; that is, $Rank(Q) = n$.

Now, suppose that there exists a vector $v \in \mathbb{R}^n$ such that $vA = \lambda v$ and $vB = 0$. Then observe that

$$v\left[B \ AB \ \ldots \ A^{n-1}B\right] = 0$$

and hence $Rank\left(\left[B \ AB \ \ldots \ A^{n-1}B\right]\right) < n$ which implies that system (7.18) is not controllable. Thus for the controllability of system (7.18), no vector $v \in \mathbb{R}^n$ with $vA = \lambda v$ should be orthogonal to the columns of B. This method is known after the mathematicians *Vasile M. Popov (1928-), Vitold Belevitch (1921–1999)* and *Malo L. J. Hautus (1940–).*

Theorem 7.3 *System* (7.18) *is controllable if and only if one of the following is satisfied.*

(a) ***PBH Rank Condition:*** $Rank[sI_n - A, B] = n, \ \forall\, s \in \mathbb{C}$.
(b) ***PBH Eigenvector Condition:*** *the relationship* $v^T A = \lambda v^T$ *implies* $v^T B \neq 0$, *where* v *is a left eigenvector of* A *associated with the eigenvalue* λ.

Using Theorem 7.2, we can verify the controllability of the satellite system. The controllability matrix Q for the given satellite system is

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 2\omega & -\omega^2 & 0 \\ 1 & 0 & 0 & 2\omega & -\omega^2 & 0 & 0 & -2\omega^3 \\ 0 & 0 & 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 \\ 0 & 1 & -2\omega & 0 & 0 & -4\omega^2 & 2\omega^3 & 0 \end{bmatrix}$$

and we can observe that Q has rank 4. This means that the given satellite system is controllable.

It is interesting to ask the following question: What happens when one of the controllers or thrusters become inoperative? Suppose that the tangential thruster fails. Then $u_2 = 0$ and hence B in (7.17) will reduce to $B_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$. The controllability matrix is given by

$$Q_1 = \begin{bmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{bmatrix}$$

We can observe that Q_1 has rank 3 and hence the system is not controllable. What if the Radial Thruster become inoperative? Try to find it out yourself!

Linear algebra is also essential in analyzing the observability and stability of dynamical systems. Linear algebraic techniques aid in determining whether all system states can be accurately determined from available measurements, allowing monitoring of the entire system's activity. Conditions for observability can be derived along the same lines as that of controllability. Stability analysis to evaluate the behavior of the system over time can also be done using tools from linear algebra. Stability conditions can be derived by studying the eigenvalues of the system matrix, indicating if the system converges or diverges. With the help of these tools, we may examine the behavior, predictability, and robustness of systems, which will enable us to make well-informed decisions in numerous real-life applications.

7.6 Artificial Neural Network as Linear Regression Model

Linear algebra is foundational in artificial neural networks and machine learning, serving as the mathematical framework that underpins their operations. Matrices and vectors represent data, weights, etc., enabling efficient information manipulation and transformation. Matrix multiplication, dot products, and vector addition are crucial to neural network training, allowing signals to propagate through layers of neurons and model parameters to be adjusted during optimization. Linear algebra allows complex mathematical operations to be expressed clearly and serves as the foundation for understanding the fundamental concepts and behavior of these robust learning systems, making it a vital component of the subject.

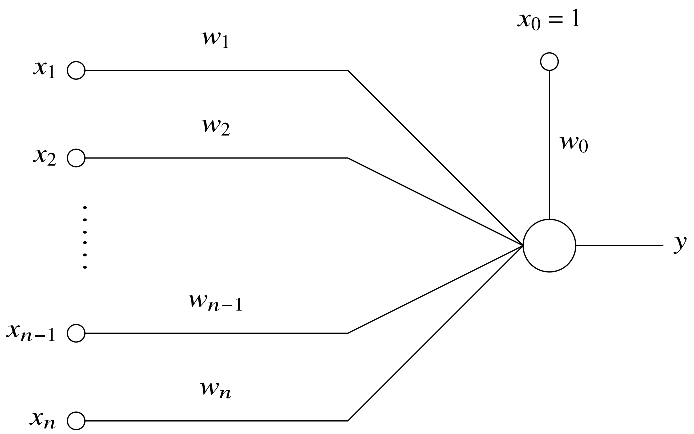


Fig. 7.6 Linear regression network

Let us consider a network that works as regression network for mapping input-output data from an experiment. The data can be used to train a linear regression network as shown in the following figure (Fig. 7.6).

The output of the neuron is given by

$$y = \sum_{i=0}^{m} x_i w_i$$

Let $\left\{\left(x^1, y^1\right), \left(x^2, y^2\right), \ldots, (x^m, y^m)\right\},\ x^i \in \mathbb{R}^{n+1}, y^i \in \mathbb{R}$ be a set of training data for determining a linear regression model in a feature space. Here $x^i = \begin{pmatrix} 1 \\ x_1^i \\ x_2^i \\ \vdots \\ x_n^i \end{pmatrix}$ is the input including constant input 1 for the bias w_0 and y_i is the corresponding output. We use the following notations;

$$W = \begin{pmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{pmatrix},\ y = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix}$$

Using the training data define a matrix $X = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \dots & x_n^1 \\ 1 & x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_1^m & x_2^m & \dots & x_n^m \end{bmatrix}$. Our objective is to find a weight vector W satisfying $XW = y$. Here, X is a rectangular matrix and hence we look for a W which minimizes the error(residue) given by,

$$\epsilon = \|y - XW\|^2 = \langle y - XW, y - XW \rangle$$

Now, by using least square techniques W can be computed as follows:

$$\begin{aligned} \frac{\partial \epsilon}{\partial W} = 0 &\Rightarrow \langle -Xh, y - XW \rangle + \langle y - XW, -Xh \rangle = 0 \text{ for } h \in \mathbb{R}^{n+1} \\ &\Rightarrow -\langle Xh, y \rangle + \langle Xh, XW \rangle - \langle y, Xh \rangle + \langle XW, Xh \rangle = 0 \\ &\Rightarrow -2\langle h, X^T y \rangle + 2\langle h, X^T XW \rangle = 0 \text{ for } h \in \mathbb{R}^{n+1} \\ &\Rightarrow X^T XW = X^T y \end{aligned}$$

If $X^T X$ is invertible, then

$$W = \left(X^T X\right)^{-1} X^T y = X^\dagger y$$

Hence

$$W = X^T X \left(X^T X\right)^{-2} X^T y = X^T \alpha$$

where $\alpha = X \left(X^T X\right)^{-2} X^T y$. That is, the weight vector W can be written as a linear combination of the input vector X.

Solutions for Selected Problems

Chapter 1

2. not an equivalence relation.
4. (a) not a bijection. (b) not a bijection. (c) not a bijection. (d) is a bijection.
15. (c) $\mathcal{Z}(GL_2(\mathbb{F}))$ = Set of all non-zero scalar matrices and $\mathcal{Z}(S_3) = \left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right\}$.
16. (a) $O\left(\begin{bmatrix}1&0\\0&-1\end{bmatrix}\right) = 2$ (b) $O\left(\begin{bmatrix}1&0\\1&1\end{bmatrix}\right) = \infty$
20. $\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}\right\}$, A_3, S_3
22. $\mathbb{Z}_p$ for any prime p.
26. $Rank(A) = 3$
28. $x_1 = 1, x_2 = 1, x_3 = 2$

Chapter 2

2. (a) $\{(x_1, 0) \in \mathbb{R}^2 \mid x_1 \geq 0\}$ (b) $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = 0 \; or \; x_2 = 0\}$
3. not a vector space.
4. (a) linearly independent (b) linearly independent (c) linearly independent (d) linearly dependent (e) linearly independent (f) linearly independent (g) linearly independent (h) linearly dependent
6. (a) not a subspace (b) is a subspace, $B = \{(1, -1)\}$, dimension is 1 (c) not a subspace (d) not a subspace (e) is a subspace, dimension is 0
7. (a) is a subspace, $B = \left\{\begin{bmatrix}1&-1\\0&0\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&0\\0&1\end{bmatrix}\right\}$, dimension is 3 (b) not a subspace (c) not a subspace (d) not a subspace (e) is a subspace, $B = \left\{\begin{bmatrix}1&0\\0&1\end{bmatrix}, \begin{bmatrix}0&0\\1&0\end{bmatrix}, \begin{bmatrix}0&1\\0&0\end{bmatrix}\right\}$, dimension is 3
8. (a) is a subspace, $B = \{x, x^2\}$, dimension is 2 (b) not a subspace (c) is a subspace, $B = \{x - x^2\}$, dimension is 2 (d) not a subspace (e) is a subspace, $B = \{1, x^2\}$, dimension is 2
9. (a) True (b) False (c) False (d) False (e) False (f) False (g) False
11. Row space $= \mathbb{R}^2$, Column space $= \mathbb{R}^2$
14. (a) doesn't span $\mathbb{R}^2$ (b) span $\mathbb{R}^2$ (c) span $\mathbb{R}^2$ (d) doesn't span $\mathbb{R}^2$ (e) span $\mathbb{R}^2$
15. (a) doesn't span $\mathbb{R}^3$ (b) doesn't span $\mathbb{R}^3$ (c) span $\mathbb{R}^3$ (d) doesn't span $\mathbb{R}^3$ (e) span $\mathbb{R}^3$
16. (a) span $\mathbb{P}_2$ (b) span $\mathbb{P}_2$ (c) doesn't span $\mathbb{P}_2$ (d) span $\mathbb{P}_2$ (e) doesn't span $\mathbb{P}_2$
20. (a) is a basis (b) is a basis (c) not a basis (d) is a basis

21. (a) is a basis, $(1, 2, 3) = 3(1, 1, 1) - 1(1, 1, 0) - 1(1, 0, 0)$ (b) is a basis $(1, 2, 3) = \frac{1}{2}(1, 2, 1) - \frac{1}{2}(2, 1, 1) + \frac{3}{2}(1, 1, 2)$ (c) not a basis
23. (a) dimension is 1, span is the line $y = -2x$, basis $= \{(1, -2)\}$
(b) dimension is 2, span is $\mathbb{R}^2$, basis $= \{(-2, 3), (1, 2)\}$
(c) dimension is 3, span is $\mathbb{R}^3$, basis $= \{(0, 3, 1), (-1, 2, 3), (2, 3, 0)\}$
(d) dimension is 2, span is $\{a_0 + a_1x + a_2x^2 \in \mathbb{P}_2 \mid a_0 = a_1\}$, basis $= \{1 + x, x^2 + x + 1\}$
(e) dimension is 2, span is $\{a_0 + a_1x + a_2x^2 \in \mathbb{P}_2 \mid a_0 = -a_1\}$, basis $= \{1 - x, x^2\}$
(f) dimension is 4, span is $\mathbb{M}_2(\mathbb{R})$, basis $= \left\{\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}1 & -1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\-1 & 1\end{bmatrix}\right\}$
(g) dimension is 4, span is $\mathbb{M}_2(\mathbb{R})$, basis $= \left\{\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}\right\}$
24. (a) $W_1 + W_2 = \mathbb{R}^4$ and $W_1 \cap W_2 = \{0\}$
(b) $dim\,(W_1 + W_2) = 4$ and $dim\,(W_1 \cap W_2) = 0$

Chapter 3

1. (a) not a linear transformation. (b) is a linear transformation. (c) not a linear transformation. (d) not a linear transformation. (e) is a linear transformation. (f) not a linear transformation.
2. (a) is a linear transformation. (b) is a linear transformation. (c) not a linear transformation. (d) is a linear transformation. (e) not a linear transformation. (f) not a linear transformation. (g) is a linear transformation.
3. Take $A = I_2$ and $B = -I_2$, where I_2 is the identity matrix of order 2. Check whether, $det(A + B) = det(A) + det(B)$ and $det(\lambda A) = \lambda det(A)$.
5. $T_1 + T_2$ will always be a non-linear map.
6. As $(x_1, y_1, z_1) = (x_1 - y_1)(1, 0, 0) + (y_1 - z_1)(1, 1, 0) + z_1(1, 1, 1)$,

$$\begin{aligned}T(x_1, y_1, z_1) &= (x_1 - y_1)T(1, 0, 0) + (y_1 - z_1)T(1, 1, 0) + z_1T(1, 1, 1)\\ &= (x_1 - y_1, 2y_1 - 2z_1, x_1)\end{aligned}$$

7. (a) $\mathcal{R}(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_2 = 0\}$ and $\mathcal{N}(T) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = x_3 = 0\}$. (b) $\mathcal{R}(T) = span\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}\right\}$ and $\mathcal{N}(T) = \{zero\ polynomial\}$.
(c) $\mathcal{R}(T) = span\left\{\begin{bmatrix}1 & 0\\-1 & 0\end{bmatrix}, \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\right\}$ and $\mathcal{N}(T) =$ Set of all 2×2 symmetric matrices.
8. Choose $v_3, v_4 \in \mathbb{R}^4$ such that $B = \{v_1 = (1, 1, 1, 1), v_2 = (1, 0, 0, 1), v_3, v_4\}$ forms a basis for $\mathbb{R}^4$. Then define $T(v_1) = T(v_2) = 0$ and $T(v_3) = (1, 1, 0)$, $T(v_4) = (1, 0, 1)$.

10. (a) Yes. $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_1(x_1, x_2) = (x_1, 0)$ and $T_2(x_1, x_2) = (2x_1, 0)$. (b) Yes. $T_1, T_2 : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T_1(x_1, x_2) = (x_1, 0)$ and $T_2(x_1, x_2) = (0, x_2)$.
11. (b) $\mathcal{R}(T) = span\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_4 = 0\}$ and $\mathcal{N}(T) = span\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 \mid x_1 = x_2 = x_3 = x_4\}$.
12. (a) False (b) False (c) False (d) True (e) True (f) False (g) True (h) False
20. $[T]_{B_1}^{B_2} = \begin{bmatrix} 2 & -1 & -3 \\ -5 & 0 & 1 \\ 3 & 1 & 2 \end{bmatrix}$
21. $[T]_B = \begin{bmatrix} 1 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
23. (a) $\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 1 & -3 & -4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{bmatrix}$
24. (a) $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$ c) $\begin{bmatrix} 3 & 1 \\ -2 & 0 \end{bmatrix}$
27. $\{x_1 - x2, x_2 - x_3, x_3\}$
30. $\begin{bmatrix} 5 & -9 \\ -5 & 15 \\ -3 & 8 \end{bmatrix}$

Chapter 4

2. (a) $\lambda^2 - 4\lambda - 32$ (b) $\lambda^2 - 11\lambda + 39$ c) $-\lambda^3 + 9\lambda^2 - 3\lambda - 30$
3. $\begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$
8. 3,5
11. (a) True (b) False (c) False (d) True (e) True (f) False (g) False (h) True (i) True (j) True
13. 0
14. (a) diagonalizable (b) not diagonalizable (c) diagonalizable (d) diagonalizable (e) diagonalizable
18. Characteristic polynomial = $\lambda^{n-1}(\lambda - n)$ and Minimal polynomial = $\lambda(\lambda - n)$.
25. (a) The eigenvalues are $2, 2, 1$ with eigenvectors $(-1, 0, 1), (0, 1, 0)$ and $(0, -1, 1)$ respectively. (b) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$ (c) $\{(0, 0, 0)\}$, $\mathbb{R}^3$, span$\{(-1, 0, 1)\}$, span$\{(0, 1, 0)\}$,
span$\{(0, -1, 1)\}$, span$\{(-1, 0, 1), (0, 1, 0)\}$, span$\{(-1, 0, 1), (0, -1, 1)\}$ and span$\{(0, 1, 0), (0, -1, 1)\}$.

26. $\begin{bmatrix} 3&1&0&0&0&0&0&0&0\\ 0&3&0&0&0&0&0&0&0\\ 0&0&3&0&0&0&0&0&0\\ 0&0&0&2&0&0&0&0&0\\ 0&0&0&0&2&0&0&0&0\\ 0&0&0&0&0&2&0&0&0\\ 0&0&0&0&0&0&-1&1&0\\ 0&0&0&0&0&0&0&-1&0\\ 0&0&0&0&0&0&0&0&-1 \end{bmatrix}$

Chapter 5

20. (a) $\left\{\frac{1}{3}(1,2,2), \frac{1}{3}(-2,-1,2), \frac{1}{3}(2,-2,1)\right\}$ (b) $\left\{1, 2\sqrt{3}\left(x-\frac{1}{2}\right), 6\sqrt{5}\left(x^2-x+\frac{1}{6}\right)\right\}$
22. (a) $\{v \mid v = \lambda(-u_2, u_1), \lambda \in \mathbb{R}\}$ (b) $\{(0,0)\}$
24. (a) $\left\{\frac{1}{\sqrt{10}}(1,0,3), \frac{1}{\sqrt{14}}(-3,2,1)\right\}$ (b) $\left(-\frac{1}{35}, -1, -\frac{13}{35}\right)$
27. (a) $\frac{1}{2}(-1,1)$ (b) $\frac{1}{15}(11,25,47)$ (c) $\frac{8}{3}+\sqrt{2}x$

Chapter 6

3. (a) True (b) False (c) False (d) False (e) False (f) False
10. No
20. (a) $\begin{bmatrix} -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{40} & 0 \\ 0 & \sqrt{10} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

(b) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$

(c) $\begin{bmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & 0 \\ -\frac{2}{3} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -3\sqrt{10} & 1\sqrt{10} \\ 1\sqrt{10} & 3\sqrt{10} \end{bmatrix}$

(d) $\begin{bmatrix} -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \sqrt{\frac{2}{3}} & 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$

26. (a) $\frac{1}{6}\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ -1 & -1 \end{bmatrix}$ (b) $\frac{1}{50}\begin{bmatrix} 5 & 16 & -13 \\ 10 & 2 & 14 \end{bmatrix}$ (c) $\frac{1}{12}\begin{bmatrix} 1 & 1 \\ 3 & -3 \\ 1 & 1 \end{bmatrix}$ (d) $\frac{1}{4}\begin{bmatrix} 1 & -1 & -1 \\ 2 & 2 & 2 \\ 2 & -2 & 2 \\ -1 & 1 & 1 \end{bmatrix}$

28. (a) inconsistent, $\frac{1}{225}\begin{bmatrix} 93 \\ -35 \\ -26 \\ -35 \end{bmatrix}$ (b) consistent, $\frac{1}{15}\begin{bmatrix} 39 \\ -3 \end{bmatrix}$

29. $\begin{bmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots \\ 2x_n & 2y_n & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ \vdots \\ x_n^2 + y_n^2 \end{bmatrix}$, where $c = r^2 - a^2 - b^2$.

30. (a) $y = 1.3x + 1.1$ (b) $y = 1.7x + 1.9$ (c) $y = 8.4x + 11.6$

31. $x_3 = 5.2171 + 0.4053x_1 + 0.6039x_2$

32. $y = -0.3937 sin\ \theta + 0.8535 cos\ \theta$

33. $V = 18.7 + 47.9I$

Part II
Solved Problems

Chapter 8
Solved Problems—Preliminaries

(1) Let $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & 1 \end{bmatrix}$ and $X = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$. Then $\lim_{n\to\infty} A^n X$

(a) does not exist (b) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (c) $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ (d) $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

Ans. ***Option c***
We have

$$AX = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ 4 \end{bmatrix} = \begin{bmatrix} 2 + \frac{1}{2} \\ 4 \end{bmatrix}$$

$$A^2 X = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{5}{2} \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{9}{4} \\ 4 \end{bmatrix} = \begin{bmatrix} 2 + \frac{1}{4} \\ 4 \end{bmatrix}$$

$$A^3 X = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{9}{4} \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{17}{8} \\ 4 \end{bmatrix} = \begin{bmatrix} 2 + \frac{1}{8} \\ 4 \end{bmatrix}$$

Proceeding like this we get $A^n X = \begin{bmatrix} 2 + \frac{1}{2^n} \\ 4 \end{bmatrix}$ *and hence* $\lim_{n\to\infty} A^n X = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$

(2) Let $A = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & \alpha \\ 2 & -\alpha & 0 \end{bmatrix}$, $\alpha \in \mathbb{R} \setminus \{0\}$ and b a non-zero vector such that $Ax = b$ for some $x \in \mathbb{R}^3$. Then the value of $x^T b$ is
(a) $-\alpha$ (b) α (c) 0 (d) 1

Ans. ***Option c***

Let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *and* $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. *Then*

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_8

$$Ax = b \Rightarrow x_2 - 2x_3 = b_1, -x_1 + \alpha x_3 = b_2, 2x_1 - \alpha x_2 = b_3$$
$$\Rightarrow x_1x_2 - 2x_1x_3 = b_1x_1, -x_1x_2 + \alpha x_2x_3 = b_2x_2, 2x_1x_3 - \alpha x_2x_3 = b_3x_3$$

Adding these, we get $b_1x_1 + b_2x_2 + b_3x_3 = 0 = x^T b$.

(3) Let

$$S = \left\{ A : A = \left[a_{ij}\right]_{5\times 5}, a_{ij} = 0 \text{ or } 1 \ \forall\, i, j, \sum_j a_{ij} = 1 \ \forall\, i \text{ and } \sum_i a_{ij} = 1 \ \forall\, j \right\}$$

Then the number of elements of S is
(a) 5^2 (b) 5^5 (c) $5!$ (d) 55

Ans. ***Option c***
Since $a_{ij} = 0$ *or* $1 \ \forall\, i, j, \sum_j a_{ij} = 1 \ \forall\, i$ *and* $\sum_i a_{ij} = 1 \ \forall\, j$ *the first row has 5 possibilities.*

$$5 \text{ possibilities} \begin{cases} 1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0 \\ 0\,0\,1\,0\,0 \\ 0\,0\,0\,1\,0 \\ 0\,0\,0\,0\,1 \end{cases}$$

Then the second row has 4 possibilities, third row has 3 possibilities, fourth row has 2 possibilities and the fifth row has only 1 possibility. Thus the number of elements of S is $5!$.

(4) Let p be a prime number and let $\mathbb{Z}_p$ denote the field of integers modulo p. Find the number of 2×2 invertible matrices with entries from this field.

Ans. *Let* $GL_2\left(\mathbb{Z}_p\right)$ *denote the set of all* 2×2 *invertible matrices with entries from* $\mathbb{Z}_p$. *Let* $M \in GL_2\left(\mathbb{Z}_p\right)$. *Then the first row of M has* $p^2 - 1$ *possibilities as first row cannot be* $\begin{bmatrix} 0 & 0 \end{bmatrix}$. *Now the second row has* $p^2 - p$ *possibilities since it cannot be a scalar multiple of the first row. So the number of* 2×2 *invertible matrices with entries from* $\mathbb{Z}_p$ *is* $(p^2 - 1)(p^2 - p)$.

(5) The order of the matrix $\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ in $GL_2\left(\mathbb{Z}_5\right)$ is ......

Ans. *We know that the order of the matrix* $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$ *in* $GL_2\left(\mathbb{Z}_5\right)$ *is the least positive integer n such that* $A^n = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. *We have*

$$A^2 = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 7 & 12 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 2 \end{bmatrix} \text{ in } GL_2\left(\mathbb{Z}_5\right)$$

and

$$A^3 = \begin{bmatrix} 2 & 2 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 10 \\ 10 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ in } GL_2(\mathbb{Z}_5)$$

Therefore order of H is 3.

(6) Let G be a subgroup of $GL_2(\mathbb{R})$ generated by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$. Then the order of G is ......

Ans. *Let* $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$. *We have*

$$A^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = I \text{ and } B^3 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} = I$$

Also

$$B^2 = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, AB = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} = B^2 A \text{ and } AB^2 = \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} = BA$$

Thus $G = \{I, A, B, B^2, AB, A^2B\}$ *and hence order of G is 6.*

(7) Consider the following group under matrix multiplication:

$$G = \left\{ \begin{bmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{bmatrix} : p, q, r \in \mathbb{R} \right\}$$

Then the center of the group, $\mathcal{Z}(G)$ is isomorphic to
(a) $(\mathbb{R} \setminus \{0\}, \times)$ (b) $(\mathbb{R}, +)$ (c) $(\mathbb{R}^2, +)$ (d) $(\mathbb{R}, +) \otimes (\mathbb{R} \setminus \{0\}, \times)$

Ans. ***Option b***
Let $\begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix} \in \mathcal{Z}(G)$. *Then for any* $\begin{bmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{bmatrix} \in G$, *we have*

$$\begin{bmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & p & q \\ 0 & 1 & r \\ 0 & 0 & 1 \end{bmatrix}$$

This implies that

$$\begin{bmatrix} 1 & x+p & q+xr+y \\ 0 & 1 & z+r \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & x+p & q+pz+y \\ 0 & 1 & z+r \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore $pz = xr$ for any $p, r \in \mathbb{R}$, which is possible only if $x = z = 0$. Thus center of G is given by

$$\mathcal{Z}(G) = \left\{ \begin{bmatrix} 1 & 0 & y \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} : y \in \mathbb{R} \right\}$$

Now define $\phi : \mathcal{Z}(G) \to (\mathbb{R}, +)$ by $\phi\left(\begin{bmatrix} 1 & 0 & y \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = y$. Clearly ϕ is both one-one and onto. Take $A = \begin{bmatrix} 1 & 0 & y_1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & y_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathcal{Z}(G)$. Then

$$\phi(AB) = \begin{bmatrix} 1 & 0 & y_1 + y_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = y_1 + y_2 = \phi(A) + \phi(B)$$

Thus, ϕ is a homomorphism. Therefore ϕ is an isomorphism.

(8) Let $F = \{\omega \in \mathbb{C} : \omega^{2020} = 1\}$. Consider the group

$$G = \left\{ \begin{pmatrix} \omega & z \\ 0 & 1 \end{pmatrix} : \omega \in F, z \in \mathbb{C} \right\} \text{ and } H = \left\{ \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} : z \in \mathbb{C} \right\}$$

under matrix multiplication. Then the number of cosets of H in G is
(a) 1010 (b) 2019 (c) 2020 (d) infinite

Ans. ***Option c***
Define $\phi : G \to F$ by $\phi(A) = det(A)$. Since

$$\phi(AB) = det(AB) = det(A)det(B) = \phi(A)\phi(B) \ \forall \ A, B \in G$$

ϕ is a homomorphism from G to F with

$$Ker(\phi) = \{A \in G : det(A) = 1\} = H$$

Then by ***First Isomorphism Theorem*** *$G \setminus H \cong Im(\phi) = F$. Therefore the number of cosets of H in G is 2020.*

(9) Let G denote the group of all 2×2 invertible matrices with entries from $\mathbb{R}$. Let

$$H_1 = \{A \in G : det\ (A) = 1\} \text{ and } H_2 = \{A \in G : A \text{ is upper triangular}\}$$

Consider the following statements:
P : H_1 is a normal subgroup of G
Q : H_2 is a normal subgroup of G
Then
(a) Both P and Q are true (b) P is true but Q is not true
(c) P is false and Q is true (d) Both P and Q are false

Ans. ***Option b***
Let $H \in H_1$, and $A \in G$. Then $AHA^{-1} \in H_1$ since

$$det(AHA^{-1}) = det(A)det(H)\frac{1}{det(A)} = 1$$

Therefore H_1 is a normal subgroup of G.
Consider the matrix $B = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in G$ and $K = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \in H_2$. Then

$$BKB^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -2 & 3 \end{pmatrix} \notin H_2$$

Therefore H_2 is not a normal subgroup of G.

(10) Let A, B be $n \times n$ matrices. Which of the following equals $tr(A^2B^2)$?
(a) $(tr(AB))^2$ (b) $tr(AB^2A)$ (c) $tr\left((AB)^2\right)$ (d) $tr(BABA)$

Ans. ***Option b***
Since $tr(AB) = tr(BA)$, $tr(A^2B^2) = tr(BA^2B) = tr(AB^2A)$.

(11) Pick out the true statements:
(a) Let A and B be two arbitrary $n \times n$ matrices. Then $(A + B)^2 = A^2 + 2AB + B^2$.
(b) There exist $n \times n$ matrices A and B such that $AB - BA = I$.
(c) Let A and B be two arbitrary $n \times n$ matrices. If B is invertible, then $tr(A) = tr(B^{-1}AB)$.

Ans. ***Option c***
(a) $(A + B)^2 = (A + B)(A + B) = A^2 + AB + BA + B^2$ since $AB \neq BA$, $(A + B)^2$ need not be equal to $A^2 + 2AB + B^2$.
(b) Suppose there exists $n \times n$ matrices A and B such that $AB - BA = I$. Then $tr\ (AB - BA)$ must be equal to $tr(I) = n$. Since $tr(AB) = tr(BA)$,$tr(AB - BA) = 0$. So there does not exists such matrices A and B.
(c) Since $tr(AB) = tr(BA)$, $tr(B^{-1}AB) = tr(BB^{-1}A) = tr(A)$.

(12) If $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ with $det(A + I) = 1 + det(A)$, then we can conclude that
(a) $det(A) = 0$ (b) $A = 0$ (c) $Tr(A) = 0$ (d) A is non-singular

Ans. ***Option c***

(a) & (b) Take $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *Then* $A + I = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $det(A + I) = 1 + det(A)$. *Here* $A \neq 0$ *and* $det(A) \neq 0$.

(c) Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Then* $A + I = \begin{bmatrix} a+1 & b \\ c & d+1 \end{bmatrix}$.

$$\begin{aligned} det(A + I) = 1 + det(A) &\Rightarrow (a + 1)(d + 1) - bc = 1 + ad - bc \\ &\Rightarrow ad + a + d + 1 - bc = 1 + ad - bc \\ &\Rightarrow a + d = 0 \end{aligned}$$

This gives, $Tr(A) = a + d = 0$.

(d) Take $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *Then* $A + I = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ *and* $det(A + I) = 1 + det(A)$. *But A is singular.*

(13) It is known that $X = X_0 \in \mathbb{M}_{2\times 2}(\mathbb{Z})$ is a solution of $AX - XA = A$ for some $A \in \left\{ \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \right\}$. Which of the following values are not possible for the determinant of X_0?
(a) 0 (b) 2 (c) 6 (d) 10

Ans. ***Option d***

Consider the equation $AX - XA = A$. *As* $tr(AX) = tr(XA)$,

$$tr(AX - XA) = tr(A) \Rightarrow tr(A) = 0$$

Therefore, $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$. *Take* $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Then*

$$\begin{aligned} AX - XA = A &\Rightarrow \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} b + c & 2b + d - a \\ -a - 2c + d & -(b + c) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \\ &\Rightarrow b + c = 1,\ 2b + d - a = 1,\ -a - 2c + d = -1 \end{aligned}$$

The equations $2b + d - a = 1$ *and* $-a - 2c + d = -1$ *are the same as* $c = 1 - b$. *Then,*

$$b = 1, d = 0, a = 1, c = 0 \Rightarrow det(X_0) = 0$$

$$b = 1, d = 1, a = 2, c = 0 \Rightarrow det(X_0) = 2$$

$$b = 1, d = 2, a = 3, c = 0 \Rightarrow det(X_0) = 6$$

Thus $det(X_0) = 10$ *is not possible.*

(14) Let A and B be $n \times n$ matrices. Suppose that the sum of elements of A in any row is 2 and the sum of elements in any column of B is 2. Which of the following matrices is necessarily singular?
(a) $I - \frac{1}{2}BA^T$ (b) $I - \frac{1}{2}AB$ (c) $I - \frac{1}{4}AB$ (d) $I - \frac{1}{4}BA^T$

Ans. ***Option d***
Take $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$. *Then,*

$$I - \frac{1}{2}BA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & \frac{-3}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

$$I - \frac{1}{2}AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -1 & 0 \end{bmatrix}$$

$$I - \frac{1}{4}AB = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{-1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}$$

$$I - \frac{1}{4}BA^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{-3}{4} \\ 0 & \frac{3}{4} \end{bmatrix}$$

It can be observed that $I - \frac{1}{4}BA^T$ *is singular.*

(15) The number of distinct real values of x for which $A = \begin{bmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{bmatrix}$ is singular is
(a) 1 (b) 2 (c) 3 (d) infinite

Ans. ***Option b***
Let $A = \begin{bmatrix} x & 1 & 1 \\ 1 & x & 1 \\ 1 & 1 & x \end{bmatrix}$. *We have A is singular if and only if* $det(A) = 0$. *Now,*

$$det\ (A) = 0 \Rightarrow x^3 - 3x + 2 = (x-1)^2(x+2) = 0 \Rightarrow x = 1, 1, -2$$

Hence, A is singular, when x is either 1 *or* -2.

(16) Let A be a 3×3 matrix with integer entries such that $det(A) = 1$. What is the maximum possible number of entries of A that are even?
(a) 2 (b) 3 (c) 6 (d) 8

Ans. ***Option c***
Consider the 3×3 *matrix* $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. *Then,*

$$det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

Clearly, the maximum number of even entries cannot be equal to 8 as the determinant would become an even number. Now consider the 3×3 *identity matrix,* I*. We know that* $det(I) = 1$*. Hence the maximum possible number of entries of* A *that are even is 6.*

(17) For $t \in \mathbb{R}$, define $A(t) = \begin{pmatrix} 1 & t & 0 \\ 1 & 1 & t^2 \\ 0 & 1 & 1 \end{pmatrix}$. Then which of the following statements is true?
(a) $det\ (A(t))$ is a polynomial function of degree 3 in t.
(b) $det\ (A(t)) = 0$ for all $t \in \mathbb{R}$.
(c) $det\ (A(t))$ is zero for infinitely many $t \in \mathbb{R}$.
(d) $det\ (A(t)) = 0$ for exactly two $t \in \mathbb{R}$.

Ans. ***Option d***
We have $det\ (A(t)) = 1 - t^2 - t = 0$*. Since* $discriminant = 5, det\ (A(t)) = 0$ *for exactly two* $t \in \mathbb{R}$*.*

(18) A permutation matrix A is a non-singular square matrix in which each row has exactly one entry equals 1, the other entries being all zeros. If A is an $n \times n$ permutation matrix, what are the possible values of determinant of A?

Ans. *A permutation matrix is obtained by interchanging rows(columns) of identity matrix. If odd number of interchanges are made, then the determinant of the permutation matrix is* -1 *and if even number of interchanges are made, then the determinant of the permutation matrix is* 1.

(19) The determinant of the $n \times n$ permutation matrix $\begin{bmatrix} & & & 1 \\ & & 1 & \\ & \cdot^{\cdot^{\cdot}} & & \\ & 1 & & \\ 1 & & & \end{bmatrix}$
(a) $(-1)^n$ (b) $(-1)^{\lfloor \frac{n}{2} \rfloor}$ (c) -1 (d) 1
Here $\lfloor x \rfloor$ denotes the greatest integer not exceeding x.

Ans. ***Option b***
Consider 2×2 *and* 5×5 *permutation matrices of the given form. All the options except Option b are false.*

(20) For $j = 1, 2, \ldots, 5$ let A_j be the matrix of order 5×5 obtained by replacing the jth column of the identity matrix of order 5×5 with the column vector $v = \begin{bmatrix} 5 & 4 & 3 & 2 & 1 \end{bmatrix}^T$. Then the determinant of the matrix product $A_1A_2A_3A_4A_5$ is ......

Ans. *We have*

$$\begin{aligned} det(A_1A_2A_3A_4A_5) &= det(A_1)det(A_2)det(A_3)det(A_4)det(A_5) \\ &= 5 \times 4 \times 3 \times 2 \times 1 \\ &= 120 \end{aligned}$$

(We can calculate the determinant easily as the first row in each of these matrices has only two non-zero entry and the determinant with the second non zero entry will be zero since the 4×4 *matrix has one column zero)*

(21) Let $A = (a_{ij}) \in \mathbb{M}_{n\times n}(\mathbb{R})$, where $a_{ij} = \begin{cases} 1, & \text{if } i + j = n + 1 \\ 0, & \text{otherwise} \end{cases}$
What is the value of $det(A)$ when (i) $n = 10$ and (ii) $n = 100$?

Ans. *The given matrix is*

$$A = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 \end{bmatrix}_{n\times n}$$

The given matrix is obtained by interchanging the rows of the identity matrix and hence $det\ (A) = (-1)^{\lfloor \frac{n}{2} \rfloor}$ *where* $\lfloor \frac{n}{2} \rfloor$ *denotes the greatest integer less than or equal to* $\frac{n}{2}$. *Therefore* $det(A)$ *when* (i) $n = 10$ *is* -1 *and* (ii) $n = 100$ *is 1.*

(22) Let A be a 4×4 matrix whose determinant is 10. Then $det(-3A)$ is
(a) −810 (b) −30 (c) 30 (d) 810

Ans. ***Option d***
As $det(\lambda A) = \lambda^n det(A)$ *for an* $n \times n$ *matrix A, we have*

$$det(-3A) = (-3)^4 det\ (A) = 810$$

(23) Let $A = \begin{bmatrix} 9 & 2 & 7 & 1 \\ 0 & 7 & 2 & 1 \\ 0 & 0 & 11 & 6 \\ 0 & 0 & -5 & 0 \end{bmatrix}$. Then the value of $det\left((8I - A)^3\right)$ is ......

Ans. *We have* $8I - A = \begin{bmatrix} -1 & -2 & -7 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & -3 & -6 \\ 0 & 0 & 5 & 8 \end{bmatrix}$. *Then*

$$det\left((8I - A)^3\right) = [det(8I - A)]^3 = -216$$

(24) The determinant of the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(a) 0 (b) −9 (c) −27 (d) 1

Ans. ***Option c***

Let $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, *where* A, B, C *and* D *are square matrices. Then*

$$det(M) = det\left(A - BD^{-1}C\right) det(D)$$

Take $A = D = I_3$ *and* $B = C = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$. *Then determinant of the given matrix is* -27.

(25) Let $D_1 = det \begin{bmatrix} a & b & c \\ x & y & z \\ p & q & r \end{bmatrix}$ and $D_2 = det \begin{bmatrix} -x & a & -p \\ y & -b & q \\ z & -c & r \end{bmatrix}$. Then

(a) $D_1 = D_2$ (b) $D_1 = 2D_2$ (c) $D_1 = -D_2$ (d) $2D_1 = D_2$

Ans. ***Option c***

$$A = \begin{bmatrix} a & b & c \\ x & y & z \\ p & q & r \end{bmatrix} \qquad \begin{matrix} R_1 \leftrightarrow C_1 \\ R_2 \leftrightarrow C_2 \\ R_3 \leftrightarrow C_3 \end{matrix}$$

$$\sim \begin{bmatrix} a & x & p \\ b & y & q \\ c & z & r \end{bmatrix} \qquad R_1 \to -R_1$$

$$\sim \begin{bmatrix} -a & -x & -p \\ b & y & q \\ c & z & r \end{bmatrix} \qquad C_1 \to -C_1$$

$$\sim \begin{bmatrix} a & -x & -p \\ -b & y & q \\ -c & z & r \end{bmatrix} \qquad C_1 \longleftrightarrow C_2$$

$$\sim \begin{bmatrix} -x & a & -p \\ y & -b & q \\ z & -c & r \end{bmatrix}$$

As $det(A) = det\left(A^T\right)$, taking transpose of a does not make any changes in D_1. However, multiplying by -1 on row/ columns and interchanging the columns will result in multiplying D_1 by $(-1)^3$(as 3 such changes are made). Therefore $D_2 = -D_1$.

Or

Let $D_1 = det \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, *then* $D_2 = det \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Clearly,* $D_1 = -D_2$. *Options a,b and d are incorrect.*

(26) If $ad - bc = 2$ and $ps - qr = 1$, then the determinant

$$\begin{vmatrix} a & b & 0 & 0 \\ 3 & 10 & 2p & q \\ c & d & 0 & 0 \\ 2 & 7 & 2r & s \end{vmatrix}$$

equals . . . . . .

Ans. *We have*

$$\begin{vmatrix} a & b & 0 & 0 \\ 3 & 10 & 2p & q \\ c & d & 0 & 0 \\ 2 & 7 & 2r & s \end{vmatrix} = a \begin{vmatrix} 10 & 2p & q \\ d & 0 & 0 \\ 7 & 2r & s \end{vmatrix} - b \begin{vmatrix} 3 & 2p & q \\ c & 0 & 0 \\ 2 & 2r & s \end{vmatrix}$$

$$= a[-2pds + 2qdr] - b[-2pcs + 2qcr]$$
$$= 2ad[qr - ps] - 2bc[-ps + qr]$$
$$= 2[ad - bc][qr - ps] = -4$$

(27) The determinant $\begin{vmatrix} 1 & 1+x & 1+x+x^2 \\ 1 & 1+y & 1+y+y^2 \\ 1 & 1+z & 1+z+z^2 \end{vmatrix}$ is equal to

(a) $(z-y)(z-x)(y-x)$ (b) $(x-y)(x-z)(y-z)$
(c) $(x-y)^2(y-z)^2(z-x)^2$ (d) $(x^2-y^2)(y^2-z^2)(z^2-x^2)$

Ans. ***Option a***
We have

$$\begin{vmatrix} 1 & 1+x & 1+x+x^2 \\ 1 & 1+y & 1+y+y^2 \\ 1 & 1+z & 1+z+z^2 \end{vmatrix} \sim \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} \sim (y-x)(z-x) \begin{vmatrix} 1 & x & x^2 \\ 0 & 1 & y+x \\ 0 & 1 & z+x \end{vmatrix}$$

Therefore $\begin{vmatrix} 1 & 1+x & 1+x+x^2 \\ 1 & 1+y & 1+y+y^2 \\ 1 & 1+z & 1+z+z^2 \end{vmatrix} = (z-y)(z-x)(y-x).$

(28) Let $a_1, a_2, \dots, a_n \in \mathbb{R}$. Evaluate the determinant:

$$\begin{vmatrix} 1+a_1 & a_2 & \cdots & a_n \\ a_1 & 1+a_2 & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_1 & a_2 & \cdots & 1+a_n \end{vmatrix}$$

Ans. *We have*

$$\begin{aligned} \begin{vmatrix} 1+a_1 & a_2 & \cdots & a_n \\ a_1 & 1+a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & 1+a_n \end{vmatrix} &= \begin{vmatrix} 1+a_1+a_2+\cdots+a_n & a_2 & \cdots & a_n \\ 1+a_1+a_2+\cdots+a_n & 1+a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ 1+a_1+a_2+\cdots+a_n & a_2 & \cdots & 1+a_n \end{vmatrix} \\ &= (1+a_1+a_2+\cdots+a_n) \begin{vmatrix} 1 & a_2 & \cdots & a_n \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} \\ &= 1+a_1+a_2+\cdots+a_n \end{aligned}$$

(29) Let $f_1(x), f_2(x), g_1(x), g_2(x)$ be differentiable functions on $\mathbb{R}$. Let $F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix}$ be the determinant of the matrix $\begin{bmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{bmatrix}$. Then $F'(x)$ is equal to

(a) $\begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & g_1'(x) \\ f_2'(x) & g_2(x) \end{vmatrix}$ (b) $\begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & g_1'(x) \\ f_2(x) & g_2'(x) \end{vmatrix}$

(c) $\begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} - \begin{vmatrix} f_1(x) & g_1'(x) \\ f_2(x) & g_2'(x) \end{vmatrix}$ (d) $\begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}$

Ans. ***Option b***

We have

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix} = f_1(x)g_2(x) - g_1(x)f_2(x)$$

Then

$$\begin{aligned} F'(x) &= f_1(x)g_2'(x) + f_1'(x)g_2(x) - g_1(x)f_2'(x) - g_1'(x)f_2(x) \\ &= f_1'(x)g_2(x) - g_1(x)f_2'(x) + f_1(x)g_2'(x) - g_1'(x)f_2(x) \\ &= \begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & g_1'(x) \\ f_2(x) & g_2'(x) \end{vmatrix} \end{aligned}$$

(30) Let $A = \left((a_{i_j})\right)$ be a 3×3 complex matrix. Identify correct statements.
(a) $det\left((-1)^{i+j}A\right) = det(A)$ (b) $det\left((-1)^{i+j}A\right) = -det(A)$
(c) $det\left(\left((\sqrt{-1})^{i+j}\right)A\right) = det(A)$ (d) $det\left(\left((\sqrt{-1})^{i+j}\right)A\right) = -det(A)$

Ans. ***Options a and c***
Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. *Then*
$det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$.
Let $B = \begin{bmatrix} a_{11} & -a_{12} & a_{13} \\ -a_{21} & a_{22} & -a_{23} \\ a_{31} & -a_{32} & a_{33} \end{bmatrix}$. *Then*

$$\begin{aligned} det(B) &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) + a_{12}(-a_{21}a_{33} + a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= det(A) \end{aligned}$$

Let $C = \begin{bmatrix} -a_{11} & -ia_{12} & a_{13} \\ -ia_{21} & a_{22} & ia_{23} \\ a_{31} & ia_{32} & -a_{33} \end{bmatrix}$. *Then*

$$\begin{aligned} det(C) &= -a_{11}(-a_{22}a_{33} + a_{32}a_{23}) + ia_{12}(ia_{21}a_{33} - ia_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \\ &= det(A) \end{aligned}$$

(31) The number of matrices in $GL_2(\mathbb{Z}_3)$ with determinant 1 is ......
(a) 24 (b) 60 (c) 20 (d) 30

Ans. ***Option a***
Let $SL_2(\mathbb{Z}_3)$ *denote* 2×2 *matrices over* $\mathbb{Z}_3$ *with determinant 1.*
Define $\phi : GL_2(\mathbb{Z}_3) \to \mathbb{Z}_3$ *by*

$$\phi(A) = det(A)$$

Then $kernel\ (\phi) = SL_2(\mathbb{Z}_3)$ *and hence by* ***First Isomorphism Theorem,***

$$GL_2(\mathbb{Z}_3)/SL_2(\mathbb{Z}_3) \cong \mathbb{Z}_3 \setminus \{0\}$$

This implies $\dfrac{\mathcal{O}(GL_2(\mathbb{Z}_3))}{\mathcal{O}(SL_2(\mathbb{Z}_3))} = 2$. *Also we have*

$$\mathcal{O}(GL_2(\mathbb{Z}_3)) = (3^2 - 1)(3^2 - 3) = 48$$

Therefore $\mathcal{O}(SL_2(\mathbb{Z}_3)) = 24$.

(32) Write down the inverse of the matrix $\begin{bmatrix} 0&1&0&0&0 \\ 1&0&0&0&0 \\ 0&0&0&0&1 \\ 0&0&1&0&0 \\ 0&0&0&1&0 \end{bmatrix}$

Ans. *The given matrix is an elementary matrix. The inverse is its transpose.*

(33) Write down the inverse of the following matrix: $\begin{bmatrix} 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&2&1&0 \\ 0&0&3&0&1 \end{bmatrix}$

Ans. *The given matrix is an* ***atomic triangular matrix****. An atomic triangular matrix is a special type of upper(lower) triangular matrix where all its diagonal entries are 1's and all off-diagonal entries except a single column are zeros. The inverse of an atomic triangular matrix is again an atomic triangular matrix where the signs of the entries in the non-zero column are reversed. Therefore the inverse of the given matrix is* $\begin{bmatrix} 1&0&0&0&0 \\ 0&1&0&0&0 \\ 0&0&1&0&0 \\ 0&0&-2&1&0 \\ 0&0&-3&0&1 \end{bmatrix}$.

(34) Let $\mathcal{M}$ be the set of all invertible 5×5 matrices with entries 0 and 1. For each $A \in \mathcal{M}$, let $n_1(A)$ and $n_0(A)$ denote the number of 1's and 0's in A, respectively. Then $\min\limits_{A \in \mathcal{M}} |n_1(A) - n_0(A)| =$
(a) 1 (b) 3 (c) 5 (d) 15

Ans. ***Option a***

Consider the matrix $A = \begin{bmatrix} 1&1&1&0&1 \\ 0&1&1&1&0 \\ 0&0&1&1&1 \\ 0&0&0&1&1 \\ 0&0&0&0&1 \end{bmatrix}$. *Then,* $|n_1(A) - n_0(A)| = 1$ *which is the minimum.*

(35) Let α, β, γ be real numbers such that $\beta \neq 0$ and $\gamma \neq 0$. Suppose $A = \begin{bmatrix} \alpha & \beta \\ \gamma & 0 \end{bmatrix}$ and $A^{-1} = A$. Then
(a) $\alpha = 0$ and $\beta\gamma = 1$ (b) $\alpha \neq 0$ and $\beta\gamma = 1$
(c) $\alpha = 0$ and $\beta\gamma = 2$ (d) $\alpha = 0$ and $\beta\gamma = -1$

Ans. ***Option a***
We have

$$A^{-1} = \frac{1}{det(A)} adj\ (A) = \frac{1}{-\beta\gamma}\begin{bmatrix} 0 & -\beta \\ -\gamma & \alpha \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{\gamma} \\ \frac{1}{\beta} & -\frac{\alpha}{\beta\gamma} \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & 0 \end{bmatrix}$$

This implies $\alpha = 0$ *and* $\beta\gamma = 1$.

(36) Let $A = (a_{ij})$ be a 2×2 lower triangular matrix with diagonal entries $a_{11} = 1$ and $a_{22} = 3$. If $A^{-1} = (b_{ij})$, what are the values of b_{11} and b_{22}?

Ans. *The inverse of a* 2×2 *matrix* $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ *is given by* $\frac{1}{det(A)}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$.
The determinant of a lower triangular matrix is product of its diagonal entries. Hence $det(A) = 3$. *Therefore* $b_{11} = 1$ *and* $b_{22} = \frac{1}{3}$.

(37) Let $A = [a_{ij}]$ be an $n \times n$ matrix such that a_{ij} is an integer for all i, j. Let $AB = I$ with $B = [b_{ij}]$, where I is the identity matrix. Which of the following statements is true?
(a) If $det(A) = 1$ then $det(B) = 1$.
(b) A sufficient condition for each b_{ij} to be an integer is that $det(A)$ is an integer.
(c) B is always an integer matrix.
(d) A necessary condition for each b_{ij} to be an integer is $det(A) \in \{+1, -1\}$.

Ans. ***Options a and d***
(a) $AB = I \Rightarrow det(AB) = det(A)det(B) = 1$. *Therefore if* $det(A) = 1$ *then* $det(B) = 1$.
(b) & (c) Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ *and* $B = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$. *Then* $AB = I$ *and hence options b and c are false.*
(d) Let $det(A) \in \{+1, -1\}$. *Since A and B are square matrices* $AB = I$ *implies B is the inverse of A. Therefore* $B = \frac{1}{det(A)} Adj(A) = \pm Adj(A)$. *Since A is an integer matrix,* $Adj(A)$ *is an integer matrix and hence B is also an integer matrix.*

(38) Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Which one of the following matrices are of the form $P^T AP$ for a suitable 2×2 invertible matrix P over $\mathbb{Q}$?
(a) $\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (d) $\begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$

Ans. ***Options a, c and d***
Let $P^T AP = B$, *then applying determinant on both sides, we get*

$$det(P^T AP) = (det(P))^2(det(A)) = det(B)$$

As $det(A)$ is negative, $det(B)$ must also be negative. Therefore option b is false. Now let $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $P^T AP = \begin{pmatrix} 2ab & ad+bc \\ ad+bc & 2cd \end{pmatrix}$. From this, if we take $P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, then $P^T AP = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$. If we take $P = \begin{pmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & -1 \end{pmatrix}$, then $P^T AP = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. If we take $P = \begin{pmatrix} 1 & \frac{3}{2} \\ 1 & \frac{5}{2} \end{pmatrix}$, then $P^T AP = \begin{pmatrix} 3 & 4 \\ 4 & 5 \end{pmatrix}$.

(39) Given the permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 2 & 5 & 4 \end{pmatrix}$, the matrix A is defined to be the one whose ith column is the $\sigma(i)$th column of the identity matrix I. Which of the following is correct?
(a) $A = A^{-2}$ (b) $A = A^{-4}$ (c) $A = A^{-5}$ (d) $A = A^{-1}$

Ans. ***Option c***
Since σ is of order 6, $\sigma^6 = e$, where σ is the identity permutation and hence $\sigma = \sigma^{-5}$. Hence, $A = A^{-5}$.

(40) The matrix $A = \begin{pmatrix} 5 & 9 & 8 \\ 1 & 8 & 2 \\ 9 & 1 & 0 \end{pmatrix}$ satisfies
(a) A is invertible and the inverse has all integer entries.
(b) $det(A)$ is odd.
(c) $det(A)$ is divisible by 13.
(d) $det(A)$ has at least two prime divisors.

Ans. ***Options c and d***
We have $det(A) = -416 = 2^5 \times 13$. Therefore A is invertible but the inverse does not have integer entries.

(41) Let A, B be $n \times n$ matrices such that $BA + B^2 = I - BA^2$ where I is the $n \times n$ identity matrix. Which of the following is always true?
(a) A is non-singular (b) B is non-singular
(c) $A + B$ is non-singular (d) AB is non-singular

Ans. ***Option b***
(a) & (d) Let $A = 0$ and $B = I$, then $BA + B^2 = I - BA^2$, but A and AB are singular.
(b) Now $BA + B^2 = I - BA^2 \Rightarrow B(A^2 + A + B) = I$. Taking determinant on both sides,

$$det(B)det(A^2 + A + B) = 1 \Rightarrow det(B) \neq 0$$

Therefore B is non-singular.
(c) Let $A = -I$ and $B = I$, then $BA + B^2 = I - BA^2$, but $A + B$ is singular.

(42) Let S denote the set of all prime numbers p such that the following matrix is invertible when considered as a matrix with entries in $\mathbb{Z}_p$.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & -1 \\ -2 & 0 & 2 \end{pmatrix}$$

which of the following statements are true?
(a) S contains all the prime numbers.
(b) S contains all the prime numbers greater than 10.
(c) S contains all the prime numbers other than 2 and 5.
(d) S contains all the odd prime numbers.

Ans. ***Options b and c***
Since the determinant of the given matrix is $det(A) = 10 = 2 \times 5$, the matrix is invertible when considered as a matrix with entries from other than $\mathbb{Z}_2$ and $\mathbb{Z}_5$.

(43) Let S denote the set of all the prime numbers p with the property that the matrix $\begin{bmatrix} 91 & 31 & 0 \\ 29 & 31 & 0 \\ 79 & 23 & 59 \end{bmatrix}$ has an inverse in the field $\mathbb{Z}_p$. Then
(a) $S = \{31\}$ (b) $S = \{31, 59\}$ (c) $S = \{7, 13, 59\}$ (d) S is infinite

Ans. ***Option d***
Determinant of the given matrix is $2 \times 31 \times 31 \times 59$ is zero only when $p = 2, 31$ or 59. Therefore S is infinite.

(44) Let m, n, r be natural numbers. Let A be an $m \times n$ matrix with real entries such that $(AA^t)^r = I$, where I is the $m \times m$ identity matrix and A^t is the transpose of the matrix A. we conclude that
(a) $m = n$ (b) AA^t is invertible
(c) A^tA is invertible (d) if $m = n$, then A is invertible

Ans. ***Options b and d***
Consider the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Then $A^t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $AA^t = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $A^tA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Therefore m need not be equal to n and A^tA need not be invertible.
Since $(AA^t)^r = I$, $det(AA^t) \neq 0$. Therefore AA^t is invertible and if $m = n$, then A is invertible.

(45) Let A, B be $n \times n$ real matrices such that $det(A) > 0$ and $det(B) < 0$. For $0 \leq t \leq 1$, consider $C(t) = tA + (1 - t)B$. Then
(a) $C(t)$ is invertible for each $t \in [0, 1]$.
(b) There is a $t_0 \in (0, 1)$ such that $C(t_0)$ is not invertible.
(c) $C(t)$ is not invertible for each $t \in [0, 1]$.
(d) $C(t)$ is invertible for only finitely many $t \in [0, 1]$.

Ans. ***Option b***
Define $f : [0, 1] \to \mathbb{R}$ *by* $f(t) = det\,[C(t)]$. *Then*

$$f(0) = det\,[C(0)] = det(B) < 0 \text{ and } f(1) = det\,[C(1)] = det(A) > 0$$

Since determinant is a continuous function, f *is continuous. Then by* ***Intermediate Value Theorem***[1] *there exists* $t_0 \in (0, 1)$ *such that* $C(t_0) = 0$. *Thus,* $C(t_0)$ *is not invertible.*

Or

Take $n = 3$, $A = I$ *and* $B = -I$. *Then* $det(A) > 0$ *and* $det(B) < 0$ *and* $C(t) = (2t - 1)I$. $C(t)$ *is invertible for every* $t \in [0, 1]$ *except for* $t = \frac{1}{2}$. *Option* a, c *and* d *are false.*

(46) The rank of the matrix $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 2 & 5 & 6 & 4 \\ 2 & 6 & 8 & 5 \end{bmatrix}$ is ......

Ans. *We have* $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 2 & 5 & 6 & 4 \\ 2 & 6 & 8 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Therefore* $Rank(A) = 3$.

(47) Let

$A = \begin{bmatrix} 1 & 2 & 0 & 2 \\ -1 & -2 & 1 & 1 \\ 1 & 2 & -3 & -7 \\ 1 & 2 & -2 & -4 \end{bmatrix}$. Then $Rank(A)$ equals
(a) 4 (b) 3 (c) 2 (d) 1

Ans. ***Option c***
We have $\begin{bmatrix} 1 & 2 & 0 & 2 \\ -1 & -2 & 1 & 1 \\ 1 & 2 & -3 & -7 \\ 1 & 2 & -2 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Therefore* $Rank(A) = 2$.

(48) Let $A = \begin{bmatrix} 1 & 5 & 7 & 9 \\ 0 & 1 & 3 & 5 \\ 1 & 6 & 10 & 14 \\ 1 & 4 & 4 & \gamma \end{bmatrix}$. Find γ so that the rank of A is two.

[1] **Intermediate Value Theorem**: Let $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. If $k \in \mathbb{R}$ satisfies $f(a) < k < f(b)$, then there exists a point $c \in [a, b]$ such that $f(c) = k$.

Ans. *We have* $\begin{bmatrix} 1 & 5 & 7 & 9 \\ 0 & 1 & 3 & 5 \\ 1 & 6 & 10 & 14 \\ 1 & 4 & 4 & \gamma \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 7 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & -3 & \gamma - 9 \end{bmatrix}$. *Then A has rank 2 if* $\gamma = 4$.

(49) Let $A = \sum_{i=1}^{4} X_i X_i^T$, where $X_1^T = \begin{bmatrix} 1 & -1 & 1 & 0 \end{bmatrix}$, $X_2^T = \begin{bmatrix} 1 & 1 & 0 & 1 \end{bmatrix}$, $X_3^T = \begin{bmatrix} 1 & 3 & 1 & 0 \end{bmatrix}$, $X_4^T = \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix}$. Then $Rank(A)$ equals ……

Ans. *We have,* $A = \begin{bmatrix} 4 & 4 & 3 & 1 \\ 4 & 12 & 3 & 1 \\ 3 & 3 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 4 & 4 & 3 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Thus Rank* $(A) = 3$.

(50) Let A, B be $n \times n$ real matrices. Which of the following statements is correct?
(a) $Rank(A + B) = Rank(A) + Rank(B)$
(b) $Rank(A + B) \leq Rank(A) + Rank(B)$
(c) $Rank(A + B) = min\{Rank(A), Rank(B)\}$
(d) $Rank(A + B) = max\{Rank(A), Rank(B)\}$

Ans. ***Option b***
Take $A = I$ *and* $B = -I$. *Then option* a, c *and* d *are false.*

(51) Let A be a non-zero $n \times n$ real matrix with $n \geq 2$. Which of the following implications is valid?
(a) $det(A) = 0$ implies $Rank(A) = 0$.
(b) $det(A) = 1$ implies $Rank(A) \neq 1$.
(c) $Rank(A) = 1$ implies $det(A) \neq 0$.
(d) $Rank(A) = n$ implies $det(A) \neq 1$.

Ans. ***Option b*** *(a) Consider the matrix* $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. $det(A) = 0$ *but* $Rank(A) \neq 0$. *(b)* $det(A) \neq 0$ *implies A has full rank. Clearly* $Rank(A) \neq 1$. *(c)* $Rank(A) = 1$ *implies* $det(A) = 0$. *(d) Consider the* $n \times n$ *identity matrix.* $Rank(A) = n$ *but* $det(A) = 1$.

(52) Let A be an $n \times m$ matrix with each entry equal to $+1, -1$ or 0 such that every column has exactly one $+1$ and exactly one -1. We can conclude that
(a) Rank$(A) \leq n - 1$ (b) Rank$(A) = m$ (c) $n \leq m$ (d) $n - 1 \leq m$

Ans. ***Option a***

Consider the matrix $\begin{bmatrix} 1 & 1 \\ -1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$. *Then options* b, c *and* d *are false.*

(53) Let A be a 5×5 matrix and let B be obtained by changing one element of A. Let r and s be the ranks of A and B respectively. Which of the following statements is/are correct?
(a) $s \leq r + 1$ (b) $r - 1 \leq s$ (c) $s = r - 1$ (d) $s \neq r$

Ans. ***Options a and b***
By changing one element of A we can either reduce the rank of A by 1, increase by 1 or the rank will be the same. Therefore $r - 1 \leq s \leq r + 1$.

(54) Let $A = \left[a_{ij}\right]$ be a 50×50 matrix, where $a_{ij} = min(i, j); i, j = 1, \dots, 50$. Then the rank of A equals
(a) 1 (b) 2 (c) 25 (d) 50

Ans. ***Option d***

We have $A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & 50 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$. *Therefore* $Rank(A)$ *is 50.*

(55) Let J denote the matrix of order $n \times n$ with all entries 1 and let B be a $(3n) \times (3n)$ matrix given by $B = det \begin{bmatrix} 0 & 0 & J \\ 0 & J & 0 \\ J & 0 & 0 \end{bmatrix}$. Then the rank of B is
(a) $2n$ (b) $3n - 1$ (c) 2 (d) 3

Ans. ***Option d***

By suitable row transformations $B = det \begin{bmatrix} 0 & 0 & K \\ 0 & K & 0 \\ K & 0 & 0 \end{bmatrix}$, *where* $K = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.
Therefore B has rank 3.

(56) Let A_n be the real $n \times n$ matrix $(n \geq 2)$ whose entry in position (i, j) is $i - j$. What is the rank of A_n as a function of n?

Ans. *We have* $A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}, \dots,$
$A_n = \begin{bmatrix} 0 & -1 & \dots & -(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ n-1 & n-2 & \dots & 0 \end{bmatrix}$. *Reducing to row-echelon form we can identify that* A_n *has rank 2 for all* n.

(57) Let x and y in $\mathbb{R}^n$ be non-zero column vectors and $A = xy^t$, where y^t is the transpose of y. Then the rank of A is
(a) 2 (b) 0 (c) at least $\frac{n}{2}$. (d) None of the above.

Ans. ***Option d***
Since x and y are non-zero column vectors in $\mathbb{R}^n$*, both have* $rank$ 1*. Then by Sylvester's inequality* $Rank(A)$ *is less than or equal to 1.*

(58) Let A and B be two real matrices of size 4×6 and 5×4, respectively. If $Rank(B) = 4$ and $Rank(BA) = 2$, then $Rank(A)$ is equal to ......

Ans. *By Sylvester's inequality, we have*

$$4 + Rank(A) - 4 \leq Rank(BA) = 2 \leq min\,\{4, Rank(A)\}$$

This implies that

$$Rank\ (A) \leq 2 \leq Rank\ (A)$$

Therefore, $Rank(A) = 2$*.*

(59) Let A be a $m \times n$ matrix and B be a $n \times m$ matrix over real numbers with $m < n$. Then
(a) AB is always non-singular
(b) AB is always singular
(c) BA is always non-singular
(d) BA is always singular

Ans. ***Option d***
We have

$$Rank(A), Rank(B) \leq min\{m, n\} = m$$

Since BA *is an* $n \times n$ *matrix, by Sylvester's inequality,*

$$Rank(BA) \leq min\{Rank(B), Rank(A)\} \leq m < n$$

Therefore BA *is always singular.*

(60) Let $m, n \in \mathbb{N},\ m < n,\ A \in M_{n\times m}(\mathbb{R}),\ B \in M_{m\times n}(\mathbb{R})$.Then which of the following is(are) NOT possible?
(a) $Rank(AB) = n$
(b) $Rank(BA) = m$
(c) $Rank(AB) = m$
(d) $Rank(BA) = \lceil \frac{m+n}{2} \rceil$, the smallest integer larger than or equal to $\frac{m+n}{2}$.

Ans. ***Option a and d***
The maximum possible rank of both A and B are m since

$$Rank(A), Rank(B) \leq min\,\{m, n\} < n$$

Then by Sylvester's inequality, if $Rank(A) = Rank(B) = m$*,* $Rank(AB) = m$ *and* $Rank(BA) = m$*. Clearly,* $Rank(AB) = n$ *is not possible. Also* $Rank\ (BA) = \lceil \frac{m+n}{2} \rceil$ *is not possible since* $m < \lceil \frac{m+n}{2} \rceil$*.*

(61) Let A be a 3×3 non zero real matrix. If there exists a 3×2 real matrix B and a 2×3 real matrix C such that $A = BC$, then
(a) $Ax = 0$ has a unique solution, where $0 \in \mathbb{R}^3$
(b) there exists $b \in \mathbb{R}^3$ such that $Ax = b$ has no solution
(c) there exists a non zero $b \in \mathbb{R}^3$ such that $Ax = b$ has a unique solution
(d) there exists a non zero $b \in \mathbb{R}^3$ such that $A^T x = b$ has a unique solution

Ans. ***Option b***
We have

$$Rank\ (A) \leq min\,\{Rank\ (B), Rank\ (C)\} \leq 2$$

Therefore A is not invertible and hence the homogeneous system has infinite number of solutions. The system $Ax = b$ has a unique solution if

$$Rank\ [A \mid b] = Rank\ (A) = 3$$

which is not possible. Similarly for the system $A^T x = b$. As we can choose $Rank\ [A \mid b] = 3$ there exists $b \in \mathbb{R}^3$ such that $Ax = b$ has no solution.

(62) If A is a 5×4 matrix with real entries such that $Ax = 0$ if and only if $x = 0$ where x is a 4×1 vector and 0 is a null vector. Then $Rank(A)$ is
(a) 5 (b) 4 (c) 2 (d) 1

Ans. ***Option b***
The system $Ax = b$ has unique solution if and only if

$$Rank[A \mid b] = Rank(A) = number\ of\ unknowns = 4$$

(63) Let A be a $n \times m$ matrix and b be a $n \times 1$ vector(with real entries). Suppose the equation $Ax = b$, $x \in \mathbb{R}^m$ admits a unique solution. Then we can conclude that (a) $m \geq n$ (b) $n \geq m$ (c) $n = m$ (d) $n > m$

Ans. ***Option b***
The system has a unique solution if and only if

$$Rank\,[A \mid b] = Rank(A) = m$$

Also

$$m = Rank(A) \leq min\,\{m, n\} \leq n$$

(64) Let A be an $m \times n$ matrix of rank n with real entries, Choose the correct statement.
(a) $Ax = b$ has a solution for any b.
(b) $Ax = 0$ does not have a solution.
(c) If $Ax = b$ has a solution, then it is unique.
(d) $x^T A = 0$ for some non zero x.

Ans. ***Option c***
If the system is consistent $Rank(A) = n$ *imply unique solution. Consider* $A = I_2$ *and* $x = \begin{bmatrix} a \\ b \end{bmatrix}$ *for option d.*

(65) Let $A = \begin{bmatrix} 2 & 0 & 5 \\ 1 & 2 & 3 \\ -1 & 5 & 1 \end{bmatrix}$. The system of linear equations $Ax = b$ has a solution

(a) only for $b = \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix}$, $x_1 \in \mathbb{R}$.
(b) only for $b = \begin{pmatrix} 0 \\ x_2 \\ 0 \end{pmatrix}$, $x_2 \in \mathbb{R}$.
(c) only for $b = \begin{pmatrix} 0 \\ x_2 \\ x_3 \end{pmatrix}$, $x_2, x_3 \in \mathbb{R}$.
(d) for all $b \in \mathbb{R}^3$.

Ans. ***Option d***
As $\begin{vmatrix} 2 & 0 & 5 \\ 1 & 2 & 3 \\ -1 & 5 & 1 \end{vmatrix} \neq 0$, *the given system has unique solution for any* $b \in \mathbb{R}^3$.

(66) The equations

$$x_1 + 2x_2 + 3x_3 = 1,\ x_1 + 4x_2 + 9x_3 = 1,\ x_1 + 8x_2 + 27x_3 = 1$$

have
(a) only one solution.
(b) two solutions.
(c) infinitely many solutions.
(d) no solutions.

Ans. ***Option a***
The above system of equations can be written in the form of $Ax = b$ *as*

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Since A is invertible the given system has unique solution.

(67) Consider the system of equations

$$x_1 + x_2 + x_3 = 3,\ x_1 - x_2 + x_3 = 4,\ x_1 - 5x_2 + \lambda x_3 = 6$$

Then the value of λ for which this system has an infinite number of solutions is (a) $\lambda = -5$ (b) $\lambda = 0$ (c) $\lambda = 1$ (d) $\lambda = 3$

Ans. ***Option c***

The above system can be represented in the form $Ax = b$ as

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -5 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}$$

Then $[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 3 \\ 1 & -1 & 1 & 4 \\ 1 & -5 & \lambda & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 3 \\ 0 & -2 & 0 & 1 \\ 0 & 0 & \lambda - 1 & 0 \end{bmatrix}$

The given system has an infinite number of solutions, when $Rank[A \mid b] = Rank(A) = 2$. Therefore $\lambda = 1$.

(68) The system of equations:

$$\begin{aligned} 1.x + 2.x^2 + 3.xy + 0.y &= 6 \\ 2.x + 1.x^2 + 3.xy + 1.y &= 5 \\ 1.x - 1.x^2 + 0.xy + 1.y &= 7 \end{aligned}$$

(a) has solutions in rational numbers (b) has solutions in real numbers
(c) has solutions in complex numbers (d) has no solution

Ans. ***Option d***

When we add the first and third equation, we get the LHS of the second equation but the RHS is different. Therefore the system has no solution.

(69) Let $A = \begin{bmatrix} 2 & 0 & 3 & 2 & 0 & -2 \\ 0 & 1 & 0 & -1 & 3 & 4 \\ 0 & 0 & 1 & 0 & 4 & 4 \\ 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$, $b_1 = \begin{bmatrix} 5 \\ 1 \\ 1 \\ 4 \end{bmatrix}$ and $b_2 = \begin{bmatrix} 5 \\ 1 \\ 3 \\ 3 \end{bmatrix}$. Then which of the following are true?

(a) Both systems $Ax = b_1$ and $Ax = b_2$ are inconsistent.
(b) Both systems $Ax = b_1$ and $Ax = b_2$ are consistent.
(c) The system $Ax = b_1 - b_2$ is consistent.
(d) The system $Ax = b_1 - b_2$ is inconsistent.

Ans. ***Options a and c***

The augmented matrix $[A \mid b_1] = \begin{bmatrix} 2 & 0 & 3 & 2 & 0 & -2 & 5 \\ 0 & 1 & 0 & -1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 0 & 4 & 4 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 3 & 2 & 0 & -2 & 5 \\ 0 & 1 & 0 & -1 & 3 & 4 & 1 \\ 0 & 0 & 1 & 0 & 4 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

$\Rightarrow$ Rank $[A \mid b_1] \neq Rank(A)$. Thus, the system $Ax = b_1$ is inconsistent. Similarly, we can show that the system $Ax = b_2$ is inconsistent.

Let $b = b_1 - b_2 = \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}$. *Then the augmented matrix* $[A \mid b] =$

$$\begin{bmatrix} 2 & 0 & 3 & 2 & 0 & -2 & 0 \\ 0 & 1 & 0 & -1 & 3 & 4 & 0 \\ 0 & 0 & 1 & 0 & 4 & 4 & -2 \\ 1 & 1 & 1 & 0 & 1 & 1 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & 3 & 2 & 0 & -2 & 0 \\ 0 & 1 & 0 & -1 & 3 & 4 & 0 \\ 0 & 0 & 1 & 0 & 4 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow Rank\,[A \mid b] = Rank(A).$$

Thus, the system $Ax = b_1 - b_2$ *is consistent.*

(70) Let A be 4×5 real matrix. Consider the system $Ax = b$ of linear equations where x is a 5×1 column matrix of indeterminates and b is some fixed 4×1 column matrix with real entries. Given that

♦ A is row equivalent to the matrix M below (which means that the rows of A are all linear combinations of the rows of M and vice versa), and
♦ C and D below are both solutions to $Ax = b$,

what is the value of y?

$$M = \begin{bmatrix} 1 & -2 & -1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, D = \begin{bmatrix} y \\ 3 \\ 4 \\ 5 \\ 5 \end{bmatrix}$$

Ans. *Since C and D below are both solutions to* $Ax = b$, $y - 25 = -18 \Rightarrow y = 7$.

(71) Consider the system of linear equations

$$x_1 + x_2 + 5x_3 = 3,\ x_1 + 2x_2 + \mu x_3 = 5,\ x_1 + 2x_2 + 4x_3 = \lambda$$

The system is consistent if
(a) $\mu \neq 4$ (b) $\lambda \neq 5$ (c) $\mu = 4$ (d) $\lambda = 5$

Ans. ***Option a and d***
The above system can be represented in the form of $Ax = b$ *as*

$$\begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & \mu \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ \lambda \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 5 & 3 \\ 1 & 2 & \mu & 5 \\ 1 & 2 & 4 & \lambda \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 5 & 3 \\ 0 & 1 & \mu - 5 & 2 \\ 0 & 0 & 4 - \mu & \lambda - 5 \end{bmatrix}$$

(a) When $\mu \neq 4$, $Rank[A \mid b] = Rank(A) = 3$, the system is consistent.
(b) & (c) When $\lambda \neq 5$, $\mu = 4$, $Rank[A \mid b] = 3 \neq Rank(A) = 2$, the system is not consistent.
(d) When $\lambda = 5$, $Rank[A \mid b] = Rank(A) = 2$ or 3 which depends upon the value of μ, then the system is consistent.

(72) Suppose $\alpha, \beta, \gamma \in \mathbb{R}$. Consider the following system of linear equations.

$$x_1 + x_2 + x_3 = \alpha,\ x_1 + \beta x_2 + x_3 = \gamma,\ x_1 + x_2 + \alpha x_3 = \beta$$

If this system has atleast one solution, then which of the following statements is(are) TRUE?

(a) If $\alpha = 1$ then $\gamma = 1$ (b) If $\beta = 1$ then $\gamma = \alpha$
(c) If $\beta \neq 1$ then $\alpha = 1$ (d) If $\gamma = 1$ then $\alpha = 1$

Ans. ***Option a and b***
The above system can be represented in the form of $Ax = b$ as

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & \beta & 1 \\ 1 & 1 & \alpha \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ \gamma \\ \beta \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & \alpha \\ 1 & \beta & 1 & \gamma \\ 1 & 1 & \alpha & \beta \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & \alpha \\ 0 & \beta - 1 & 0 & \gamma - \alpha \\ 0 & 0 & \alpha - 1 & \beta - \alpha \end{bmatrix}$$

(a) If $\alpha = 1$, we have

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & \beta - 1 & 0 & \gamma - 1 \\ 0 & 0 & 0 & \beta - 1 \end{bmatrix}$$

As the given system is consistent, $Rank[A \mid b] = Rank(A)$ for any $\alpha, \beta, \gamma \in \mathbb{R}$. Therefore if $\beta \neq 1$, then the given system does not have a solution as otherwise,

$$Rank[A \mid b] = 3 \neq Rank(A) = 2$$

Therefore $\beta = 1$.Then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & \gamma - 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \gamma = 1$$

(b) If $\beta = 1$, then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & \alpha \\ 0 & 0 & 0 & \gamma - \alpha \\ 0 & 0 & \alpha - 1 & 1 - \alpha \end{bmatrix}$$

If $\gamma \neq \alpha$, then $Rank[A \mid b] \neq Rank(A)$. Now suppose that $\alpha = \gamma$. Then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha - 1 & 1 - \alpha \end{bmatrix}$$

The given system is consistent.
(c) If $\beta \neq 1$ and $\alpha = 1$, then

$$Rank[A \mid b] = 3 \neq Rank(A) = 2$$

Then the system is not consistent.
(d) If $\gamma = \alpha = 1$, then

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & \beta - 1 & 0 & 0 \\ 0 & 0 & 0 & \beta - 1 \end{bmatrix}$$

Clearly, the system need not be consistent.

(73) For two non zero real numbers λ and μ, consider the system of linear equations

$$\begin{bmatrix} \lambda & \mu \\ \mu & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{\mu}{2} \\ \frac{\lambda}{2} \end{bmatrix}$$

Which of the following statements is(are) TRUE?
(a) If $\lambda = \mu$, the solutions of the system lie on the line $x_1 + x_2 = \frac{1}{2}$.
(b) If $\lambda = -\mu$, the solutions of the system lie on the line $x_2 - x_1 = \frac{1}{2}$.
(c) If $\lambda \neq \pm\mu$, the system has no solution.
(d) If $\lambda \neq \pm\mu$, the system has a unique solution.

Ans. ***Options a, b and d***
(a) If $\lambda = \mu$ then the given system written in the form $Ax = b$ is

$$\begin{bmatrix} \mu & \mu \\ \mu & \mu \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{\mu}{2} \\ \frac{\mu}{2} \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} \mu & \mu & \frac{\mu}{2} \\ \mu & \mu & \frac{\mu}{2} \end{bmatrix} \sim \begin{bmatrix} \mu & \mu & \frac{\mu}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Clearly, $Rank[A \mid b] = Rank(A) = 1$. *The system is consistent and the solutions of the system lies on* $x_1 + x_2 = \frac{1}{2}$
(b) If $\lambda = -\mu$ *then the given system written in the form* $Ax = b$ *is*

$$\begin{bmatrix} -\mu & \mu \\ \mu & -\mu \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{\mu}{2} \\ \frac{-\mu}{2} \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} -\mu & \mu & \frac{\mu}{2} \\ \mu & -\mu & \frac{-\mu}{2} \end{bmatrix} \sim \begin{bmatrix} -\mu & \mu & \frac{\mu}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

Clearly, $Rank[A \mid b] = Rank(A) = 1$. *The system is consistent and the solutions of the system lies on* $x_2 - x_1 = \frac{1}{2}$.

(c) If $\lambda \neq \pm\mu$, *then* $det(A) = \lambda^2 - \mu^2 \neq 0$. *Therefore*

$$Rank[A \mid b] = Rank(A) = 2 = number\ of\ unknowns$$

Thus the system has a unique solution.

(74) The system of equations

$$x_1 + x_2 + 2x_3 = 2,\ 2x_1 + 3x_2 - x_3 = 5,\ 4x_1 + 7x_2 + \lambda x_3 = 6$$

does NOT have a solution. Then, the value of λ must be equal to ......

Ans. *The given system can be represented in the form* $Ax = b$ *as*

$$\begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & -1 \\ 4 & 7 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$$

We have

$$\begin{bmatrix} 1 & 1 & 2 & 2 \\ 2 & 3 & -1 & 5 \\ 4 & 7 & \lambda & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & 2 \\ 0 & 1 & -5 & 1 \\ 0 & 0 & \lambda + 7 & -5 \end{bmatrix}$$

Therefore, the system has no solution if $\lambda + 7 = 0$. *That is, if* $\lambda = -7$.

(75) Consider the following system of three linear equations in four unknowns x_1, x_2, x_3 and x_4

$$x_1 + x_2 + x_3 + x_4 = 4,\ x_1 + 2x_2 + 3x_3 + 4x_4 = 5,\ x_1 + 3x_2 + 5x_3 + \lambda x_4 = 5$$

If the system has no solutions, then $\lambda = \cdots\cdots$

Ans. *The given system can be represented in the form $Ax = b$ as*

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 5 \end{bmatrix}$$

We have

$$[A \mid b] = \begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & \lambda & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 4 \\ 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 0 & \lambda - 7 & -1 \end{bmatrix}$$

When $\lambda = 7$, $Rank[A \mid b] = 2 \neq Rank(A) = 3$. Therefore the given system has no solution.

(76) Consider the following system of linear equations

$$\lambda x_1 + 2x_2 + x_3 = 0, \ x_2 + 5x_3 = 1, \ \mu x_2 - 5x_3 = -1$$

Which one of the following statements is true?
(a) The system has unique solution for $\lambda = 1, \mu = -1$
(b) The system has unique solution for $\lambda = -1, \mu = 1$
(c) The system has no solution for $\lambda = 1, \mu = 0$
(d) The system has infinitely many solutions for $\lambda = 0, \mu = 0$

Ans. ***Option b***
The above system can be represented in the form $Ax = b$ as

$$\begin{bmatrix} \lambda & 2 & 1 \\ 0 & 1 & 5 \\ 0 & \mu & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} \lambda & 2 & 1 & 0 \\ 0 & 1 & 5 & 1 \\ 0 & \mu & -5 & -1 \end{bmatrix} \sim \begin{bmatrix} \lambda & 2 & 1 & 0 \\ 0 & 1 & 5 & 1 \\ 0 & 0 & -5 - 5\mu & -1 - \mu \end{bmatrix}$$

(a) When $\lambda = 1, \mu = -1$,

$$Rank(A) = Rank[A \mid b] = 2 < \textit{number of unknowns}$$

Therefore the system has infinitely many solution.
(b) When $\lambda = -1, \mu = 1$,

$$Rank(A) = Rank[A \mid b] = 3 = \textit{number of unknowns}$$

Therefore the system has unique solution.
(c) When $\lambda = 1, \mu = 0$,

$$Rank(A) = Rank[A \mid b] = 3 = \textit{number of unknowns}$$

Therefore the system has unique solution.
(d) When $\lambda = 0, \mu = 0$,

$$Rank[A \mid b] = 3 \neq Rank(A) = 2$$

Therefore the system has no solution.

(77) In which case the system of equations

$$x_1 - 2x_2 + x_3 = 3,\ 2x_1 - 5x_2 + 2x_3 = 2,\ x_1 + 2x_2 + \lambda x_3 = \mu$$

has infinite number of solutions? (a) $\lambda = 1, \mu = -19$ (b) $\lambda = -1, \mu = 19$ (c) $\lambda = 2, \mu = 18$ (d) $\lambda = 1, \mu = 19$

Ans. ***Option d***
The above system can be represented in the form $Ax = b$ as

$$\begin{bmatrix} 1 & -2 & 1 \\ 2 & -5 & 2 \\ 1 & 2 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ \mu \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 1 & -2 & 1 & 3 \\ 2 & -5 & 2 & 2 \\ 1 & 2 & \lambda & \mu \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & \lambda - 1 & \mu - 19 \end{bmatrix}$$

(a) When $\lambda = 1, \mu = -19$,

$$Rank[A \mid b] = 3 \neq Rank(A) = 2$$

Therefore the system has no solutions.
(b) When $\lambda = -1, \mu = 19$,

$$Rank[A \mid b] = Rank(A) = 3 = \textit{number of unknowns}$$

Therefore the system has unique solution.
(c) When $\lambda = 2, \mu = 18$,

$$Rank[A \mid b] = Rank(A) = 3 = \textit{number of unknowns}$$

Therefore the system has unique solution.
(d) When $\lambda = 1, \mu = 19,$

$$Rank[A \mid b] = Rank(A) = 2 < \textit{number of unknowns}$$

Therefore the system has infinite number of solutions.

(78) Consider the linear system

$$x_1 + x_2 + 2x_3 = \lambda,\ x_1 + 4x_2 + x_3 = 4,\ 3x_2 - x_3 = \mu$$

If the above system always has a solution then the value of $\lambda + \mu$ is equal to ......

Ans. *The above system can be represented in the form* $Ax = b$ *as*

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 4 & 1 \\ 0 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \lambda \\ 4 \\ \mu \end{bmatrix}$$

Then

$$\begin{bmatrix} 1 & 1 & 2 & \lambda \\ 1 & 4 & 1 & 4 \\ 0 & 3 & -1 & \mu \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 2 & \lambda \\ 0 & 3 & -1 & 4-\lambda \\ 0 & 0 & 0 & \lambda+\mu-4 \end{bmatrix}$$

The given system has a solution when $Rank[A \mid b] = Rank(A) = 2$. *That is, when* $\lambda + \mu - 4 = 0$. *Therefore* $\lambda + \mu = 4$.

(79) The system of equations

$$x_1 + 3x_2 + 2x_3 = \lambda,\ 2x_1 + x_2 - 4x_3 = 4,\ 5x_1 - 14x_3 = 10$$

(a) has unique solution for $\lambda = 2$ (b) has infinitely many solutions for $\lambda = 2$
(c) has no solution for $\lambda = 2$ (d) has unique solution for any $\lambda \neq 2$

Ans. ***Option b***
The above system can be represented in the form $Ax = b$ *as*

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 1 & -4 \\ 5 & 0 & -14 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \lambda \\ 4 \\ 10 \end{bmatrix}$$

Then

$$[A \mid B] = \begin{bmatrix} 1 & 3 & 2 & \lambda \\ 2 & 1 & -4 & 4 \\ 5 & 0 & -14 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 2 & \lambda \\ 0 & -5 & -8 & 4-2\lambda \\ 0 & 0 & 0 & \lambda-2 \end{bmatrix}$$

When $\lambda = 2$, $Rank[A \mid b] = Rank(A) = 2 <$ number of unknowns. Therefore the given system has infinitely many solutions.
When $\lambda \neq 2$, $Rank[A \mid b] = 3 \neq Rank(A) = 2$. Therefore the system has no solution.

(80) Find the value(s) of λ for which the following system of linear equations

$$\begin{pmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

(a) has a unique solution
(b) has infinitely many solutions
(c) has no solution

Ans. *Since the determinant of the given matrix is $\lambda^3 - 3\lambda + 2 = (\lambda - 1)^2(\lambda + 2)$, the given matrix is invertible for all λ other than 1 and -2. When $\lambda = 1$, the given system is*

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Then $Rank[A \mid b] = Rank(A) = 1 <$ number of unknowns. So the system has infinitely many solutions when $\lambda = 1$.
When $\lambda = -2$, the given system is

$$\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Then

$$[A \mid b] = \begin{pmatrix} -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \\ 1 & 1 & -2 & 1 \end{pmatrix} \sim \begin{pmatrix} -2 & 1 & 1 & 1 \\ 0 & -3 & 3 & 3 \\ 0 & 0 & 0 & 6 \end{pmatrix}$$

Clearly, $Rank[A \mid B] = 3 \neq Rank(A) = 2$ and hence the given system has no solution when $\lambda = -2$.

(81) Let $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 2 & 3 & \lambda \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ 3 \\ \mu \end{bmatrix}$. Then the system $Ax = b$ over the real numbers has
(a) no solution whenever $\mu \neq 7$.
(b) an infinite number of solutions whenever $\lambda \neq 2$.
(c) an infinite number of solutions if $\lambda = 2$ and $\mu \neq 7$.
(d) a unique solution if $\lambda \neq 2$.

Ans. ***Option d***
Consider the augmented matrix

$$[A \mid b] = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 3 \\ 2 & 3 & \lambda & \mu \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 5 & \lambda-2 & \mu-2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & \lambda-2 & \mu-7 \end{bmatrix}$$

When $\lambda \neq 2$, *Rank*$[A \mid b] =$ *Rank*$(A) = 3$, *the given system has a unique solution.*
When $\lambda = 2$ *and* $\mu \neq 7$, *Rank*$[A \mid b] = 3 \neq$ *Rank*$(A) = 2$ *implies that the system is inconsistent.*

(82) Consider the system of simultaneous equations

$$2x_1 - 2x_2 - 2x_3 = \alpha, \ -2x_1 + 2x_2 - 3x_3 = \beta, \ 4x_1 - 4x_2 + 5x_3 = \gamma$$

Write down the condition to be satisfied by α, β, γ for this system NOT to have a solution.

Ans. *The above system can be written in the form of* $Ax = b$ *as*

$$\begin{bmatrix} 2 & -2 & -2 \\ -2 & 2 & -3 \\ 4 & -4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 2 & -2 & -2 & \alpha \\ -2 & 2 & -3 & \beta \\ 4 & -4 & 5 & \gamma \end{bmatrix} \sim \begin{bmatrix} 2 & -2 & -2 & \alpha \\ 0 & 0 & -5 & \alpha+\beta \\ 0 & 0 & 0 & \dfrac{-\alpha+9\beta+5\gamma}{9} \end{bmatrix}$$

The system has no solution when Rank$[A \mid b] \neq$ *Rank*(A). *Here Rank*$(A) = 2$. *So when* $-\alpha + 9\beta + 5\gamma \neq 0$ *the system has no solution.*

(83) Let $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ be a singular matrix. Let x_0 and b be vectors in $\mathbb{R}^n$ such that $Ax_0 = b$. Which of the following statements are true?
(a) There exists $y_0 \in \mathbb{R}^n$ such that $A^T y_0 = b$.
(b) There exist infinitely many solutions to the equation $Ax = b$.
(c) If $A^T x = 0$, then it follows that $b^T x = 0$.

Ans. ***Option b and c***
(a) Consider the system

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

Then $\begin{bmatrix} 0 & 2 & 1 \end{bmatrix}^T$ *is a solution for the given system. But the system*

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

does not have solution as the rank of the augmented matrix is 3 and the rank of the coefficient matrix is 2.
(b) Since A is singular,

$$Rank[A \mid b] = Rank(A) < \textit{number of unknowns}$$

Therefore there exist infinitely many solutions to the equation $Ax = b$.
(c) $A^T x = 0 \Rightarrow\ x_0{}^T A^T x = 0 \Rightarrow b^T x = 0$.

(84) The system of equations

$$x_1 - x_2 + 2x_3 = \alpha,\ x_1 + 2x_2 - x_3 = \beta,\ 2x_2 - 2x_3 = \gamma$$

is inconsistent when (α, β, γ) equals
(a) $(2, 2, 0)$ (b) $(0, 3, 2)$ (c) $(2, 2, 1)$ (d) $(2, -1, -2)$

Ans. ***Option c***
The above system can be represented in the form of $Ax = b$ as

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & -1 \\ 0 & 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 1 & -1 & 2 & \alpha \\ 1 & 2 & -1 & \beta \\ 0 & 2 & -2 & \gamma \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 & \alpha \\ 0 & 1 & -1 & \dfrac{\beta - \alpha}{3} \\ 0 & 0 & 0 & \dfrac{3\gamma - 2\beta + 2\alpha}{3} \end{bmatrix}$$

Since the system is inconsistent $Rank[A \mid b] \neq Rank(A) = 2$. Therefore $Rank[A \mid b] = 3$. Hence, $\dfrac{3\gamma - 2\beta + 2\alpha}{3}$ must be non-zero.

(85) Let $u, v \in \mathbb{R}^4$ be such that $u = \begin{bmatrix} 1 & 2 & 3 & 5 \end{bmatrix}^T$ and $v = \begin{bmatrix} 5 & 3 & 2 & 1 \end{bmatrix}^T$. Then the equation $uv^T x = v$ has
(a) infinitely many solutions (b) no solution
(c) exactly one solution (d) exactly two solutions

Ans. ***Option b***
We have

$$[uv^T \mid v] = \begin{bmatrix} 5 & 3 & 2 & 1 & 5 \\ 10 & 6 & 4 & 2 & 3 \\ 15 & 9 & 6 & 3 & 2 \\ 25 & 15 & 10 & 5 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & 3 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 0 & -13 \\ 0 & 0 & 0 & 0 & -24 \end{bmatrix}$$

Then

$$Rank\ [uv^T \mid v] = 2 \neq Rank\ (uv^T) = 1$$

Thus the equation $uv^T x = v$ has no solutions.

(86) If x_1, x_2 and x_3 are real numbers such that $4x_1 + 2x_2 + x_3 = 31$ and $2x_1 + 4x_2 - x_3 = 19$, then the value of $9x_1 + 7x_2 + x_3$
(a) equals $\frac{821}{3}$ (b) equals $\frac{281}{3}$ (c) equals $\frac{182}{3}$ (d) equals $\frac{218}{3}$

Ans. ***Option d***
The given system can be written in the form $Ax = b$ as

$$\begin{bmatrix} 4 & 2 & 1 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 31 \\ 19 \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 4 & 2 & 1 & 31 \\ 2 & 4 & -1 & 19 \end{bmatrix} \sim \begin{bmatrix} 4 & 2 & 1 & 31 \\ 0 & 6 & -3 & 7 \end{bmatrix}$$

and

$$Rank\ [A \mid b] = Rank\ (A) = 2 < \textit{number of unknowns}$$

Therefore the system has infinite number of solutions. The general form of the solutions is given by $z = \lambda$, $y = \dfrac{7 + 3\lambda}{6}$ and $x = \dfrac{43 - 3\lambda}{6}$. If we take $\lambda = 0$, we have $x = \frac{43}{6}$, $y = \frac{7}{6}$, $z = 0$. Then $9x_1 + 7x_2 + x_3 = \frac{218}{3}$.

(87) Let $A = \begin{bmatrix} \alpha & 1 & 1 \\ 1 & \beta & 1 \\ 1 & 1 & \gamma \end{bmatrix}$, $\alpha\beta\gamma = 1, \alpha, \beta, \gamma \in \mathbb{R}$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$. Then $Ax = 0$ has infinitely many solutions if $tr(A)$ is ......

Ans. *The homogeneous system has infinitely many solutions when $det(A) = 0$. Here*

$$det(A) = 3 - (\alpha + \beta + \gamma) = 0 \Rightarrow tr(A) = \alpha + \beta + \gamma = 3$$

(88) Let a unit vector $v = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$ be such that $Av = 0$ where
$A = \begin{bmatrix} \frac{5}{6} & \frac{-1}{3} & \frac{-1}{6} \\ \frac{-1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{-1}{6} & \frac{-1}{3} & \frac{5}{6} \end{bmatrix}$. Then the value of $\sqrt{6}(|v_1| + |v_2| + |v_3|)$ equals ......

Ans. *We have*

$$A = \begin{bmatrix} \frac{5}{6} & \frac{-1}{3} & \frac{-1}{6} \\ \frac{-1}{3} & \frac{1}{3} & \frac{-1}{3} \\ \frac{-1}{6} & \frac{-1}{3} & \frac{5}{6} \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 5 \end{bmatrix}$$

Then

$$\begin{bmatrix} 5 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 5 \end{bmatrix} \sim \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -2 \\ -1 & -2 & 5 \end{bmatrix} \Rightarrow v_2 = 2v_3,\, v_1 = v_3$$

Since v is a unit vector, $\sqrt{|v_1^2| + |v_2^2| + |v_3^2|} = 1$. *Consider* $v = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T$. *Then* $\sqrt{6}(|v_1| + |v_2| + |v_3|) = 4$.

(89) Let p be a prime and consider the field $\mathbb{Z}_p$. List the primes for which the following system of linear equations DOES NOT have a solution in $\mathbb{Z}_p$:

$$5x_1 + 3x_2 = 4$$

$$3x_1 + 6x_2 = 1$$

Ans. *The system can be written as*

$$\begin{bmatrix} 5 & 3 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

Then

$$[A \mid b] = \begin{bmatrix} 5 & 3 & 4 \\ 6 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 5 & 3 & 4 \\ 0 & 21 & -7 \end{bmatrix}$$

The system of equations $Ax = b$ *is inconsistent when* $Rank[A \mid b] \neq Rank(A)$. *If* $Rank(A) = 2$ *then clearly* $Rank[A \mid b] = Rank(A)$. *So here* $Rank(A)$ *must be equal to 1. Hence* $det(A) = 0 \Rightarrow p = 3$ *or* $p = 7$. *When* $p = 7$, $Rank[A \mid b] = Rank(A) = 1$ *and when* $p = 3$, $Rank[A \mid b] = 2 \neq Rank(A) = 1$. *Therefore* $p = 3$ *is the only possibility.*

(90) Check whether the following statements are true or false.

(a) If A and B are 3×3 matrices and A is invertible, then there exists an integer n such that $A + nB$ is invertible.

(b) The 10×10 matrix $\begin{pmatrix} v_1 w_1 & \cdots & v_1 w_{10} \\ v_2 w_1 & \cdots & v_2 w_{10} \\ \vdots & \ddots & \vdots \\ v_{10} w_1 & \cdots & v_{10} w_{10} \end{pmatrix}$ has rank 2, where $v_i, w_i \neq 0 \in \mathbb{C}$.

(c) Let S be the set of all $n \times n$ real matrices whose entries are only 0, 1, or 2. Then the average determinant of a matrix in S is greater than or equal to 1.

(d) Every 2×2 matrix over $\mathbb{C}$ is a square of some matrix.
(e) For all positive integers m and n, if A is an $m \times n$ real matrix, and B is an $n \times m$ real matrix such that $AB = I$, then $BA = I$.
(f) Suppose $A_1, \ldots, A_m$ are distinct $n \times n$ real matrices such that $A_i A_j = 0$ for all $i \neq j$. Then $m \leq n$.
(g) Let $A, B, C \in \mathbb{M}_{3\times 3}(\mathbb{R})$ be such that A commutes with B, B commutes with C and B is not a scalar matrix. Then A commutes with C.
(h) Let $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$ be such that $A + B = AB$. Then $AB = BA$.
(i) Suppose A, B, C are 3×3 real matrices with $Rank(A) = 2$, $Rank(B) = 1$,$Rank(C) = 2$. Then $Rank(ABC) = 1$.

Ans. *(a) True.* $det(A + xB)$ *is a polynomial of degree 3. It is either zero for all* x *or equal to zero for all but finitely many* x. *When* $x = 0$, $det(A + xB) \neq 0$ *as* A *is invertible. So there exists an integer n such that* $A + nB$ *is invertible.*
(b) False. The ith *row of the given matrix can be obtained by multiplying* $\dfrac{v_i}{v_1}$ *to the first row. So the given matrix is of rank 1.*
(c) False. The given set contains 3^{n^2} *elements as each entry has three possibilities. Now suppose that* $A \in S$ *and* $det(A) = \lambda$ *then by interchanging one row we get another matrix in* S *which has determinant* $-\lambda$. *Then the average determinant of a matrix in* S *is zero.*
(d) False. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *Suppose that there exists a matrix* $B = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ *such that* $B^2 = A$. *This gives,*

$$\begin{bmatrix} a_{11}^2 + a_{12}a_{21} & a_{12}(a_{11} + a_{22}) \\ a_{21}(a_{11} + a_{22}) & a_{21}a_{12} + a_{22}^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Then $a_{11}^2 + a_{12}a_{21} = a_{21}a_{12} + a_{22}^2 = 0$ *which implies that* $a_{11} = \pm a_{22}$. *And* $a_{12}(a_{11} + a_{22}) = 1 \Rightarrow a_{11} = a_{22}$ *for if* $a_{11} = -a_{22}$ *then* $a_{12}(a_{11} + a_{22}) = 0$. *Now* $a_{21}(a_{11} + a_{22}) = 2a_{21}a_{11} = 0 \Rightarrow a_{21} = 0$ *or* $a_{11} = 0$. *Now suppose* $a_{21} = 0$, *then* $a_{21}a_{12} + a_{22}^2 = 0 \Rightarrow a_{22} = a_{11} = 0$. *But this is not possible since* $a_{12}(a_{11} + a_{22}) = 1$. *By the same reason* $a_{11} = 0$ *is not possible. So there does not exist a matrix* B *such that* $B^2 = A$.
(e) False. Consider the matrix $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, *Then* $AB = [1] = I_{1\times 1}$. *But* $BA = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} \neq I$.
(f) False. Consider the set $\left\{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$. *Then* $A_i A_j = 0$ *for all* $i \neq j$. *But* $m > n$.
(g) False. Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. *Then* A *commutes with* B, B *commutes with* C *and* B *is not a scalar matrix. But* A *does not commutes with* C.

(h) True. For, we have

$$\begin{aligned} A + B = AB &\Rightarrow A + (B - I) + I = AB \\ &\Rightarrow I = AB - A - (B - I) \\ &\Rightarrow I = A(B - I) - (B - I) \\ &\Rightarrow I = (A - I)(B - I) \end{aligned}$$

i.e., $(A - I)$ *and* $(B - I)$ *are inverses of each other. Now*

$$\begin{aligned} I = (B - I)(A - I) &\Rightarrow I = BA - B - A + I \\ &\Rightarrow BA = A + B \\ &\Rightarrow AB = BA \end{aligned}$$

(i) False. Consider the matrices $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, *and* $C = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$.

Then ABC *is the zero matrix.*

Chapter 9
Solved Problems—Vector Spaces

(1) Let $p(x) = \alpha x^2 + \beta x + \gamma$ be a polynomial, where $\alpha, \beta, \gamma \in \mathbb{R}$. Fix $x_0 \in \mathbb{R}$. Let $S = \{(a, b, c) \in \mathbb{R}^3 \mid p(x) = a(x - x_0)^2 + b(x - x_0) + c \; \forall \, x \in \mathbb{R}\}$. Then the number of elements in S is
(a) 0 (b) 1 (c) strictly greater than 1 but finite (d) infinite

Ans. ***Option b***
Fix $x_0 \in \mathbb{R}$. Then

$$p(x) = a(x - x_0)^2 + b(x - x_0) + c \Rightarrow \alpha x^2 + \beta x + \gamma = a(x - x_0)^2 + b(x - x_0) + c$$
$$\Rightarrow a = \alpha,\; b = \beta + 2\alpha x_0,\; c = \gamma + (\beta + 2\alpha x_0)x_0 + x_0^2$$

(2) If V is a vector space over the field $\mathbb{Z}_5$ and $dim_{\mathbb{Z}_5}(V) = 3$, then V has
(a) 125 elements. (b) 15 elements. (c) 243 elements. (d) None of the above.

Ans. ***Option a***
Let $\{v_1, v_2, v_3\}$ be a basis of V. Then

$$V = \{a_1 v_1 + a_2 v_2 + a_3 v_3 : a_1, a_2, a_3 \in \mathbb{Z}_5\}$$

Therefore V has 125 elements.

(3) Which of the following matrices has the same row space as the matrix $\begin{pmatrix} 4 & 8 & 4 \\ 3 & 6 & 1 \\ 2 & 4 & 0 \end{pmatrix}$?

(a) $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

Ans. ***Option a***

By suitable elementary transformations, we get $\begin{pmatrix} 4 & 8 & 4 \\ 3 & 6 & 1 \\ 2 & 4 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_9

(4) Let A be an $m \times n$ matrix with rank r. If the linear system $Ax = b$ has a solution for each $b \in \mathbb{R}^m$, then

(a) $m = r$
(b) the column space of A is a proper subspace of $\mathbb{R}^m$.
(c) the null space of A is a non trivial subspace of $\mathbb{R}^m$ whenever $m = n$.
(d) $m \geq n$ implies $m = n$.

Ans. ***Options a and d***
The system $Ax = b$ has a solution for each $b \in \mathbb{R}^m$ implies that the column space of A is $\mathbb{R}^m$ and since $Rank(A) = r$, $m = r$. As $Rank(A) \leq min\{m, n\}$, $m \geq n$ implies $m = n$. If $m = n$ the null space of A is the trivial subspace of $\mathbb{R}^m$.

(5) Let V be the set of 2×2 matrices $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ with complex entries such that $a_{11} + a_{22} = 0$. Let W be the set of matrices in V with $a_{12} + \overline{a_{21}} = 0$. Then, under usual matrix addition and scalar multiplication, which of the following is (are) true?
(a) V is a vector space over $\mathbb{C}$ (b) W is a vector space over $\mathbb{C}$
(c) V is a vector space over $\mathbb{R}$ (d) W is a vector space over $\mathbb{R}$

Ans. ***Options a, c and d***
An element in V is of the form $\begin{bmatrix} a+ib & c+id \\ e+if & -a-ib \end{bmatrix}$. Let $A = \begin{bmatrix} a_1+ib_1 & c_1+id_1 \\ e_1+if_1 & -a_1-ib_1 \end{bmatrix}$, $B = \begin{bmatrix} a_2+ib_2 & c_2+id_2 \\ e_2+if_2 & -a_2-ib_2 \end{bmatrix} \in V$. Then $\lambda A + B \in V$ as

$$\lambda(a_1 + ib_1) + (a_2 + ib_2) + \lambda(-a_1 - ib_1) + (-a_2 - ib_2) = 0$$

for any $\lambda \in \mathbb{C}$(or $\mathbb{R}$) and $A, B \in V$.
An element in W is of the form $\begin{bmatrix} a+ib & c+id \\ -c+id & -a-ib \end{bmatrix}$. Then

$$i\begin{bmatrix} a+ib & c+id \\ -c+id & -a-ib \end{bmatrix} = \begin{bmatrix} ia-b & ic-d \\ -ic-d & -ia+b \end{bmatrix} \notin W$$

But W is a vector space over $\mathbb{R}$.

(6) Let A be an $m \times n$ real matrix and $b \in \mathbb{R}^m$ with $b \neq 0$.

(a) The set of all real solutions of $Ax = b$ is a vector space.
(b) If x_1 and x_2 are two solutions of $Ax = b$, then $\lambda x_1 + (1 - \lambda)x_2$ is also a solution of $Ax = b$ for any $\lambda \in \mathbb{R}$
(c) For any two solutions x_1 and x_2 of $Ax = b$, the linear combination $\lambda x_1 + (1 - \lambda)x_2$ is also a solution of $AX = B$ only when $0 \leq \lambda \leq 1$.
(d) If rank of A is n, then $Ax = b$ has at most one solution

Ans. ***Options b and d***

(a) *Since* 0 *vector is not a solution of the given system Option a is false.*

(b) & (c) x_1 *and* x_2 *are solutions* $\Rightarrow Ax_1 = b$ *and* $Ax_2 = b$. *Now*

$$A(\lambda x_1 + (1-\lambda)x_2) = \lambda Ax_1 + (1-\lambda)Ax_2 = b$$

for any $\lambda \in \mathbb{R}$.

(d) *If the system is consistent* $Rank(A) = n$ *imply unique solution.*

(7) Which of the following is a subspace of the vector space $\mathbb{R}^3$?

(a) $\{(x, y, z) \in \mathbb{R}^3 : x + 2y = 0, 2x + 3z = 0\}$
(b) $\{(x, y, z) \in \mathbb{R}^3 : 2x + 3y + 4z - 3 = 0, z = 0\}$
(c) $\{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0\}$
(d) $\{(x, y, z) \in \mathbb{R}^3 : x - 1 = 0, y = 0\}$

Ans. ***Option a***

(a) *Let*

$$W_1 = \{(x, y, z) \in \mathbb{R}^3 : x + 2y = 0, 2x + 3z = 0\}$$

Take $(x_1, y_1, z_1), (x_2, y_2, z_2) \in A$. *Then* $x_1 + 2y_1 = 0, 2x_1 + 3z_1 = 0$ *and* $x_2 + 2y_2 = 0, 2x_2 + 3z_2 = 0$. *For* $\lambda \in \mathbb{R}$, *we have*

$$\lambda(x_1, y_1, z_1) + (x_2, y_2, z_2) = (\lambda x_1 + x_2, \lambda y_1 + y_2, \lambda z_1 + z_2) \in W_1$$

as

$$\lambda x_1 + x_2 + 2(\lambda y_1 + y_2) = \lambda(x_1 + 2y_1) + (x_2 + 2y_2) = 0$$

and

$$2(\lambda x_1 + x_2) + 3(\lambda z_1 + z_2) = \lambda(2x_1 + 3z_1) + (2x_2 + 3z_2) = 0$$

Therefore W_1 *is a subspace.*

(b) *Let*

$$W_2 = \{(x, y, z) \in \mathbb{R}^3 : 2x + 3y + 4z - 3 = 0, z = 0\}$$

Then W_2 *is not a subspace since* $(0, 0, 0) \notin W_2$.

(c) *Let*

$$W_3 = \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0\}$$

Then $(1, 1, 0) \in W_3$. *Take* $-1 \in \mathbb{R}$. $-1(1, 1, 0) = (-1, -1, 0) \notin W_3$. *Therefore* W_3 *is not a subspace.*

(d) *Let*

$$W_4 = \{(x, y, z) \in \mathbb{R}^3 : x - 1 = 0, y = 0\}$$

Then W_4 is not a subspace since $(0, 0, 0) \notin W_4$.

(8) Which of the following are subspaces of the vector space $\mathbb{R}^3$?
(a) $\{(x, y, z) : x + y = 0\}$ (b) $\{(x, y, z) : x - y = 0\}$
(c) $\{(x, y, z) : x + y = 1\}$ (d) $\{(x, y, z) : x - y = 1\}$

Ans. ***Options a and b***
Since the possible subspaces of $\mathbb{R}^3$ are either $\{0\}$, lines passing through origin, planes passing through origin and $\mathbb{R}^3$, options c and d are not subspaces as they does not pass through origin and a and b represents planes passing through origin.

(9) Which of the following sets of functions from $\mathbb{R}$ to $\mathbb{R}$ is a vector space over $\mathbb{R}$?

$$W_1 = \{f \mid \lim_{x\to 3} f(x) = 0\}, \quad W_2 = \{g \mid \lim_{x\to 3} g(x) = 1\}, \quad W_3 = \{h \mid \lim_{x\to 3} h(x)\ exists\}$$

(a) Only W_1
(b) Only W_2
(c) W_1 and W_3 but not W_2
(d) All the three are vector spaces

Ans. ***Option c***
Let $f_1, f_2 \in W_1$ and $\lambda \in \mathbb{R}$. Since

$$\lim_{x\to 3} \lambda(f_1(x) + f_2(x)) = \lambda \lim_{x\to 3} f_1(x) + \lim_{x\to 3} f_2(x) = 0$$

W_1 is a vector space over $\mathbb{R}$.
Let $h_1, h_2 \in W_3$ and $\lambda \in \mathbb{R}$. Since

$$\lim_{x\to 3} \lambda(h_1(x) + h_2(x)) = \lambda \lim_{x\to 3} h_1(x) + \lim_{x\to 3} h_2(x)$$

exists, W_3 is a vector space over $\mathbb{R}$. Since 0 element does not belong to W_2, it is not a vector space over $\mathbb{R}$.

(10) Consider $\mathbb{M}_{n\times n}(\mathbb{R})$. Among the following subsets of $\mathbb{M}_{n\times n}(\mathbb{R})$, decide which are linear subspaces.

(a) $W_1 = \{A \in \mathbb{M}_{n\times n}(\mathbb{R}) : A\ is\ nonsingular\}$
(b) $W_2 = \{A \in \mathbb{M}_{n\times n}(\mathbb{R}) : det(A) = 0\}$
(c) $W_3 = \{A \in \mathbb{M}_{n\times n}(\mathbb{R}) : tr(A) = 0\}$
(d) $W_4 = \{BA : A \in \mathbb{M}_{n\times n}(\mathbb{R})\}$ where B is some fixed matrix in $\mathbb{M}_{n\times n}(\mathbb{R})$.

Ans. ***Options c and d***

(a) We have $I, -I \in W_1$, where I is the identity matrix. But their sum is the zero matrix which does not belong to W_1. Therefore W_1 is not a linear subspace.

(b) *Take* $A_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$ *(*$a_{ii} = 1 \ \forall \ i = 1, 2, \dots, n-1$, $a_{nn} = 0$ *and all other entries are zero) and* $A_2 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$ *(*$a_{ii} = 0 \ \forall \ i = 1, 2, \dots, n-1$, $a_{nn} = 1$ *and all other entries are zero). Then* $A_1, A_2 \in W_2$ *and* $A_1 + A_2 = I \notin W_2$*. Therefore* W_2 *is not a linear subspace.*

(c) *Let* $A_1, A_2 \in W_3$ *and* $\lambda \in \mathbb{R}$*. Then*

$$tr(\lambda A_1 + A_2) = \lambda tr(A_1) + tr(A_2) = 0$$

Therefore W_3 *is a linear subspace.*

(d) *Let* $BA_1, BA_2 \in W_4$ *where* $A_1, A_2 \in \mathbb{M}_{n\times n}(\mathbb{R})$ *and let* $\lambda \in \mathbb{R}$*. Then*

$$\lambda(BA_1 + BA_2) = B(\lambda A_1 + A_2) \in W_4$$

Therefore W_4 *is a linear subspace.*

(11) Fix $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \in \mathbb{M}_{3\times 3}(\mathbb{R})$. Which of the following are subspaces of $\mathbb{M}_{3\times 3}(\mathbb{R})$?

(a) $W_1 = \{B \in \mathbb{M}_{3\times 3}(\mathbb{R}) : BA = AB\}$
(b) $W_2 = \{B \in \mathbb{M}_{3\times 3}(\mathbb{R}) : B + A = A + B\}$
(c) $W_3 = \{B \in \mathbb{M}_{3\times 3}(\mathbb{R}) : tr(AB) = 0\}$
(d) $W_4 = \{B \in \mathbb{M}_{3\times 3}(\mathbb{R}) : det(AB) = 0\}$

Ans. ***Options a, b and c***

(a) *Let* $B_1, B_2 \in W_1$*, then* $AB_1 = B_1A$ *and* $AB_2 = B_2A$*. Now for any* $\lambda \in \mathbb{R}$,

$$(\lambda B_1 + B_2)A = \lambda B_1 A + B_2 A = \lambda AB_1 + AB_2 = A(\lambda B_1 + B_2)$$

Hence, $\lambda B_1 + B_2 \in W_1$ *for any* $B_1, B_2 \in W_1$ *and* $\lambda \in \mathbb{R}$*. Therefore* W_1 *is a subspace.*

(b) *Let* $B_1, B_2 \in W_2$*, then* $B_1 + A = A + B_1$ *and* $B_2 + A = A + B_2$*. Now for any* $\lambda \in \mathbb{R}$*. Then*

$$(\lambda B_1 + B_2) + A = \lambda B_1 + (B_2 + A) = (\lambda B_1 + A) + B_2 = A + (\lambda B_1 + B_2)$$

Hence, $\lambda B_1 + B_2 \in W_2$ *for any* $B_1, B_2 \in W_2$ *and* $\lambda \in \mathbb{R}$*. Therefore* W_2 *is a subspace.*

(c) Let $B_1, B_2 \in W_3$, then $tr\,(AB_1) = tr\,(AB_2) = 0$. Now for any $\lambda \in \mathbb{R}$,

$$tr\,[A\,(\lambda B_1 + B_2)] = tr\,(\lambda AB_1 + AB_2) = \lambda[tr\,(AB_1)] + tr\,(AB_2) = 0$$

Hence, $\lambda B_1 + B_2 \in W_3$ for any $B_1, B_2 \in W_3$ and $\lambda \in \mathbb{R}$. Therefore W_3 is a subspace.

(d) Let $B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then $B_1, B_2 \in W_4$. But $B_1 + B_2 = I \notin W_4$. Therefore W_4 is not a subspace.

(12) Let V be a finite-dimensional vector space and let W_1, W_2 and W_3 be subspaces of V. Which of the following statements are true?

(a) $W_1 \cap (W_2 + W_3) = W_1 \cap W_2 + W_1 \cap W_3$
(b) $W_1 \cap (W_2 + W_3) \subset W_1 \cap W_2 + W_1 \cap W_3$
(c) $W_1 \cap (W_2 + W_3) \supset W_1 \cap W_2 + W_1 \cap W_3$

Ans. *Option c*
Let $V = \mathbb{R}^2$, $W_1 = \{(x, y) \in \mathbb{R}^2 : y = x\}$, $W_2 = x-$ axis, and $W_3 = y-$ axis. Then W_1, W_2 and W_3 are subspaces of $\mathbb{R}^2$. Also

$$W_1 \cap (W_2 + W_3) = W_1 \cap \mathbb{R}^2 = W_1$$

and

$$W_1 \cap W_2 + W_1 \cap W_3 = \{(0, 0)\}$$

Therefore (a) and (b) are false. Now let $v \in W_1 \cap W_2 + W_1 \cap W_3$, then

$$\begin{aligned} v \in W_1 \cap W_2 + W_1 \cap W_3 &\Rightarrow v = v_1 + v_2, v_1 \in W_1 \cap W_2 \text{ and } v_2 \in W_1 \cap W_3 \\ &\Rightarrow v_1, v_2 \in W_1, v_1 \in W_2 \text{ and } v_2 \in W_3 \\ &\Rightarrow v = v_1 + v_2 \in W_1 \text{ and } v = v_1 + v_2 \in W_2 + W_3 \\ &\Rightarrow v \in W_1 \cap (W_2 + W_3) \end{aligned}$$

Therefore $W_1 \cap (W_2 + W_3) \supset W_1 \cap W_2 + W_1 \cap W_3$.

(13) For arbitrary subspaces U, V and W of a finite-dimensional vector space, which of the following hold:

(a) $U \cap (V + W) \subset U \cap V + U \cap W$
(b) $U \cap (V + W) \supset U \cap V + U \cap W$
(c) $(U \cap V) + W \subset (U + W) \cap (V + W)$
(d) $(U \cap V) + W \supset (U + W) \cap (V + W)$

Ans. ***Options b and c***
Consider $\mathbb{R}^2$. *Let* $U = x$-*axis,* $V = y$-*axis and* $W = \{(x, y) \in \mathbb{R}^2 : y = x\}$. *Then* $U \cap (V + W) = U$, *but* $U \cap V + U \cap W = \{0\}$. *Also* $(U \cap V) + W = W$ *and* $(U + W) \cap (V + W) = \mathbb{R}^2$. *Therefore options a and d are false.*

$$\begin{aligned} v \in U \cap V + U \cap W &\Rightarrow v = v_1 + v_2, \text{ where } v_1 \in U \cap V \text{ and } v_2 \in U \cap W \\ &\Rightarrow v_1 \in U, V \text{ and } v_2 \in U, W \\ &\Rightarrow v_1 + v_2 \in U \text{ and } v_1 + v_2 \in V + W \\ &\Rightarrow v \in U \cap (V + W) \end{aligned}$$

Therefore $U \cap (V + W) \supset U \cap V + U \cap W$.

$$\begin{aligned} v \in (U \cap V) + W &\Rightarrow v = v_1 + v_2, \text{ where } v_1 \in (U \cap V) \text{ and } v_2 \in W \\ &\Rightarrow v = v_1 + v_2 \in (U + W) \text{ and } (V + W) \\ &\Rightarrow v \in (U + W) \cap (V + W) \end{aligned}$$

Therefore $(U \cap V) + W \subset (U + W) \cap (V + W)$.

(14) Let $\{v_1, v_2, v_3\}$ is a linearly independent set of vectors in a vector space over $\mathbb{R}$, then which one of the following sets is also linearly independent?

(a) $\{v_1 + v_2 - v_3, 2v_1 + v_2 + 3v_3, 5v_1 + 4v_2\}$
(b) $\{v_1 - v_2, v_2 - v_3, v_3 - v_1\}$
(c) $\{v_1 + v_2 - v_3, v_2 + v_3 - v_1, v_3 + v_1 - v_2, v_1 + v_2 + v_3\}$
(d) $\{v_1 + v_2, v_2 + 2v_3, v_3 + 3v_1\}$

Ans. ***Option d***
Suppose that $\{v_1, v_2, v_3\}$ *is linearly independent. Then*

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0 \Leftrightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$$

Now

$$\begin{aligned} &\alpha(v_1 + v_2 - v_3) + \beta(2v_1 + v_2 + 3v_3) + \gamma(5v_1 + 4v_2) = 0 \\ &\Rightarrow (\alpha + 2\beta + 5\gamma)\, v_1 + (\alpha + \beta + 4\gamma)\, v_2 + (-\alpha + 3\beta)\, v_3 = 0 \end{aligned}$$

Since $\{v_1, v_2, v_3\}$ *is linearly independent,*

$$\alpha + 2\beta + 5\gamma,\ \alpha + \beta + 4\gamma = 0,\ -\alpha + 3\beta = 0$$

We have to check whether $\alpha = \beta = \gamma = 0$. *It is enough to check whether the homogeneous system has unique solution or not. The homogeneous system has unique solution when the determinant of the coefficient matrix is not equal to zero.*

Since $\begin{vmatrix} 1 & 2 & 5 \\ 1 & 1 & 4 \\ -1 & 3 & 0 \end{vmatrix} = 0$, *option a is not correct. Similarly, since* $\begin{vmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{vmatrix} = 0$, *option b is not correct. Since* $\begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{vmatrix} \neq 0$ *option d is correct. For option c we get* 4×3 *coefficient matrix which ensures the existence of infinite number of solutions.*

(15) Let n be an integer, $n \geq 3$, and $u_1, u_2, \ldots, u_n$ be n linearly independent elements in a vector space over $\mathbb{R}$. Set $u_0 = 0$ and $u_{n+1} = u_1$. Define $v_i = u_i + u_{i+1}$ and $w_i = u_{i-1} + u_i$ for $i = 1, 2, \ldots, n$. Then

(a) $v_1, v_2, \ldots, v_n$ are linearly independent, if $n = 2010$.
(b) $v_1, v_2, \ldots, v_n$ are linearly independent, if $n = 2011$.
(c) $w_1, w_2, \ldots, w_n$ are linearly independent, if $n = 2010$.
(d) $w_1, w_2, \ldots, w_n$ are linearly independent, if $n = 2011$.

Ans. ***Options b, c and d***
Let $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ *such that* $\sum_{i=1}^{n} \lambda_i v_i = 0$.

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i v_i = 0 \Rightarrow \sum_{i=1}^{n} \lambda_i (u_i + u_{i+1}) &= \sum_{i=1}^{n} \lambda_i u_i + \sum_{i=1}^{n} \lambda_i u_{i+1} = 0 \\
&\Rightarrow (\lambda_1 + \lambda_n)\, u_1 + \sum_{i=1}^{n-1} (\lambda_i + \lambda_{i+1})\, u_{i+1} = 0 \\
&\Rightarrow \lambda_1 = -\lambda_n \text{ and } \lambda_{i+1} = -\lambda_i \text{ since } \{u_i \mid i = 1, \ldots, n\} \text{ is } LI \\
&\Rightarrow \lambda_1 = -\lambda_n \text{ and } \lambda_i = (-1)^{i-1}\lambda_1 \text{ for } i = 2, 3, \ldots, n
\end{aligned}$$

When n is odd, this gives $\lambda_1 = -\lambda_n = \lambda_n$. *Therefore,* $\lambda_1 = \lambda_n = 0$. *This implies* $\lambda_i = 0 \; \forall \; i = 1, 2, \ldots, n$ *when n is odd. Therefore* $v_1, v_2, \ldots, v_n$ *are linearly independent, if* $n = 2011$ *and linearly dependent, if* $n = 2010$.
Now let $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{R}$ *such that* $\sum_{i=1}^{n} \mu_i w_i = 0$.

$$\begin{aligned}
\sum_{i=1}^{n} \mu_i w_i = 0 \Rightarrow \sum_{i=1}^{n} \mu_i (u_{i-1} + u_i) &= \sum_{i=1}^{n} \mu_i u_{i-1} + \sum_{i=1}^{n} \mu_i u_i = 0 \\
&\Rightarrow \sum_{i=1}^{n-1} (\mu_i + \mu_{i+1})\, u_i + \mu_n u_n = 0 \\
&\Rightarrow \mu_i = 0 \; \forall \; i = 1, 2, \ldots, n
\end{aligned}$$

Therefore $w_1, w_2, \ldots, w_n$ *are linearly independent, if* $n = 2010$ *and if* $n = 2011$.

(16) Let A be a 4×3 real matrix and let $\{e_1, e_2, e_3\}$ be the standard basis of $\mathbb{R}^3$. Which of the following is true?

(a) If $Rank(A) = 1$, then $\{Ae_1, Ae_2\}$ is a linearly independent set.
(b) If $Rank(A) = 2$, then $\{Ae_1, Ae_2\}$ is a linearly independent set.
(c) If $Rank(A) = 2$, then $\{Ae_1, Ae_3\}$ is a linearly independent set.
(d) If $Rank(A) = 3$, then $\{Ae_1, Ae_2\}$ is a linearly independent set.

Ans. ***Option d***

(a) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. *Then A is matrix of Rank* 1. *But* $\{Ae_1, Ae_2\}$ *is not a linearly independent set as* Ae_2 *is the zero vector.*

(b) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. *Then A is matrix of Rank* 2. *But* $\{Ae_1, Ae_2\}$ *is not a linearly independent set as* Ae_2 *is the zero vector.*

(c) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, *Then A is matrix of Rank* 2. *But* $\{Ae_1, Ae_3\}$ *is not a linearly independent set as* Ae_3 *is the zero vector.*

(d) Now consider the matrix B with $\{e_1, e_3\}$ *as its columns. Then B is matrix with Rank* 2. *Then by Sylvester's inequality,*

$$Rank(A) + Rank(B) - 3 \leq Rank(AB) \leq min\,\{Rank(A), Rank(B)\}$$

which implies $Rank(AB) = 2$.*Therefore* $\{Ae_1, Ae_3\}$ *is a linearly independent set.*

(17) Let $A = \begin{bmatrix} 1 & 3 & 5 & \lambda & 13 \\ 0 & 1 & 7 & 9 & \mu \\ 0 & 0 & 1 & 11 & 15 \end{bmatrix}$ where $\lambda, \mu \in \mathbb{R}$. Choose the correct statement.

(a) There exists values of λ and μ for which the columns of A are linearly independent.
(b) There exists values of λ and μ for which $Ax = 0$ has $x = 0$ as the only solution.
(c) For all values of λ and μ the rows of A span a three-dimensional subspace of $\mathbb{R}^5$.
(d) There exists values of λ and μ for which $Rank(A) = 2$.

Ans. ***Option c***
Since $\mathbb{R}^3$ cannot have a linearly independent set of 5 vectors, there does not exists values of λ and μ for which the columns of A are linearly independent. Since A contains the sub matrix $\begin{bmatrix} 1 & 3 & 5 \\ 0 & 1 & 7 \\ 0 & 0 & 1 \end{bmatrix}$ of rank 3, $Rank(A) = 3 <$ number of unknowns, the system has infinitely many solutions for any values of λ and μ. Since $Rank(A) = 3$, for all values of λ and μ the rows of A span a three-dimensional subspace of $\mathbb{R}^5$.

(18) Let $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$ be linearly independent. Let $\delta_1 = x_2y_3 - y_2x_3$, $\delta_2 = x_1y_3 - y_1x_3$, $\delta_3 = x_1y_2 - y_1x_2$. If V is the span of x, y, then
(a) $V = \{(u, v, w) : \delta_1 u - \delta_2 v + \delta_3 w = 0\}$
(b) $V = \{(u, v, w) : -\delta_1 u + \delta_2 v + \delta_3 w = 0\}$
(c) $V = \{(u, v, w) : \delta_1 u + \delta_2 v - \delta_3 w = 0\}$
(d) $V = \{(u, v, w) : \delta_1 u + \delta_2 v + \delta_3 w = 0\}$

Ans. ***Option a***
Take $x = (1, 0, 0)$ and $y = (0, 1, 1)$. Then

$$V = span\{x, y\} = \{(u, v, w) \in \mathbb{R}^3 : v = w\}$$

and $\delta_1 = 0$, $\delta_2 = 1$ and $\delta_3 = 1$. Then options b and d are incorrect.
Now take $x = (1, 1, 0)$ and $y = (0, 0, 1)$. Then

$$V = span\{x, y\} = \{(u, v, w) \in \mathbb{R}^3 : u = v\}$$

and $\delta_1 = 1$, $\delta_2 = 1$ and $\delta_3 = 0$. Then options c is incorrect.

(19) Let V denote a vector space over a field $\mathbb{K}$ and with a basis $B = \{e_1, e_2, \ldots, e_n\}$. Let $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{K}$. Let $C = \{\lambda_1 e_1, \lambda_1 e_1 + \lambda_2 e_2, \ldots, \lambda_1 e_1 + \lambda_2 e_2 + \ldots + \lambda_n e_n\}$. Then

(a) C is a linearly independent set implies that $\lambda_i \neq 0$ for every $i = 1, 2, \ldots, n$.
(b) $\lambda_i \neq 0$ for every $i = 1, 2, \ldots, n$ implies that C is a linearly independent set.
(c) The linear span of C is V implies that $\lambda_i \neq 0$ for every $i = 1, 2, \ldots, n$.
(d) $\lambda_i \neq 0$ for every $i = 1, 2, \ldots, n$ implies that the linear span C is V.

Ans. ***Options a, b, c and d***

(a) Suppose that C is a linearly independent. Then $\lambda_1 e_1 \neq 0 \Rightarrow \lambda_1 \neq 0$. Likewise,

$$\lambda_1 e_1 + \lambda_2 e_2 \neq 0 \Rightarrow \lambda_2 \neq 0$$

For if $\lambda_2 = 0$,

$$\lambda_1 e_1 + \lambda_2 e_2 = \lambda_1 e_1$$

which implies C is linearly dependent. Proceeding like this we get $\lambda_i \neq 0$ *for every* $i = 1, 2, \ldots, n$.

(b) Let $\mu_1, \mu_2, \ldots, \mu_n \in \mathbb{K}$ *and consider*

$$\mu_1 \lambda_1 e_1 + \mu_2 (\lambda_1 e_1 + \lambda_2 e_2) + \cdots + \mu_n (\lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n) = 0$$

That is,

$$(\mu_1 + \ldots + \mu_n) \lambda_1 e_1 + (\mu_2 + \cdots + \mu_n) \lambda_2 e_2 + \cdots + \mu_n \lambda_n e_n = 0$$

Since B is a basis,

$$(\mu_1 + \cdots + \mu_n) \lambda_1 = (\mu_2 + \cdots + \mu_n) \lambda_2 = \cdots = \mu_n \lambda_n = 0$$

As $\lambda_i \neq 0$ *for every* $i = 1, 2, \ldots, n$ *this implies* $\mu_n = 0$. *Now,*

$$(\mu_{n-1} + \mu_n) \lambda_{n-1} = 0 \Rightarrow \mu_{n-1} + \mu_n = 0 \Rightarrow \mu_{n-1} = 0$$

Proceeding like this we get $\mu_i = 0$ *for every* $i = 1, 2, \ldots, n$.

(c) Suppose that the linear span of C is V. Since V is of dimension n, every element in C is non-zero. Then $\lambda_1 e_1 \neq 0 \Rightarrow \lambda_1 \neq 0$. *Also,*

$$\lambda_1 e_1 + \lambda_2 e_2 \neq 0 \Rightarrow \lambda_2 \neq 0$$

For if, $\lambda_2 = 0$

$$\lambda_1 e_1 + \lambda_2 e_2 = \lambda_1 e_1$$

which implies C is linearly dependent. Proceeding like this we get $\lambda_i \neq 0$ *for every* $i = 1, 2, \ldots, n$.

(d) As $\lambda_i \neq 0$ *for every* $i = 1, 2, \ldots, n$ *implies C is a linearly independent set, we get the linear span of C is V.*

(20) Let $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ be a non-constant polynomial of degree $n \geq 1$. Consider the polynomial

$$q(x) = \int_0^x p(t)dt, r(x) = \frac{d}{dx} p(x)$$

Let V denote the real vector space of all polynomials in x. Then which of the following are true?

(a) q and r are linearly independent in V.
(b) q and r are linearly dependent in V.

(c) x^n belongs to the linear span of q and r.
(d) x^{n+1} belongs to the linear span of q and r.

Ans. ***Option a***
Let $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. *Then* $q(x) = a_0 x + a_1 \frac{x^2}{2} + \cdots + a_n \frac{x^{n+1}}{n+1}$ *and* $r(x) = a_1 + 2a_2 x + + \cdots + n a_n x^{n-1}$. *Since* $q(x)$ *and* $r(x)$ *are polynomials of different degrees they are linearly independent.*
Let $p(x) = x^n + 1$. *Then* $q(x) = \frac{x^{n+1}}{n+1} + x$ *and* $r(x) = n x^{n-1}$. *Clearly both* x^n *and* x^{n+1} *cannot be written as a linear combination of* q *and* r.

(21) (a) Find a value of λ such that the following system of linear equations has no solution: $x + 2y + 3z = 1,\ 3x + 7y + \lambda z = 2,\ 2x + \lambda y + 12z = 3$
(b) Let V be the vector space of all polynomials with real coefficients of degree at most n, where $n \geq 2$. Considering elements of V as functions from $\mathbb{R}$ to $\mathbb{R}$, define

$$W = \left\{ p \in V : \int_0^1 p(x)dx = 0 \right\}$$

Show that W is a subspace of V and $dim\ (W) = n$.

Ans. *(a) The above system can be represented in the form* $Ax = b$ *as*

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 7 & \lambda \\ 2 & \lambda & 12 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Then

$$[A \mid B] = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 3 & 7 & \lambda & 2 \\ 2 & \lambda & 12 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & \lambda - 4 & (\lambda - 4)(\lambda - 9) & \lambda - 4 \\ 0 & 0 & 6 - (\lambda - 4)(\lambda - 9) & 5 - \lambda \end{bmatrix}$$

Now

$$6 - (\lambda - 4)(\lambda - 9) = 0 \Rightarrow \lambda^2 - 13\lambda + 30 = 0 \Rightarrow \lambda = 10 \ or\ \lambda = 3$$

When $\lambda = 10$ *or* $\lambda = 3$, $Rank[A \mid B] = 2$ *and* $Rank(A) = 3$, *the system has no solution.*
(b) Let $p_1(x), p_2(x) \in W$, *then*

$$\int_0^1 (\lambda p_1(x) + p_2(x))dx = 0$$

for any $\lambda \in \mathbb{R}$. *Therefore* W *is a subspace. Consider the set*

$$\left\{x - \frac{1}{2}, x^2 - \frac{1}{3}, x^3 - \frac{1}{4}, \ldots, x^n - \frac{1}{n+1}\right\}$$

which is a subset of W, as

$$\int_0^1 \left(x^k - \frac{1}{k+1}\right) dx = \left[\frac{x^{k+1}}{k+1} - \frac{1}{k+1}\right]_0^1 = 0 \; \forall \, k = 1, \ldots, n$$

Also the set is linearly independent, for

$$\lambda_1 \left(x - \frac{1}{2}\right) + \lambda_2 \left(x^2 - \frac{1}{3}\right) + \lambda_3 \left(x^3 - \frac{1}{4}\right) + \cdots + \lambda_n \left(x^n - \frac{1}{n+1}\right) = 0$$
$$\Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = \cdots = \lambda_n = 0$$

For constant polynomials $\int_0^1 p(x)dx \neq 0$. Therefore $dim\ (W) = n$.

(22) Which of the following sets of vectors form a basis for $\mathbb{R}^3$?

(a) $\{(-1, 0, 0), (1, 1, 1)(1, 2, 3)\}$
(b) $\{(0, 1, 2), (1, 1, 1)(1, 2, 3)\}$
(c) $\{(-1, 1, 0), (2, 0, 0)(0, 1, 1)\}$

Ans. ***Options a and c***
Any linearly independent set which has cardinality as that of dimension is a basis. So it is enough to check whether the set is linearly independent or not.

(a) We have $\begin{vmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = -1 \neq 0$. *Therefore the given set is linearly independent and hence is a basis.*

(b) We have $\begin{vmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 3 \end{vmatrix} = 0$. *Therefore the given set is linearly dependent and hence is not a basis.*

(c) We have $\begin{vmatrix} 2 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = 2 \neq 0$. *Therefore the given set is linearly independent and hence is a basis.*

(23) Which of the following subsets of $\mathbb{R}^4$ is a basis of $\mathbb{R}^4$?

$$B_1 = \{(1, 0, 0, 0), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$$

$$B_2 = \{(1, 0, 0, 0), (1, 2, 0, 0), (1, 2, 3, 0), (1, 2, 3, 4)\}$$

$$B_3 = \{(1, 2, 0, 0), (0, 0, 1, 1), (2, 1, 0, 0), (-5, 5, 0, 0)\}$$

(a) B_1 and B_2 but not B_3 (b) B_1, B_2 and B_3
(c) B_1 and B_3 but not B_2 (d) Only B_1

Ans. ***Option a***
It is enough to check whether which of the above sets are linearly independent. Since the elements of B_1*, when written as the rows of a* 4×4 *matrix forms a lower triangular matrix with non-zero diagonal entries, the matrix is invertible and hence the rows are linearly independent. Similarly,* B_2 *is also a linearly independent set. As,* $5(1,2,0,0) + (-5)(2,1,0,0) = (-5,5,0,0)$*,* B_3 *is linearly dependent.*

(24) A basis of

$$V = \{(x, y, z, w) \in \mathbb{R}^4 : x + y - z = 0,\ y + z + w = 0,\ 2x + y - 3z - w = 0\}$$

is
(a) $\{(1, -1, 0, 1)\}$ (b) $\{(1, 1, -1, 0), (0, 1, 1, 1), (2, 1, -3, 1)\}$
(c) $\{(1, 0, 1, -1)\}$ (d) $\{(1, -1, 0, 1), (1, 0, 1, -1)\}$

Ans. ***Option d***
We have

$$\begin{aligned} V &= \{(x, y, z, w) \in \mathbb{R}^4 : x + y - z = 0,\ y + z + w = 0,\ 2x + y - 3z - w = 0\} \\ &= \{(x, y, z, w) \in \mathbb{R}^4 : x = z - y,\ w = -y - z\} \\ &= span\ \{(1, -1, 0, 1), (1, 0, 1, -1)\} \end{aligned}$$

Clearly the set $\{(1, -1, 0, 1), (1, 0, 1, -1)\}$ *is linearly independent.*

(25) Write down a basis for the following subspace of $\mathbb{R}^4$:

$$V = \left\{(x, y, z, t) \in \mathbb{R}^4 : z = x + y, x + y + t = 0\right\}$$

Ans. *We have*

$$\begin{aligned} V &= \left\{(x, y, z, t) \in \mathbb{R}^4 : z = x + y, x + y + t = 0\right\} \\ &= \left\{(x, y, z, t) \in \mathbb{R}^4 : z = x + y, t = -(x + y)\right\} \\ &= span\ \{(1, 0, 1, -1), (0, 1, 1, -1)\} \end{aligned}$$

Therefore $\{(1, 0, 1, -1), (0, 1, 1, -1)\}$ *is a basis.*

(26) Let $W \subset \mathbb{R}^4$ be the subspace defined by $W = \left\{x \in \mathbb{R}^4 : Ax = 0\right\}$, where $A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$. Write down a basis for W.

Ans. *By suitable elementary transformations we get*

$$A = \begin{bmatrix} 2 & 1 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & 0 & -2 & 6 \\ 0 & 1 & 4 & -3 \end{bmatrix}$$

and hence

$$W = \left\{(x, y, z, w) \in \mathbb{R}^4 : x - z + 3w = 0, y + 4z - 3w = 0\right\}$$

Therefore $\{(-3, 3, 0, 1), (1, -4, 1, 0)\}$ *forms a basis for* W.

(27) Let W be the subspace of $\mathbb{M}_2(\mathbb{R})$ consisting of matrices such that the entries of the first row add up to zero. Write down a basis for W.

Ans. *The general form of a matrix in W is* $\begin{bmatrix} a_{11} & -a_{11} \\ a_{21} & a_{22} \end{bmatrix}$ *where* $a_{11}, a_{21}, a_{22} \in \mathbb{R}$. *Therefore* $\left\{\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ *is a basis for* W.

(28) Let W be the subspace of $\mathbb{M}_2(\mathbb{R})$ consisting of all matrices with trace zero and such that the entries of the first row add up to zero. Write down a basis for W.

Ans. *The general form of a matrix in W is* $\begin{bmatrix} a_{11} & -a_{11} \\ a_{21} & -a_{11} \end{bmatrix}$ *where* $a_{11}, a_{21} \in \mathbb{R}$. *Therefore* $\left\{\begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right\}$ *is a basis for* W.

(29) Let $W \subset \mathbb{M}_2(\mathbb{R})$ be the subspace of all matrices such that the entries of the first column add up to zero. Write down a basis for W.

Ans. *The general form of an element in W is* $\begin{bmatrix} a_{11} & a_{12} \\ -a_{11} & a_{22} \end{bmatrix}$ *where* $a_{11}, a_{12}, a_{22} \in \mathbb{R}$. *Therefore* $\left\{\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ *is a basis for* W.

(30) Let $W = \left\{A \in \mathbb{M}_3(\mathbb{R}) : A = A^T \text{ and } tr(A) = 0\right\}$. Write down a basis for W.

Ans. *The general form of a matrix in W is* $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & -(a_{11} + a_{22}) \end{bmatrix}$. *Then* $\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right\}$ *forms a basis for* W.

(31) Let $\{v_1, v_2, v_3\}$ be be a basis of $\mathbb{R}^3$. Consider the following statements P and Q:
$P : \{v_1 + v_2, v_2 + v_3, v_1 - v_3\}$ is a basis of $\mathbb{R}^3$.
$Q : \{v_1 + v_2 + v_3, v_1 + 2v_2 - v_3, v_1 - 3v_3\}$ is a basis of $\mathbb{R}^3$.
Which of the above statements hold TRUE?
(a) both P and Q (b) only P
(c) only Q (d) Neither P nor Q

Ans. ***Option c***

P: Since

$$v_1 + v_2 - (v_2 + v_3) = v_1 - v_3$$

the given set is linearly dependent and hence is not a basis of $\mathbb{R}^3$.

Q: It is enough to check whether the given set is linearly independent or not, as the given set has cardinality same as the dimension of $\mathbb{R}^3$. *For* $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$,

$$\begin{aligned} &\lambda_1(v_1 + v_2 + v_3) + \lambda_2(v_1 + 2v_2 - v_3) + \lambda_3(v_1 - 3v_3) = 0 \\ &\Rightarrow (\lambda_1 + \lambda_2 + \lambda_3)\, v_1 + (\lambda_1 + 2\lambda_2)\, v_2 + (\lambda_1 - \lambda_2 - 3\lambda_3)\, v_3 = 0 \\ &\Rightarrow (\lambda_1 + \lambda_2 + \lambda_3) = 0,\ (\lambda_1 + 2\lambda_2) = 0,\ (\lambda_1 - \lambda_2 - 3\lambda_3) = 0 \\ &\Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0 \end{aligned}$$

Thus the given set is linearly independent and hence is a basis for $\mathbb{R}^3$.

(32) Let V be a vector space of dimension $d < \infty$, over $\mathbb{R}$. Let W be a vector subspace of V. Let S be a subset of V. Identify which of the following statements is true:

(a) If S is a basis of V then $W \cap S$ is a basis of W.
(b) If $W \cap S$ is a basis of W and $\{s + W \in V \setminus W \mid s \in S\}$ is a basis of $V \setminus W$. Then S is a basis of V.
(c) If S is a basis of W as well as V then the dimension of W is d.

Ans. ***Option c***

(a) Let $V = \mathbb{R}^2$ *and* $W = x{-}$*axis. Then* W *is a subspace of* V. *Consider* $S = \{(1, 1), (1, -1)\}$. *Then clearly* S *is a basis for* $\mathbb{R}^2$. *But* $W \cap S = \phi$ *is not a basis for* W.
(b) Let $V = \mathbb{R}^2$, $W = x{-}$*axis and* $S = \{(1, 0), (1, 1), (2, 1)\}$. *Clearly* $W \cap S = \{(1, 0)\}$ *is a basis of* W. *Also* $\{s + W \in V \setminus W \mid s \in S\}$ *is a basis of* $V \setminus W$. *But* S *is not a basis of* V.
(c) W *and* V *have same basis means they have same dimension. Therefore,* $dim(W) = dim(V) = d$.

(33) Consider the following row vectors:

$$v_1 = (1, 1, 0, 1, 0, 0) \;\; v_2 = (1, 1, 0, 0, 1, 0) \;\; v_3 = (1, 1, 0, 0, 0, 1)$$

$$v_4 = (1, 0, 1, 1, 0, 0) \;\; v_5 = (1, 0, 1, 0, 1, 0) \;\; v_6 = (1, 0, 1, 0, 0, 1)$$

The dimension of the vector space spanned by these row vectors is
(a) 6 (b) 5 (c) 4 (d) 3

Ans. ***Option c***
By suitable elementary transformations, we have

$$\begin{bmatrix} 1&1&0&1&0&0\\ 1&1&0&0&1&0\\ 1&1&0&0&0&1\\ 1&0&1&1&0&0\\ 1&0&1&0&1&0\\ 1&0&1&0&0&1 \end{bmatrix} \sim \begin{bmatrix} 1&1&0&1&0&0\\ 0&0&0&-1&1&0\\ 0&0&0&-1&0&1\\ 1&0&1&1&0&0\\ 0&0&0&0&0&0\\ 0&0&0&0&0&0 \end{bmatrix}$$

Therefore the dimension of the vector space spanned by these row vectors is 4.

(34) Consider the subspace

$$W = \left\{(x_1, x_2, \dots x_{10}) \in \mathbb{R}^{10} : x_n = x_{n-1} + x_{n-2} \text{ for } 3 \leq n \leq 10\right\}$$

of the vector space $\mathbb{R}^{10}$. The dimension W is
(a) 2 (b) 3 (c) 9 (d) 10

Ans. ***Option a***
Since every vector can be written in terms of x_1 and x_2, the dimension W is 2. (For example, $x_4 = x_2 + x_3 = 2x_2 + x_1$)

(35) Consider the real vector space $\mathbb{P}_{2020} = \left\{\sum_{i=0}^{n} a_i x^i : a_i \in \mathbb{R} \textit{ and } 0 \leq n \leq 2020\right\}$.
Let W be the subspace given by

$$W = \left\{\sum_{i=0}^{n} a_i x^i \in \mathbb{P}_{2020} : a_i = 0 \text{ for all odd } i\right\}$$

Then, the dimension of W is ......

Ans. $\mathbb{P}_{2020}$ *is a vector space of dimension 2021. There are 1011 even integers upto 2020 starting from 0. Therefore the dimension of W is 1011.*

(36) Consider the subspace $W = \left\{[a_{ij}] : a_{ij} = 0 \text{ if } i \text{ is even}\right\}$ of all 10×10 real matrices. Then the dimension of W is
(a) 25 (b) 50 (c) 75 (d) 100

Ans. ***Option b***
Since the entries are zero on even rows there are only 50 non-zero entries and no other restrictions are given. Hence the dimension of W is 50.

(37) Let V be the vector space of all 6×6 real matrices over the field $\mathbb{R}$. Then the dimension of the subspace of V consisting of all symmetric matrices is
(a) 15 (b) 18 (c) 21 (d) 35

Ans. ***Option a***
The general form of a 6×6 *symmetric matrix is*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{12} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{13} & a_{23} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{14} & a_{24} & a_{34} & a_{44} & a_{45} & a_{46} \\ a_{15} & a_{25} & a_{35} & a_{45} & a_{55} & a_{56} \\ a_{16} & a_{26} & a_{36} & a_{46} & a_{56} & a_{66} \end{bmatrix}$$

Therefore V is of dimension 21. In general, if V is the vector space of all $n \times n$ *real matrices over the field* $\mathbb{R}$*, the dimension of the subspace of V consisting of all symmetric matrices is* $\frac{n(n+1)}{2}$.

(38) The dimension of the vector space

$$V = \left\{ A = \left(a_{ij}\right)_{n \times n} : a_{ij} \in \mathbb{C}, a_{ij} = -a_{ji} \right\}$$

over the field $\mathbb{R}$ is
(a) n^2 (b) $n^2 - 1$ (c) $\dfrac{n(n-1)}{2}$ (d) $\frac{n^2}{2}$

Ans. ***Option c***
Since $a_{ii} = -a_{ii}\ \forall\, i$, $a_{ii} = 0\ \forall\, i$ *and hence the general form of an element in V*
is $\begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ -a_{12} & 0 & a_{23} & \cdots & a_{2n} \\ -a_{13} & -a_{23} & 0 & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & -a_{3n} & \cdots & 0 \end{bmatrix}$*. So the dimension of V is the number of elements above the main diagonal which is* $\dfrac{n(n-1)}{2}$.

(39) The dimension of the vector space of all symmetric matrices of order $n \times n (n \geq 2)$ with real entries and trace equal to zero is
(a) $\dfrac{(n^2 - n)}{2} - 1$ (b) $\dfrac{(n^2 - 2n)}{2} - 1$
(c) $\dfrac{(n^2 + n)}{2} - 1$ (d) $\dfrac{(n^2 + 2n)}{2} - 1$

Ans. ***Option c***
The general form of an n × n matrix is

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{13} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & -\left(a_{11} + \cdots a_{(n-1)(n-1)}\right) \end{bmatrix}$$

Therefore dimension of the given subspace is $\dfrac{(n^2+n)}{2} - 1$.

(40) The dimensions of the vector space of all symmetric matrices $A = \left(a_{ij}\right)$ of order $n \times n (n \geq 2)$ with real entries, $a_{11} = 0$ and trace zero is
(a) $\dfrac{(n^2+n-4)}{2}$ (b) $\dfrac{(n^2-n+4)}{2}$ (c) $\dfrac{(n^2+n-3)}{2}$ (d) $\dfrac{(n^2-n+3)}{2}$

Ans. ***Option b***
The general form of an n × n matrix is

$$\begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ a_{12} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{13} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & -\left(a_{22} + \cdots + a_{(n-1)(n-1)}\right) \end{bmatrix}$$

Therefore dimension of the given subspace is $\dfrac{(n^2+n-4)}{2}$.

(41) Let

$$W = \left\{ \left(a_{ij}\right) \in M_{4\times 4}(\mathbb{R}) \mid \sum_{i+j=k} a_{ij} = 0, \;\; for\ k = 2, 3, 4, 5, 6, 7, 8, \right\}$$

then $dim(W)$ is
(a) 7 (b) 8 (c) 9 (d) 10

Ans. ***Option c***
The general form of an element in W is

$$\begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & a_{22} & a_{23} & a_{24} \\ -(a_{13}+a_{22}) & a_{32} & a_{33} & a_{34} \\ -(a_{14}+a_{23}+a_{32}) & -(a_{24}+a_{33}) & -a_{34} & 0 \end{bmatrix}$$

Therefore $dim(W) = 9$.

(42) Let $V \subset \mathbb{M}_{n\times n}(\mathbb{R})$ be the subspace of all matrices such that the entries in every row add up to zero and the entries in every column also add up to zero. What is the dimension of V?

Ans. *The general form of a matrix in V is*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & -(a_{11}+\cdots+a_{1(n-1)}) \\ a_{21} & a_{22} & \cdots & -(a_{21}+\cdots+a_{2(n-1)}) \\ \vdots & \vdots & \ddots & \vdots \\ -(a_{11}+\cdots+a_{(n-1)1}) & -(a_{12}+\cdots+a_{(n-1)2}) & \cdots & (a_{11}+\cdots+a_{1(n-1)}+\cdots+a_{(n-1)1}+\cdots+a_{(n-1)(n-1)} \end{bmatrix}.$$

As the last row and column of the above matrix can be represented as a linear combination of other elements, dimension of V is $(n-1)^2$.

(43) Let W be the subset of $\mathbb{M}_{n\times n}(\mathbb{R})$ consisting $\{(a_{ij}) \mid a_{11}+a_{22}+\cdots+a_{nn}=0\}$. Is it true that W is a vector subspace of V over $\mathbb{R}$? If so what is its dimension?

Ans. *W consists of set of all* $n \times n$ *matrices with trace* 0. *Let* $A, B \in W$ *and* $\lambda \in \mathbb{R}$, *then*

$$tr(\lambda A + B) = \lambda tr(A) + tr(B) = 0 \Rightarrow \lambda A + B \in W$$

So W is a vector subspace of V over $\mathbb{R}$. *General form of a matrix in W is* $\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{11} & \dots & -(a_{11}+a_{22}+\cdots+a_{(n-1)(n-1)}) \end{bmatrix}$ *and hence the* $dim(W) = n^2 - 1$.

(44) Consider the following subspace of $\mathbb{R}^3$

$$W = \left\{(x, y, z) \in \mathbb{R}^3 : 2x + 2y + z = 0,\ 3x + 3y - 2z = 0,\ x + y - 3z = 0\right\}$$

The dimension of W is
(a) 0 (b) 1 (c) 2 (d) 3

Ans. ***Option b***
W is the solution space of the system

$$2x + 2y + z = 0,\ 3x + 3y - 2z = 0,\ x + y - 3z = 0$$

The above system can be written in the form $AX = B$ *as*

$$\begin{bmatrix} 2 & 2 & 1 \\ 3 & 3 & -2 \\ 1 & 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Then

$$\begin{bmatrix} 2 & 2 & 1 \\ 3 & 3 & -2 \\ 1 & 1 & -3 \end{bmatrix} \sim \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -7 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore A *has* $rank$ 2 *and hence the dimension of* W *is* $3 - 2 = 1$.

(45) Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ x & y & z \end{bmatrix}$ and let $V = \{(x, y, z) \in \mathbb{R}^3 : det(A) = 0\}$. Then the dimension of V equals
(a) 0 (b) 1 (c) 2 (d) 3

Ans. ***Option c***
We have

$$\begin{aligned} V &= \{(x, y, z) \in \mathbb{R}^3 : det(A) = 0\} \\ &= \{(x, y, z) \in \mathbb{R}^3 : x = y\} \end{aligned}$$

Clearly V *is a two-dimensional subspace.*

(46) Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \\ 1 & 5 & 3 \end{bmatrix}$ and V be the vector space of all $x \in \mathbb{R}^3$ such that $Ax = 0$. Then $dim\ (V)$ is
(a) 0 (b) 1 (c) 2 (d) 3

Ans. ***Option b***
We have

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \\ 1 & 5 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -4 & -2 \\ 1 & 0 & 0 \end{bmatrix}$$

is of $rank$ 2. *Therefore* $dim\ (V) = 3 - 2 = 1$.

(47) Let V be a subspace of $M_{2\times 2}(\mathbb{R})$ defined by

$$V = \left\{ A \in M_{2\times 2}(\mathbb{R}) : A \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 3 & 1 \end{bmatrix} A \right\}$$

Then the dimension of V is ......

Ans. *Let* $A = \begin{bmatrix} x & y \\ z & w \end{bmatrix}$. *Then*

$$A\begin{bmatrix}0 & 2\\ 3 & 1\end{bmatrix} = \begin{bmatrix}0 & 2\\ 3 & 1\end{bmatrix} A \Rightarrow \begin{bmatrix}x & y\\ z & w\end{bmatrix}\begin{bmatrix}0 & 2\\ 3 & 1\end{bmatrix} = \begin{bmatrix}0 & 2\\ 3 & 1\end{bmatrix}\begin{bmatrix}x & y\\ z & w\end{bmatrix}$$
$$\Rightarrow 3y = 2z,\ 2x + y = 2w,\ 3w = 3x + z$$
$$\Rightarrow y = \frac{2}{3}z,\ w = x + \frac{1}{3}z$$

Therefore the dimension of V is 2.

(48) Consider the set $V = \{[x\ y\ z] \in \mathbb{R}^3 : \alpha x + \beta y + z = \gamma,\ \alpha, \beta, \gamma \in \mathbb{R}\}$. For which of the following choice(s) the set V becomes a two dimensional subspace of $\mathbb{R}^3$ over $\mathbb{R}$?

(a) $\alpha = 0,\ \beta = 1,\ \gamma = 0$ (b) $\alpha = 0,\ \beta = 1,\ \gamma = 1$
(c) $\alpha = 1,\ \beta = 0,\ \gamma = 0$ (d) $\alpha = 1,\ \beta = 1,\ \gamma = 0$

Ans. ***Options a, c and d***
When $\alpha = 0,\ \beta = 1,\ \gamma = 0,\ V = \{[x\ y\ z] \in \mathbb{R}^3 : z = -y\}$ *which is a 2 - dimensional subspace.*
When $\alpha = 0,\ \beta = 1,\ \gamma = 1,\ V = \{[x\ y\ z] \in \mathbb{R}^3 : y + z = 1\}$ *which is not a subspace since* $(0, 0, 0) \in V$.
When $\alpha = 1,\ \beta = 0,\ \gamma = 0,\ V = \{[x\ y\ z] \in \mathbb{R}^3 : z = -x\}$ *which is a 2 - dimensional subspace.*
When $\alpha = 1,\ \beta = 1,\ \gamma = 0,\ V = \{[x\ y\ z] \in \mathbb{R}^3 : z = -(x + y)\}$ *which is a 2 - dimensional subspace.*

(49) Let $v_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}$ and $v_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}$. Let A be the matrix whose columns are $v_1, v_2, 2v_1 - v_2, v_1 + 2v_2$ in that order. Then the number of linearly independent solutions of the homogeneous system of linear equations $Ax = 0$ is ......

Ans. v_1 *and* v_2 *are linearly independent vectors and hence A has rank* 2*(as* $2v_1 - v_2$ *and* $v_1 + 2v_2$ *can be written as a linear combination of both* v_1 *and* v_2*). A is of order* 3×4*. So it has 4 unknowns and hence the solution space has dimension* $4 - Rank(A) = 4 - 2 = 2$.

(50) Consider a homogeneous system of linear equations $Ax = 0$, where A is an $m \times n$ real matrix and $n > m$. Then which of the following statements are always true?

(a) $Ax = 0$ has a solution.
(b) $Ax = 0$ has no non-zero solution.
(c) $Ax = 0$ has a non-zero solution.
(d) Dimension of the space of all solution is atleast $n - m$.

Ans. ***Options a, c and d***
A homogeneous system is always consistent. Since Number of unknowns is greater than number of equations the given system has infinite number of solutions and dimension of the space of all solution is atleast $n - m$.

(51) Let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ be a non-zero vector and $A = \dfrac{xx^T}{x^T x}$. Then the dimension of the vector space $\{y \in \mathbb{R}^3 : Ay = 0\}$ over $\mathbb{R}$ is ......

Ans. *Since x and x^T are of rank 1, by Sylvester's inequality $A = \dfrac{xx^T}{x^T x}$ has rank 1. Therefore the solution space of homogeneous system has dimension* $3 - 1 = 2$.

(52) Let W be the subspace of $C[0, 1]$ spanned by $\{sin(x), cos(x), tan(x)\}$. Then the dimension of W over $\mathbb{R}$ is
(a) 1 (b) 2 (c) 3 (d) infinite

Ans. ***Option c***
The set $S = \{sin(x), cos(x), tan(x)\}$ is linearly independent. For, let

$$p(x) = \lambda_1 sin(x) + \lambda_2 cos(x) + \lambda_3 tan(x) = 0 \;\forall\; x \in [0, 1]$$

Then $p(0) = 0 \Rightarrow \lambda_2 = 0$. Now,

$$p\left(\frac{\pi}{6}\right) = 0 \Rightarrow \lambda_1 sin\left(\frac{\pi}{6}\right) + \lambda_3 tan\left(\frac{\pi}{6}\right) = 0 \Rightarrow \frac{1}{2}\lambda_1 + \frac{1}{\sqrt{3}}\lambda_3 = 0$$
$$p\left(\frac{\pi}{4}\right) = 0 \Rightarrow \lambda_1 sin\left(\frac{\pi}{4}\right) + \lambda_3 tan\left(\frac{\pi}{4}\right) = 0 \Rightarrow \frac{1}{\sqrt{2}}\lambda_1 + \lambda_3 = 0$$

Thus we get a system of two equations in two unknowns, namely λ_1 and λ_3. This system has unique solution $\lambda_1 = \lambda_3 = 0$, as the coefficient matrix $\begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 1 \end{bmatrix}$ is invertible. Therefore $S = \{sin(x), cos(x), tan(x)\}$ is linearly independent and hence the dimension of W over $\mathbb{R}$ is 3.

(53) Let V denote the vector space $C^5[a, b]$ over $\mathbb{R}$ and

$$W = \left\{ f \in V : \frac{d^4 f}{dt^4} + 2\frac{d^2 f}{dt^2} - f = 0 \right\}$$

Then
(a) $dim(V) = \infty$ and $dim(W) = \infty$ (b) $dim(V) = \infty$ and $dim(W) = 4$
(c) $dim(V) = 6$ and $dim(W) = 5$ (d) $dim(V) = 5$ and $dim(W) = 4$

Ans. ***Option b***
$C^5[a, b]$ is an infinite-dimensional space and the given differential equation is a fourth order ODE with constant coefficients. It has 4 linearly independent solutions. Therefore $dim(V) = \infty$ and $dim(W) = 4$.

(54) Consider the subspaces

$$W_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = x_2 + 2x_3\}$$

$$W_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 = 3x_2 + 2x_3\}$$

of $\mathbb{R}^3$. Then the dimension of $W_1 + W_2$ equals ......

Ans. *Clearly* $dim(W_1) = 2$ *and* $dim(W_2) = 2$. *Now*

$$W_1 \cap W_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_2 = 0, x_1 = 2x_3\}$$

Therefore $dim(W_1 + W_2) = 2 + 2 - 1 = 3$.

(55) Let

$$W_1 = \{(u, v, w, x) \in \mathbb{R}^4 \mid u + v + w = 0, 2v + x = 0, 2u + 2w - x = 0\}$$

$$W_2 = \{(u, v, w, x) \in \mathbb{R}^4 \mid u + w + x = 0, u + w - 2x = 0, v - x = 0\}$$

Then which among the following is true?
(a) $dim(W_1) = 1$ (b) $dim(W_2) = 2$ (c) $dim(W_1 \cap W_2) = 1$ (d) $dim(W_1 + W_2) = 3$

Ans. ***Option c***
We have

$$\begin{aligned} W_1 &= \{(u, v, w, x) \in \mathbb{R}^4 \mid u + v + w = 0, 2v + x = 0, 2u + 2w - x = 0\} \\ &= \{(u, v, w, x) \in \mathbb{R}^4 \mid v = -(u + w), x = 2(u + w)\} \\ &= span\{(1, -1, 0, 2), (0, -1, 1, 2)\} \end{aligned}$$

$$\begin{aligned} W_2 &= \{(u, v, w, x) \in \mathbb{R}^4 \mid u + w + x = 0, u + w - 2x = 0, v - x = 0\} \\ &= \{(u, v, w, x) \in \mathbb{R}^4 \mid v = x = 0, u = -w\} \\ &= span\{(1, 0, -1, 0)\} \end{aligned}$$

Therefore $dim(W_1) = 2$ *and* $dim(W_2) = 1$. *Since*

$$(1, 0, -1, 0) = (1, -1, 0, 2) - (0, -1, 1, 2)$$

W_2 *is a subspace of* W_1. *Hence* $dim(W_1 \cap W_2) = 1$ *and* $dim(W_1 + W_2) = 2$.

(56) Let W_1, W_2, W_3 be three distinct subspaces of $\mathbb{R}^{10}$ such that each W_i has dimension 9. Let $W = W_1 \cap W_2 \cap W_3$. Then we can conclude that

(a) $dim(W) \leq 3$ (b) $dim(W) \leq 8$
(c) $dim(W) \geq 7$ (d) W may not be a subspace of $\mathbb{R}^{10}$

Ans. ***Options b and c***
Since

$$dim(W_1 \cap W_2) = dim\ W_1 + dim\ W_2 - dim(W_1 + W_2)$$

and the possible dimension of $W_1 + W_2$ *is 10 (since* W_1 *and* W_2 *are distinct), we get* $dim(W_1 \cap W_2) = 8$. *Now*

$$dim(W_1 \cap W_2 \cap W_3) = dim(W_1 \cap W_2) + dim\ W_3 - dim(W_1 \cap W_2 + W_3)$$

and the possible dimension of $W_1 \cap W_2 + W_3$ *are 9 and 10. Therefore the possible dimensions of* W *is 7 and 8.*

(57) Let $v_i = (v_i^{(1)}, v_i^{(2)}, v_i^{(3)}, v_i^{(4)})$ be four vectors in $\mathbb{R}^4$ such that $\sum_{i=1}^4 v_i^{(j)} = 0$, for each $j = 1, 2, 3, 4$. Let W be the subspace of $\mathbb{R}^4$ spanned by $\{v_1, v_2, v_3, v_4\}$. Then the dimension of W over $\mathbb{R}$ is always
(a) either equal to 1 or equal to 4. (b) less than or equal to 3.
(c) greater than or equal to 2. (d) either equal to 0 or equal to 4.

Ans. ***Option b***
Since $\sum_{i=1}^4 v_i^{(j)} = 0$, *for each* $j = 1, 2, 3, 4$, *the set* $\{v_1, v_2, v_3, v_4\}$ *is linearly dependent as we can write* v_4 *as a linear combination of* $\{v_1, v_2, v_3\}$. *Therefore the dimension of* W *over* $\mathbb{R}$ *is always less than or equal to 3.*

(58) Let V be the vector space of all polynomials of degree at most equal to $2n$ with real coefficients. Let V_0 stand for the vector subspace

$$V_0 = \{P \in V : P(1) + P(-1) = 0\}$$

and V_e stand for the subspace of polynomials which have terms of even degree alone. Find $dim(V_0)$ and $dim\ (V_0 \cap V_e)$.

Ans. *Let* $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{2n}x^{2n} \in V$. *Then*

$$\begin{aligned} p(1) + p(-1) = 0 &\Rightarrow a_0 + a_1 + a_2 + \cdots + a_{2n} + a_0 - a_1 + a_2 - \cdots + a_{2n} = 0 \\ &\Rightarrow 2a_0 + 2a_2 + \cdots + 2a_{2n} = 0 \\ &\Rightarrow a_{2n} = -(a_0 + a_2 + \cdots + a_{2n-2}) \end{aligned}$$

Since V *is vector space of dimension* $2n + 1$, V_0 *is a subspace of* V *of dimension* $2n$. *Now* $V_0 \cap V_e$ *contains all those polynomials in* V_0 *which have terms of even degree alone. The dimension of* V_e *is* $n + 1$ *and hence* $dim(V_0 \cap V_e) = n + 1 - 1 = n$.

(59) Let V be the real vector space of all polynomials in one variable with real coefficients and of degree less than, or equal to, 5. Let W be the subspace defined by

$$W = \{p \in V : p(1) = p'(2) = 0\}$$

What is the dimension of W?

Ans. *Let* $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 + a_5x^5 \in V$. *Then,* $p'(x) = a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3 + 5a_5x^4$.

$$\begin{aligned} p'(2) = 0 &\Rightarrow a_1 + 4a_2 + 12a_3 + 32a_4 + 80a_5 = 0 \\ &\Rightarrow a_1 = -(4a_2 + 12a_3 + 32a_4 + 80a_5) \end{aligned}$$

$$\begin{aligned} p(1) = 0 &\Rightarrow a_0 + a_1 + a_2 + a_3 + a_4 + a_5 = 0 \\ &\Rightarrow a_0 = -(a_1 + a_2 + a_3 + a_4 + a_5) \\ &\Rightarrow a_0 = 3a_2 + 11a_3 + 31a_4 + 79a_5 \end{aligned}$$

Therefore dimension of W *is 4.*

(60) Let V be the vector space of all polynomials in one variable with real coefficients and having degree at most 20. Define the subspaces

$$W_1 = \left\{p \in V : p(1) = 0,\ p\left(\frac{1}{2}\right) = 0,\ p(5) = 0,\ p(7) = 0\right\}$$

$$W_2 = \left\{p \in V : p\left(\frac{1}{2}\right) = 0,\ p(3) = 0,\ p(4) = 0,\ p(7) = 0\right\}$$

Then the dimension of $W_1 \cap W_2$ is ......

Ans. *We have*

$$W_1 \cap W_2 = \left\{p \in V : p(1) = p\left(\frac{1}{2}\right) = p(3) = p(4) = p(5) = p(7) = 0\right\}$$

V *is a vector space of dimension 21. Since six independent conditions are given, we can represent 6 coefficients in terms of the remaining 15 coefficients. Therefore the dimension of* $W_1 \cap W_2 = 15$.

(61) If U and V are the null spaces of $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix}$ respectively. Then the dimension of the subspace $U + V$ equals ....

Ans. *Since both matrices have rank 2, by Rank-Nullity theorem, we have* $dim(U) = dim(V) = 2$. *Now the dimension of* $U \cap V$ *is the dimension of the null space of the matrix* $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix}$. *Since this matrix has rank 3,* $dim(U \cap V) = 1$. *Then the dimension of the subspace* $U + V$ *equals 3.*

(62) Let V be the vector space of all 3×3 matrices with complex entries over the real field. If

$$W_1 = \{A \in V : A = A^*\} \text{ and } W_2 = \{A \in V : tr(A) = 0\}$$

then the dimension of $W_1 + W_2$ is equal to ......

Ans. *The general form of an element in* W_1 *is*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \overline{a_{11}} & a_{22} & a_{23} \\ \overline{a_{13}} & \overline{a_{23}} & a_{13} \end{bmatrix}$$

where a_{11}, a_{22}, a_{33} *are real numbers as the diagonal entries of a Hermitian matrix is always real and the other entries are complex numbers. Therefore* $dim(W_1) = 9$.
Now general form of an element in W_2 *is*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & -(a_{11} + a_{22}) \end{bmatrix}$$

where all the entries are complex numbers. Therefore $dim(W_2) = 16$.
Also an element in $W_1 \cap W_2$ *will be of the form*

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \overline{a_{11}} & a_{22} & a_{23} \\ \overline{a_{13}} & \overline{a_{23}} & -(a_{11} + a_{22}) \end{bmatrix}$$

where a_{11}, a_{22} *are real numbers and the other entries are complex numbers. Therefore* $dim(W_1 \cap W_2) = 8$ *Then the dimension of* $W_1 + W_2$ *is* $16 + 9 - 8 = 17$.

(63) Let W_1 be be the real vector space of all 5×2 matrices such that the sum of the entries in each row is zero. Let W_2 be the real vector space of all 5×2 matrices such that the sum of the entries in each column is zero. Then the dimension of the space $W_1 \cap W_2$ is ......

Ans. *The elements of W_1 and W_2 are respectively of the form*

$$\begin{bmatrix} a_{11} & -a_{11} \\ a_{21} & -a_{21} \\ a_{31} & -a_{31} \\ a_{41} & -a_{41} \\ a_{51} & -a_{51} \end{bmatrix} \text{ and } \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \\ -(a_{11}+a_{21}+a_{31}+a_{41}) & -(a_{12}+a_{22}+a_{32}+a_{42}) \end{bmatrix}$$

Therefore elements of $W_1 \cap W_2$ are of the form

$$\begin{bmatrix} a_{11} & -a_{11} \\ a_{21} & -a_{21} \\ a_{31} & -a_{31} \\ a_{41} & -a_{41} \\ -(a_{11}+a_{21}+a_{31}+a_{41}) & (a_{11}+a_{21}+a_{31}+a_{41}) \end{bmatrix}$$

Then the dimension of the space $W_1 \cap W_2$ is 4.

(64) Let V be the vector space of all 2×2 matrices over $\mathbb{R}$. Consider the subspaces

$$W_1 = \left\{ \begin{bmatrix} a & -a \\ c & d \end{bmatrix} : a, c, d \in \mathbb{R} \right\} \text{ and } W_2 = \left\{ \begin{bmatrix} a & b \\ -a & d \end{bmatrix} : a, b, d \in \mathbb{R} \right\}$$

If $m = dim\ (W_1 \cap W_2)$ and $n = dim\ (W_1 + W_2)$, then the pair (m, n) is
(a) $(2, 3)$ (b) $(2, 4)$ (c) $(3, 4)$ (d) $(1, 3)$

Ans. ***Option b***
Clearly, $dim\ (W_1) = 3$ and $dim\ (W_2) = 3$. Every element of $W_1 \cap W_2$ is of the form $\begin{bmatrix} a & -a \\ -a & d \end{bmatrix}$. Therefore $m = dim\ (W_1 \cap W_2) = 2$ and hence

$$n = dim\ (W_1 + W_2) = dim\ (W_1) + dim\ (W_2) - dim\ (W_1 \cap W_2) = 4$$

(65) Let $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix}$. If

$$V = \left\{ (x, y, 0) \in \mathbb{R}^3 : A \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \text{ and } W = \left\{ (x, y, z) \in \mathbb{R}^3 : A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

Then
(a) the dimension of V equals 2 (b) the dimension of W equals 2
(c) the dimension of V equals 1 (d) $V \cap W = \{(0, 0, 0)\}$

Ans. ***Option c***
We have

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix}$$

Then

$$A \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x - y = 0 \Rightarrow x = y$$

Hence, $V = \{(x, y, 0) : x = y, x, y \in \mathbb{R}\}$. *Therefore* V *has dimension 1. Now*

$$A \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x - y + z = 0,\ -2z = 0 \Rightarrow z = 0,\ x = y$$

Thus, $W = \{(x, y, z) : x = y\, z = 0, x, y, z \in \mathbb{R}\} = V$.

(66) Let $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ be such that $A \neq 0$ but $A^2 = 0$. Which of the following statements are true?

(a) If n is even, then $dim(Col(A)) > dim(Null(A))$.
(b) If n is even, then $dim(Col(A)) \leq dim(Null(A))$.
(c) If n is odd, then $dim(Col(A)) < dim(Null(A))$.
(d) If n is odd, then $dim(Col(A)) > dim(Null(A))$.

Ans. ***Options b and c***
Let $v \in Col(A)$. *Then there exists* $u \in \mathbb{R}^n$ *such that* $Au = v$. *Now,*

$$0 = A^2u = A(Au) = Av$$

That is, $v \in Null(A)$. *Therefore* $dim(Col(A)) \leq dim(Null(A))$. *As, rank of A is same as the dimension of the column space of A, by Rank-Nullity Theorem,* $dim(Col(A)) < dim(Null(A))$. *For if,* $dim(Col(A)) = dim(Null(A))$, *n must be even.*

(67) Let V be a vector space of dimension 4 over the field $\mathbb{Z}_3$ with 3 elements. What is the number of one-dimensional vector subspaces of V?

Ans. *V is a vector space with cardinality* $3^4 = 81$. *Consider the non-zero elements in V, each non-zero vector spans a one-dimensional subspace and each subspace has 2 non-zero elements and hence the number of one-dimensional vector subspaces of V is* $\frac{80}{2} = 40$.

(68) Let V be a vector space such that $dim(V) = 5$. Let W_1 and W_2 be subspaces of V such that $dim(W_1) = 3$ and $dim(W_2) = 4$. Write down all possible values of $dim\ (W_1 \cap W_2)$.

Ans. *The possible values of* $dim(W_1 \cap W_2)$ *are 0,1,2 and 3. Suppose* $dim(W_1 \cap W_2) = 0$*, then* $W_1 + W_2$ *will be a subspace of dimension* $4 + 3 = 7$ *which is not possible since V is a vector space of dimension 5. Suppose* $dim\ (W_1 \cap W_2) = 1$*, then* $W_1 + W_2$ *will be a subspace of dimension* $4 + 3 - 1 = 6$ *which is not possible since V is a vector space of dimension 5. But if* $dim\ (W_1 \cap W_2) = 2$*, then* $W_1 + W_2$ *will be a subspace of dimension* $4 + 3 - 2 = 5$*, which is possible. Similarly if* $dim\ (W_1 \cap W_2) = 3$*, then* $W_1 + W_2$ *will be a subspace of dimension* $4 + 3 - 3 = 4$*, which is possible. Therefore the possible dimensions of* $W_1 \cap W_2$ *are 2 and 3.*

(69) Let $\{v_1, v_2, \ldots, v_n\}$ be a linearly independent subset of a vector space V where $n \geq 4$. Set $w_{ij} = v_i - v_j$. Let W be the span of $\{w_{ij} : 1 \leq i, j \leq n\}$. Then

(a) $\{w_{ij} : 1 \leq i, j \leq n\}$ spans W.
(b) $\{w_{ij} : 1 \leq i, j \leq n\}$ is a linearly independent subset of W.
(c) $\{w_{ij} : 1 \leq i \leq n - 1, j = i + 1\}$ spans W.
(d) dim $W = n$.

Ans. ***Options a and c***
Since $w_{ji} = v_j - v_i = -(v_i - v_j) = -w_{ij}$, $\{w_{ij} : 1 \leq i, j \leq n\}$ *is a linearly dependent subset of* W *and* $\{w_{ij} : 1 \leq i \leq n - 1, j = i + 1\}$ *spans* W*. Also* $w_{ij} = w_{1j} - w_{1i}$ *and since* $\{v_1, v_2, \ldots, v_n\}$ *is a linearly independent set,* $dim(W) = n - 1$.

(70) Check whether the following statements are true or false.

(a) Let S be a finite subset of $\mathbb{R}^3$ such that any three elements in S span a two dimensional subspace. Then S spans a two dimensional space.
(b) The polynomials $(x - 1)(x - 2)$, $(x - 2)(x - 3)$, $(x - 3)(x - 4)$, $(x - 4)(x - 6) \in \mathbb{R}[x]$ are linearly independent.
(c) There exists an infinite subset $S \subset \mathbb{R}^3$ such that any three vectors in S are linearly independent.
(d) The set of nilpotent matrices in $\mathbb{M}_{n\times n}(\mathbb{R})$ spans $\mathbb{M}_{n\times n}(\mathbb{R})$ considered as an $\mathbb{R}$-vector space(a matrix A is said to be nilpotent if there exists $n \in \mathbb{N}$ such that $A^n = 0$).

Ans. *(a) True. Clearly* $span(S)$ *has dimension* ≥ 2*. Suppose* $span(S)$ *has dimension 3, then there exists a subset W of S with three linearly independent vectors and this is not possible as any three elements in S span a two dimensional subspace.*
(b) False. Since the dimension of set of all polynomials of degree at most 2 is 3. So we cannot find a collection of 4 linearly independent second degree polynomials.
(c) True. Consider the set $S = \{(1, x, x^2) : x \in \mathbb{R}\}$*. Then take any three vectors* $\{(1, r, r^2), (1, s, s^2), (1, t, t^2) : r, s, t \in \mathbb{R}\}$*. Now*

$$\begin{vmatrix} 1 & r & r^2 \\ 1 & s & s^2 \\ 1 & t & t^2 \end{vmatrix} = (t-r)(s-r)(t-s)$$

Since $(t-r)(s-r)(t-s) = 0$ *only if any two of them are equal or* $r = s = t$*. So any three vectors in S are linearly independent.*

(d) *False. Since we cannot span matrices with non-zero trace using the set of nilpotent matrices in* $\mathbb{M}_{n\times n}(\mathbb{R})$*, the given statement is false.*

Chapter 10
Solved Problems—Linear Transformations

(1) Which of the following is NOT a linear transformation?
(a) $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T(x, y, z) = (x, z)$
(b) $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (x, y - 1, z)$
(c) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (2x, y - x)$
(d) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (y, x)$.

Ans. ***Option b***

(a) Let $v_1 = (x_1, y_1, z_1)$, $v_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$, *and* $\alpha \in \mathbb{R}$. *Then*

$$\begin{aligned} T(\alpha v_1 + v_2) &= (\alpha x_1 + x_2, \alpha y_1 + y_2, \alpha z_1 + z_2) \\ &= (\alpha x_1 + x_2, \alpha z_1 + z_2) \\ &= \alpha(x_1, z_1) + (x_2, z_2) \\ &= \alpha T(v_1) + T(v_2) \end{aligned}$$

T *is a linear transformation.*

(b) $T(0, 0, 0) = (0, -1, 0)$. *Therefore* T *is not a linear transformation. (Since a linear transformation maps identity to identity.)*

(c) Let $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2) \in \mathbb{R}^2$, *and* $\alpha \in \mathbb{R}$. *Then*

$$\begin{aligned} T(\alpha v_1 + v_2) &= (2(\alpha x_1 + x_2), \alpha y_1 + y_2 - \alpha x_1 - x_2) \\ &= \alpha(2x_1, y_1 - x_1) + (2x_2, y_2 - x_2) \\ &= \alpha T(v_1) + T(v_2) \end{aligned}$$

T *is a linear transformation.*

(d) Let $v_1 = (x_1, y_1)$, $v_2 = (x_2, y_2) \in \mathbb{R}^2$, *and* $\alpha \in \mathbb{R}$. *Then*

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_10

$$\begin{aligned} T(\alpha v_1 + v_2) &= T(\alpha x_1 + x_2, \alpha y_1 + y_2) \\ &= (\alpha y_1 + y_2, \alpha x_1 + x_2) \\ &= \alpha (y_1, x_1) + (y_2, x_2) \\ &= \alpha T(v_1) + T(v_2) \end{aligned}$$

T is a linear transformation.

(2) Let the mappings T_1, T_2, T_3, T_4 from $\mathbb{R}^3$ to $\mathbb{R}^3$ be defined by

$$T_1(x, y, z) = (x^2 + y^2, x + z, x + y + z) \qquad T_2(x, y, z) = (y + z, x + z, x + y)$$

$$T_3(x, y, z) = (x + y, xy, x - z) \qquad T_4(x, y, z) = (x, 2y, 3z)$$

Then which of these are linear transformations of $\mathbb{R}^3$ over $\mathbb{R}$?
(a) T_1 and T_2 (b) T_2 and T_3 (c) T_2 and T_4 (d) T_3 and T_4.

Ans. ***Option c***

(a) We have $T_1(1, 0, 0) = (1, 1, 1)$. As

$$T_1(2, 0, 0) = (4, 2, 2) \neq 2T_1(1, 0, 0)$$

T_1 is not a linear transformation.

(b) Let $v_1 = (x_1, y_1, z_1), v_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$, and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} T_2(\alpha v_1 + v_2) &= T_2(\alpha x_1 + x_2, \alpha y_1 + y_2, \alpha z_1 + z_2) \\ &= (\alpha(y_1 + z_1) + y_2 + z_2, \alpha(x_1 + z_1) + x_2 + z_2, \alpha(x_1 + y_1) + x_2 + y_2) \\ &= \alpha(y_1 + z_1, x_1 + z_1, x_1 + y_1) + (y_2 + z_2, x_2 + z_2, x_2 + y_2) \\ &= \alpha T_2(v_1) + T_2(v_2) \end{aligned}$$

T_2 is a linear transformation.

(c) We have $T_3(1, 0, 0) = (1, 0, 1)$ and $T_3(0, 1, 0) = (1, 0, 0)$. As

$$T_3(1, 1, 0) = (2, 1, 1) \neq T_3(1, 0, 0) + T_3(0, 1, 0)$$

T_3 is not a linear transformation.

(d) Let $v_1 = (x_1, y_1, z_1), v_2 = (x_2, y_2, z_2) \in \mathbb{R}^3$, and $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} T_4(\alpha v_1 + v_2) &= T_4(\alpha x_1 + x_2, \alpha y_1 + y_2, \alpha z_1 + z_2) \\ &= (\alpha x_1 + x_2, 2(\alpha y_1 + y_2), 3(\alpha z_1 + z_2)) \\ &= \alpha(x_1, 2y_1, 3z_1) + (x_2, 2y_2, 3z_2) \\ &= \alpha T_4(v_1) + T_4(v_2) \end{aligned}$$

T_4 is a linear transformation.

(3) Which of the following is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$?
(1) $T_1\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}4\\x+y\end{pmatrix}$ (2) $T_2\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}xy\\x+y\end{pmatrix}$ (3) $T_3\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}z-x\\x+y\end{pmatrix}$
(a) only T_1 (b) only T_2 (c) only T_3 (d) all the transformations T_1, T_2, and T_3.

Ans. ***Option c***
Since $T_1\begin{pmatrix}0\\0\\0\end{pmatrix}=\begin{pmatrix}4\\0\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}$, *$T_1$ is not a linear transformation. We have* $T_2\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$ *and* $T_2\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$. *But* $T_2\begin{pmatrix}1\\1\\0\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}\neq T_2\begin{pmatrix}1\\0\\0\end{pmatrix}+T_2\begin{pmatrix}0\\1\\0\end{pmatrix}$. *$T_2$ is not a linear transformation. But*

$$T_3\begin{pmatrix}\lambda x_1+x_2\\\lambda y_1+y_2\\\lambda z_1+z_2\end{pmatrix}=\begin{pmatrix}\lambda(z_1-x_1)+(z_2-x_2)\\\lambda(x_1+y_1)+x_2+y_2\end{pmatrix}=\lambda T_3\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix}+T_3\begin{pmatrix}x_2\\y_2\\z_2\end{pmatrix}$$

Therefore T_3 is a linear transformation.

(4) Let $a, b, c, d\in\mathbb{R}$ and let $T:\mathbb{R}^2\to\mathbb{R}^2$ be the linear transformation defined by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}ax+by\\cx+dy\end{bmatrix}\ for\ \begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2$$

Let $S:\mathbb{C}\to\mathbb{C}$ be the corresponding map defined by

$$S(x+iy)=(ax+by)+i(cx+dy)\ for\ x, y\in\mathbb{R}$$

Then
(a) S is always $\mathbb{C}-$ linear, that is $S(z_1+z_2)=S(z_1)+S(z_2)$ for all $z_1, z_2\in\mathbb{C}$ and $S(\alpha z)=\alpha S(z)$ for all $\alpha\in\mathbb{C}$ and $z\in\mathbb{C}$.
(b) S is $\mathbb{C}-$ linear if $b=-c$ and $d=a$.
(c) S is $\mathbb{C}-$ linear only if $b=-c$ and $d=a$.
(d) S is $\mathbb{C}-$ linear if and only if T is the identity transformation.

Ans. ***Options b and c***
Let $z_1 = x_1 + iy_1$ *and* $z_2 = x_2 + iy_2 \in \mathbb{C}$. *Then*

$$\begin{aligned} S(z_1 + z_2) &= S\left((x_1 + x_2) + i(y_1 + y_2)\right) \\ &= a(x_1 + x_2) + b(y_1 + y_2) + i\left(c(x_1 + x_2) + d(y_1 + y_2)\right) \\ &= ax_1 + by_1 + i\left(cx_1 + dy_1\right) + ax_2 + by_2 + i\left(cx_2 + dy_2\right) \\ &= S(z_1) + S(z_2) \end{aligned}$$

Also

$$S(iz) = S\left(i(x + iy)\right) = S(-y + ix) = -ay + bx + i\left(-cy + dx\right)$$

and

$$iS(z) = i\left(ax + by + i\left(cx + dy\right)\right) = i(ax + by) - (cx + dy) \neq S(iz)$$

Therefore, S need not be $\mathbb{C}-$ *linear. Now let* $\alpha = c_1 + ic_2 \in \mathbb{C}$ *and* $b = -c$ *and* $d = a$.

$$\begin{aligned} S(\alpha z) &= S\left(c_1x - c_2y + i(c_2x + c_1y)\right) \\ &= a(c_1x - c_2y) + b(c_2x + c_1y) + i\left(-b(c_1x - c_2y) + a(c_2x + c_1y)\right) \\ &= c_1(ax + by + i(-bx + ay)) + ic_2(ax + by + i(-bx + ay)) \\ &= \alpha S(z) \end{aligned}$$

Now suppose that $S(\alpha z) = \alpha S(z)$.

$$\begin{aligned} S(\alpha z) &= S\left(c_1x - c_2y + i(c_2x + c_1y)\right) \\ &= a(c_1x - c_2y) + b(c_2x + c_1y) + i\left(c(c_1x - c_2y) + d(c_2x + c_1y)\right) \\ &= c_1(ax + by + i(-bx + ay)) + ic_2(ax + by + i(-bx + ay)) \end{aligned}$$

$$\begin{aligned} \alpha S(z) &= (c_1 + ic_2)\left[(ax + by + i(cx + dy))\right] \\ &= c_1\left[(ax + by + i(cx + dy))\right] + ic_2\left[(ax + by + i(cx + dy))\right] \end{aligned}$$

Comparing, we get $b = -c$ *and* $d = a$.

(5) Consider the vector space $C[0, 1]$ over $\mathbb{R}$. Consider the following statements:

P : If the set $\{xf_1, x^2f_2, x^3f_3\}$ is linearly independent, then the set $\{f_1, f_2, f_3\}$ is linearly independent, where $f_1, f_2, f_3 \in C[0, 1]$ and x^n represents the polynomial function $x \mapsto x^n$, $n \in \mathbb{N}$.

Q : If $T : [0, 1] \to \mathbb{R}$ is given by $T(f) = \int_0^1 f(x^2)dx$ for each $f \in C[0, 1]$, then T is a linear map.

Which of the above statements hold true?
(a) Only P (b) Only Q (c) Both P and Q (d) Neither P nor Q.

Ans. ***Option b***
Let $f_1 = f_2 = f_3 = 1$. *Then the set* $\{xf_1, x^2f_2, x^3f_3\} = \{x, x^2, x^3\}$ *is linearly independent. But the set* $\{f_1, f_2, f_3\}$ *is not linearly independent. For* $f, g \in C[0, 1]$ *and* $\lambda \in \mathbb{R}$,

$$T(\lambda f + g) = \int_0^1 (\lambda f + g)(x^2)dx = \lambda \int_0^1 f(x^2)dx + \int_0^1 g(x^2)dx = \lambda T(f) + T(g)$$

Therefore P *is false and* Q *is true.*

(6) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that $T(1, 2) = (2, 3)$ and $T(0, 1) = (1, 4)$. Then $T(5, 6)$ is

(a) $(6, -1)$ (b) $(-6, 1)$ (c) $(-1, 6)$ (d) $(1, -6)$.

Ans. ***Option a***
Since $(5, 6) = 5(1, 2) + (-4)(0, 1)$, *we have*

$$T(5, 6) = 5\,T(1, 2) + (-4)\,T(0, 1) = (6, -1)$$

(7) Let $T : \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation. If $T(1, 1, 0) = (2, 0, 0, 0)$, $T(1, 0, 1) = (2, 4, 0, 0)$, $T(0, 1, 1) = (0, 0, 2, 0)$, then $T(1, 1, 1)$ equals

(a) $(1, 1, 1, 0)$ (b) $(0, 1, 1, 1)$ (c) $(2, 2, 1, 0)$ (d) $(0, 0, 0, 0)$.

Ans. ***Option c***
Since $(1, 1, 1) = \frac{1}{2}(1, 1, 0) + \frac{1}{2}(1, 0, 1) + \frac{1}{2}(0, 1, 1)$

$$\begin{aligned} T(1, 1, 1) &= \frac{1}{2}T(1, 1, 0) + \frac{1}{2}T(1, 0, 1) + \frac{1}{2}T(0, 1, 1) \\ &= \frac{1}{2}(2, 0, 0, 0) + \frac{1}{2}(2, 4, 0, 0) + \frac{1}{2}(0, 0, 2, 0) \\ &= (1, 0, 0, 0) + (1, 2, 0, 0) + (0, 0, 1, 0) \\ &= (2, 2, 1, 0) \end{aligned}$$

(8) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that $T\,(1, 2) = (1, 0)$ and $T\,(2, 1) = (0, 1)$. Suppose that $(3, -2) = \alpha\,(1, 2) + \beta\,(2, 1)$ and $T\,(3, -2) = (a, b)$. Then $\alpha + \beta + a + b$ equals

(a) $\frac{2}{3}$ (b) $\frac{4}{3}$ (c) $\frac{5}{3}$ (d) $\frac{7}{3}$.

Ans. ***Option a***
We have

$$(3, -2) = \alpha\,(1, 2) + \beta\,(2, 1) = (\alpha + 2\beta, 2\alpha + \beta) \Rightarrow \alpha = \frac{-7}{3}, \beta = \frac{8}{3}$$

Since T is linear transformation,

$$T\,(3, -2) = \frac{-7}{3}T\,(1, 2) + \frac{8}{3}T\,(2, 1) = \left(\frac{-7}{3}, \frac{8}{3}\right) = (a, b)$$

Therefore $\alpha + \beta + a + b = \frac{2}{3}$.

(9) Consider the vector space $\mathbb{P}_2[x] = \{a_0 + a_1x + a_2x^2 : a_i \in \mathbb{R},\ for\ i = 0, 1, 2\}$ of polynomials of degree at most 2. Let $f : \mathbb{P}_2[x] \to \mathbb{R}$ be a linear functional such that $f(1 + x) = 0, f(1 - x^2) = 0$ and $f(x^2 - x) = 2$. Then $f(1 + x + x^2)$ equals ......

Ans. *Since* $(1 + x + x^2) = \frac{3}{2}(1 + x) + \left(-\frac{1}{2}\right)(1 - x^2) + \frac{1}{2}(x^2 - x)$,

$$f\,(1 + x + x^2) = \frac{3}{2}f\,(1 + x) + \left(-\frac{1}{2}\right)f\,(1 - x^2) + \frac{1}{2}f\,(x^2 - x) = 1$$

(10) Let $S = \{T : \mathbb{R}^3 \to \mathbb{R}^3 : T(1, 0, 1) = (1, 2, 3), T(1, 2, 3) = (1, 0, 1)\}$ where T denotes a linear transformation. Then S is
(a) a singleton set
(b) a finite set containing more than one element
(c) a countably infinite set
(d) an uncountable set.

Ans. ***Option d***
The set $\{(1, 0, 1), (1, 2, 3)\}$ is a linearly independent set in $\mathbb{R}^3$. Then we choose a third vector $v \in \mathbb{R}^3$ such that $\{(1, 0, 1), (1, 2, 3), v\}$ forms a basis for $\mathbb{R}^3$ and we can assign any vector in $\mathbb{R}^3$ for v. So S is an uncountable set.

(11) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation, where $n \geq 2$. For $k \leq n$, let

$$E = \{v_1, v_2, \ldots, v_k\} \subseteq \mathbb{R}^n \textit{ and } F = \{Tv_1, Tv_2, \ldots, Tv_k\}$$

Then
(a) If E is linearly independent, then F is linearly independent.
(b) If F is linearly independent, then E is linearly independent.
(c) If E is linearly independent, then F is linearly dependent.
(d) If F is linearly independent, then E is linearly dependent.

Ans. ***Option b***
Suppose that E is linearly independent, consider the zero transformation, then F is not linearly independent. If we consider the identity transformation F is linearly independent. Now suppose that F is linearly independent, i.e.,

$$\lambda_1 Tv_1 + \lambda_2 Tv_2 + \cdots + \lambda_k Tv_k = 0 \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$$

Now for $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}$,

$$\begin{aligned}\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k = 0 &\Rightarrow T(\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_k v_k) = 0\\ &\Rightarrow \mu_1 Tv_1 + \mu_2 Tv_2 + \cdots + \mu_k Tv_k = 0\\ &\Rightarrow \mu_1 = \mu_2 = \cdots = \mu_k = 0\end{aligned}$$

Therefore E is linearly independent.

(12) Let V be a non-zero vector space over a field $\mathbb{K}$. Let $S \subset V$ be a non-empty set. Consider the following properties of S :

(I) For any vector space W over $\mathbb{K}$, any map $T_1 : S \to W$ extends to a linear map from V to W.
(II) For any vector space W over $\mathbb{K}$ and any two linear maps $T_1, T_2 : V \to W$ satisfying $T_1(s) = T_2(s)$ for all $s \in S$, we have $T_1(v) = T_2(v)$ for all $v \in V$.
(III) S is linearly independent.
(IV) The span of S is V.

Which of the following statement(s) is(are) true?

(a) (I) implies (IV) (b) (I) implies (III)

(c) (II) implies (III) (d) (II) implies (IV).

Ans. ***Option d***
(I) need not imply (IV)
Consider $V = W = \mathbb{R}^2,\ S = \{(1, 0)\}$. *Define* $T_1 : S \to W$ *by* $T_1(1, 0) = (1, 0)$. *This map can be extended to a linear map from V to W as* $T_1(x, y) = (x, y)$. *But the span of S is not* V.
(I) need not imply (III)
Consider $V = W = \mathbb{R}^2,\ S = \{(1, 0), (0, 1), (1, 1)\}$. *Define* $T_1 : S \to W$ *by* $T_1(1, 0) = (1, 0), T_1(0, 1) = (0, 1), T_1(1, 1) = (1, 1)$. *This map can be extended to a linear map from V to W as* $T_1(x, y) = (x, y)$. *But S is linearly dependent.*
If (II) is satisfied S need not be a linearly independent set. For example, if $V = W = \mathbb{R}^2,\ S = \{(1, 0), (0, 1), (1, 1)\}$. *We can define* T_1 *and* T_2 *satisfying the above conditions. But S is not linearly independent. Also for* T_1 *and* T_2 *to be equal for all* $v \in V$, *S must span* V.

(13) Consider the basis $\{v_1, v_2, v_3\}$ of $\mathbb{R}^3$, where $v_1 = (1, 0, 0),\ v_2 = (1, 1, 0),\ v_3 = (1, 1, 1)$. Let $\{f_1, f_2, f_3\}$ be the dual basis of $\{v_1, v_2, v_3\}$ and f be a linear functional defined by $f(x, y, z) = x + y + z,\ (x, y, z) \in \mathbb{R}^3$. If $f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3$, then $(\lambda_1, \lambda_2, \lambda_3)$ is

(a) (1, 2, 3) (b) (1, 3, 2) (c) (2, 3, 1) (d) (3, 2, 1).

Ans. ***Option a***
Since f_i is linear

$$f_i(x, y, z) = xf_i(1, 0, 0) + yf_i(0, 1, 0) + zf_i(0, 0, 1)$$

Also we have $f_1(1, 0, 0) = 1, f_1(1, 1, 0) = 0, f_1(1, 1, 1) = 0$
$\Rightarrow f_1(0, 1, 0) = -1$ *and* $f_1(0, 0, 1) = 0$. *Therefore,* $f_1(x, y, z) = x - y$.
Similarly, $f_2(1, 0, 0) = 0, f_2(1, 1, 0) = 1, f_2(1, 1, 1) = 0$
$\Rightarrow f_2(0, 1, 0) = 1$ *and* $f_2(0, 0, 1) = -1$. *Therefore,* $f_2(x, y, z) = y - z$,
and $f_3(1, 0, 0) = 0, f_3(1, 1, 0) = 0, f_3(1, 1, 1) = 1$
$\Rightarrow f_3(0, 1, 0) = 0$ *and* $f_3(0, 0, 1) = 1$. *Therefore,* $f_3(x, y, z) = z$. *Now,*

$$\begin{aligned} f = \lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 &\Rightarrow f(x, y, z) = \lambda_1 f_1(x, y, z) + \lambda_2 f_2(x, y, z) + \lambda_3 f_3(x, y, z) \\ &\Rightarrow x + y + z = \lambda_1(x - y) + \lambda_2(y - z) + \lambda_3(z) \\ &\Rightarrow \lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3 \end{aligned}$$

(14) Let S be the set of all 2×3 real matrices each of whose entries is $1, 0$, or -1 (there are 3^6 matrices in S). Recall that the column space of a matrix A in S is the subspace of $\mathbb{R}^2$ spanned by the three columns of A. For two elements A and B in S, let us write $A \sim B$ if A and B have the same column space. Note that $\sim$ is an equivalence relation. How many equivalence classes are there in S?

Ans. *The possible columns of elements in S are* $\begin{bmatrix}0 & 0\end{bmatrix}^T, \begin{bmatrix}0 & 1\end{bmatrix}^T, \begin{bmatrix}0 & -1\end{bmatrix}^T, \begin{bmatrix}1 & 0\end{bmatrix}^T, \begin{bmatrix}-1 & 0\end{bmatrix}^T, \begin{bmatrix}1 & 1\end{bmatrix}^T, \begin{bmatrix}1 & -1\end{bmatrix}^T, \begin{bmatrix}-1 & 1\end{bmatrix}^T, \begin{bmatrix}-1 & -1\end{bmatrix}^T$. *So the only possible column spaces are* $\{0\}, \mathbb{R}^2, \langle(1, 0)\rangle, \langle(0, 1)\rangle, \langle(1, 1)\rangle, \langle(1, -1)\rangle$, *where* $\langle v\rangle$ *denote the span of v. Therefore there exist 6 equivalence classes.*

(15) Let V be a finite-dimensional vector space over $\mathbb{R}$, and $W \subset V$ a subspace. $W \cap T(W) \neq 0$ for every linear isomorphism $T : V \to V$ if and only if

(a) $W = V$ (b) $dim\ W < \frac{1}{2} dim\ V$
(c) $dim\ W = \frac{1}{2} dim\ V$ (d) $dim\ W > \frac{1}{2} dim\ V$.

Ans. ***Option d***
Since T is an isomorphism we have $dim\ W = dim\ T(W)$. Let $dim\ V = n$. Suppose that $dim\ W \leq \frac{n}{2}$. Then define a linear isomorphism which maps W to its complement W'. Then $W \cap T(W) = \{0\}$. Now suppose that $dim\ W > \frac{1}{2} dim\ V$. Then $W \cap T(W) \neq 0$ for if $W \cap T(W) = \{0\}$,

$$\begin{aligned} dim\ (W + T(W)) &= dim\ W + dim\ T(W) - dim\ (W \cap T(W)) \\ &= dim\ W + dim\ W,\ since\ dim\ W = dim\ T(W) \\ &= 2dim\ W \\ &> dimV,\ since\ dim\ W > \frac{1}{2}dim\ V \end{aligned}$$

which is not possible.

(16) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by $T(x, y) = (-x, y)$. Then
(a) $T^{2k} = T$ for all $k \geq 1$
(b) $T^{2k+1} = -T$ for all $k \geq 1$
(c) $\mathcal{R}(T^2)$ is a proper subspace of $\mathcal{R}(T)$
(d) $\mathcal{R}(T^2) = \mathcal{R}(T)$.

Ans. ***Option d***
We have

$$T^2(x, y) = T\ (T(x, y)) = T(-x, y) = (x, y) \neq T(x, y)$$

Since $T^2 = I$, *we have* $T^{2k} = I$ *and* $T^{2k+1} = T$ *for all* $k \geq 1$. $T^2 = I \Rightarrow \mathcal{R}(T^2) = \mathbb{R}^2$ *and* $T(1, 0) = (-1, 0), T(0, 1) = (0, 1) \Rightarrow \mathcal{R}(T) = \mathbb{R}^2$.

(17) Let V be a vector space over $\mathbb{R}$ and let $T : \mathbb{R}^6 \to V$ be a linear transformation such that $S = \{Te_2, Te_4, Te_6\}$ spans V. Which one of the following must be true?

(a) S is a basis of V. (b) $T\ (\mathbb{R}^6) \neq V$.
(c) $\{Te_1, Te_3, Te_5\}$ spans V. (d) $\mathcal{N}(T)$ contains more than one element.

Ans. ***Option d***
S need not be a basis. S should be linearly independent also. For example, let $T : \mathbb{R}^6 \to \mathbb{R}^2$, *defined by*

$$T(e_2) = (1, 0),\ T(e_4) = (0, 1),\ T(e_6) = (1, 1)$$

and

$$T(e_1) = T(e_3) = T(e_5) = 0$$

Then $\{Te_2, Te_4, Te_6\}$ *spans* $V = \mathbb{R}^2$. *But S is not a basis of W as S is not linearly independent. Also* $\{Te_1, Te_3, Te_5\}$ *does not span* V *and* $T(\mathbb{R}^6) = W$. *By Rank-nullity theorem,* $Rank(T) + Nullity(T) = 6$ *and* $Rank(T) \leq 3$. *Therefore* $\mathcal{N}(T)$ *contains more than one element.*

(18) If $T : \mathbb{M}_{n\times n}(\mathbb{R}) \to \mathbb{M}_{n\times n}(\mathbb{R})$ is a linear transformation such that $T(A) = 0$ whenever $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ is symmetric or skew-symmetric, then the rank of T is
(a) $\frac{n(n+1)}{2}$ (b) $\frac{n(n-1)}{2}$ (c) n (d) 0.

Ans. ***Option d***
Since every matrix $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ can be written as a sum of a symmetric and skew-symmetric matrix and T is a linear transformation $T(A) = 0$ for all $A \in \mathbb{M}_{n\times n}(\mathbb{R})$.

(19) If $T : \mathbb{P}_n[x] \to \mathbb{P}_{n+1}[x]$ is defined by

$$(Tp)\,(x) = p'(x) - \int_0^x p(t)dt$$

then the dimension of null space of T is

(a) 0 (b) 1 (c) n (d) $n + 1$.

Ans. ***Option a***
Let $p(x) \in \mathcal{N}(T)$, then

$$(Tp)\,(x) = 0 \Rightarrow p'(x) = \int_0^x p(t)dt \Rightarrow p(x) = 0$$

Therefore $\mathcal{N}(T) = \{0\}$ and hence dimension of null space of T is 0.

(20) If the nullity of the matrix $\begin{bmatrix} \lambda & 1 & 2 \\ 1 & -1 & -2 \\ 1 & 1 & 4 \end{bmatrix}$ is 1, then the value of λ is

(a) -1 (b) 0 (c) 1 (d) 2.

Ans. ***Option a***
Since the nullity of the given matrix is 1, determinant of the given matrix is zero. Therefore, $-2\lambda - 2 = 0 \Rightarrow \lambda = -1$.

(21) Let A be a 5×4 matrix with real entries such that the space of all solutions of the linear system $AX^T = \begin{bmatrix} 1 \; 2 \; 3 \; 4 \; 5 \end{bmatrix}^T$ is given by

$$\left\{ \begin{bmatrix} 1 + 2s \;\; 2 + 3s \;\; 3 + 4s \;\; 4 + 5s \end{bmatrix}^t : s \in \mathbb{R} \right\}$$

Then the rank of A is equal to

(a) 4 (b) 3 (c) 2 (d) 1.

Ans. ***Option b***
Since the solution space is of dimension 1, the rank of the matrix is 3.

(22) Let $A = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 2 & 6 & 9 & 18 \\ 1 & 2 & 6 & 12 \end{bmatrix}$. Find a basis for the null space of A.

Ans. *By suitable elementary transformations, we have*

$$A = \begin{bmatrix} 1 & 2 & 3 & 6 \\ 2 & 6 & 9 & 18 \\ 1 & 2 & 6 & 12 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 3 & 6 \\ 0 & 4 & 3 & 6 \\ 0 & 0 & 3 & 6 \end{bmatrix}$$

Therefore

$$\mathcal{N}(A) = \left\{(x, y, z, w) \in \mathbb{R}^4 : z = -2w\right\} = span\,\{(0, 0, -2, 1)\}$$

Hence $\{(0, 0, -2, 1)\}$ *is a basis for* $\mathcal{N}(A)$.

(23) Let V be an $n-$ dimensional vector space and let $T : V \to V$ be a linear transformation such that $Rank\ (T) \leq Rank\ \left(T^3\right)$. Then which of the following statement is necessarily true?
(a) $\mathcal{N}(T) = \mathcal{R}((T)$
(b) $\mathcal{N}(T) \cap \mathcal{R}(T) = \{0\}$
(c) There exists a non-zero subspace W of V such that $\mathcal{N}(T) \cap \mathcal{R}(T) = W$
(d) $\mathcal{N}(T) \subseteq \mathcal{R}(T)$.

Ans. ***Option b***
Since $\mathcal{R}(T) \supset \mathcal{R}(T^2)$, $Rank(T) \leq Rank\left(T^3\right) \Rightarrow Rank(T) = Rank\left(T^3\right)$. *Therefore* $dim\,[\mathcal{N}(T)] = dim\left[\mathcal{N}(T^3)\right]$. *Since* $\mathcal{N}(T) \subset \mathcal{N}(T^3)$, $\mathcal{N}(T) = \mathcal{N}(T^3)$. *Now, let* $v \in \mathcal{R}(T) \cap \mathcal{N}(T)$, *then there exists* w *such that* $Tw = v$ *and* $Tv = 0$. *Then,*

$$T^3w = T^2(Tw) = T^2v = 0 \Rightarrow w \in \mathcal{N}(T^3) = \mathcal{N}(T)$$

Therefore $Tw = v = 0$. *Hence* $\mathcal{N}(T) \cap \mathcal{R}(T) = \{0\}$.

(24) Let V be a finite-dimensional vector space over $\mathbb{R}$. Let $T : V \to V$ be a linear transformation such that $Rank(T^2) = Rank(T)$. Then,

(a) $\mathcal{N}(T^2) = \mathcal{N}(T)$ (b) $\mathcal{R}(T^2) = \mathcal{R}(T)$
(c) $\mathcal{N}(T) \cap \mathcal{R}(T) = \{0\}$ (d) $\mathcal{N}(T^2) \cap \mathcal{R}(T^2) = \{0\}$.

Ans. ***Options a, b, c, and d***

(a) By Rank-Nullity theorem, $Nullity(T^2) = Nullity(T)$. *Since* $\mathcal{N}(T) \subseteq \mathcal{N}(T^2)$ *we get* $\mathcal{N}(T^2) = \mathcal{N}(T)$.
(b) Since $\mathcal{R}(T^2) \subseteq \mathcal{R}(T)$ *and* $Rank(T^2) = Rank(T)$, *we get* $\mathcal{R}(T^2) = \mathcal{R}(T)$.

(c) Let $w \in \mathcal{N}(T) \cap \mathcal{R}(T)$

$$\begin{aligned} w \in \mathcal{N}(T) \cap \mathcal{R}(T) &\Rightarrow T(w) = 0 \text{ and } \exists\, v \in V \text{ such that } T(v) = w \\ &\Rightarrow T^2(v) = T(T(v)) = T(w) = 0 \\ &\Rightarrow v \in \mathcal{N}(T^2) = \mathcal{N}(T) \\ &\Rightarrow Tv = 0 = w \end{aligned}$$

(d) From (a), (b), *and* (c), $\mathcal{N}(T^2) \cap \mathcal{R}(T^2) = \{0\}$.

(25) Let $\mathcal{L}(\mathbb{R}^n)$ be the space of $\mathbb{R}-$ linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then which of the following are true?
(a) There exists $T \in \mathcal{L}(\mathbb{R}^5) \setminus \{0\}$ such that $\mathcal{R}(T) = \mathcal{N}(T)$.
(b) There does not exist $T \in \mathcal{L}(\mathbb{R}^5) \setminus \{0\}$ such that $\mathcal{R}(T) = \mathcal{N}(T)$.
(c) There exists $T \in \mathcal{L}(\mathbb{R}^6) \setminus \{0\}$ such that $\mathcal{R}(T) = \mathcal{N}(T)$.
(d) There does not exist $T \in \mathcal{L}(\mathbb{R}^6) \setminus \{0\}$ such that $\mathcal{R}(T) = \mathcal{N}(T)$.

Ans. ***Options b and c***
By Rank-Nullity theorem,

$$dim(V) = dim\,[\mathcal{R}(T)] + dim\,[\mathcal{N}(T)]$$

If $dim\,[\mathcal{R}(T)] = dim\,[\mathcal{N}(T)]$, $dim(V) = 2dim\mathcal{R}(T)$ *is even. Therefore there does not exist* $T \in \mathcal{L}(\mathbb{R}^5) \setminus \{0\}$ *such that* $\mathcal{R}(T) = \mathcal{N}(T)$.
Now define $T : \mathbb{R}^6 \to \mathbb{R}^6$ *by*

$$T(e_1) = T(e_2) = T(e_3) = 0, T(e_4) = e_1, T(e_5) = e_2, T(e_6) = e_3$$

Then $\mathcal{R}(T) = \mathcal{N}(T) = span\{e_1, e_2, e_3\}$. *That is, there exists* $T \in \mathcal{L}(\mathbb{R}^6) \setminus \{0\}$ *such that* $\mathcal{R}(T) = \mathcal{N}(T)$.

26) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 5 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 3 & 1 \end{bmatrix}$. Then the dimension of $\mathcal{N}(A) \cap \mathcal{R}(B)$ over $\mathbb{R}$ is ......

Ans. *The null space of A is*

$$\begin{aligned} \mathcal{N}(A) &= \left\{x \in \mathbb{R}^3 : Ax = 0\right\} \\ &= \left\{x \in \mathbb{R}^3 : x + 2y = 0, 7y + 2z = 0\right\} \\ &= span\,\{(-4, 2, -7)\} \end{aligned}$$

$\mathcal{R}(B) = span\,\{(1, -1, 3), (2, 0, 1)\}$. *Also the set* $\{(-4, 2, -7), (1, -1, 3), (2, 0, 1)\}$ *is linearly dependent. Therefore* $\mathcal{N}(A) \cap \mathcal{R}(B) = \mathcal{N}(A)$ *and hence the dimension of* $\mathcal{N}(A) \cap \mathcal{R}(B) = 1$.

(27) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Which of the following statements implies that T is bijective?

(a) $Nullity(T) = n$ (b) $Rank(T) = Nullity(T) = n$
(c) $Rank(T) + Nullity(T) = n$ (d) $Rank(T) - Nullity(T) = n.$

Ans. ***Option d***

(a) & (b) *A linear transformation T is one-one if and only if* $\mathcal{N}(T) = \{0\}$. *That is, if and only if* $Nullity(T) = 0$.

(c) *Consider* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *given by* $T(x, y) = (x, 0)$. *Clearly* $Rank(T) + Nullity(T) = 2$, *but T is not bijective.*

(d) *By Rank-Nullity theorem, we have* $Rank(T) + Nullity(T) = n$ *and for T to be one-one,* $Nullity(T) = 0 \Rightarrow Rank(T) - Nullity(T) = n$.

(28) Consider non-zero vector spaces V_1, V_2, V_3, V_4 and linear transformations $T_1 : V_1 \to V_2$, $T_2 : V_2 \to V_3$, $T_3 : V_3 \to V_4$, such that $\mathcal{N}(T_1) = \{0\}, \mathcal{R}(T_1) = \mathcal{N}(T_2), \mathcal{R}(T_2) = \mathcal{N}(T_3), \mathcal{R}(T_2) = V_4$. Then

(a) $\sum_{i=1}^{4}(-1)^i dim\ V_i = 0$ (b) $\sum_{i=2}^{4}(-1)^i dim\ V_i > 0$
(c) $\sum_{i=1}^{4}(-1)^i dim\ V_i < 0$ (d) $\sum_{i=i}^{4}(-1)^i dim\ V_i \neq 0.$

Ans. ***Options a and b***
Since $\mathcal{N}(T_1) = \{0\}$, *we have* $\mathcal{N}(T_1) = \{0\}$. *As* $\mathcal{R}(T_1) = \mathcal{N}(T_2)$ *and* $\mathcal{R}(T_2) = \mathcal{N}(T_3)$, *we get* $Rank(T_1) = Nullity(T_2)$ *and* $Rank(T_2) = Nullity(T_3)$. *Also* $dim(V_4) = Rank(T_2)$ *as* $\mathcal{R}(T_2) = V_4$. *Now,*

$$\begin{aligned}
dim(V_1) &= Rank(T_1) + Nullity(T_1)\\
&= Rank(T_1) \text{ since } \mathcal{N}(T_1) = \{0\}\\
dim(V_2) &= Rank(T_2) + Nullity(T_2)\\
&= Rank(T_2) + dim(V_1) \text{ since } Rank(T_1) = Nullity(T_2)\\
dim(V_3) &= Rank(T_3) + Nullity(T_3)\\
&= Rank(T_3) + dim(V_2) - dim(V_1) \text{ since } Rank(T_2) = Nullity(T_3)\\
dim(V_4) &= Rank(T_3) = dim(V_3) - dim(V_2) + dim(V_1)
\end{aligned}$$

(29) Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ be a linear map defined by

$$T(x, y, z, w) = (x + z, 2x + y + 3z, 2y + 2z, w)$$

Then the rank of T is equal to ......

Ans. *We have*

$$T(x, y, z, w) = (x + z, 2x + y + 3z, 2y + 2z, w) = (0, 0, 0, 0)$$
$$\Rightarrow x = y = -z, \; w = 0$$

Therefore

$$\mathcal{N}(T) = \{(x, y, z, w) \in \mathbb{R}^4 : x = y = -z \text{ and } w = 0\}$$
$$= span\{(1, 1, -1, 0)\}$$

Since $dim\,[\mathcal{N}(T)] = 1$ *by Rank-Nullity theorem, rank of T is 3.*

(30) Let T be a 4×4 real matrix such that $T^4 = 0$. Let $k_i = dim\left[\mathcal{N}\left(T^i\right)\right]$ for $1 \leq i \leq 4$. Which of the following is NOT a possibility for the sequence $k_1 \leq k_2 \leq k_3 \leq k_4$?

(a) $3 \leq 4 \leq 4 \leq 4$ (b) $1 \leq 3 \leq 4 \leq 4$
(c) $2 \leq 4 \leq 4 \leq 4$ (d) $2 \leq 3 \leq 4 \leq 4$.

Ans. ***Option b***
Suppose that $k_1 = dim\,[\mathcal{N}(T)] = 1$, *then* $dim\,[\mathcal{R}(T)] = 3$. *Then by Sylvester's inequality,* $3 + 3 - 4 = 2 \leq Rank(T^2) \leq 3$ *and hence* $1 \leq k_2 \leq 2$.

(31) Let $T : \mathbb{R}^7 \to \mathbb{R}^7$ be a linear transformation with $nullity(T) = 2$. Then, the minimum possible value for $Rank(T^2)$ is ......

Ans. *By Rank-Nullity theorem, we have* $Rank(T) = 7 - 2 = 5$ *and by Sylvester's inequality,*

$$5 + 5 - 7 = 3 \leq Rank(T^2) \leq min\,\{5, 5\} = 5$$

That is, minimum possible value for $Rank(T^2)$ *is 3.*

(32) Let $T_1, T_2 : \mathbb{R}^5 \to \mathbb{R}^3$ be linear transformations such that $Rank(T_1) = 3$ and $Nullity(T_2) = 3$. Let $T_3 : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation such that $T_3 \circ T_1 = T_2$. Then $Rank(T_3)$ is ......

Ans. *Since* $Nullity(T_2) = 3$, $Rank(T_2) = 2 = Rank(T_3 \circ T_1)$ *(By Rank-Nullity theorem). By Sylvester's inequality,*

$$Rank(T_3) + Rank(T_1) - 3 \leq Rank(T_3 \circ T_1) = 2 \leq min\{Rank(T_3), Rank(T_1)\}$$

Since $Rank(T_1) = 3$, *this gives* $Rank(T_3) \leq 2 \leq Rank(T_3)$. *Thus,* $Rank(T_3) = 2$.

(33) Let V be a vector space (over $\mathbb{R}$) of dimension 7 and let $f : V \to \mathbb{R}$ be a non-zero linear functional. Let W be a linear subspace of V such that $V = \mathcal{N}(f) \oplus W$. What is the dimension of W?

Ans. *By Rank-Nullity theorem,*

$$7 = dim(V) = Rank(f) + Nullity(f) = 1 + Nullity(f) \Rightarrow Nullity(f) = 6$$

Since $V = \mathcal{N}(f) \oplus W$,

$$7 = dim(V) = Nullity(f) + dim(W) = 6 + dim(W) \Rightarrow dim(W) = 1$$

(34) Let A and B be $n \times n$ real matrices such that $AB = BA = 0$ and $A + B$ is invertible. Which of the following is always true?
(a) $Rank(A) = Rank(B)$. (b) $Rank(A) + Rank(B) = n$.
(c) $Nullity(A) + Nullity(B) = n$. (d) $A - B$ is invertible.

Ans. ***Option b, c, and d***

(a) *Let* $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Then* $AB = BA = 0$ *and* $A + B$ *is invertible, but* $Rank(A) \neq Rank(B)$.

(b) *Since* $A + B$ *is invertible,*

$$n = Rank(A + B) \leq Rank(A) + Rank(B)$$

By Sylvester's inequality,

$$Rank(A) + Rank(B) - n \leq Rank(AB) = 0 \Rightarrow Rank(A) + Rank(B) \leq n$$

Therefore $Rank(A) + Rank(B) = n$.

(c) *By Rank-Nullity theorem,*

$$Nullity(A) = n - Rank(A) \text{ and } Nullity(B) = n - Rank(B)$$

Therefore

$$Nullity(A) + Nullity(B) = 2n - (Rank(A) + Rank(B)) = n$$

(d) *Since* $AB = BA = 0$,

$$(A - B)^2 = A^2 - AB - BA + B^2 = A^2 + B^2 = (A + B)^2$$

Therefore $A + B$ *is invertible implies that* $A - B$ *is invertible.*

(35) Let A be a 4×7 real matrix and B be a 7×4 real matrix such that $AB = I_4$. Which of the following is/are always true?

(a) $Rank(A) = 4$ (b) $Rank(B) = 7$
(c) $Nullity(B) = 0$ (d) $BA = I_7$.

Ans. ***Options a and c***
We have $Rank(A), Rank(B) \leq min\{4, 7\} = 4$. *By Sylvester's inequality,*

$$Rank(AB) = 4 \leq min\{Rank(A), Rank(B)\} \Rightarrow Rank(A) = Rank(B) = 4$$

By Rank-Nullity theorem, $Nullity(B) = 0$ *and again by Sylvester's inequality,* $Rank(BA) \leq 4$.

(36) Let $\mathbb{M}_{n\times n}(\mathbb{R})$ be the set of all $m \times n$ matrices with real entries. Which of the following statements is correct?
(a) There exists $A \in \mathbb{M}_{2\times 5}(\mathbb{R})$ such that the dimension of null space of A is 2.
(b) There exists $A \in \mathbb{M}_{2\times 5}(\mathbb{R})$ such that the dimension of null space of A is 0.
(c) There exists $A \in \mathbb{M}_{2\times 5}(\mathbb{R})$ and $B \in \mathbb{M}_{5\times 2}(\mathbb{R})$ such that AB is the 2×2 identity matrix.
(d) There exists $A \in \mathbb{M}_{2\times 5}(\mathbb{R})$ whose null space is $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 = x_2, x_3 = x_4 = x_5\}$.

Ans. ***Option c***
$A \in \mathbb{M}_{2\times 5}(\mathbb{R})$ *implies that it is a linear transformation from a five-dimensional space to a two-dimensional space. By Rank-Nullity theorem, the dimension of the null space of a cannot be 2 and 0. For, then the dimension of the range space will be 3 and 5 respectively. Since a two-dimensional space cannot have subspaces of dimension 3 and 5 this is not possible. Take* $A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$ *and* $B = A^T$, *then* $AB = I_2$. *Now the set* $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 = x_2, x_3 = x_4 = x_5\} = span\{(1, 1, 0, 0, 0), (0, 0, 1, 1, 1)\}$ *has dimension 2 and it cannot be the null space of A.*

(37) Let A be a 4×4 matrix. Suppose that

$$\mathcal{N}(A) = \{(x, y, z, w) \in \mathbb{R}^4 : x + y + z = 0,\ x + y + w = 0\}$$

Then

(a) $dim(column\ space(A)) = 1$ (b) $Rank(A) = 1$
(c) $dim(column\ space(A)) = 2$ (d) $S = \{(1, 1, 1, 0), (1, 1, 0, 1)\}$ is a basis of $\mathcal{N}(A)$.

Ans. ***Option c***

Since

$$\mathcal{N}(A) = \{(x, y, z, w) \in \mathbb{R}^4 : x + y + z = 0,\ x + y + w = 0\}$$
$$= span\{(1, 0, -1, -1), (0, 1, -1, -1)\}$$

As $\{(1, 0, -1, -1), (0, 1, -1, -1)\}$ *is a linearly independent set, it is a basis for* $\mathcal{N}(A)$. *Therefore* $Nullity(A) = 2$. *By Rank-Nullity theorem,*

$$dim(column\ space(A)) = 4 - 2 = 2$$

That is, $Rank(A) = 2$.

(38) Let U, V, and W be finite-dimensional real vector spaces, and $T_1 : U \to V, T_2 : V \to W$, and $T_3 : W \to U$ be linear transformations. If $\mathcal{R}(T_2T_1) = \mathcal{N}(T_3)$, $\mathcal{N}(T_2T_1) = \mathcal{R}(T_3)$, and $Rank(T_1) = Rank(T_2)$, then which one of the following is true?
(a) $\mathcal{N}(T_1) = \mathcal{N}(T_2)$.
(b) $dim(U) \neq dim(W)$.
(c) If $dim(V) = 3$, $dim(U) = 4$, then T_3 is not identically zero.
(d) If $dim(V) = 4$, $dim(U) = 3$ and T_1 is one-one, then T_3 is identically zero.

Ans. ***Option c***

(a) We know that $\mathcal{N}(T_1)$ *is a subspace of* U *and* $\mathcal{N}(T_2)$ *is a subspace of* V. *Since* U *and* V *need not be the same space,* $\mathcal{N}(T_1)$ *need not be equal to* $\mathcal{N}(T_2)$.

(b) Since $T_2T_1 : U \to W$, *by Rank-Nullity theorem,*

$$\begin{aligned} dim\ U &= dim\ range(T_2T_1) + Nullity(T_2T_1) \\ &= Nullity(T_3) + Rank(T_3) \\ &= dim\ W \end{aligned}$$

(c) Let $dim(V) = 3$ *and* $dim(U) = 4$. *Then from option* (b), $dim(W) = 4$. *We have* $\mathcal{N}(T_2T_1) = \mathcal{R}(T_3)$. *We know that* $Nullity(T_2T_1)$ *is minimum when* $Rank(T_2T_1)$ *is maximum. Since the maximum possible dimension for* $\mathcal{R}(T_2)$ *and* $\mathcal{R}(T_1)$ *is 3, the maximum possible dimension of* $\mathcal{R}(T_2T_1)$ *is 3. Therefore by Rank-Nullity theorem,* $\mathcal{N}(T_2T_1)$ *has dimension* ≥ 1. *Hence,* $\mathcal{R}(T_3)$ *is of dimension* ≥ 1. *Hence* T_3 *is not identically zero.*

(d) As in option (c), $\mathcal{R}(T_3)$ *is of dimension* ≥ 1. *Hence* T_3 *is not identically zero.*

(39) If A is a 5×5 matrix and the dimension of the solution space of $Ax = 0$ is at least two, then

(a) $Rank(A^2) \leq 3$ (b) $Rank(A^2) \geq 3$ (c) $Rank(A^2) = 3$ (d) $det(A^2) = 0$.

Ans. ***Options a and d***
If $Nullity(A) = 2$, by Rank-Nullity theorem, $Rank(A) = 5 - 2 = 3$. Then by Sylvester's inequality, $1 \leq Rank(A^2) \leq 3$. Therefore if $Nullity(A) > 2$, $Rank(A)$ will be less than 3 and hence $Rank(A^2) \leq 3$. Since $Rank(A)$ is always less than 5, $det(A) = 0$. Therefore $det(A^2) = 0$.

(40) The row space of a 20×50 matrix A has dimension 13. What is the dimension of the space of solutions of $Ax = 0$?
(a) 7 (b) 13 (c) 33 (d) 37.

Ans. ***Option d***
A is the matrix representation of a linear transformation from $\mathbb{R}^{50}$ to $\mathbb{R}^{20}$. The dimension of the space of solutions of $Ax = 0$ is the dimension of null space of A. Then by Rank-Nullity theorem, the dimension of the space of solutions of $Ax = 0$ is 37.

(41) Let A be an $m \times n$ matrix with rank r. If the linear system $Ax = b$ has a solution for each $b \in \mathbb{R}^m$, then
(a) $m = r$.
(b) the column space of A is a proper subspace of $\mathbb{R}^m$.
(c) the null space of A is a non-trivial subspace of $\mathbb{R}^m$ whenever $m = n$.
(d) $m \geq n$ implies $m = n$.

Ans. ***Options a and d***
The system $Ax = b$ has a solution for each $b \in \mathbb{R}^m$ implying that the column space of A is $\mathbb{R}^m$ and since $Rank(A) = r$, $m = r$. As $Rank(A) \leq min\{m, n\}$, $m \geq n$ implies $m = n$. If $m = n$, the null space of A is the trivial subspace of $\mathbb{R}^m$.

(42) For $n \neq m$, let $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $T_2 : \mathbb{R}^m \to \mathbb{R}^n$ be linear transformations such that T_1T_2 is bijective

(a) $Rank(T_1) = n$ and $Rank(T_2) = m$ (b) $Rank(T_1) = m$ and $Rank(T_2) = n$
(c) $Rank(T_1) = n$ and $Rank(T_2) = n$ (d) $Rank(T_1) = m$ and $Rank(T_2) = m$.

Ans. ***Option d***
If $T_1 : \mathbb{R}^n \to \mathbb{R}^m$ and $T_2 : \mathbb{R}^m \to \mathbb{R}^n$, then $T_1T_2 : \mathbb{R}^m \to \mathbb{R}^m$. Since T_1T_2 is bijective $Rank(T_1T_2) = m$. By Sylvester's inequality,

$$Rank(T_1T_2) = m \leq min\{Rank(T_1), Rank(T_2)\}$$

This implies that $Rank(T_1), Rank(T_2) \geq m$. Also, $Rank(T_1), Rank(T_2) \leq min\{m, n\}$ implies $Rank(T_1), Rank(T_2) \leq m$. Hence $Rank(T_1) = m$ and $Rank(T_2) = m$.

(43) Let $T_1 : \mathbb{R}^3 \to \mathbb{R}^4$ and $T_2 : \mathbb{R}^4 \to \mathbb{R}^3$ be linear transformations such that $T_2 \circ T_1$ is the identity map of $\mathbb{R}^3$. Then

(a) $T_1 \circ T_2$ is the identity map of $\mathbb{R}^4$. (b) $T_1 \circ T_2$ is one-one, but not onto.
(c) $T_1 \circ T_2$ is onto, but not one-one. (d) $T_1 \circ T_2$ is neither one-one nor onto.

Ans. ***Option d***
The matrix of T_1 *is of order* 4×3 *and* T_2 *is of order* 3×4. *Therefore* $Rank(T_1)$, $Rank(T_2) \leq 3$. *Then by Sylvester's inequality,* $Rank(T_1T_2) \leq 3$. *Therefore* $T_1 \circ T_2$ *is neither one-one nor onto.*

(44) If a linear transformation $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{M}_{2\times2}(\mathbb{R})$ is defined as

$$T(p) = \begin{bmatrix} p(0) - p(2) & 0 \\ 0 & p(1) \end{bmatrix}$$

then

(a) T is one-one but not onto (b) T is onto but not one-one
(c) $\mathcal{R}(T) = span \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ (d) $\mathcal{N}(T) = span \left\{ x^2 - 2x,\ 1 - x \right\}$.

Ans. ***Option c***
Consider the standard basis for $\mathbb{P}_2(\mathbb{R})$. *We have*

$$T(1) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},\ T(x) = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}, T(x^2) = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix}$$

Then the matrix of T *is* $\begin{bmatrix} 0 & -2 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, *which is of rank* 2. *Therefore* $\mathcal{R}(T) = span \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ *and* $\mathcal{N}(T)$ *has dimension 1 by Rank-Nullity theorem. Hence* T *is not one-one.*

(45) Let $T : \mathbb{C} \to \mathbb{C}$ be the map defined by $T(z) = z + \bar{z}$. For a $\mathbb{C}$-vector space V, consider the map

$$\phi : \{f : V \to \mathbb{C} \mid f \text{ is } \mathbb{C} - linear\} \to \{g : V \to \mathbb{R} \mid g \text{ is } \mathbb{R} - linear\}$$

defined by $\phi(f) = T \circ f$. Then this map is

(a) injective, but not surjective. (b) surjective, but not injective.
(c) bijective. (d) neither injective nor surjective.

Ans. ***Option c***
Let $f \in \mathcal{N}(\phi)$.

$$\begin{aligned} f \in \mathcal{N}(\phi) &\Rightarrow (T \circ f)(z) = 0 \; \forall \; z \in V \\ &\Rightarrow T(f(z)) = 0 \; \forall \; z \in V \\ &\Rightarrow f(z) + \overline{f(z)} = 0 \; \forall \; z \in V \Rightarrow Re(f(z)) = 0 \end{aligned}$$

Therefore $f(z)$ *is of the form* $f(z) = ih(z)$. *Now* $f(iz) = -1h(z) = -h(z)$. *Since* $Re(f) = 0$, *we get* $h(z) = 0$. *Therefore* $f(z) = 0$ *and hence* ϕ *is one-one.*
Now let $g \in \{g : V \to \mathbb{R} \mid g \text{ is } \mathbb{R} - linear\}$. *Define*

$$f(x) = \frac{g(x) - ig(ix)}{2}$$

Since

$$f(ix) = \frac{g(ix) - ig(-x)}{2} = \frac{g(ix) + ig(x)}{2} = if(x)$$

f *is* $\mathbb{C} - linear$. *Also* $f(x) + \overline{f(x)} = g(x) \Rightarrow \phi$ *is onto. Therefore* ϕ *is bijective.*

(46) Let $T : \mathbb{C} \to \mathbb{M}_{2\times 2}(\mathbb{R})$ be the map given by

$$T(z) = T(x + iy) = \begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

Then which of the following statements is false?
P : $T(z_1 z_2) = T(z_1)T(z_2)$ for all $z_1, z_2 \in \mathbb{C}$.
Q : $T(z)$ is singular if and only if $z = 0$.
R : There does not exist non-zero $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ such that the trace of $T(z)A$ is zero for all $z \in \mathbb{C}$.
S : $T(z_1 + z_2) = T(z_1) + T(z_2)$ for all $z_1, z_2 \in \mathbb{C}$.

(a) P (b) Q (c) R (d) S.

Ans. ***Option c***
Let $z_1 = x_1 + iy_1, z_2 = x_2 + iy_2 \in \mathbb{C}$. *Then* $T(z_1) = \begin{bmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{bmatrix}$ *and* $T(z_2) = \begin{bmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{bmatrix}$.

(a) We have

$$T(z_1z_2) = T\left((x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)\right)$$
$$= \begin{bmatrix} x_1x_2 - y_1y_2 & x_1y_2 + x_2y_1 \\ -(x_1y_2 + x_2y_1) & x_1x_2 - y_1y_2 \end{bmatrix} = T(z_1)T(z_2)$$

P is true.

(b) $T(z)$ is singular $\Rightarrow x^2 + y^2 = 0 \Rightarrow x = 0, y = 0 \Rightarrow z = 0$. The converse is trivial. Q is true.

(c) Take $A = \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$. Then

$$T(z)A = \begin{bmatrix} x & y \\ -y & x \end{bmatrix}\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

Clearly, trace of $T(z)A$ is zero for all $z \in \mathbb{C}$. R is false.

(d) We have

$$T(z_1 + z_2) = T\left((x_1 + x_2) + i(y_1 + y_2)\right)$$
$$= \begin{bmatrix} x_1 + x_2 & y_1 + y_2 \\ -(y_1 + y_2) & x_1 + x_2 \end{bmatrix}$$
$$= \begin{bmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{bmatrix} + \begin{bmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{bmatrix} = T(z_1) + T(z_2)$$

S is true.

(47) Let V be a finite-dimensional vector space over $\mathbb{R}$ and $T : V \to V$ be a linear map. Can you always write $T = T_2 \circ T_1$ for some linear maps $T_1 : V \to W, T_2 : W \to V$, where W is some finite-dimensional vector space and such that

(a) both T_1 and T_2 are onto. (b) both T_1 and T_2 are one-one.
(c) T_1 is onto, T_2 is one-one. (d) T_1 is one-one, T_2 is onto.

Ans. ***Options c and d***

(a) If both T_1 and T_2 are onto, then $T = T_2 \circ T_1$ must also be onto. So it need not be true.

(b) If both T_1 and T_2 are one-one, then $T = T_2 \circ T_1$ must also be one-one. So it need not be true.

(c) Let $W = V \setminus \mathcal{N}(T)$. Define $T_1 : V \to W$ by $T_1(v) = v + \mathcal{N}(T)$ and $T_2 : W \to V$ by $T_2(v + \mathcal{N}(T)) = T(v)$. Then

$$(T_2 \circ T_1)(v) = T_2(v + \mathcal{N}(T)) = T(v)$$

Clearly T_1 is onto. Now

$$\begin{aligned} T_2(v + \mathcal{N}(T)) = 0 &\Rightarrow T(v) = 0 \\ &\Rightarrow v \in \mathcal{N}(T) \\ &\Rightarrow v + \mathcal{N}(T) = \mathcal{N}(T) \\ &\Rightarrow \mathcal{N}(T_2) = \mathcal{N}(T) \\ &\Rightarrow T_2 \text{ is one} - \text{one} \end{aligned}$$

(d) *Let $W = V \oplus V$. Define $T_1 : V \to W$ by $T_1(v) = (v, T(v))$ and $T_2 : W \to V$ by $T_2(v_1, v_2) = v_2$. Then*

$$(T_2 \circ T_1)(v) = T_2\,(v, T(v)) = T(v)$$

Since T_2 is the projection map, clearly T_2 is onto. Now

$$T_1(v_1) = T_1(v_2) \Rightarrow (v_1, T(v_1)) = (v_2, T(v_2)) \Rightarrow v_1 = v_2 \Rightarrow T_1 \text{ is one} - \text{one}$$

(48) On $\mathbb{R}^2$, consider the linear transformation which maps the point (x, y) to the point $(2x + y, x - 2y)$. Write down the matrix of this transformation with respect to the basis $\{(1, 1), (1, -1)\}$.

Ans. *We have*

$$T(1, 1) = (3, -1) = 1(1, 1) + 2(1, -1)$$

and

$$T(1, -1) = (1, 3) = 2(1, 1) + (-1)(1, -1)$$

Therefore the matrix of T is $\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$.

(49) A linear transformation T rotates each vector in $\mathbb{R}^2$ clockwise through $90°$. The matrix of T with respect to the standard ordered basis $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ is

(a) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

Ans. ***Option b***

Since T rotates the vectors through $90°$, $T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Therefore the matrix of T with respect to the standard basis is $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

(50) Let $B = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$ be a basis of $\mathbb{R}^2$ and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ x - 2y \end{pmatrix}$$

Then the matrix of T with respect to B is $[T]_B = \cdots\cdots$

(a) $\begin{bmatrix} -3 & -2 \\ 3 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & -2 \\ -3 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} -3 & -1 \\ 3 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 3 & -1 \\ -3 & 2 \end{bmatrix}$.

Ans. ***Option c***
We have

$$T \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix} = (-3) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

and

$$T \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} = (-1) \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Therefore the matrix of T with respect to the basis B is given by $[T]_B = \begin{bmatrix} -3 & -1 \\ 3 & 2 \end{bmatrix}$.

(51) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$T(x, y, z) = (x + y - z, x + y + z, y - z)$$

Then the matrix of the linear transformation T with respect to ordered basis $B = \{(0, 1, 0), (0, 0, 1), (1, 0, 0)\}$ of $\mathbb{R}^3$ is

(a) $\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}$.

Ans. ***Option c***
We have

$$\begin{aligned} T(0, 1, 0) &= (1, 1, 1) = 1(0, 1, 0) + 1(0, 0, 1) + 1(1, 0, 0) \\ T(0, 0, 1) &= (-1, 1, -1) = 1(0, 1, 0) + (-1)(0, 0, 1) + (-1)(1, 0, 0) \\ T(1, 0, 0) &= (1, 1, 0) = 1(0, 1, 0) + 0(0, 0, 1) + 1(1, 0, 0) \end{aligned}$$

Therefore the matrix of T with respect to B is $\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$.

(52) Let $T : \mathbb{P}_3[0, 1] \to \mathbb{P}_2[0, 1]$ be defined by

$$(Tp)(x) = \frac{d^2}{dx^2}\left[p(x)\right] + \frac{d}{dx}\left[p(x)\right]$$

Then the matrix representation of T with respect to bases $\{1, x, x^2, x^3\}$ and $\{1, x, x^2\}$ of $\mathbb{P}_3[0, 1]$ and $\mathbb{P}_2[0, 1]$ respectively is

(a) $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 6 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 2 & 1 & 0 \\ 6 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & 2 \\ 3 & 6 & 0 \end{bmatrix}$.

Ans. ***Option b***
We have

$$T(1) = 0 = 0.1 + 0x + 0x^2$$

$$T(x) = 1 = 1.1 + 0x + 0x^2$$

$$T(x^2) = 2 + 2x = 2.1 + 2x + 0x^2$$

$$T(x^3) = 6x + 3x^2 = 0.1 + 6x + 3x^2$$

Therefore the matrix of T with respect to the given bases is $\begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 0 & 3 \end{bmatrix}$.

(53) Define the linear transformation $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = (a_0 + a_1 + a_2 + a_3) + (a_1 + 2a_2 + 3a_3)x + (a_2 + 3a_3)x^2 + a_3x^3$$

Write down the matrix of T with respect to the basis $\left\{1, 1 + x, 1 + x^2, 1 + x^3\right\}$.

Ans. *We have*

$$T(1) = 1 = 1.1 + 0(1 + x) + 0(1 + x^2) + 0(1 + x^3)$$

$$T(1 + x) = 2 + x = T(1) = 1 = 1.1 + 1(1 + x) + 0(1 + x^2) + 0(1 + x^3)$$

$$T(1 + x^2) = 2 + 2x + x^2 = (-1).1 + 2(1 + x) + 1(1 + x^2) + 0(1 + x^3)$$

$$T(1 + x^3) = 2 + 3x + 3x^2 + x^3 = (-5).1 + 3(1 + x) + 3(1 + x^2) + 1(1 + x^3)$$

Therefore the matrix of T with respect to the given basis is $\begin{bmatrix} 1 & 1 & -1 & -5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(54) Consider the vector space $\mathbb{P}_3(\mathbb{R})$ over $\mathbb{R}$. Define

$$(Tp)(x) = a_0 + a_1(x+1) + a_2(x+1)^2 + a_3(x+1)^3$$

where $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$. Write down the matrix representing the linear transformation T with respect to the standard basis.

Ans. *We have*

$$T(1) = 1 = 1.1 + 0x + 0x^2 + 0x^3$$

$$T(x) = x + 1 = 1.1 + 1x + 0x^2 + 0x^3$$

$$T(x^2) = (x+1)^2 = x^2 + 2x + 1 = 1.1 + 2x + 1x^2 + 0x^3$$

$$T(x^3) = (x+1)^3 = x^3 + 3x^2 + 3x + 1 = 1.1 + 3x + 3x^2 + 1x^3$$

Therefore the matrix of T with respect to the standard basis is $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$.

(55) Define $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ by

$$(Tp)(x) = \int_0^x p(t)dt + \frac{d}{dx}\left[p(x)\right]$$

Then the matrix representation of T with respect to the bases $\{1, x, x^2\}$ and $\{1, x, x^2, x^3\}$ is

(a) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 0 & \frac{1}{3} \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{3} \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & \frac{1}{2} \\ 0 & 2 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$.

Ans. ***Option b***
We have

$$T(1) = x = 0.1 + 1x + 0x^2 + 0x^3$$

$$T(x) = \frac{x^2}{2} + 1 = 1.1 + 0x + \frac{1}{2}x^2 + 0x^3$$

$$T(x^2) = \frac{x^3}{3} + 2x = 0.1 + 2x + 0x^2 + \frac{1}{3}x^3$$

The matrix of T with respect to the given bases is $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$.

(56) Consider the linear transformation $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ defined by

$$(Tp)(x) = x\frac{d^2}{dx^2}\left[p(x)\right] + 3x\frac{d}{dx}\left[p(x)\right] + 2p(x), \ \forall\, p \in \mathbb{P}_3(\mathbb{R})$$

Write down the corresponding matrix of T with respect to the standard basis.

Ans. *Consider the standard ordered basis* $\{1, x, x^2, x^3\}$. *Then*

$$T(1) = 2 = 2.1 + 0x + 0x^2 + 0x^3$$

$$T(x) = 5x = 0.1 + 5x + 0x^2 + 0x^3$$

$$T(x^2) = 8x^2 + 2x = 0.1 + 2x + 8x^2 + 0x^3$$

$$T(x^3) = 11x^3 + 6x^2 = 0.1 + 0x + 6x^2 + 11x^3$$

The matrix of T with respect to the standard basis is given by $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 5 & 2 & 0 \\ 0 & 0 & 8 & 6 \\ 0 & 0 & 0 & 11 \end{bmatrix}$.

(57) With the notations and definitions of the problem above, find $p \in \mathbb{P}_3(\mathbb{R})$ such that

$$x\frac{d^2}{dx^2}\left[p(x)\right]) + 3x\frac{d}{dx}\left[p(x)\right] + 2p(x) = 11x^3 + 14x^2 + 7x + 2$$

Ans. *Let* $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \in V$. *Then*

$$x\frac{d^2}{dx^2}\left[p(x)\right] + 3x\frac{d}{dx}\left[p(x)\right] + 2p(x) = 2a_0 + (5a_1 + 2a_2)x + (8a_2 + 6a_3)x^2 + 11a_3x^3$$
$$\Rightarrow a_0 = 1,\ 5a_1 + 2a_2 = 7,\ 8a_2 + 6a_3 = 14,\ a_3 = 1$$
$$\Rightarrow a_0 = a_1 = a_2 = a_3 = 1$$

Therefore $p(x) = 1 + x + x^2 + x^3$.

(58) Define $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ by

$$(Tp)(x) = p(x+1),\ p \in \mathbb{P}_3(\mathbb{R})$$

Then the matrix of T in the basis $\{1, x, x^2, x^3\}$ is given by

(a) $\begin{bmatrix} 1&0&0&0\\ 0&2&0&0\\ 0&0&3&0\\ 0&0&0&4 \end{bmatrix}$ (b) $\begin{bmatrix} 1&1&1&1\\ 0&1&2&3\\ 0&0&1&3\\ 0&0&0&1 \end{bmatrix}$ (c) $\begin{bmatrix} 1&1&2&3\\ 1&1&2&3\\ 2&2&2&3\\ 3&3&3&3 \end{bmatrix}$ (d) $\begin{bmatrix} 0&0&0&0\\ 1&0&0&0\\ 0&1&0&0\\ 0&0&1&0 \end{bmatrix}$.

Ans. ***Option b***

We have

$$T(1) = 1 = 1.1 + 0x + 0x^2 + 0x^3$$

$$T(x) = x + 1 = 1.1 + 1x + 0x^2 + 0x^3$$

$$T(x^2) = x^2 + 2x + 1 = 1.1 + 2x + 1x^2 + 0x^3$$

$$T(x^3) = x^3 + 3x^2 + 3x + 1 = 1.1 + 3x + 3x^2 + 1x^3$$

The matrix of T with respect to standard basis is $\begin{bmatrix} 1&1&1&1\\ 0&1&2&3\\ 0&0&1&3\\ 0&0&0&1 \end{bmatrix}$.

(59) Let $A = \begin{bmatrix} 1&0&0\\ 1&1&0\\ 1&1&1 \end{bmatrix}$. Then the matrix of the linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x) = Ax$ with respect to the basis $B = \left\{ \begin{pmatrix} 1\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\1 \end{pmatrix} \right\}$ over $\mathbb{R}$ is

(a) $\begin{bmatrix} 0&-1&-1\\ 0&0&-1\\ 1&2&3 \end{bmatrix}$ (b) $\begin{bmatrix} 0&1&1\\ 0&0&1\\ 1&2&3 \end{bmatrix}$ (c) $\begin{bmatrix} 0&1&1\\ 1&0&1\\ 1&2&3 \end{bmatrix}$ (d) $\begin{bmatrix} 0&-1&1\\ -1&0&2\\ 3&2&1 \end{bmatrix}$.

Ans. ***Option a***

We have

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right)=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\1\\1\end{bmatrix}=0\begin{bmatrix}1\\0\\0\end{bmatrix}+0\begin{bmatrix}1\\1\\0\end{bmatrix}+1\begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$T\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}1\\2\\2\end{bmatrix}=(-1)\begin{bmatrix}1\\0\\0\end{bmatrix}+0\begin{bmatrix}1\\1\\0\end{bmatrix}+2\begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)=\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}=(-1)\begin{bmatrix}1\\0\\0\end{bmatrix}+(-1)\begin{bmatrix}1\\1\\0\end{bmatrix}+3\begin{bmatrix}1\\1\\1\end{bmatrix}$$

The matrix of T *with respect to* B *is* $\begin{bmatrix}0&-1&-1\\0&0&-1\\1&2&3\end{bmatrix}$.

(60) If $T:\mathbb{P}_3[x]\to\mathbb{P}_4[x]$ is the linear transformation defined by

$$(Tp)\,(x)=x^2\frac{d}{dx}\left[p(x)\right]+\int_0^x p(t)dt$$

and $A=\left[a_{ij}\right]_{5\times 4}$ is the matrix of T with respect to standard bases of $\mathbb{P}_3[x]$ and $\mathbb{P}_4[x]$, then

(a) $a_{32}=\frac{3}{2}$ and $a_{33}=\frac{7}{3}$. (b) $a_{32}=\frac{3}{2}$ and $a_{33}=0$.
(c) $a_{32}=0$ and $a_{33}=\frac{7}{3}$. (d) $a_{32}=0$ and $a_{33}=0$.

Ans. ***Option b***

Consider the standard ordered basis for both $\mathbb{P}_3[x]$ *and* $\mathbb{P}_4[x]$*:*

$$T(1)=x=0.1+1x+0x^2+0x^3+0x^4$$

$$T(x)=\frac{3x^2}{2}=0.1+0x+\frac{3}{2}x^2+0x^3+0x^4$$

$$T(x^2)=\frac{7x^3}{3}=0.1+0x+0x^2+\frac{7}{3}x^3+0x^4$$

$$T(x^3)=\frac{13x^4}{4}=0.1+0x+0x^2+0x^3+\frac{13}{4}x^4$$

The matrix of T with respect to the given bases is given by $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{7}{3} & 0 \\ 0 & 0 & 0 & \frac{13}{4} \end{bmatrix}$.

(61) Consider the linear transformation $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ defined by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_3 + a_2x + a_1x^2 + a_0x^3$$

Then the matrix representation A of T with respect to the ordered basis $\{1, x, x^2, x^3\}$ satisfies

(a) $A^2 + I_4 = 0$ (b) $A^2 - I_4 = 0$ (c) $A - I_4 = 0$ (d) $A + I_4 = 0$.

Ans. ***Option b***
We have

$$T(1) = x^3 = 0.1 + 0x + 0x^2 + 1x^3$$

$$T(x) = x^2 = 0.1 + 0x + 1x^2 + 0x^3$$

$$T(x^2) = x = 0.1 + 1x + 0x^2 + 0x^3$$

$$T(x^3) = 1 = 1.1 + 0x + 0x^2 + 0x^3$$

Then $A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ *and* $A^2 = I_4$.

(62) Let $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ be the map given by

$$(Tp)(x) = \int_1^x p'(t)dt$$

where $p'(t)$ denotes the derivative of $p(t)$. If the matrix of T relative to the standard basis of $\mathbb{P}_3(\mathbb{R})$ is A, then $A + A^T$ is

(a) $\begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} -1 & 0 & 0 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 2 & 0 & 2 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 0 & 0 & -1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ -1 & 0 & -1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix}$.

Ans. ***Option d***
We have

$$T(1) = 0 = 0.1 + 0x + 0x^2 + 0x^3$$

$$T(x) = x - 1 = (-1).1 + 1x + 0x^2 + 0x^3$$

$$T(x^2) = x^2 - 1 = (-1).1 + 0x + 1x^2 + 0x^3$$

$$T(x^3) = x^3 - 1 = (-1).1 + 0x + 0x^2 + 1x^3$$

Therefore $A = \begin{bmatrix} 0 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ *and* $A^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$. *Hence* $A + A^T = \begin{bmatrix} 0 & -1 & -1 & -1 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 2 & 0 \\ -1 & 0 & 0 & 2 \end{bmatrix}$.

(63) Define a linear transformation $T : \mathbb{M}_{2\times2}(\mathbb{R}) \to \mathbb{M}_{2\times2}(\mathbb{R})$ by

$$T(A) = 2A + A^t$$

where A^t denotes the transpose of the matrix A. Then the trace of T equals ......

Ans. *We have*

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} = 3\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} = 0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 1\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix} = 0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 1\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)=\begin{bmatrix}0&0\\0&3\end{bmatrix}=0\begin{bmatrix}1&0\\0&0\end{bmatrix}+0\begin{bmatrix}0&1\\0&0\end{bmatrix}+0\begin{bmatrix}0&0\\1&0\end{bmatrix}+3\begin{bmatrix}0&0\\0&1\end{bmatrix}$$

Therefore the matrix of T is $\begin{bmatrix}3&0&0&0\\0&2&1&0\\0&1&2&0\\0&0&0&3\end{bmatrix}$ *and hence trace of T is 10.*

(64) Let $T:\mathbb{M}_{2\times2}(\mathbb{R})\to\mathbb{M}_{2\times2}(\mathbb{R})$ be the linear transformation defined by

$$T(A)=2A+3A^t$$

Write down the matrix of this transformation with respect to the basis $\{E_i : 1\le i\le 4\}$ where $E_1=\begin{bmatrix}1&0\\0&0\end{bmatrix}$, $E_2=\begin{bmatrix}0&1\\0&0\end{bmatrix}$, $E_3=\begin{bmatrix}0&0\\1&0\end{bmatrix}$, $E_4=\begin{bmatrix}0&0\\0&1\end{bmatrix}$.

Ans. *We have*

$$T\,(E_1)=2E_1+3{E_1}^t=\begin{bmatrix}5&0\\0&0\end{bmatrix}=5E_1+0E_2+0E_3+0E_4$$

$$T\,(E_2)=2E_2+3{E_2}^t=\begin{bmatrix}0&2\\3&0\end{bmatrix}=0E_1+2E_2+3E_3+0E_4$$

$$T\,(E_3)=2E_3+3{E_3}^t=\begin{bmatrix}0&3\\2&0\end{bmatrix}=0E_1+3E_2+2E_3+0E_4$$

$$T\,(E_4)=2E_1+3{E_1}^t=\begin{bmatrix}0&0\\0&5\end{bmatrix}=0E_1+0E_2+0E_3+5E_4$$

The matrix of T with respect to the given basis is $\begin{bmatrix}5&0&0&0\\0&2&3&0\\0&3&2&0\\0&0&0&5\end{bmatrix}$.

65) Let $B_1=\{(1,2),(2,-1)\}$ and $B_2=\{(1,0),(0,1)\}$ be ordered bases of $\mathbb{R}^2$. If $T:\mathbb{R}^2\to\mathbb{R}^2$ is a linear transformation such that $[T]_{B_1}^{B_2}=\begin{bmatrix}4&3\\2&-4\end{bmatrix}$, then $T(5,5)$ is equal to ....

(a) $(-9,8)$ (b) $(9,8)$ (c) $(-15,-2)$ (d) $(15,2)$.

Ans. ***Option d***

$[T]_{B_1}^{B_2}=\begin{bmatrix}4&3\\2&-4\end{bmatrix}\Rightarrow T(x,y)=aT(1,2)+bT(2,-1)$ *where* $(x,y)=a(1,2)+b(2,-1)$. *Then*

$$T(1, 2) = 4(1, 0) + 2(0, 1) = (4, 2)$$

and

$$T(2, -1) = 3(1, 0) - 4(0, 1) = (3, -4)$$

Thus $(5, 5) = 3(1, 2) + 1(2, -1) \Rightarrow T(5, 5) = 3(4, 2) + 1(3, -4) = (15, 2)$.

(66) Let $D : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ and $T : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_{n+1}(\mathbb{R})$ be the linear transformations defined by

$$D(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$$

$$T(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_0x + a_1x^2 + a_2x^3 + \cdots + a_nx^{n+1}$$

respectively. If A is the matrix representation of the transformation $DT - TD : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ with respect to the standard basis of $\mathbb{P}_n(\mathbb{R})$, then $tr(A) = \cdots\cdots$.

(a) $-n$ (b) n (c) $n + 1$ (d) $-(n + 1)$.

Ans. ***Option c***
We have

$$\begin{aligned} DT(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) &= D(a_0x + a_1x^2 + a_2x^3 + \cdots + a_nx^{n+1}) \\ &= a_0 + 2a_1x + \cdots + (n + 1)a_nx^n \end{aligned}$$

and

$$\begin{aligned} TD(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) &= T(a_1 + 2a_2x + \cdots + na_nx^{n-1}) \\ &= a_1x + 2a_2x^2 + \cdots + na_nx^n \end{aligned}$$

Therefore

$$(DT - TD)(a_0 + a_1x + \cdots + a_nx^n) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = I$$

Hence $tr(A) = tr(I) = n + 1$.

(67) Let V be a vector space of dimension 3 over $\mathbb{R}$. Let $T : V \to V$ be a linear transformation, given by the matrix $A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -4 & 3 \\ -2 & 5 & -3 \end{pmatrix}$ with respect to an ordered basis $\{v_1, v_2, v_3\}$ of V. Then which of the following statements are true?

(a) $T(v_3) = 0$ (b) $T(v_1 + v_2) = 0$
(c) $T(v_1 + v_2 + v_3) = 0$ (d) $T(v_1 + v_3) = T(v_2)$.

Ans. ***Option c***
Since A is the matrix of the linear transformation T, $T(v_1) = v_1 + v_2 - 2v_3$, $T(v_2) = -v_1 - 4v_2 + 5v_3$, *and* $T(v_3) = 3v_2 - 3v_3$.

(a) $T(v_3) = 0 \Rightarrow 3v_2 - 3v_3 = 0 \Rightarrow v_2 = v_3$, *which is not possible.*
(b) $T(v_1 + v_2) = 0 \Rightarrow T(v_1) + T(v_2) = -3v_2 + 3v_3 = 0 \Rightarrow v_2 = v_3$, *which is not possible.*
(c) $T(v_1 + v_2 + v_3) = T(v_1) + T(v_2) + T(v_3) = 0$.
(d) $T(v_1 + v_3) = -T(v_2)$.

(68) Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map which maps each point in $\mathbb{R}^2$ to its reflection on the $x-$axis. What is the determinant of T? What is its trace?

Ans. $T : \mathbb{R}^2 \to \mathbb{R}^2$ *is defined by* $T(x, y) = (x, -y)$. *Since* $T(1, 0) = (1, 0)$, $T(0, 1) = (0, -1)$, *the matrix of T is given by* $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *Therefore* $tr(T) = 0$ *and* $det(T) = -1$.

(69) Let $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ be defined by

$$(Tp)(x) = p(x) - x\frac{d}{dx}\left[p(x)\right]$$

Then the $Rank(T)$ is

(a) 1 (b) 2 (c) 3 (d) 4.

Ans. ***Option c***
Consider the standard basis $\left\{1, x, x^2, x^3\right\}$ *for V. We have*

$$T(1) = 1,\ T(x) = 0,\ T(x^2) = -x^2,\ T(x^3) = -2x^3$$

Therefore the matrix of T with respect to the standard basis is $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$ *which has rank*$= 3$.

(70) Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation defined by

$$T(x, y, z) = (x + y, x - z)$$

Then $dim\,[\mathcal{N}(T)]$ is

(a) 0 (b) 1 (c) 2 (d) 3.

Ans. ***Option b***
We have

$$T(1,0,0) = (1,1), T(0,1,0) = (1,0), T(0,0,1) = (0,-1)$$

Then the matrix of T is $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$. *Clearly,* $Rank(T) = 2$. *Hence by Rank-Nullity theorem,* $dim\,[\mathcal{N}(T)] = 1$.

(71) Consider the linear transformation

$$T(x, y, z) = (2x + y + z, x + z, 3x + 2y + z)$$

The rank of T is ......
Ans. *Consider the standard ordered basis of* $\mathbb{R}^3$. *Then*

$$T(1,0,0) = (2,1,3),\ T(0,1,0) = (1,0,2)\ T(0,0,1) = (1,1,1)$$

Therefore the matrix of T is $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix}$. *Clearly,* $Rank(T) = 2$.

(72) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x, y, z) = (x + y, y + z, z + x)$$

for all $(x, y, z) \in \mathbb{R}^3$. Then
(a) $Rank(T) = 0,\ Nullity(T) = 3$
(b) $Rank(T) = 2,\ Nullity(T) = 1$
(c) $Rank(T) = 1,\ Nullity(T) = 2$
(d) $Rank(T) = 3,\ Nullity(T) = 0$.

Ans. ***Option d***
Consider the standard ordered basis for $\mathbb{R}^3$. *Then we have*

$$T(1,0,0) = (1,0,1),\ T(0,1,0) = (1,1,0),\ T(0,0,1) = (0,1,1)$$

Therefore the matrix of T is $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ *which is of rank 3.By Rank-Nullity theorem,* $Nullity(T) = 0$.

(73) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x, y, z) = (x + 3y + 2z, 3x + 4y + z, 2x + y - z)$$

(i) The dimension of the $\mathcal{R}\left(T^2\right)$ is

(a) 0 (b) 1 (c) 2 (d) 3.

(ii) The dimension of the $\mathcal{N}\left(T^3\right)$ is

(a) 0 (b) 1 (c) 2 (d) 3.

Ans. *We have*

$$T(1,0,0) = (1,3,2) = 1(1,0,0) + 3(0,1,0) + 2(0,0,1)$$

$$T(0,1,0) = (3,4,1) = 3(1,0,0) + 4(0,1,0) + 1(0,0,1)$$

$$T(0,0,1) = (2,1,-1) = 2(1,0,0) + 1(0,1,0) + (-1)(0,0,1)$$

The matrix of T is given by $\begin{bmatrix} 1 & 3 & 2 \\ 3 & 4 & 1 \\ 2 & 1 & -1 \end{bmatrix}$.

(i) The matrix of T^2 *is given by* $\begin{bmatrix} 14 & 17 & 3 \\ 17 & 26 & 9 \\ 3 & 9 & 6 \end{bmatrix}$. *Therefore* $dim\mathcal{R}\left(T^2\right) = 2$.

Option c

(ii) The matrix of T^3 *is given by* $\begin{bmatrix} 71 & 113 & 42 \\ 113 & 164 & 51 \\ 42 & 51 & 9 \end{bmatrix}$. *Therefore* $dim\mathcal{R}\left(T^3\right) = 2$.

Hence by Rank-Nullity theorem, $Nullity(T^3) = 1$.

Option b

(74) Let $T : \mathbb{P}_n(\mathbb{R}) \to \mathbb{P}_n(\mathbb{R})$ be the map $(Tp)(x) = p'(1), x \in \mathbb{C}$. Which of the following are correct?

(a) $Nullity(T) = n$ (b) $Rank(T) = 1$
(c) $Nullity(T) = 1$ (d) $Rank(T) = n + 1$.

Ans. ***Options a and b***

We have

$$\begin{aligned} T(1) &= 0 = 0.1 + 0x + 0x^2 + \cdots + 0x^n \\ T(x) &= 1 = 1.1 + 0x + 0x^2 + \cdots + 0x^n \\ T(x^2) &= 2 = 2.1 + 0x + 0x^2 + \cdots + 0x^n \\ &\vdots \\ T(x^n) &= n = n.1 + 0x + 0x^2 + \cdots + 0x^n \end{aligned}$$

The matrix of T is $\begin{bmatrix} 0 & 1 & 2 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$ *which is of dimension* $(n+1) \times (n+1)$.
Clearly $Rank(T) = 1$ *and* $Nullity(T) = n$.

(75) Consider the map $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_4(\mathbb{R})$ defined by $(Tp)(x) = p(x^2)$. Then
(a) T is a linear transformation and $Rank(T) = 5$.
(b) T is a linear transformation and $Rank(T) = 3$.
(c) T is a linear transformation and $Rank(T) = 2$.
(d) T is not a linear transformation.

Ans. ***Option b***
Let $p, q \in \mathbb{P}_2(\mathbb{R})$ *and* $\lambda \in \mathbb{R}$. *Then*

$$\left[T(\lambda p + q)\right](x) = (\lambda p + q)(x^2) = \lambda p(x^2) + q(x^2) = \lambda(Tp)(x) + (Tq)(x)$$

Therefore T is a linear transformation. Now consider the standard order basis for both $\mathbb{P}_2(\mathbb{R})$ *and* $\mathbb{P}_4(\mathbb{R})$:

$$T(1) = 1 = 1.1 + 0x + 0x^2 + 0x^3 + 0x^4$$

$$T(x) = x^2 = 0.1 + 0x + 1x^2 + 0x^3 + 0x^4$$

$$T(x^2) = x^4 = 0.1 + 0x + 0x^2 + 0x^3 + 1x^4$$

Then the matrix of T is given by $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Since rank of matrix of T is 3,* $Rank(T) = 3$.

(76) Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ be the linear map satisfying

$$T(e_1) = e_2, \ T(e_2) = e_3, \ T(e_3) = 0, \ T(e_4) = e_3$$

where $\{e_1, e_2, e_3, e_4\}$ is the standard basis of $\mathbb{R}^4$. Then

(a) T is idempotent. (b) T is invertible.
(c) $Rank\ T = 3$. (d) T is nilpotent.

Ans. ***Option d***
We have

$$T(e_1) = e_2 = 0e_1 + 1e_2 + 0e_3 + 0e_4$$

$$T(e_2) = e_3 = 0e_1 + 0e_2 + 1e_3 + 0e_4$$

$$T(e_3) = 0 = 0e_1 + 0e_2 + 0e_3 + 0e_4$$

$$T(e_4) = e_3 = 0e_1 + 0e_2 + 1e_3 + 0e_4$$

Therefore the matrix of T with respect to the standard basis is $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Since* $T^2 \neq T$, *T is not idempotent. T has rank 2 and hence is not invertible. Since* $T^3 = 0$, *T is nilpotent.*

(77) Let n be a positive integer and V be an $(n+1)$-dimensional vector space over $\mathbb{R}$. If $\{e_1, e_2, \ldots, e_{n+1}\}$ is a basis of V and $T : V \to V$ is the linear transformation satisfying

$$T(e_i) = e_{i+1} \text{ for } i = 1, 2, \ldots, n \text{ and } T(e_{n+1}) = 0$$

then

(a) $tr(T) \neq 0$ (b) $Rank(T) = n$ (c) $Nullity(T) = 1$ (d) T^n is the zero map.

Ans. ***Options b and c***
We have

$$T(e_1) = e_2 = 0e_1 + 1e_2 + 0e_3 + \cdots + 0e_{n+1}$$

$$T(e_2) = e_3 = 0e_1 + 0e_2 + 1e_3 + \cdots + 0e_{n+1}$$

$$\vdots$$

$$T(e_n) = e_{n+1} = 0e_1 + 0e_2 + 0e_3 + \cdots + 1e_{n+1}$$

$$T(e_{n+1}) = 0 = 0e_1 + 0e_2 + 0e_3 + \cdots + 0e_{n+1}$$

Therefore the matrix of T is of the form, $T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$. *Then* $Rank(T) = n$ *and by Rank-Nullity theorem,* $Nullity(T) = 1$. *Clearly,* $tr(T) = 0$ *and* $T^n \neq 0$.

(78) Given a 4×4 real matrix A. For which choices of A given below do $\mathcal{R}(T)$ and $\mathcal{R}(T^2)$ have respective dimensions 2 and 1? ($*$ denotes a non-zero entry)

(a) $\begin{bmatrix} 0&0&*&* \\ 0&0&*&* \\ 0&0&0&* \\ 0&0&0&0 \end{bmatrix}$ (b) $\begin{bmatrix} 0&0&*&0 \\ 0&0&*&0 \\ 0&0&0&* \\ 0&0&0&* \end{bmatrix}$ (c) $\begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&* \\ 0&0&*&0 \end{bmatrix}$ (d) $\begin{bmatrix} 0&0&0&0 \\ 0&0&0&0 \\ 0&0&*&* \\ 0&0&*&* \end{bmatrix}$.

Ans. ***Options a and b***

(a) Since $Rank(A) = 2$ and $Rank(A^2) = 1$, $\mathcal{R}(T)$ and $\mathcal{R}(T^2)$ have respective dimensions 2 and 1.
(b) Since $Rank(A) = 2$ and $Rank(A^2) = 1$, $\mathcal{R}(T)$ and $\mathcal{R}(T^2)$ have respective dimensions 2 and 1.
(c) Since $Rank(A) = 2$ and $Rank(A^2) = 2$, $\mathcal{R}(T)$ and $\mathcal{R}(T^2)$ have respective dimension 2 each.
(d) Since $Rank(A) = 1$ and $Rank(A^2) = 1$, $\mathcal{R}(T)$ and $\mathcal{R}(T^2)$ have respective dimension 1 each.

(79) Fix a non-singular matrix $A = \left(a_{ij}\right) \in \mathbb{M}_{n\times n}(\mathbb{K})$, and consider the linear map $T : \mathbb{M}_{n\times n}(\mathbb{K}) \to \mathbb{M}_{n\times n}(\mathbb{K})$ given by $T(X) = AX$. Then

(a) $tr(T) = n\sum_{i=1}^{n} a_{ii}$. (b) $tr(T) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}$.
(c) Rank of T is n^2. (d) T is non-singular.

Ans. ***Options a, c, and d***

Take $n = 2$ and fix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Then

$$T\left(\begin{pmatrix} 1&0 \\ 0&0 \end{pmatrix}\right) = \begin{pmatrix} a_{11}&0 \\ a_{21}&0 \end{pmatrix} = a_{11}\begin{pmatrix} 1&0 \\ 0&0 \end{pmatrix} + 0\begin{pmatrix} 0&1 \\ 0&0 \end{pmatrix} + a_{21}\begin{pmatrix} 0&0 \\ 1&0 \end{pmatrix} + 0\begin{pmatrix} 0&0 \\ 0&1 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 0&1 \\ 0&0 \end{pmatrix}\right) = \begin{pmatrix} 0&a_{11} \\ 0&a_{21} \end{pmatrix} = 0\begin{pmatrix} 1&0 \\ 0&0 \end{pmatrix} + a_{11}\begin{pmatrix} 0&1 \\ 0&0 \end{pmatrix} + 0\begin{pmatrix} 0&0 \\ 1&0 \end{pmatrix} + a_{21}\begin{pmatrix} 0&0 \\ 0&1 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 0&0 \\ 1&0 \end{pmatrix}\right) = \begin{pmatrix} a_{12}&0 \\ a_{22}&0 \end{pmatrix} = a_{12}\begin{pmatrix} 1&0 \\ 0&0 \end{pmatrix} + 0\begin{pmatrix} 0&1 \\ 0&0 \end{pmatrix} + a_{22}\begin{pmatrix} 0&0 \\ 1&0 \end{pmatrix} + 0\begin{pmatrix} 0&0 \\ 0&1 \end{pmatrix}$$

$$T\left(\begin{pmatrix} 0&0 \\ 0&1 \end{pmatrix}\right) = \begin{pmatrix} 0&a_{12} \\ 0&a_{22} \end{pmatrix} = 0\begin{pmatrix} 1&0 \\ 0&0 \end{pmatrix} + a_{12}\begin{pmatrix} 0&1 \\ 0&0 \end{pmatrix} + 0\begin{pmatrix} 0&0 \\ 1&0 \end{pmatrix} + a_{22}\begin{pmatrix} 0&0 \\ 0&1 \end{pmatrix}$$

Therefore the matrix of T is $\begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & a_{21} & 0 & a_{22} \end{bmatrix}$. Then

$$det(A) = a_{11}^2 a_{22}^2 - 2a_{11}a_{22}a_{12}a_{21} + a_{12}^2 a_{21}^2 = (a_{11}a_{22} - a_{12}a_{21})^2 \neq 0$$

and

$$tr(T) = 2a_{11} + 2a_{22} = 2\,(a_{11} + a_{22}) = 2\sum_{i=1}^{2} a_{ii}$$

Therefore rank of T is n^2 and T is non-singular.

(80) Let $T \in \mathbb{M}_{m\times n}(\mathbb{R})$. Let V be the subspace of $\mathbb{M}_{n\times p}(\mathbb{R})$ defined by

$$V = \left\{X \in \mathbb{M}_{n\times p}(\mathbb{R}) : TX = 0\right\}$$

Then the dimension of V is

(a) $pn - Rank(T)$ (b) $mn - pRank(T)$
(c) $p\,(m - Rank\,(T))$ (d) $p\,(n - Rank\,(T))$.

Ans. *Option d*
Consider $X = \left[x_1 \mid x_2 \mid \cdots \mid x_p\right] \in \mathbb{M}_{n\times p}(\mathbb{R})$ *where* $x_i \in \mathbb{M}_{n\times 1}, i = 1, 2, \ldots, p$. *Then*

$$TX = 0 \Leftrightarrow Tx_i = 0\ \forall\ i = 1, 2, \ldots, p \Leftrightarrow x_i \in \mathcal{N}(T)\ \forall\ i = 1, 2, \ldots, p$$

As $\mathcal{N}(T)$ has dimension $n - Rank(T)$, V has dimension $p\,(n - Rank\,(T))$.

OR

Take $\mathbb{M}_{m\times n}(\mathbb{R}) = \mathbb{M}_{3\times 2}(\mathbb{R})$ *and* $\mathbb{M}_{n\times p}(\mathbb{R}) = \mathbb{M}_{2\times 2}(\mathbb{R})$. *Fix* $T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. *Let* $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}_{2\times 2}(\mathbb{R})$. *Then* $TX = 0 \Rightarrow X = 0$. *Here* $Rank(T) = 2$ *and* $dim(V) = 0$. *Therefore options* (a), (b), *and* (c) *are false.*

(81) Let a linear transformation $T : \mathbb{P}_2\,(\mathbb{R}) \to \mathbb{P}_2\,(\mathbb{R})$ be defined by

$$T(a_0 + a_1x + a_2x^2) = (a_0 + a_1) + (a_1 - a_2)x + (a_0 + a_2)x^2$$

Consider the following statements:

I. The null space of T is $\{\lambda(-1 + x + x^2); \lambda \in \mathbb{R}\}$.
II. The range space of T is spanned by the set $\{1 + x^2, 1 + x\}$.
III. $T(T(1 + x)) = 1 + x^2$.
IV. If A is the matrix representation of T with respect to the standard basis $\{1, x, x^2\}$ of $\mathbb{P}_2\,(\mathbb{R})$, then the trace of the matrix A is 3.

Which of the following statements are TRUE?

(a) *I* and *II* only. (b) *I*, *III*, and *IV* only.
(c) *I*, *II*, and *IV* only. (d) *II* and *IV* only.

Ans. ***Option c***

We have

$$T(1) = 1 + x^2 = 1.1 + 0x + 1x^2$$

$$T(x) = 1 + x = 1.1 + 1x + 0x^2$$

$$T(x^2) = -x + x^2 = 0.1 + (-1)x + 1x^2$$

The matrix of T is given by $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$.

I. *Observe that* $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$. *Therefore the null space of T is given by*

$$\begin{aligned} \mathcal{N}(T) &= \{a_0 + a_1x + a_2x^2 \in \mathbb{P}_2(\mathbb{R}) : -a_0 = a_1 = a_2,\ a_0, a_1, a_2 \in \mathbb{R}\} \\ &= \{\lambda(-1 + x + x^2); \lambda \in \mathbb{R}\} \end{aligned}$$

II. *Since the third column of A can be spanned by the first two columns, the range space of T is spanned by the set* $\{1 + x^2, 1 + x\}$.

III. *We have*

$$T(T(1 + x)) = T(2 + x + x^2) = 3 + 3x^2 = 3(1 + x^2)$$

IV. *Clearly,* $tr(A) = 3$.

(82) For $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{P}_n[x]$, define a linear transformation $T : \mathbb{P}_n[x] \to \mathbb{P}_n[x]$ by

$$(Tp)(x) = a_n + a_{n-1}x + \cdots + a_0x^n$$

Then which of the following are correct?

(a) T is one-one. (b) T is onto.
(c) T is invertible. (d) $det(T) = \pm 1$.

Ans. ***Options a, b, c, and d***

We have

$$T(1) = x^n = 0.1 + 0x + \cdots + 0x^{n-1} + 1x^n$$

$$T(x) = x^{n-1} = 0.1 + 0x + \cdots + 1x^{n-1} + 0x^n$$

$$\vdots$$

$$T(x^n) = 1 = 1 + 0x + \cdots + 0x^{n-1} + 0x^n$$

The matrix of T *is* $\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$. *Since* T *can be obtained by interchanging the rows of identity matrix,* $det(T) = \pm 1$. T *is one-one, onto, and invertible.*

(83) For $p(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{P}_n[x]$, define a linear transformation $T : \mathbb{P}_n[x] \to \mathbb{P}_n[x]$ by

$$(Tp)(x) = a_0 - a_1x + a_2x^2 - \cdots + (-1)^n a_nx^n$$

Then which of the following are correct?

(a) T is one-one. (b) T is onto.
(c) T is invertible. (d) $det(T) = 0$.

Ans. ***Options a, b, and c***

We have

$$T(1) = 1 = 1.1 + 0x + \cdots + 0x^{n-1} + 0x^n$$

$$T(x) = -x = 0.1 + (-1)x + \cdots + 1x^{n-1} + 0x^n$$

$$\vdots$$

$$T(x^n) = (-1)^n = 0.1 + 0x + \cdots + 0x^{n-1} + (-1)^n x^n$$

Therefore the matrix of T *is*

$$\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & (-1)^n \end{bmatrix}_{(n+1)\times(n+1)}$$

Clearly T is one-one, onto, and invertible. Also $det(T) = \pm 1$.

(84) Let V be the space of twice differentiable functions on $\mathbb{R}$ satisfying

$$f'' - 2f' + f = 0$$

Define $T : V \to \mathbb{R}^2$ by $T(f) = \big(f'(0), f(0)\big)$. Then T is

(a) one-one and onto. (b) one-one but not onto.
(c) onto but not one-one. (d) neither one-one nor onto.

Ans. ***Option a***
The auxiliary equation of the given matrix is $m^2 - 2m + 1 = 0 \Rightarrow m = 1, 1$. *Therefore the general solution is given by* $y(x) = \lambda_1 e^x + \lambda_2 x e^x$. *That is, the solution space V is spanned by* $\{e^x, xe^x\}$. *As*

$$T(e^x) = (1, 1) = 1(1, 0) + 1(0, 1)$$

$$T(xe^x) = (1, 0) = 1(1, 0) + 0(0, 1)$$

the matrix of T is given by $T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. *Since T has rank 2,* $\mathcal{N}(T) = \{0\}$. *Therefore T is both one-one and onto.*

(85) The least positive integer n such that $\begin{bmatrix} cos\ \frac{\pi}{4} & sin\ \frac{\pi}{4} \\ -sin\ \frac{\pi}{4} & cos\ \frac{\pi}{4} \end{bmatrix}^n$ is the identity matrix of order 2 is

(a) 4 (b) 8 (c) 12 (d) 16.

Ans. ***Option a***
Since $\begin{bmatrix} cos\ \theta & sin\ \theta \\ -sin\ \theta & cos\ \theta \end{bmatrix}^n = \begin{bmatrix} cos\ n\theta & sin\ n\theta \\ -sin\ n\theta & cos\ n\theta \end{bmatrix}$, $n = 4$.

(86) Let A be the 2×2 matrix $\begin{pmatrix} sin\ \frac{\pi}{18} & -sin\ \frac{4\pi}{9} \\ sin\ \frac{4\pi}{9} & sin\ \frac{\pi}{18} \end{pmatrix}$. Then the smallest number $n \in \mathbb{N}$ such that $A^n = I$ is

(a) 3 (b) 9 (c) 18 (d) 27.

Ans. ***Option b***
Since $cos\ (\frac{\pi}{2} - \theta) = sin\ \theta$, $sin\ \frac{\pi}{18} = cos\ \frac{4\pi}{9}$ *and hence* $A = \begin{pmatrix} cos\ \frac{4\pi}{9} & -sin\ \frac{4\pi}{9} \\ sin\ \frac{4\pi}{9} & cos\ \frac{4\pi}{9} \end{pmatrix}$.
We know that

$$\begin{pmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{pmatrix}^n = \begin{pmatrix} cos\ n\theta & -sin\ n\theta \\ sin\ n\theta & cos\ n\theta \end{pmatrix}$$

Therefore $n = 9$.

(87) Consider the matrix $A = \begin{pmatrix} cos\ \theta & sin\ \theta \\ -sin\ \theta & cos\ \theta \end{pmatrix}$, where $\theta = \dfrac{2\pi}{31}$. Then A^{2015} equals

(a) A (b) I (c) $\begin{pmatrix} cos\ 13\theta & sin\ 13\theta \\ -sin\ 13\theta & cos\ 13\theta \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Ans. ***Option b***

Since $\begin{pmatrix} cos\ \theta & sin\ \theta \\ -sin\ \theta & cos\ \theta \end{pmatrix}^n = \begin{pmatrix} cos\ n\theta & sin\ n\theta \\ -sin\ n\theta & cos\ n\theta \end{pmatrix}$, *we have*

$$A^{2015} = \begin{pmatrix} cos\ 130\theta & sin\ 130\theta \\ -sin\ 130\theta & cos\ 130\theta \end{pmatrix} = I$$

(88) For the matrix A as given below, which of the them satisfy $A^6 = I$?

(a) $A = \begin{pmatrix} cos\ \frac{\pi}{4} & sin\ \frac{\pi}{4} & 0 \\ -sin\ \frac{\pi}{4} & cos\ \frac{\pi}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\ \frac{\pi}{3} & sin\ \frac{\pi}{3} \\ 0 & -sin\ \frac{\pi}{3} & cos\ \frac{\pi}{3} \end{pmatrix}$

(c) $A = \begin{pmatrix} cos\ \frac{\pi}{6} & 0 & sin\ \frac{\pi}{6} \\ 0 & 1 & 0 \\ -sin\ \frac{\pi}{6} & 0 & cos\ \frac{\pi}{6} \end{pmatrix}$ (d) $A = \begin{pmatrix} cos\ \frac{\pi}{2} & sin\ \frac{\pi}{2} & 0 \\ -sin\ \frac{\pi}{2} & cos\ \frac{\pi}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Ans. ***Option c***

$A = \begin{pmatrix} cos\ \frac{\pi}{6} & 0 & sin\ \frac{\pi}{6} \\ 0 & 1 & 0 \\ -sin\ \frac{\pi}{6} & 0 & cos\ \frac{\pi}{6} \end{pmatrix}$ *is the rotation matrix in* $\mathbb{R}^3$ *about y-axis. Therefore* $A^6 = I$.

(89) Let A be a 5×3 real matrix of rank 2. Let $b \in \mathbb{R}^5$ be a non-zero vector that is in the column space of A. Let $S = \{x \in \mathbb{R}^3 : Ax = b\}$. Define the translation of a subspace V of $\mathbb{R}^3$ by $x_0 \in \mathbb{R}^3$ as the set $x_0 + V = \{x_0 + v : v \in V\}$. Then
(a) S is the empty set
(b) S has only one element
(c) S is a translation of a one-dimensional subspace
(d) S is a translation of a two-dimensional subspace.

Ans. ***Option c***

Given that the system has a solution. Therefore Rank $[A \mid b] = Rank\ (A) = 2 <$ *number of unknowns. Therefore dim* $(S) = 3 - 2 = 1$.

(90) Check whether the following statements are true or false?

(a) Consider the map $T : \mathbb{P}_5 \to \mathbb{R}^2$, given by $(Tp)(x) = (p(3), p'(3))$. Then $Nullity(T) = 3$.

(b) Consider the map $T : \mathbb{P}_2 \to \mathbb{P}_2$, given by $(Tp)(x) = p(x) + p'(x)$. Then T is invertible.

(c) Let V be the vector space over $\mathbb{R}$ consisting of polynomials of degree less than or equal to 3. Let $T : V \to V$ be the operator sending $f(x)$ to $f(x+1)$, and $D : V \to V$ the operator sending $f(x)$ to $\frac{d}{dx}\left[f(x)\right]$. Then T is a polynomial in D.

(d) An invertible linear map from $\mathbb{R}^2$ to itself takes parallel to parallel lines.

Ans. *(a) False. We know that $\mathbb{P}_5$ is of dimension 6. Suppose that $Nullity(T) = 3$. Then by Rank-Nullity Theorem,*

$$6 = dim(V) = Rank(T) + Nullity(T) \Rightarrow Rank(T) = 3$$

which is not possible as $\mathbb{R}^2$ is a vector space of dimension 2 and it cannot have a subspace of dimension 3.

(b) True. Consider the standard ordered basis for $\mathbb{P}_2$. Then

$$T(1) = 1,\ T(x) = x + 1,\ T(x^2) = x^2 + 2x$$

Hence the matrix of T is given by $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$. Clearly T is invertible.

(c) True. Consider the standard ordered basis for V. Then

$$T(1) = 1,\ T(x) = x + 1,\ T(x^2) = x^2 + 2x + 1,\ T(x^3) = x^3 + 3x^2 + 3x + 1$$

and the matrix of T is $T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Also

$$D(1) = 0,\ D(x) = 1,\ D(x^2) = 2x,\ D(x^3) = 3x^2$$

and the matrix of D is $D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Then $T = \frac{1}{6}D^3 + \frac{1}{2}D^2 + D + I$.

(d) *True. Consider two parallel lines in* $\mathbb{R}^2$ *given by* $L_1 = \{\lambda \overrightarrow{x} + \overrightarrow{y_1} \mid \lambda \in \mathbb{R}\}$ *and* $L_2 = \{\lambda \overrightarrow{x} + \overrightarrow{y_2} \mid \lambda \in \mathbb{R}\}$ *passing through* $\overrightarrow{y_1}, \overrightarrow{y_2} \in \mathbb{R}^2$ *and* $\overrightarrow{x} \in \mathbb{R}^2$ *is its direction vector.*
$T(L_1) = \{\lambda T(\overrightarrow{x}) + T(\overrightarrow{y_1}) \mid \lambda \in \mathbb{R}\}$ *and* $T(L_2) = \{\lambda T(\overrightarrow{x}) + T(\overrightarrow{y_2}) \mid \lambda \in \mathbb{R}\}$*. Since* T *is invertible* $T(\overrightarrow{x}) \neq 0$ *if* $\overrightarrow{x} \neq 0$ *and* $T(L_1)$ *and* $T(L_2)$ *are parallel lines with direction vector* $T(\overrightarrow{x})$ *and passing through* $T(\overrightarrow{y_1})$ *and* $T(\overrightarrow{y_2})$*, respectively.*

Chapter 11
Solved Problems—Eigenvalues and Eigenvectors

(1) Which of the following is an eigenvector of the matrix $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$?

(a) $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ (b) $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ (c) $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ (d) $\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$.

Ans. ***Option a***
As A is an upper triangular matrix, the eigenvalues of A are its diagonal entries. Therefore the only eigenvalue of A is 2. Also $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

(2) The imaginary parts of the eigenvalues of the matrix $A = \begin{bmatrix} 3 & 2 & 5 \\ 2 & -3 & 6 \\ 0 & 0 & -3 \end{bmatrix}$ are

(a) 0, 0, 0 (b) 2, −2, 0 (c) 1, −2, 0 (d) 3, −3, 0.

Ans. ***Option a***
The characteristic equation of A is given by $(3 + \lambda)(\lambda^2 - 13) = 0$. *Therefore it has no complex eigenvalues.*

(3) The number of linearly independent eigenvectors of the matrix $\begin{bmatrix} 2 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 \end{bmatrix}$ is

(a) 1 (b) 2 (c) 3 (d) 4.

Ans. ***Option d***
Since the given matrix is a block diagonal matrix, the eigenvalues of the given matrices are precisely the eigenvalues of the sub-block matrices. That is, in

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_11

this case the eigenvalues of the given matrix are eigenvalues of the matrices $\begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}$ *and* $\begin{bmatrix} 3 & 0 \\ 1 & 4 \end{bmatrix}$. *Therefore the given matrix has 4 distinct eigenvalues. Since eigenvectors corresponding to different eigenvalues are linearly independent, there are 4 linearly independent eigenvectors for the given matrix.*

(4) Let $p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n$ be the characteristic polynomial of a $n \times n$ matrix A with entries in $\mathbb{R}$. Then which of the following statements is true?

(a) $p(\lambda)$ has no repeated roots.
(b) $p(\lambda)$ can be expressed as a product of linear polynomials with real coefficients.
(c) If $p(\lambda)$ can be expressed as a product of linear polynomials with real coefficients, then there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A.

Ans. ***Option c***

(a) Consider the identity matrix. Then the characteristic polynomial is $(\lambda - 1)^n$ *which has repeated roots.*
(b) Consider the matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. *Then the characteristic polynomial is* $\lambda^2 + 1$ *which cannot be expressed as a product of linear polynomials with real coefficients.*
(c) $p(\lambda)$ *can be expressed as a product of linear polynomials with real coefficients implying that it has n distinct eigenvalues. Since eigenvalues corresponding to distinct eigenvalues are linearly independent, the collection of all eigenvectors is a linearly independent set of cardinality n in* $\mathbb{R}^n$. *Therefore* $\mathbb{R}^n$ *has a basis consisting of eigenvectors of A.*

(5) If $\begin{bmatrix} 2 \\ y \\ z \end{bmatrix}$ $(y, z \in \mathbb{R})$ is an eigenvector corresponding to a real eigenvalue of the matrix $\begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & -4 \\ 0 & 1 & 3 \end{bmatrix}$, then $z - y$ is equal to ....

Ans. *The characteristic equation of the above matrix is*

$$\lambda^3 - 3\lambda^2 + 4\lambda - 2 = (\lambda - 1)(\lambda^2 - 2\lambda + 2) = 0 \Rightarrow \lambda = 1 \text{ or } \lambda = 1 \pm i$$

The only real eigenvalue of A is 1. Then,

$$\begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & -4 \\ 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ y \\ z \end{bmatrix} = 1 \begin{bmatrix} 2 \\ y \\ z \end{bmatrix} \Rightarrow 2z = 2 \text{ and } 2 - 4z = y \Rightarrow z = 1, y = -2$$

Therefore $z - y = 3$.

(6) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\,\theta & -sin\,\theta \\ 0 & sin\,\theta & cos\,\theta \end{bmatrix}$, where $0 < \theta < \frac{\pi}{2}$. Let $V = \{v \in \mathbb{R}^3 : Av^T = v^T\}$. Then the dimension of V is
(a) 0 (b) 1 (c) 2 (d) 3.

Ans. ***Option b***
V is the eigenspace of A with respect to the eigenvalue 1. Thus,

$$\begin{aligned} V &= \{v \in \mathbb{R}^3 : Av^T = v^T\} \\ &= \{v \in \mathbb{R}^3 : (A - I)v^T = 0\} \\ &= \mathcal{N}(A - I) \end{aligned}$$

Since $0 < \theta < \frac{\pi}{2}$, *Rank*$(A - I) = 2$ *and hence dim* $[\mathcal{N}(A - I)] = 3 - 2 = 1$.

(7) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation such that

$$T(1, 2, 3) = (1, 2, 3), T(1, 5, 0) = (2, 10, 0) \text{ and } T(-1, 2, -1) = (-3, 6, -3)$$

Then dimension of the vector space spanned by all the eigenvectors of T is
(a) 0 (b) 1 (c) 2 (d) 3.

Ans. ***Option d***
Since

$$T(1, 2, 3) = (1, 2, 3) = 1(1, 2, 3)$$

$$T(1, 5, 0) = (2, 10, 0) = 2(1, 5, 0)$$

$$T(-1, 2, -1) = (-3, 6, -3) = 3(-1, 2, -1)$$

$1, 2,$ *and* 3 *are eigenvalues of* T. *Since* T *has three distinct eigenvalues, eigenvectors of* T *are linearly independent. As* $\mathbb{R}^3$ *has dimension 3, the eigenvectors of* T *spans* $\mathbb{R}^3$.

(8) Let $D : \mathbb{P}_4[x] \to \mathbb{P}_4[x]$ be the linear operator that takes any polynomial $p(x)$ to its derivative $p'(x)$. Then the characteristic polynomial $f(x)$ of D is
(a) x^4 (b) x^5 (c) $x^3(x - 1)$ (d) $x^4(x - 1)$.

Ans. ***Option b***
Consider the standard basis $\{1, x, x^2, x^3, x^4\}$. *Then*

$$D(1) = 0 = 0.1 + 0x + 0x^2 + 0x^3 + 0x^4$$

$$D(x) = 1 = 1.1 + 0x + 0x^2 + 0x^3 + 0x^4$$

$$D(x^2) = 2x = 0.1 + 2x + 0x^2 + 0x^3 + 0x^4$$

$$D(x^3) = 3x^2 = 0.1 + 0x + 3x^2 + 0x^3 + 0x^4$$

$$D(x^4) = 4x^3 = 0.1 + 0x + 0x^2 + 4x^3 + 0x^4$$

Then the matrix of T is $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. *Then the characteristic polynomial* $f(x)$ *of D is* x^5.

(9) Let n be an odd number ≥ 7. Let $A = [a_{ij}]$ be an $n \times n$ matrix with $a_{i(i+1)} = 1$ for all $i = 1, 2, \ldots, n-1$ and $a_{n1} = 1$. Let $a_{ij} = 0$ for all other pairs (i, j). Then we can conclude that
(a) A has 1 as an eigenvalue (b) A has -1 as an eigenvalue
(c) A has no real eigenvalues (d) A has at least one eigenvalue with $AM \geq 2$.

Ans. ***Option a***
We have

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix}$$

The matrix is obtained by interchanging the rows of the identity matrix of order n and hence the characteristic polynomial is $\lambda^n - 1$. *The eigenvalues of A are precisely the* nth *roots of unity.*

(10) Let $P_A(\lambda)$ denote the characteristic polynomial of a matrix A. Then for which of the following matrices $P_A(\lambda) - P_{A^{-1}}(\lambda)$ is a constant?
(a) $\begin{pmatrix} 3 & 3 \\ 2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 4 & 3 \\ 2 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$ (d) $\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix}$.

Ans. ***Option c***
Let λ_1 *and* λ_2 *be two eigenvalues of A, then the eigenvalues of* A^{-1} *are* $\frac{1}{\lambda_1}$ *and* $\frac{1}{\lambda_2}$ *(since A is invertible). Hence the characteristic polynomials for A and* A^{-1} *are given by* $P_A(\lambda) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$ *and* $P_{A^{-1}}(\lambda) = \lambda^2 - (\frac{1}{\lambda_1} + \frac{1}{\lambda_2})\lambda + \frac{1}{\lambda_1\lambda_2}$. *Now* $P_A(\lambda) - P_{A^{-1}}(\lambda)$ *is a constant only if*

$$\begin{aligned} \lambda_1 + \lambda_2 = \frac{1}{\lambda_1} + \frac{1}{\lambda_2} &\Rightarrow (\lambda_1 + \lambda_2)\left(1 - \frac{1}{\lambda_1\lambda_2}\right) = 0 \\ &\Rightarrow \lambda_1 + \lambda_2 = 0 \; or \; \lambda_1\lambda_2 = 1 \\ &\Rightarrow trace = 0 \; or \; determinant = 1 \end{aligned}$$

Only $\begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$ *satisfy one of these conditions.*

(11) Let ω be a complex number such that $\omega^3 = 1$, but $\omega \neq 1$. If $A = \begin{bmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega^2 & \omega & 1 \end{bmatrix}$, then which of the following statements are true?

(a) A is invertible.
(b) $Rank(A) = 2$.
(c) 0 is an eigenvalue of A.
(d) there exist linearly independent vectors $v, w \in \mathbb{C}^3$ such that $Av = Aw = 0$.

Ans. ***Options b and c***

Since $det(A) = 0$, A is not invertible and 0 is an eigenvalue of A. As $\begin{vmatrix} 1 & \omega \\ \omega & \omega^2 \end{vmatrix} \neq 0$, $Rank(A) = 2$ and hence $Ax = 0$ has only one linearly independent solution.

(12) Let $M = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \text{ and the eigenvalues of } A \text{ be } in\ \mathbb{Q} \right\}$. Then

(a) M is empty
(b) $M = \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$
(c) if $A \in M$ then the eigenvalues of A are in $\mathbb{Z}$
(d) if $A, B \in M$ are such that $AB = I$ then $det(A) \in \{+1, -1\}$.

Ans. ***Options c and d***

(a) Let $A = I$. Then clearly $A \in M$. Thus, M is non-empty.
(b) Consider the matrix $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$. Clearly, $B \notin M$. Therefore $M \neq \left\{ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$.
(c) The characteristic polynomial has rational roots implying that it is reducible over $\mathbb{Q}$ which gives the reducibility of the characteristic polynomial over $\mathbb{Z}$. Therefore if $A \in M$ then the eigenvalues of A are in $\mathbb{Z}$.
(d) If $A, B \in M$, then their determinant must be an integer (as $det(A) = ad - bc$). Now $AB = I \Rightarrow det(AB) = det(A)det(B) = 1$. Therefore $det(A) \in \{+1, -1\}$.

(13) Let $\{u_1, \ldots, u_n\}$ and $\{v_1, \ldots, v_n\}$ be two bases of $\mathbb{R}^n$. Let A be an $n \times n$ matrix with real entries such that $Au_i = v_i,\ i = 1, 2, \ldots, n$. Suppose that every eigenvalue of A is either -1 or 1. Let $B = I + 2A$. Then which of the following statements are true?

(a) $\{u_i + 2v_i \mid i = 1, 2, \ldots, n\}$ is also a basis of V.
(b) B is invertible.
(c) Every eigenvalue of B is either 3 or -1.
(d) $det(B) > 0$ if $det(A) > 0$.

Ans. ***Options a, b, c, and d***
Let $c_1, \ldots, c_n \in \mathbb{R}$ *be such that*

$$c_1(u_1 + 2v_1) + \cdots + c_n(u_n + 2v_n) = 0$$

Since $Au_i = v_i$ *this implies that*

$$c_1u_1 + \cdots + c_nu_n + 2A(c_1u_1 + \cdots + c_nu_n) = 0$$

$$\Rightarrow A(c_1u_1 + \cdots + c_nu_n) = \frac{-1}{2}(c_1u_1 + \cdots + c_nu_n)$$

Since every eigenvalue of P is either -1 *or 1, this is possible only when* $c_1u_1 + \cdots + c_nu_n = 0$. $\{u_1, \ldots, u_n\}$ *is a basis implying that*

$$c_1 = c_2 = \cdots = c_n = 0$$

Hence $\{u_i + 2v_i \mid i = 1, 2, \ldots, n\}$ *is also a basis of* V. *Since every eigenvalue of A is either* -1 *or 1, the eigenvalues of B are either* $2(1) + 1 = 3$ *or* $2(-1) + 1 = -1$. *Since 0 is not an eigenvalue B is invertible. Since the algebraic multiplicity of* -1 *as an eigenvalue of A and as an eigenvalue of B are the same,* $det(B) > 0$ *if* $det(A) > 0$.

(14) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation of $\mathbb{R}^n$, where $n \geq 3$, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of T. Which of the following statements are true?

(a) If $\lambda_i = 0$, for some $i = 1, 2, \ldots, n$, then T is not surjective.
(b) If T is injective, then $\lambda_i = 1$ for some $i,\ 1 \leq i \leq n$.
(c) If there is a three-dimensional subspace W of V such that $T(W) = W$, then $\lambda_i \in \mathbb{R}$ for some $i,\ 1 \leq i \leq n$.

Ans. ***Options a and c***

(a) If one of the eigenvalues is zero, then $Rank(T) < n$. *Therefore T cannot be surjective.*
(b) T is injective need not imply $\lambda_i = 1$ *for some i. For example, consider the linear transformation* $T : \mathbb{R}^3 \to \mathbb{R}^3$ *given by* $T(x, y, z) = (2x, 3y, 4z)$. *T is injective but 1 is not an eigenvalue of T.*
(c) If we consider T as a linear transformation from W to itself, the characteristic polynomial of T will be of degree 3 with real coefficients. It always has a real root.

(15) Let A be a 5×5 matrix with real entries such that the sum of the entries in each row of A is 1. Then the sum of all the entries in A^3 is
(a) 3 (b) 15 (c) 5 (d) 125.

Ans. ***Option c***
Since the sum of all entries in each row of A is 1, 1 is an eigenvalue of A with eigenvector $\begin{bmatrix}1 & 1 & 1 & 1 & 1\end{bmatrix}^T$. *Then 1 is an eigenvalue of* A^3 *with the same eigenvector and hence the sum of entries in each row is 1. Therefore the sum of all the entries in* A^3 *is 5.*

(16) Let A be an invertible 10×10 matrix with real entries such that the sum of each row is 1. Then

(a) The sum of entries of each row of the inverse of A is 1.
(b) The sum of entries of each column of the inverse of A is 1.
(c) The trace of the inverse of A is non-zero.
(d) None of the above.

Ans. ***Option a***

(a) We have

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1_{10}} \\ a_{21} & a_{22} & \cdots & a_{2_{10}} \\ \vdots & \vdots & \ddots & \vdots \\ a_{10_1} & a_{10_2} & \cdots & a_{10_{10}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} + a_{12} + \cdots + a_{1_{10}} \\ a_{21} + a_{22} + \cdots + a_{2_{10}} \\ \vdots \\ a_{10_1} + a_{10_2} + \cdots + a_{10_{10}} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Therefore 1 is an eigenvalue of A with eigenvector $\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$. *Hence 1 is an eigenvalue of* A^{-1} *with eigenvector* $\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}^T$ *(Theorem 4.4). So sum of entries of each row of the inverse of A is 1.*

(b) Consider $B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$. *Observe that* $B^{-1} = \begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$. *Take* $A = \begin{bmatrix} B & 0_{2\times 8} \\ 0_{8\times 2} & I_8 \end{bmatrix}$ *R. Then A is invertible with sum of entries in each row being 1. Also,* $A^{-1} = \begin{bmatrix} B^{-1} & 0_{2\times 8} \\ 0_{8\times 2} & I_8 \end{bmatrix}$. *Clearly, sum of entries of each column of the inverse of A need not be 1.*

(c) Consider the matrix

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Then $A^{-1} = A$ *and hence* $tr(A^{-1}) = 0$.

(17) Let A be a 3×3 matrix with real entries which commutes with all 3×3 matrices with real entries. What is the maximum number of distinct roots that the characteristic polynomial of A can have?

Ans. *Let* $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ *be a matrix which commutes with all* 3×3 *matrices with real entries. In particular,*

$$A \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} A \Rightarrow \begin{bmatrix} a_{13} & a_{12} & a_{11} \\ a_{23} & a_{22} & a_{21} \\ a_{33} & a_{32} & a_{31} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix} \Rightarrow A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{bmatrix}$$

Also A commutes with $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *which gives*

$$A \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} A \Rightarrow \begin{bmatrix} a_{12} & a_{11} & 0 \\ a_{22} & a_{21} & 0 \\ a_{12} & a_{13} & 0 \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{21} \\ a_{11} & a_{12} & a_{13} \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{11} & 0 \\ 0 & 0 & a_{11} \end{bmatrix}$$

Hence, A must be a scalar matrix and hence it has only one eigenvalue.

(18) Let $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$ be a real matrix with eigenvalues 1, 0, and 3. If the eigenvectors corresponding to 1 and 0 are $(1, 1, 1)^T$ and $(1, -1, 0)^T$ respectively, then the value of $3f$ is equal to ....

Ans. *Since 0 is an eigenvalue of A with eigenvector* $(1, -1, 0)^t$,

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow a = b = d, \ c = e$$

And as $(1, 1, 1)^t$ *is an eigenvector of A, corresponding to the eigenvalue* 1,

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \Rightarrow 2a + c = 2c + f = 1 \Rightarrow 2a + 3c + f = 2$$

Since $tr(A) = 4, a + d + f = 4 \Rightarrow 2a + f = 4$. *Substituting this in* $2a + 3c + f = 2$, *we get* $c = \frac{-2}{3}$ *and hence* $f = \frac{7}{3}$. *Therefore* $3f = 7$.

(19) The matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$

(a) is an elementary matrix
(b) can be written as a product of elementary matrices
(c) does NOT have linearly independent eigenvectors
(d) is a nilpotent matrix.

Ans. ***Option b***
A is not an elementary matrix as an elementary matrix is obtained by a single row(column) transformation on the identity matrix. The given matrix is invertible and hence can be written as a product of elementary matrices. A is an upper triangular matrix hence the eigenvalues are its diagonal entries. Since A has three distinct eigenvalues it has three linearly independent eigenvectors. A nilpotent matrix is never invertible as zero is an eigenvalue.

(20) Let A, B be complex $n \times n$ matrices. Which of the following are true?

(a) If A, B and $A + B$ are invertible, then $A^{-1} + B^{-1}$ is invertible.
(b) If A, B and $A + B$ are invertible, then $A^{-1} - B^{-1}$ is invertible.
(c) If AB is nilpotent, then BA is nilpotent.
(d) Characteristic polynomials of AB and BA are equal if A is invertible.

Ans. ***Options a, c, and d***

(a) Since $(A + B)$ is invertible, we have

$$(A+B)(A+B)^{-1} = I \Rightarrow A\left(A^{-1}+B^{-1}\right)B(A+B)^{-1} = I$$
$$\Rightarrow A^{-1}+B^{-1} \; is \; invertible$$

(b) If $A = B = I_n$, then A, B, and $A + B$ are invertible, but $A^{-1} - B^{-1}$ is not invertible.
(c) AB is nilpotent $\Rightarrow (AB)^k = 0$ for some positive integer k. That is,

$$(AB)(AB)\ldots(AB)(k \; times) = 0$$

Multiplying by B from left and A from right, we get

$$B(AB)(AB)\ldots(AB)A = (BA)^{k+1} = 0$$

Therefore BA is nilpotent.
(d) The characteristic equation of AB is given by $det\,(\lambda I - AB) = 0$. Since A is invertible,

$$\begin{aligned} det\,(\lambda I - AB) &= det\left(\lambda AA^{-1} - ABAA^{-1}\right) \\ &= det\left[A\,(\lambda I - BA)\,A^{-1}\right] \\ &= det(A) det\,(\lambda I - BA)\, det\left(A^{-1}\right) \\ &= det\,(\lambda I - BA) \end{aligned}$$

Therefore characteristic polynomials of AB and BA are equal if A is invertible.

(21) Let a, b, c, d be distinct non-zero real numbers with $a + b = c + d$. Then an eigenvalue of the matrix $\begin{bmatrix} a & b & 1 \\ c & d & 1 \\ 1 & -1 & 0 \end{bmatrix}$ is

(a) $a + c$ (b) $a + b$ (c) $a - b$ (d) $b - d$.

Ans. ***Option b***

The characteristic polynomial of the above matrix is

$$\begin{aligned}
\lambda^3 - (a + d)\lambda^2 + (ad - bc)\lambda = 0 \Rightarrow \lambda^2 - (a + d)\lambda + (ad - bc) = 0 \\
\Rightarrow \lambda = \frac{(a + d) \pm \sqrt{(a + d)^2 - 4(ad - bc)}}{2} \\
\Rightarrow \lambda = \frac{(a + d) \pm \sqrt{(a - d)^2 - 4bc}}{2} \\
\Rightarrow \lambda = \frac{(a + d) \pm \sqrt{(c - b)^2 - 4bc}}{2} \text{ since } a - d = c - b \\
\Rightarrow \lambda = \frac{(a + d) \pm \sqrt{(b + c)^2}}{2} \\
\Rightarrow \lambda = \frac{a + b + c + d}{2} = a + b \text{ or } \lambda = \frac{a + d - c - b}{2}
\end{aligned}$$

(22) Let A and B be two $n \times n$ matrices and $C = \begin{bmatrix} A & B \\ B & A \end{bmatrix}$. Which of the following statements are true?

(a) If λ is an eigenvalue of $A + B$, then λ is an eigenvalue of C.
(b) If λ is an eigenvalue of $A - B$, then λ is an eigenvalue of C.
(c) If λ is an eigenvalue of A or B, then λ is an eigenvalue of C.
(d) All the eigenvalues of C are real.

Ans. ***Options a and b***

(a) Let λ be an eigenvalue of $A + B$. Then there exists $v \neq 0 \in \mathbb{R}^n$ such that $(A + B)v = \lambda v$. Now, consider $V = \begin{bmatrix} v \\ v \end{bmatrix} \in \mathbb{R}^{2n}$. Observe that

$$CV = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} v \\ v \end{bmatrix} = \begin{bmatrix} (A + B)v \\ (A + B)v \end{bmatrix} = \begin{bmatrix} \lambda v \\ \lambda v \end{bmatrix} = \lambda \begin{bmatrix} v \\ v \end{bmatrix} = \lambda V$$

That is, λ is an eigenvalue of C.

(b) Let λ be an eigenvalue of $A - B$. Then there exists $v \neq 0 \in \mathbb{R}^n$ such that $(A - B)v = \lambda v$. Now, consider $V = \begin{bmatrix} v \\ -v \end{bmatrix} \in \mathbb{R}^{2n}$. Observe that

$$CV = \begin{bmatrix} A & B \\ B & A \end{bmatrix} \begin{bmatrix} v \\ -v \end{bmatrix} = \begin{bmatrix} (A-B)v \\ -(A-B)v \end{bmatrix} = \begin{bmatrix} \lambda v \\ -\lambda v \end{bmatrix} = \lambda \begin{bmatrix} v \\ -v \end{bmatrix} = \lambda V$$

That is, λ is an eigenvalue of C.

(c) *Consider $n = 2$. Take $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Then the eigenvalues of A and B are 0 and 2. Observe that 2 is not an eigenvalue of C (eigenvalues of C are 0 and 4). Therefore, every eigenvalue of A or B need not be an eigenvalue of C.*

(d) *Consider $n = 2$. Take $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Then the eigenvalues of C are i and $-i$, each repeated twice. Therefore, all the eigenvalues of C need not be real.*

(23) Write down a necessary and sufficient condition, in terms of a, b, c, and d (which are assumed to be real numbers), for the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ not to have a real eigenvalue.

Ans. *The characteristic equation of a 2×2 matrix is given by $\lambda^2 - (tr(A))\lambda + det(A) = 0$. Therefore the characteristic equation of the given matrix is $\lambda^2 - (a+d)\lambda + (ad - bc) = 0$. Then the equation does not have a real eigenvalue when $(a+d)^2 - 4(ad - bc) < 0$.*

(24) Let A be a 3×3 matrix such that $A \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 0 \end{pmatrix}$ and suppose that $A^3 \begin{pmatrix} 1 \\ \frac{-1}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ for some $\alpha, \beta, \gamma \in \mathbb{R}$. Then $|\alpha|$ is equal to ....

Ans. *If λ is an eigenvalue of A with eigenvector $v \neq 0$, then λ^n is an eigenvalue of A^n with eigenvector v (Theorem 4.5). In this case, as*

$$A \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ -3 \\ 0 \end{pmatrix} = (-3) \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}$$

-3 is an eigenvalue of A with eigenspace $E = span\{(-2, 1, 0)^t\}$. Clearly $\left(1, \frac{-1}{2}, 0\right)^t \in E$. Therefore

$$A^3 \begin{pmatrix} 1 \\ \frac{-1}{2} \\ 0 \end{pmatrix} = 27 \begin{pmatrix} 1 \\ \frac{-1}{2} \\ 0 \end{pmatrix}$$

and hence $|\alpha| = 27$.

(25) If a 3×3 real skew-symmetric matrix has an eigenvalue $2i$, then one of the remaining eigenvalues is
(a) $\frac{1}{2i}$ (b) $\frac{-1}{2i}$ (c) 0 (d) 1.

Ans. ***Option c***
A is skew-symmetric implies $A^T = -A$

$$\begin{aligned} A^T = -A &\Rightarrow det(A^T) = (-1)^3 det(A) \\ &\Rightarrow det(A) = (-1) det(A) \;\; since\; det(A^T) = det(A) \\ &\Rightarrow det(A) = 0 \end{aligned}$$

Therefore 0 is one of the remaining eigenvalues. Since the complex roots of a polynomial with real coefficients occur in conjugate pairs, the third eigenvalue is $-2i$.

(26) Let A be an $n \times n$ non-null skew-symmetric matrix, where n is even. Which of the following statements is (are) always true?

(a) $Ax = 0$ has infinitely many solutions, where $0 \in \mathbb{R}^n$.
(b) $Ax = \lambda x$ has a unique solution for every non-zero $\lambda \in \mathbb{R}$.
(c) If $B = (I_n + A)(I_n - A)^{-1}$, then $B^T A = I_n$.
(d) The sum of all eigenvalues of A is zero.

Ans. ***Options c and d***
Consider the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. *Then A is skew-symmetric. Since A is invertible the system* $Ax = 0$ *has a unique solution. The characteristic equation of A is* $\lambda^2 + 1 = 0$. *As A has no real eigenvalues,* $Ax = \lambda x$ *has no solutions for every non-zero* $\lambda \in \mathbb{R}$.

$$\begin{aligned} B^T &= \left[(I_n + A)(I_n - A)^{-1}\right]^T \\ &= \left[(I_n - A)^{-1}\right]^T (I_n + A)^T \\ &= (I_n - A^T)^{-1}(I_n + A^T) \\ &= (I_n + A)^{-1}(I_n - A) \end{aligned}$$

Since $(I_n + A)$ *and* $(I_n - A)$ *commutes,* $B^T B = I_n$. *Also since the diagonal entries of a skew-symmetric matrix is zero,* $tr(A) =$ *the sum of all eigenvalues of* $A = 0$.

(27) let $A = \begin{bmatrix} 0 & 1-i \\ -1-i & i \end{bmatrix}$ and $B = A^T \bar{A}$. Then
(a) an eigenvalue of B is purely imaginary (b) an eigenvalue of A is zero
(c) all eigenvalues of B are real (d) A has a non-zero real eigenvalue.

Ans. ***Option c***

We have $B = A^T \bar{A} = \begin{bmatrix} 2 & i-1 \\ -i-1 & 3 \end{bmatrix}$. *The characteristic polynomial of B is* $\lambda^2 - 5\lambda + 2 = 0$ *and hence the eigenvalues are* $\lambda = \dfrac{5 \pm \sqrt{17}}{2}$. *The characteristic polynomial of A is* $\lambda^2 - i\lambda + 2 = 0$ *and hence the eigenvalues of A are* $-i$ *and* $2i$.

(28) Let $v \in \mathbb{R}^n$ with $v^T v \neq 0$. Let $A = I - 2\dfrac{vv^T}{v^T v}$. Then which of the following statements is(are) true?
(a) $A^{-1} = I - A$. (b) -1 and 1 are eigenvalues of A.
(c) $A^{-1} = A$. (d) $(I + A)v = v$.

Ans. ***Options b and c***

We have

$$A^2 = \left(I - 2\frac{vv^T}{v^T v}\right)\left(I - 2\frac{vv^T}{v^T v}\right) = I - 4\frac{vv^T}{v^T v} + 4\left(\frac{vv^T}{v^T v}\right)^2 = I - 4\frac{vv^T}{v^T v} + 4\frac{vv^T}{v^T v} = I$$

Thus $A^{-1} = A$. *We know that* vv^T *is a matrix of rank 1 and the eigenvalues of* vv^T *are* $v^T v$ *and 0. For,*

$$(vv^T)v = v(v^T v) = (v^T v)v$$

and if we take $w \in \mathbb{R}^n$ *with* $v^T w = 0$, *we have*

$$(vv^T)w = v(v^T w) = 0$$

As we can choose $n - 1$ *such as* $w's$ *from* $\mathbb{R}^n$, *the only eigenvalues of* vv^T *are* $v^T v$ *and 0. Therefore the eigenvalues of* $2\dfrac{vv^T}{v^T v}$ *are 2 and 0. This implies that the eigenvalues of A are* -1 *and 1. Also*

$$(I + A)v = \left(I + I - 2\frac{vv^T}{v^T v}\right)v = 2v - 2\frac{v(v^T v)}{v^T v} = 2v - 2v = 0$$

(29) Let A be an invertible 4×4 real matrix. Which of the following are not true?
(a) $Rank(A) = 4$. (b) $Ax = b$ has exactly one solution $\forall\, b \in \mathbb{R}^4$.
(c) $dim\,[\mathcal{N}(A)] \geq 1$. (d) 0 is an eigenvalue of A.

Ans. ***Options c and d***

Since A is invertible $Rank(A) = 4$ *and 0 is not an eigenvalue of A. For every vector* $B \in \mathbb{R}^4$, $Ax = b$ *has exactly one solution* $x = A^{-1}b$. *By Rank–Nullity theorem,* $dim\,[\mathcal{N}(A)] = 0$.

(30) Which of the following statements is true?

(a) Any matrix $A \in \mathbb{M}_{4\times 4}(\mathbb{R})$ has a real eigenvalue.
(b) Any matrix $A \in \mathbb{M}_{5\times 5}(\mathbb{R})$ has a real eigenvalue.
(c) Any matrix $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ has a real eigenvalue.

Ans. ***Option b***
Since any odd degree polynomial with real coefficients has at least one real root, any matrix $A \in \mathbb{M}_{5\times 5}(\mathbb{R})$ *has a real eigenvalue. Consider the matrices* $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{M}_{2\times 2}(\mathbb{R})$ *and* $\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \in \mathbb{M}_{4\times 4}(\mathbb{R})$. *Both do not have real eigenvalues as their characteristic polynomials are respectively* $\lambda^2 + 1$ *and* $(\lambda^2 + 1)^2$.

(31) Let A be a 5×5 matrix with real entries, then A has

(a) an eigenvalue which is purely imaginary.
(b) at least one real eigenvalue.
(c) at least two eigenvalues which are not real.
(d) at least 2 distinct real eigenvalues.

Ans. ***Option b***
Since any odd degree polynomial with real coefficients has at least one real root, any 5×5 *matrix with real entries has a real eigenvalue. Consider* I_5. *It has only one eigenvalue which has algebraic multiplicity 5 and is also real.*

(32) Let $A(\theta) = \begin{bmatrix} cos\ \theta & sin\ \theta \\ -sin\ \theta & cos\ \theta \end{bmatrix}$, where $\theta \in [0, 2\pi]$. Mark the correct statement below.

(a) $A(\theta)$ has eigenvectors in $\mathbb{R}^2$ for all $\theta \in (0, 2\pi)$.
(b) $A(\theta)$ does not have an eigenvector in $\mathbb{R}^2$ for any $\theta \in (0, 2\pi)$.
(c) $A(\theta)$ has eigenvectors in $\mathbb{R}^2$ for exactly one value of $\theta \in (0, 2\pi)$.
(d) $A(\theta)$ has eigenvectors in $\mathbb{R}^2$ for exactly 2 values of $\theta \in (0, 2\pi)$.

Ans. ***Option c***
The characteristic equation of $A(\theta)$ *is*

$$\lambda^2 - 2(cos\ \theta)\lambda + 1 = 0$$

Therefore, the eigenvalues of $A(\theta)$ *are*

$$\lambda_1 = \frac{2cos\ \theta + \sqrt{4cos\ ^2\theta - 4}}{2} = cos\ \theta + isin\ \theta$$

and

$$\lambda_2 = \frac{2cos\,\theta - \sqrt{4cos\,^2\theta - 4}}{2} = cos\,\theta - i sin\,\theta$$

λ_1 and λ_2 are real when $sin\,\theta = 0$. In the interval $(0, 2\pi)$, $sin\,\theta = 0$ only for $\theta = \pi$.

(33) Let $A = \begin{bmatrix} a & 2f & 0 \\ 2f & b & 3f \\ 0 & 3f & c \end{bmatrix}$ where a, b, c, f are real numbers and $f \neq 0$. The geometric multiplicity of the largest eigenvalue of A equals ....

Ans. *Take $a = b = c = 0$ and $f = 1$. Then the matrix is given by $A = \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 0 \end{bmatrix}$.*
Then the characteristic equation of the matrix is $\lambda^3 - 13\lambda = 0$. Since this matrix has 3 distinct eigenvalues, the geometric multiplicity of the largest eigenvalue of A is 1.

(34) Consider the matrix $A = I_9 - 2v^T v$ with $v = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$. If λ and μ are two distinct eigenvalues of A, then $|\lambda - \mu| = \cdots\cdots$.

Ans. *The eigenvalues of $v^T v$ where v is a row vector are vv^T and 0 with algebraic multiplicity 1 and $n - 1$ respectively (see Question 31). Therefore the eigenvalues of A are $\lambda = 1$ and $\mu = -1$ and hence $|\lambda - \mu| = 2$.*

(35) Let A be the matrix $\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}$. Which one of the following matrix equations does A satisfy?
(a) $A^2 + 3A + 5I = 0$
(b) $A^3 - A^2 - 5A = 0$
(c) $A^2 - 3A + I = 0$
(d) $A^2 - A + 5I = 0$.

Ans. ***Option b***
The given matrix A has trace 1 and determinant -5. Therefore the characteristic polynomial of A is $\lambda^2 - \lambda - 5 = 0$ and hence by Cayley–Hamilton Theorem, $A^2 - A - 5I = 0$.

(36) Let $A = \begin{bmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{3}{5} & \frac{2}{5} \end{bmatrix}$. Then
(a) $20A^2 - 13A + 7I = 0$ (b) $20A^2 - 13A - 7I = 0$
(c) $20A^2 + 13A + 7I = 0$ (d) $20A^2 + 13A - 7I = 0$.

Ans. ***Option b***
Since $tr(A) = \frac{13}{20}$ and $det(A) = -\frac{7}{20}$ the characteristic equation of A is

$$\lambda^2 - \frac{13}{20}\lambda - \frac{7}{20} = 0$$

By Cayley–Hamilton Theorem, we have $A^2 - \frac{13}{20}A - \frac{7}{20} = 0$. *Therefore,* $20A^2 - 13A - 7I = 0$.

(37) Given a $n \times n$ matrix A, define e^A by

$$e^A = \sum_{j=0}^{\infty} \frac{A^j}{j!}$$

Let p be the characteristic polynomial of A. Then the matrix $e^{p(A)}$ is
(a) I_n (b) 0_n (c) eI_n (d) πI_n.

Ans. ***Option a***
By Cayley–Hamilton theorem, every matrix satisfies its characteristic equation. Therefore $p(A) = 0_{n\times n}$ *and hence* $e^{p(A)} = I_{n\times n}$.

(38) Let A be a 4×4 matrix with entries from the set of rational numbers. If $\sqrt{2} + i$, with $i = \sqrt{-1}$, is a root of the characteristic polynomial of A and I is the 4×4 identity matrix, then
(a) $A^4 = 4A^2 + 9I$
(b) $A^4 = 4A^2 - 9I$
(c) $A^4 = 2A^2 - 9I$
(d) $A^4 = 2A^2 + 9I$.

Ans. ***Option c***
The characteristic polynomial of A is a polynomial with real coefficients. Since $\sqrt{2} + i$ *is a root, the other roots are* $\sqrt{2} - i$, $-\sqrt{2} + i$ *and* $-\sqrt{2} - i$. *Therefore the characteristic equation of A is* $\lambda^4 - 2\lambda^2 + 9 = 0$. *By Cayley–Hamilton theorem* $A^4 - 2A^2 + 9I = 0$.

(39) If the roots of the characteristic polynomial of a 4×4 matrix A are $\pm\sqrt{\dfrac{1 \pm \sqrt{5}}{2}}$, then $A^8 =$
(a) $I + A^2$ (b) $2I + A^2$ (c) $2I + 3A^2$ (d) $3I + 2A^2$.

Ans. ***Option c***

Since the roots of the characteristic polynomial of A are $\pm\sqrt{\dfrac{1 \pm \sqrt{5}}{2}}$, *the characteristic equation is*

$$\left(\lambda + \sqrt{\frac{1+\sqrt{5}}{2}}\right)\left(\lambda - \sqrt{\frac{1+\sqrt{5}}{2}}\right)\left(\lambda + \sqrt{\frac{1-\sqrt{5}}{2}}\right)\left(\lambda - \sqrt{\frac{1-\sqrt{5}}{2}}\right) = \lambda^4 - \lambda^2 - 1 = 0$$

By Cayley–Hamilton theorem, A must satisfy its characteristic equation. Then,

$$A^4 = A^2 + I \Rightarrow A^8 = (A^2 + I)(A^2 + I) = A^4 + 2A^2 + 1 = 3A^2 + 2I$$

(40) Let A be a 3×3 singular matrix such that $Av = v$ for a non-zero vector v and

$$A \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ 0 \\ \frac{-2}{5} \end{bmatrix}$$

Then
(a) $A^3 = \frac{1}{5}(7A^2 - 2A)$ (b) $A^3 = \frac{1}{4}(7A^2 - 2A)$
(c) $A^3 = \frac{1}{3}(7A^2 - 2A)$ (d) $A^3 = \frac{1}{2}(7A^2 - 2A)$.

Ans. ***Option a***

$Av = v$ for a non-zero vector v implies that 1 is an eigenvalue of A. $A \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{2}{5} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ implies that $\frac{2}{5}$ is an eigenvalue of A. Since A is singular, 0 is also an eigenvalue. Therefore characteristic polynomial of A is given by $\lambda^3 - 7/5\lambda^2 + 2/5\lambda$. Hence, by Cayley–Hamilton Theorem, $A^3 = \frac{1}{5}(7A^2 - 2A)$.

(41) Let $A \in \mathbb{M}_{10\times 10}(\mathbb{C})$. Let W_A be the subspace of $\mathbb{M}_{10\times 10}(\mathbb{C})$ spanned by $\{A^n \mid n \geq 0\}$. Choose the correct statements.
(a) For any A, $dim(W_A) \leq 10$. (b) For some A, $10 < dim(W_A) < 100$.
(c) For any A, $dim(W_A) < 10$. (d) For some A, $dim(W_A) = 100$.

Ans. ***Option a***

By Cayley–Hamilton Theorem, any matrix of order can be represented as a linear combination of $\{A^n \mid 0 \leq n \leq 9\}$, for any A, $dim(W_A) \leq 10$.

(42) Let A be an $n \times n$ real matrix. Let V be the vector space spanned by $\{I, A, A^2, \ldots, A^{2n}\}$. The dimension of the vector space V is
(a) $2n$ (b) at most n (c) n^2 (d) at most $2n$.

Ans. ***Options b and d***

By Cayley–Hamilton Theorem, every matrix of order n can be written as a linear combination of $\{I, A, A^2, \ldots, A^{n-1}\}$. The dimension of the vector space V is n.

(43) Let $V = \{p(A) : p$ is a polynomial with real coefficients$\}$, where $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. The dimension of the vector space V satisfies
(a) $4 \leq d \leq 6$ (b) $6 \leq d \leq 9$ (c) $3 \leq d \leq 8$ (d) $3 \leq d \leq 4$.

Ans. ***Options c and d***

By Cayley–Hamilton Theorem, every matrix of degree greater than 3 can be written as a linear combination of $\{I, A, A^2, A^3\}$. Here $A^3 = I$ and hence V can be spanned by $\{I, A, A^2\}$.

(44) Let A be a 3×3 upper triangular matrix whose diagonal entries are 1, 2, and -3. Express A^{-1} as a linear combination of I, A, and A^2.

Ans. *The characteristic equation of the given matrix is*

$$(\lambda - 1)(\lambda - 2)(\lambda + 3) = \lambda^3 - 7\lambda + 6 = 0$$

By Cayley–Hamilton Theorem,

$$A^3 - 7A + 6I = 0 \Rightarrow A^{-1} = \frac{7}{6}I - \frac{1}{6}A^2$$

(45) Let A be a 3×3 matrix and suppose that 1, 2, and 3 are eigenvalues of A. If

$$A^{-1} = \frac{1}{\alpha}A^2 - A + \frac{11}{\alpha}I$$

for some scalar $\alpha \neq 0$, then α is equal to ....

Ans. *Since 1, 2, and 3 are eigenvalues of A, the characteristic equation of A is*

$$(\lambda - 1)(\lambda - 2)(\lambda - 3) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$$

By Cayley–Hamilton Theorem, $A^{-1} = \frac{1}{6}A^2 - A + \frac{11}{6}I$. *Therefore* $\alpha = 6$.

(46) Let A be a 3×3 matrix with complex entries, whose eigenvalues are 1, i, and $-2i$. If $A^{-1} = aA^2 + bA + cI$, with $a, b, c \in \mathbb{C}$, what are the values of a, b, and c?

Ans. *Since* $1, i$ *and* $-2i$ *are the eigenvalues of A, the characteristic equation of A is*

$$(\lambda - 1)(\lambda - i)(\lambda + 2i) = \lambda^3 - (1 - i)\lambda^2 - (i - 2)\lambda - 2 = 0$$

By Cayley–Hamilton Theorem,

$$\lambda^3 - (1 - i)\lambda^2 - (i - 2)\lambda - 2 = 0 \Rightarrow A^3 + (i - 1)A^2 = (2 - i)A - 2I = 0$$
$$\Rightarrow A^{-1} = \frac{1}{2}A^2 + \frac{(i - 1)}{2}A + \frac{(2 - i)}{2}I$$

Therefore $a = \frac{1}{2}$, $b = \dfrac{(i - 1)}{2}$, *and* $c = \dfrac{(2 - i)}{2}$.

(47) Let A be a 3×3 upper triangular matrix with real entries. If $a_{11} = 1$, $a_{22} = 2$, and $a_{33} = 3$, determine α, β, and γ such that

$$A^{-1} = \alpha A^2 + \beta A + \gamma I$$

Ans. *The eigenvalues of an upper triangular matrices are its diagonal entries. Therefore its characteristic equation is*

$$(\lambda - 1)(\lambda - 2)(\lambda - 3) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = 0$$

By Cayley–Hamilton theorem,

$$A^3 - 6A^2 + 11A - 6I = 0 \Rightarrow A^{-1} = \frac{1}{6}A^2 - A + \frac{11}{6}I$$

Therefore $\alpha = \frac{1}{6}$, $\beta = -1$, *and* $\gamma = \frac{11}{6}$.

(48) Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$. If $6A^{-1} = aA^2 + bA + cI$ for $a, b, c \in \mathbb{R}$ then (a, b, c) equals
(a) $(1, 2, 1)$ (b) $(1, -1, 2)$ (c) $(4, 1, 1)$ (d) $(1, 4, 1)$.

Ans. ***Option d***
The characteristic equation of A is

$$\lambda^3 - (1 - 2 - 3)\lambda^2 + (6 - 3 - 2)\lambda - 6 = \lambda^3 + 4\lambda^2 + \lambda - 6 = 0$$

By Cayley–Hamilton Theorem, $6A^{-1} = A^2 + 4A + I$. *Therefore* $(a, b, c) = (1, 4, 1)$.

(49) If $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 2 & 0 & -1 \end{bmatrix}$ and $6A^{-1} = aI + bA - A^2$, then the ordered pair (a, b) is
(a) $(3, 2)$ (b) $(2, 3)$ (c) $(4, 5)$ (d) $(5, 4)$.

Ans. ***Option a***
The characteristic equation of A is

$$\lambda^3 - (1 + 2 - 1)\lambda^2 + (-2 - 3 + 2)\lambda - (-6) = \lambda^3 - 2\lambda^2 - 3\lambda + 6 = 0$$

By Cayley–Hamilton Theorem, $6A^{-1} = 3I + 2A - A^2$. *Therefore* $(a, b) = (3, 2)$.

(50) Let $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$. If $A^{-1} = \frac{5}{4}I + kA + \frac{1}{4}A^2$, find the value of k. Hence or otherwise, solve the system of equations $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$.

Ans. *The characteristic equation of A is*

$$\lambda^3 - (1+2+1)\lambda^2 + (2+1+2) - 4 = \lambda^3 - 4\lambda^2 + 5\lambda - 4 = 0$$

By Cayley–Hamilton Theorem, we have $A^3 - 4A^2 + 5A - 4I = 0$. *Now,*

$$A^3 - 4A^2 + 5A - 4I = 0 \Rightarrow A^{-1} = \frac{5}{4}I + (-1)A + \frac{1}{4}A^2 \Rightarrow k = -1$$

Also, we get $A^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{-1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{-1}{4} \\ \frac{-1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$. *Since A is invertible the system of equations is unique and the solution is* $A^{-1}\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\ \frac{1}{4}\\ -\frac{1}{2}\end{bmatrix}$.

(51) Let $A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$. Then the smallest positive integer n such that $A^n = I$ is
(a) 1 (b) 2 (c) 4 (d) 6.

Ans. ***Option d***
The characteristic equation of the given matrix is

$$\lambda^2 - tr(A) + det(A) = \lambda^2 - \lambda + 1 = 0$$

By Cayley–Hamilton Theorem, $A^2 = A - I$. *Therefore,*

$$A^4 = A^2 - 2A + I = -A \text{ and } A^6 = A^4A^2 = -A(A-I) = I$$

(52) Let A be the matrix $A = \begin{bmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & -1 \end{bmatrix}$. Compute the matrix

$$B = 3A - 2A^2 - A^3 - 5A^4 + A^6$$

Ans. *The characteristic equation of A is*

$$\lambda^2 - tr(A) + det(A) = \lambda^2 + 1 = 0$$

By Cayley–Hamilton Theorem, we have $A^2 + I = 0$. *Thus,* $A^2 = -I$ *and hence* $A^3 = -A$, $A^4 = I$, $A^6 = -I$. *Therefore*

$$B = 3A - 2A^2 - A^3 - 5A^4 + A^6 = 3A + 2I + A - 5I - I = \begin{bmatrix} 0 & 4\sqrt{2} \\ -4\sqrt{2} & -8 \end{bmatrix}$$

(53) Denote by $\mathcal{U}$ the set of all $n \times n$ complex matrices $A(n \geq 2)$ having the property that 4 is the only eigenvalue of A. Consider the following four statements.
(i) $(A - 4I)^n = 0$, (ii) $A^n = 4^n I$,
(iii) $(A^2 - 5A + 4I)^n = 0$, (iv) $A^n = 4nI$.
How many of the above statements are true for all $A \in \mathcal{U}$?
(a) 0 (b) 1 (c) 2 (d) 3.

Ans. ***Option c***
Since A is of order n and 4 is the only eigenvalue of A, the characteristic polynomial of A is $(\lambda - 4)^n$*. Then by Cayley–Hamilton theorem* $(A - 4I)^n = 0$*. As*

$$A^2 - 5A + 4I = (A - 4I)(A - I)$$

we have

$$(A^2 - 5A + 4I)^n = (A - I)^n(A - 4I)^n = 0$$

Thus statements (i) *and* (iii) *are true for all* $A \in \mathcal{U}$*.*
Let $A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$*. Then* $A^2 = \begin{bmatrix} 16 & 8 \\ 0 & 16 \end{bmatrix}$*. In this case* $A^2 \neq 16I$ *and* $A^2 \neq 8I$*. Thus statements* (ii) *and* (iv) *need not be true for all* $A \in \mathcal{U}$*.*

(54) Let V be a vector space over $\mathbb{C}$ with dimension n. Let $T : V \to V$ be a linear transformation with only 1 as eigenvalue. Then which of the following must be true?
(a) $T - I = 0$ (b) $(T - I)^{n-1} = 0$ (c) $(T - I)^n = 0$ (d) $(T - I)^{2n} = 0$.

Ans. ***Options c and d***
Since 1 is the only eigenvalue of the matrix, the characteristic polynomial is $(\lambda - 1)^n$ *and hence by Cayley–Hamilton Theorem* $(T - I)^n = 0$ *and* $(T - I)^{2n} = 0$*.*
Now for $n = 3$ *consider* T *with the matrix* $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$*. Then both options a and b are false.*

(55) Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear transformation, $n \geq 2$. Suppose 1 is the only eigenvalue of T. Which of the following statements are true?
(a) $T^k \neq I$ for any $k \in \mathbb{N}$ (b) $(T - I)^{n-1} = 0$
(c) $(T - I)^n = 0$ (d) $(T - I)^{n+1} = 0$.

Ans. ***Options c and d***

(a) Consider the identity transformation. Then $T^k = I$ *for any* $k \in \mathbb{N}$*.*

(b) *Consider the linear transformation with matrix* $T = \begin{bmatrix} 1 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$. *Clearly 1 is the only eigenvalue of* T. *Then,* $(T - I)^{n-1} \neq 0$.
(c) *Since 1 is the only eigenvalue of* T*, the characteristic equation of* T *is* $(\lambda - 1)^n = 0$. *Then by Cayley–Hamilton Theorem* $(T - I)^n = 0$.
(d) *Since* $(T - I)^n = 0$, $(T - I)^{n+1} = 0$.

(56) Consider a real vector space V of dimension n and a non-zero linear transformation $T : V \to V$. If $dim\,[T(V)] < n$ and $T^2 = \lambda T$, for some $\lambda \in \mathbb{R} \setminus \{0\}$, then which of the following statements is true?

(a) $det(T) = |\lambda|^n$.
(b) There exists a non-trivial subspace W of V such that $T(v) = 0$ for all $v \in W$.
(c) T is invertible.
(d) λ is the only eigenvalue of T.

Ans. ***Option b***
Consider the transformation $T(x_1, x_2) = (x_1, 0)$. *Then Range*$(T) = x-$ *axis. Therefore* $dim\,[T(V)] < 2 = n$. *Also*

$$T^2(x_1, x_2) = T(x_1, 0) = (x_1, 0) = T(x_1, x_2)$$

Since T *is not onto, it is not invertible. Therefore* $det(T) = 0$ *and 0 is also an eigenvalue. Hence options* (a), (c), (d) *are false. Since* $dim\,[T(V)] < n$, *by Rank-Nullity Theorem, there exists a non-trivial subspace* W *of* V *such that* $T(v) = 0$ *for all* $v \in W$.

(57) Let $A = \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix}$ be a 3×3 matrix where a, b, c, d are integers. Then, we must have

(a) If $a \neq 0$, there is a polynomial $p \in \mathbb{Q}[\lambda]$ such that $p(A)$ is the inverse of A.
(b) For each polynomial $q \in \mathbb{Z}[\lambda]$, the matrix $q(A) = \begin{pmatrix} q(a) & q(b) & q(c) \\ 0 & q(a) & q(d) \\ 0 & 0 & q(a) \end{pmatrix}$.
(c) If $A^n = 0$ for some positive integer n, then $A^3 = 0$.
(d) A commutes with every matrix of the form $\begin{pmatrix} a' & 0 & c' \\ 0 & a' & 0 \\ 0 & 0 & a' \end{pmatrix}$.

Ans. ***Options a, c, and d***

(a) Since the characteristic polynomial of the given matrix is

$$(\lambda - a)^3 = \lambda^3 - 3a\lambda^2 + 3a^2\lambda - a^3$$

By Cayley–Hamilton Theorem, we get

$$A^{-1} = \frac{1}{a^3}\left(A^2 - 3aA + 3a^2 I\right)$$

If $a \neq 0$, there is a polynomial $p \in \mathbb{Q}[x]$ such that $p(A)$ is the inverse of A.

(b) Let $q(\lambda) = \lambda + 1$. Then,

$$q(A) = A + I = \begin{pmatrix} a+1 & b & c \\ 0 & a+1 & d \\ 0 & 0 & a+1 \end{pmatrix}$$

But

$$\begin{pmatrix} q(a) & q(b) & q(c) \\ 0 & q(a) & q(d) \\ 0 & 0 & q(a) \end{pmatrix} = \begin{pmatrix} a+1 & b+1 & c+1 \\ 0 & a+1 & d+1 \\ 0 & 0 & a+1 \end{pmatrix} \neq q(A)$$

Therefore option b is false.

(c) $A^n = 0 \Rightarrow A$ is a nilpotent matrix. Since the degree of nilpotency of a nilpotent matrix is always less than or equal to its order, $A^3 = 0$.

(d) We have

$$\begin{pmatrix} a' & 0 & c' \\ 0 & a' & 0 \\ 0 & 0 & a' \end{pmatrix}\begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix} = \begin{pmatrix} a'a & a'b & a'c + c'a \\ 0 & a'a & a'd \\ 0 & 0 & a'a \end{pmatrix} = \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix}\begin{pmatrix} a' & 0 & c' \\ 0 & a' & 0 \\ 0 & 0 & a' \end{pmatrix}$$

Clearly, A commutes with every matrix of the form $\begin{pmatrix} a' & 0 & c' \\ 0 & a' & 0 \\ 0 & 0 & a' \end{pmatrix}$.

(58) Let V be the vector space of all real polynomials of degree ≤ 10. Let $(Tp)(x) = p'(x)$ for $p \in V$ be a linear transformation from V to V. Consider the basis $\{1, x, x^2, \ldots x^{10}\}$ of V. Let A be the matrix of T with respect to this basis. Then

(a) $tr(A) = 1$ (b) there is no $m \in \mathbb{N}$ such that $A^m = 0$

(c) $det(A) = 0$ (d) A has a non-zero eigenvalue.

Ans. ***Option c***

The matrix of T is $A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 10 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$. *Then* $T^{11} = 0$, $tr(A) = 0$, $det(A) = 0$ *and 0 is the only eigenvalue.*

(59) Suppose that A is a 5×5 matrix with real entries and $p(\lambda) = det(\lambda I - A)$.

(a) $p(0) = det(A)$.
(b) every eigenvalue of A is real if $p(1) + p(2) = 0 = p(2) + p(3)$.
(c) A^{-1} is necessarily a polynomial in A of degree 4 if A is invertible.
(d) A is not invertible if $A^2 - 2A = 0$.

Ans. ***Option c***

(a) $p(0) = det(-A) = (-1)^5 det(A) = -det(A)$.
(b) Let $p(\lambda) = \lambda^5 + a_4\lambda^4 + a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$. *Then*

$$p(1) + p(2) = 0 \Rightarrow 33 + 17a_4 + 9a_3 + 5a_2 + 3a_1 + 2a_0 = 0$$

and

$$p(1) + p(3) = 0 \Rightarrow 275 + 97a_4 + 35a_3 + 13a_2 + 5a_1 + 2a_0 = 0$$

Let $a_2 = a_3 = a_4 = 0$. *Then solving for* a_0 *and* a_1, *we get* $a_0 = 165$ *and* $a_1 = -121$. *Then* $p(\lambda) = \lambda^5 - 121\lambda + 165$. *Then* $p(\lambda)$ *has 2 sign changes and one sign change for* $p(-\lambda)$. *Then by Descarte's Rule of Signs,* $p(\lambda)$ *has at most 3 real roots. Therefore p has complex roots even if* $p(1) + p(2) = 0 = p(2) + p(3)$.
(c) By Cayley–Hamilton Theorem, A^{-1} *is necessarily a polynomial in A of degree 4 if A is invertible.*
(d) Consider $A = 2I$. *Then* $A^2 - 2A = 0$. *But A is invertible.*

(60) Let $\alpha, \beta, \gamma, \delta$ be the eigenvalues of the matrix $A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$. Then $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = \cdots\cdots$.

Ans. *Let* $\alpha, \beta, \gamma, \delta$ *be the eigenvalues of the matrix A, then* $\alpha^2, \beta^2, \gamma^2, \delta^2$ *be the eigenvalues of the matrix* A^2. *Then* $\alpha^2 + \beta^2 + \gamma^2 + \delta^2$ *is the trace of* A^2. *Therefore* $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 6$.

(61) Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and let α_n and β_n denote the two eigenvalues of A^n such that $\alpha_n \geq \beta_n$. Then
(a) $\alpha_n \to \infty$ as $n \to \infty$ (b) as $n \to \infty$, $\beta_n \to 0$
(c) β_n is positive if n is even. (d) β_n is negative if n is odd.

Ans. ***Options a, b, c, and d***
Let α and β denote the two eigenvalues of A with $\alpha \geq \beta$. Then $\alpha_n = (\alpha)^n$ and $\beta_n = (\beta)^n$. The characteristic polynomial of A is $\lambda^2 - \lambda - 1$ and hence $\alpha_n = \left(\frac{1+\sqrt{5}}{2}\right)^n$ and $\beta_n = \left(\frac{1-\sqrt{5}}{2}\right)^n$.

(62) Let A be a 3×3 singular matrix and suppose that 2 and 3 are eigenvalues of A. Then the number of linearly independent eigenvectors of $A^3 + 2A + I$ is equal to ....

Ans. *Since A is singular 0 is also an eigenvalue of A. Therefore the eigenvalues of $A^3 + 2A + I$ are 1, 13, and 34. Since it has distinct eigenvalues, $A^3 + 2A + I$ has three linearly independent eigenvectors.*

(63) Consider the matrix $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 1 & 4 \end{bmatrix}$. Let P be a non-singular matrix such that $P^{-1}AP$ is a diagonal matrix. Then the trace of the matrix $P^{-1}A^3P$ equals ....

Ans. *Eigenvalues of A are 5, 2, and 1. Therefore eigenvalues of A^3 are 125, 8, and 1. Since $tr(AB) = tr(BA)$,*

$$tr(P^{-1}A^3P) = tr(A^3) = 134$$

(64) Let λ, μ be distinct eigenvalues of a 2×2 matrix A. Then which of the following statements must be true?

(a) A^2 has distinct eigenvalues.
(b) $A^3 = \dfrac{\lambda^3 - \mu^3}{\lambda - \mu} A - \lambda\mu(\lambda + \mu)I$.
(c) trace of A^n is $\lambda^n + \mu^n$ for every positive integer n.
(d) A^n is not a scalar multiple of identity for any positive integer n.

Ans. ***Options b and c***
As λ, μ are the eigenvalues of A, $tr(A) = \lambda + \mu$ and $det(A) = \lambda\mu$. Then, the characteristic equation of A is

$$x^2 - tr(A)x + det(A) = x^2 - (\lambda + \mu)x + \lambda\mu = 0$$

By Cayley–Hamilton theorem, we get

$$A^2 = (\lambda + \mu)A - \lambda\mu I$$

Multiplying by A on both sides and substituting for A^2, we have

$$A^3 = (\lambda^2 + \lambda\mu + \mu^2)A - \lambda\mu(\lambda + \mu)I = \frac{\lambda^3 - \mu^3}{\lambda - \mu}A - \lambda\mu(\lambda + \mu)I$$

Since the eigenvalues of A are λ and μ, the eigenvalues of A^n are λ^n and μ^n. Hence trace of A^n is $\lambda^n + \mu^n$ for every positive integer n (Theorem 4.5 and Corollary 4.2). Consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then A has distinct eigenvalues, but $A^2 = I$ does not have distinct eigenvalues.

(65) Let A be a 3×3 matrix. Suppose that the eigenvalues of A are $-1, 0, 1$ with respective eigenvectors $(1, -1, 0)^T$, $(1, 1, -2)^T$, and $(1, 1, 1)^T$. Then $6A$ equals
(a) $\begin{bmatrix} -1 & 5 & 2 \\ 5 & -1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 5 & 3 \\ 5 & 1 & 3 \\ 3 & 3 & 3 \end{bmatrix}$ (d) $\begin{bmatrix} -3 & 9 & 0 \\ 9 & -3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$.

Ans. ***Option a***
Since the eigenvalues of A are $-1, 0, 1$ with respective eigenvectors $(1, -1, 0)^T$, $(1, 1, -2)^T$, and $(1, 1, 1)^T$, the eigenvalues of $6A$ are $-6, 0, 6$ with respective eigenvectors $(1, -1, 0)^T$, $(1, 1, -2)^T$, and $(1, 1, 1)^T$ (Theorem 4.6). Therefore $6A$ has trace 0 and determinant 0. Thus options (c) and (d) are false. Since 6 is an eigenvalue with eigenvector $(1, 1, 1)^T$, the row sum of $6A$ must be the same. Hence option (b) is false.

(66) Let $\alpha = e^{\frac{2\pi i}{5}}$. Consider the matrix

$$A = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ 0 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ 0 & 0 & \alpha^2 & \alpha^3 & \alpha^4 \\ 0 & 0 & 0 & \alpha^3 & \alpha^4 \\ 0 & 0 & 0 & 0 & \alpha^4 \end{bmatrix}$$

Then $tr\left(I + A + A^2\right)$ is ......
(a) -5 (b) 0 (c) 3 (d) 5.

Ans. ***Option d***
Since $\alpha = e^{\frac{2\pi i}{5}}$, the set $\{1, \alpha, \alpha^2, \alpha^3, \alpha^4\}$ is fifth roots of unity. Hence $tr(A) = 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4 = 0$. Also the diagonal entries of A^2 are $\{1, \alpha, \alpha^2, \alpha^3, \alpha^4\}$. Therefore $tr(A^2) = 0$ and $tr(I + A + A^2) = tr(I) = 5$.

(67) If

$$A = \begin{bmatrix} 1 & 0 & 0 \\ i & \dfrac{-1+i\sqrt{3}}{2} & 0 \\ 0 & 1+2i & \dfrac{-1-i\sqrt{3}}{2} \end{bmatrix}$$

then $tr\left(A^{102}\right)$ is ……
(a) 0 (b) 1 (c) 2 (d) 3.

Ans. ***Option d***
Since the given matrix is lower triangular, the eigenvalues of A are precisely the diagonal elements of A and the diagonal elements of A are the cube roots of unity. Therefore the only possible eigenvalue of A^{102} *is 1. Therefore* $tr\left(A^{102}\right) = 3$.

(68) Consider the matrix $A(\lambda) = \begin{pmatrix} 1+\lambda^2 & 7 & 11 \\ 3\lambda & 2\lambda & 4 \\ 8\lambda & 17 & 13 \end{pmatrix}$, $\lambda \in \mathbb{R}$. Then

(a) $A(\lambda)$ has eigenvalue 0 for some $\lambda \in \mathbb{R}$
(b) 0 is not an eigenvalue of $A(\lambda)$ for any $\lambda \in \mathbb{R}$
(c) $A(\lambda)$ has eigenvalue 0 for all $\lambda \in \mathbb{R}$
(d) $A(\lambda)$ is invertible for all $\lambda \in \mathbb{R}$.

Ans. ***Option a***
$det\,[A(\lambda)] = 26\lambda^3 + 108\lambda^2 + 538\lambda - 68$ *is a polynomial of degree 3 and hence must have a real root. Therefore* $A(\lambda)$ *has eigenvalue 0 for some* $\lambda \in \mathbb{R}$.

(69) Let A be a 2×2 complex matrix such that $tr(A) = 1$ and $det(A) = -6$. Then $tr(A^4 - A^3)$ is …….

Ans. *Let* λ_1 *and* λ_2 *be the eigenvalues of A. Then* $\lambda_1 + \lambda_2 = 1$ *and* $\lambda_1\lambda_2 = -6$. *Therefore* $\lambda_1 = 3$ *and* $\lambda_2 = -2$. *Hence, the eigenvalues of* A^4 *are 81, 16 and that of* A^3 *are 27,* -8. *This gives* $tr(A^4) = 97$ *and* $tr(A^3) = 19$. *Therefore* $tr(A^4 - A^3) = 97 - 19 = 78$.

(70) Let $A \in \mathbb{M}_{3\times 3}(\mathbb{R})$ be such that $det(A - I) = 0$. If the $tr(A) = 13$ and $det(A) = 32$, then the sum of squares of the eigenvalues of A is …….

Ans. *Since* $det(A - I) = 0$, *0 is an eigenvalue of* $(A - I)$ *and hence 1 is an eigenvalue of A. Let* λ_1 *and* λ_2 *be the other two eigenvalues of A. Then* $\lambda_1 + \lambda_2 = 12$ *and* $\lambda_1\lambda_2 = 32$. *Therefore* $\lambda_1 = 8$ *and* $\lambda_2 = 4$. *Then the sum of squares of the eigenvalues of A is 81.*

(71) The trace of the matrix $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}^{20}$ is
(a) 7^{20} (b) $2^{20} + 3^{20}$ (c) $2.2^{20} + 3^{20}$ (d) $2^{20} + 3^{20} + 1$.

Ans. ***Option c***

The matrix $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ *is upper triangular. The eigenvalues of the matrix are 2, 2, and 3. The eigenvalues of* A^{20} *are* $2^{20}, 2^{20}$*, and* 3^{20}*. Therefore the trace of the given matrix is* $2.2^{20} + 3^{20}$.

(72) The set of eigenvalues of which one of the following matrices is not equal to the set of eigenvalues of $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$?
(a) $\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$.

Ans. ***Option d***

As trace is the sum of eigenvalues and determinant is the product of eigenvalues, it is enough to check which of the given matrices do not have the same trace and determinant as that of $\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$. *The given matrix has trace 4 and determinant* -5. *All matrices except option* (d) *have trace 4 and determinant* -5.

(73) Let $A = \begin{bmatrix} a & -1 & 4 \\ 0 & b & 7 \\ 0 & 0 & 3 \end{bmatrix}$ be a matrix with real entries. If the sum and the product of all the eigenvalues of A are 10 and 30 respectively, then $a^2 + b^2$ equals
(a) 29 (b) 40 (c) 58 (d) 65.

Ans. ***Option a***

As A is an upper triangular matrix, the eigenvalues of A are a, b*, and* 3. *Then*

$$tr(A) = a + b + 3 = 10 \Rightarrow a + b = 7 \text{ and } det(A) = 3ab = 30 \Rightarrow ab = 10$$

Solving, we get $a = 5, b = 2$. *Therefore* $a^2 + b^2 = 29$.

(74) Let A be a 3×3 matrix with $tr(A) = 3$ and $det(A) = 2$. If 1 is an eigenvalue of A, then the eigenvalues of the matrix $A^2 - 2I$ are
(a) $1,\ 2(i-1),\ -2(i+1)$ (b) $-1,\ 2(i-1),\ 2(i+1)$
(c) $1,\ 2(i+1),\ -2(i+1)$ (d) $-1,\ 2(i-1),\ -2(i+1)$.

Ans. ***Option d***

Let λ_1 *and* λ_2 *be the eigenvalues of A, Then*

$$tr(A) = \lambda_1 + \lambda_2 + 1 = 3 \Rightarrow \lambda_1 + \lambda_2 = 2 \text{ and } det(A) = 2 \Rightarrow \lambda_1\lambda_2 = 2$$

Then the eigenvalues of A are $1, 1+i, 1-i$ *and hence the eigenvalues of* A^2 *are* $1, 2i, -2i$. *Therefore the eigenvalues of* $A^2 - 2I$ *are* $-1,\ 2(i-1),\ -2(i+1)$.

(75) Let A be a 3×3 matrix having characteristic roots $\lambda_1 = -\frac{2}{3}$, $\lambda_2 = 0$, and $\lambda_3 = 1$. Define $B = 3A^3 - A^2 - A + I_3$ and $C = 3A^3 - 2A$. If $a = det(B)$ and $b = tr(C)$, then $a + b$ equals ....

Ans. *Since $\lambda_1 = -\frac{2}{3}$, $\lambda_2 = 0$, and $\lambda_3 = 1$ are the eigenvalues of A, the eigenvalues of B are $\frac{1}{3}, 1, 2$ and the eigenvalues of C are $\frac{4}{9}, 1, 0$. Therefore $a = det(B) = \frac{2}{3}$ and $b = tr(C) = \frac{13}{9}$ and $a + b = \dfrac{19}{9}$.*

(76) Given that the matrix $A = \begin{pmatrix} a & 1 \\ 2 & 3 \end{pmatrix}$ has 1 as an eigenvalue, compute $tr(A)$ and its $det(A)$.

Ans. *Let λ be the second eigenvalue. We have*

$$tr(A) = a + 3 = \lambda + 1 \Rightarrow \lambda = a + 2 \text{ and } det(A) = 3a - 2 = \lambda$$

Now

$$3a - 2 = a + 2 \Rightarrow 2a = 4 \Rightarrow a = 2$$

Therefore $tr(A) = 5$ and $det(A) = 4$.

(77) Find the values of $a \in \mathbb{R}$ such that the matrix $A = \begin{bmatrix} 3 & a \\ a & 5 \end{bmatrix}$ has 2 as an eigenvalue.

Ans. *Let λ be the second eigenvalue of the given matrix. Then*

$$tr(A) = \lambda + 2 = 8 \Rightarrow \lambda = 6$$

Therefore

$$det(A) = 15 - a^2 = 2 \times 6 = 12 \Rightarrow a = \pm\sqrt{3}$$

(78) The largest eigenvalue of the matrix $A = \begin{bmatrix} 1 & 4 & 16 \\ 4 & 16 & 21 \\ 16 & 1 & 4 \end{bmatrix}$ is

(a) 16 (b) 21 (c) 48 (d) 64.

Ans. ***Option b***
Since the sum of entries in each row are 21, 21 is an eigenvalue of the given matrix. Now $tr(A) = 21$ implies that one of the eigenvalues is negative of the other. The determinant of A is $3969 < 21^3$, which implies that the largest eigenvalue is 21.

(79) Let $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -2 & 5 \\ 2 & 5 & -3 \end{bmatrix}$. Then the eigenvalues of A are

(a) $-4, 3, -3$ (b) $4, 3, 1$ (c) $4, -4 \pm \sqrt{13}$ (d) $4, -2 \pm 2\sqrt{7}$.

Ans. ***Option c***
Since $tr(A) = -4$ *and* $det(A) = 12$*, the eigenvalues of A are* $4, -4 \pm \sqrt{13}$.

(80) Let A be a 3×3 real matrix with eigenvalues 1, 2, 3 and let $B = A^{-1} + A^2$. Then the trace of the matrix B is equal to
(a) $\frac{91}{6}$ (b) $\frac{95}{6}$ (c) $\frac{97}{6}$ (d) $\frac{101}{6}$.

Ans. ***Option b***
The eigenvalues of A^{-1} *are* $1, \frac{1}{2}, \frac{1}{3}$ *and the eigenvalues of* A^2 *are* $1, 4, 9$. *Therefore* $tr(B) = tr(A^{-1}) + tr(A^2) = 1 + \frac{1}{2} + \frac{1}{3} + 1 + 4 + 9 = \frac{95}{6}$.

(81) Suppose that $A = \begin{pmatrix} 40 & -29 & -11 \\ -18 & 30 & -12 \\ 26 & 24 & -50 \end{pmatrix}$ has a certain complex number $\lambda \neq 0$. Which of the following numbers must also be an eigenvalue of A?
(a) $\lambda + 20$ (b) $\lambda - 20$ (c) $20 - \lambda$ (d) $-20 - \lambda$.

Ans. ***Option c***
Since the third column is the negative of the sum of first and second columns, determinant of the given matrix is zero and hence 0 is an eigenvalue of A. Now trace of the matrix is 20 gives that the third eigenvalue is $20 - \lambda$.

(82) If A is a 5×5 real matrix with $tr(A) = 15$ and if 2 and 3 are eigenvalues of A, each with algebraic multiplicity 2, then $det(A)$ is equal to ......
(a) 0 (b) 24 (c) 120 (d) 180.

Ans. ***Option d***
Let λ *be the fifth eigenvalue. Since* $tr(A) =$*sum of eigenvalues* $= 2 + 2 + 3 + 3 + \lambda = 15$*, we get* $\lambda = 5$. *Then* $det(A) =$ *product of eigenvalues* $= 180$.

(83) Let A be a 4×4 matrix with real entries such that $-1, 1, 2, -2$ are its eigenvalues. If $B = A^4 - 5A^2 + 5I$, then which of the following statements are correct?
(a) $det(A + B) = 0$ (b) $det(B) = 1$
(c) $tr(A - B) = 0$ (d) $tr(A + B) = 4$.

Ans. ***Options a, b, and d***
Let λ *be an eigenvalue of A, then* $p(\lambda)$ *is an eigenvalue of* $p(A)$. *Therefore 1 is the only eigenvalue of B and hence* $det(B) = 1$. *Since* -1 *is an eigenvalue of A, 0 is an eigenvalue of* $A + B$. *Therefore* $det(A + B) = 0$. *Also*

$$tr(A - B) = tr(A) - tr(B) = 0 - 4 = -4$$

and

$$tr(A + B) = tr(A) + tr(B) = 0 + 4 = 4$$

(84) Let A be a 4×4 matrix such that$-1, 1, 1, -2$ are its eigenvalues. If $B = A^4 - 5A^2 + 5I$, then $tr(A + B)$ equals
(a) 0 (b) -12 (c) 3 (d) 9.

Ans. ***Option c***
We have $A + B = A^4 - 5A^2 + A + 5I = p(A)$, as a polynomial in A. Then the eigenvalues of $A + B$ are $p(\lambda)$, where λ is an eigenvalue of A. Therefore, the eigenvalues of $A + B$ are $0, 2, 2, -1$ and hence $tr(A + B) = 3$.

(85) If the determinant of an $n \times n$ matrix A is zero, then
(a) $Rank(A) \leq n - 2$ (b) $tr(A)$ is zero
(c) 0 is an eigenvalue of A (d) $x = 0$ is the only solution of $Ax = 0$.

Ans. ***Option c***
Consider $A = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$. Then $det(A) = 0$. But $Rank(A) \not\leq n - 2$ and $tr(A) \neq 0$. Since determinant is equal to the product of eigenvalues, determinant is zero implies one eigenvalue is 0. Since A is not invertible the homogeneous system has an infinite number of solutions.

(86) Let A be a 10×10 matrix with complex entries such that all its eigenvalues are non-negative real numbers, and at least one eigenvalue is positive. Which of the following statements is always false?

(a) There exists a matrix B such that $AB - BA = B$.
(b) There exists a matrix B such that $AB - BA = A$.
(c) There exists a matrix B such that $AB + BA = A$.
(d) There exists a matrix B such that $AB + BA = B$.

Ans. ***Option b***
For any given A, take $B = 0$. Then $AB - BA = B$. Suppose there exists a matrix B such that $AB - BA = A$. Then $tr(AB - BA) = 0 = tr(A)$. But $tr(A)$ cannot be 0 since all the eigenvalues of A are non-negative. For any given A, take B as the diagonal matrix with diagonal entries $\frac{1}{2}$. Then $AB + BA = A$. For any given A, take $B = 0$. Then $AB + BA = B$.

(87) Let $A = (a_{ij})$ be a 10×10 matrix such that $a_{ij} = 1$ for $i \neq j$ and $a_{ii} = \alpha + 1$, where $\alpha > 0$. Let λ and μ be the largest and the smallest eigenvalues of A, respectively. If $\lambda + \mu = 24$, then α equals ....

Ans. *The given matrix is of the form $A = \alpha I + B$ where B is the matrix with $b_{ij} = 1 \ \forall \ i, j$. Then the eigenvalue of B are 10 and zero with algebraic multiplicities 1 and 9, respectively. Therefore the eigenvalues of A are $10 + \alpha$ and α with algebraic multiplicities 1 and 9, respectively. Since $\alpha > 0$, $\lambda = 10 + \alpha$ and $\mu = \alpha$. Therefore $\lambda + \mu = 24 \Rightarrow 10 + 2\alpha = 24 \Rightarrow \alpha = 7$.*

(88) Let v be a real $n \times 1$ vector satisfying $v^T v = 1$. Define $A = I - 2vv^T$. Which of the following statements are true?
(a) A is singular. (b) $A^2 = A$. (c) $tr(A) = n - 2$. (d) $A^2 = I$.

Ans. ***Options c and d***
Since $(vv^T)v = v(v^T v) = v$ and vv^T is a matrix of rank 1, the eigenvalues of vv^T are 0 and 1. Therefore the eigenvalues of A are 1 and -1. Therefore A is non-singular. Now,

$$A^2 = (I - 2vv^T)(I - 2vv^T) = I - 4vv^T + 4vv^T vv^T = I$$

and

$$tr(A) = tr(I - 2vv^T t) = tr(I) - 2(tr(vv^T)) = n - 2$$

(89) Let $A \in M_{2\times2}(\mathbb{R})$ such that $tr(A) = 2$ and $det(A) = 3$. Write down the characteristic polynomial of A^{-1}.

Ans. *Let λ_1 and λ_2 be the eigenvalues of A, then $tr(A) = \lambda_1 + \lambda_2 = 2$ and $det(A) = \lambda_1\lambda_2 = 3$. Then $\lambda_1 = 1 + \sqrt{2}i$ and $\lambda_2 = 1 - \sqrt{2}i$. The eigenvalues of A^{-1} are $\frac{1}{\lambda_1} = \frac{1-\sqrt{2}i}{3}$ and $\frac{1}{\lambda_2} = \frac{1+\sqrt{2}i}{3}$ and hence the characteristic polynomial of A^{-1} is $\lambda^2 - \frac{2}{3}\lambda + \frac{1}{3}$.*

(90) Let A be a real 6×6 matrix. Let 2 and 1 be two eigenvalues of A. If $A^5 = aI + bA$, where $a, b \in \mathbb{R}$, then
(a) $a = 10, b = 11$ (b) $a = -11, b = 10$
(c) $a = -10, b = 11$ (d) $a = 10, b = -11$.

Ans. ***Option a***
The given matrix satisfies the equation $A^5 = aI + bA$. Therefore $p(\lambda) = \lambda^5 - b\lambda - a$ is an annihilating polynomial for A. Since 2 and -1 are two eigenvalues of A, we have

$$p(2) = 32 - 2b - a = 0 \text{ and } p(-1) = -1 + b - a = 0$$

Solving, we get $a = 10, b = 11$.

(91) Let A be a 3×3 matrix with real entries such that $A^2 = A + 2I$. If α, β, and γ are eigenvalues of A such that $\alpha\beta\gamma = -4$, then $\alpha + \beta + \gamma$ is equal to ....

Ans. *As A satisfies the equation $A^2 = A + 2I$, $\lambda^2 - \lambda - 2 = 0$ is an annihilating polynomial of A. Then 2 and -1 are the possible eigenvalues. Since $\alpha\beta\gamma = -4$, the eigenvalues must be $2, 2, -1$. Therefore $\alpha + \beta + \gamma = 3$.*

(92) Let

$$S = \{x \in \mathbb{R} \mid x = tr(A) \text{ for some } A \in \mathbb{M}_{4\times4}(\mathbb{R}) \text{ such that } A^2 = A\}$$

Then which of the following describes S?
(a) $S = \{0, 2, 4\}$ (b) $S = \{0, \frac{1}{2}, 1, \frac{3}{2}, \frac{5}{2}, 3, \frac{7}{2}, 4\}$
(c) $S = \{0, 1, 2, 3, 4\}$ (d) $S = [0, 4]$.

Ans. ***Option c***
Since A satisfies $A^2 = A$, $\lambda^2 - \lambda$ is an annihilating polynomial of the given matrix. Hence the possible eigenvalues are 0 and 1. Since A is of order 4 and trace is the sum of eigenvalues, $S = \{0, 1, 2, 3, 4\}$.

(93) Let A be an $n \times n$ matrix (with $n > 1$) satisfying $A^2 - 7A + 12I = 0$. Then which of the following statements are true?

(a) A is invertible
(b) $\mu^2 - 7\mu + 12n = 0$ where $\mu = tr(A)$
(c) $d^2 - 7d + 12 = 0$ where $d = det(A)$
(d) $\lambda^2 - 7\lambda + 12 = 0$ where λ is an eigenvalue of A.

Ans. ***Options a and d***
Since A satisfies $A^2 - 7A + 12I = 0$, $(\lambda - 3)(\lambda - 4) = 0$ is an annihilating polynomial of the given matrix. Since 0 cannot be an eigenvalue of A, the matrix is invertible and an eigenvalue of A satisfies $\lambda^2 - 7\lambda + 12 = 0$. Consider the matrix $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$. Then A satisfies the given equation but $tr(A) = 7$ and $det(A) = 12$ are not solutions of $\lambda^2 - 7\lambda + 12 = 0$.

(94) Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map that satisfies $T^2 = T - I$. Then which of the following are true?
(a) T is invertible. (b) $T - I$ is not invertible.
(c) T has a real eigenvalue. (d) $T^3 = -I$.

Ans. ***Options a and d***
Since T satisfies $T^2 = T - I$, $\lambda^2 - \lambda + 1$ is an annihilating polynomial and hence its possible eigenvalues are $\frac{1 \pm \sqrt{3i}}{2}$. As 0 is not an eigenvalue of T, it is invertible and as 1 is not an eigenvalue of $T - I$ is also invertible. Now

$$T^2 = T - I \Rightarrow T^3 = T^2 - T = (T - I) - T = -I$$

(95) Let A be a non-zero linear transformation on a real vector space V of dimension n. Let the subspace $W \subset V$ be the image of V under A. Let $k = dim(W) < n$ and suppose that for some $\lambda \in \mathbb{R}$, $A^2 = \lambda A$. Then

(a) $\lambda = 1$.
(b) $det(A) = |\lambda|^n$.
(c) λ is the only eigenvalue of A.
(d) There is a non-trivial subspace $W_1 \subset V$ such that $Av = 0$ for all $v \in W_1$.

Ans. ***Option d***
Since A satisfies $A^2 = \lambda A$, $x^2 - \lambda x = x(x - \lambda)$ *is an annihilating polynomial of A. Since* $dim\,[\mathcal{R}(A)] < n$, *A is not of full rank. Therefore 0 is an eigenvalue of A. Also* $dim\,[\mathcal{N}(A)] \geq 1$. *Therefore there is a non-trivial subspace* $W_1 \subset V$ *such that* $Av = 0$ *for all* $v \in W_1$.

(96) Let A be an $n \times n$ matrix with real entries such that $A^3 = I$. Suppose that $Av \neq v$ for any non-zero vector v. Then which of the following statements is/are TRUE?
(a) A has real eigenvalues (b) $A + A^{-1}$ has real eigenvalues
(c) n is divisible by 2 (d) n is divisible by 3.

Ans. ***Options b and c***
Since $A^3 = I$, $\lambda^3 - 1$ *is an annihilating polynomial of A. Thus the possible eigenvalues of A are* $1, \omega, \omega^2$. *Since* $Av \neq v$ *for any non-zero vector v, 1 is not an eigenvalue of A. Hence the only possible eigenvalues of A are* ω, ω^2. $A + A^{-1}$ *has only one eigenvalue* $\omega + \omega^2 = -1$, *which is real. As A has only complex eigenvalues and complex eigenvalues occur in conjugate pairs, n is divisible by 2.*

(97) (a) Let A be a 3×3 real matrix with $det(A) = 6$. Then find $det(adj A)$.
(b) Let v_1 and v_2 be non-zero vectors in $\mathbb{R}^n$, $n \geq 3$, such that v_2 is not a scalar multiple of v_1. Prove that there exists a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ such that $T^3 = T$, $Tv_1 = v_2$, and T has at least three distinct eigenvalues.

Ans. *(a) We have*

$$A\,(adj(A)) = det(A)I$$

Taking determinant on both sides,

$$det(A)det(adj(A)) = (det(A))^3 = 216 \Rightarrow det(adj(A)) = 36$$

(b) Let $\{v_1, v_2, \ldots, v_n\}$ *be a basis for* $\mathbb{R}^n$. *Define* $T(v_1) = v_2$, $T(v_2) = v_1$, *and* $T(v_j) = v_j\ \forall\ j = 3, \ldots, n$. *Then*

$$T^3(v_1) = T^2\,(T(v_1)) = T^2(v_2) = T\,(T(v_2)) = T(v_1)$$

and

$$T^3(v_2) = T^2\,(T(v_2)) = T^2(v_1) = T\,(T(v_1)) = T(v_2)$$

Clearly $T^3(v_j) = T(v_j)\ \forall\ j = 3, \ldots, n$. *Therefore T is a linear transformation such that* $T^3 = T$ *and* $Tv_1 = v_2$. *Now since* $\lambda^3 - \lambda$ *is an annihilating polynomial of T and T is not the identity transformation. Therefore T has at least three distinct eigenvalues.*

(98) How many elements of $\mathbb{M}_{2\times 2}(\mathbb{Z}_7)$ are similar to the following matrix?

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Ans. *The characteristic polynomial of the given matrix is* $\lambda^2 - \lambda = 0$. *It has* $trace = 1$ *and* $determinant = 0$. *The general form of an element is* $\begin{bmatrix} a & b \\ c & 1-a \end{bmatrix}$ *where* $a(1-a) = bc, a, b, c \in \mathbb{Z}_7$. *There are 56 such matrices.*

(99) Let A and B be real invertible matrices such that $AB = -BA$. Then
(a) $tr(A) = tr(B) = 0$ (b) $tr(A) = tr(B) = 1$
(c) $tr(A) = 0, tr(B) = 1$ (d) $tr(A) = 1, tr(B) = 0$.

Ans. ***Option a***
Let $AB = -BA$. *Since* A *is invertible, multiplying by* A^{-1} *gives* $B = A^{-1}(-B)A$. *Thus,* B *and* $-B$ *are similar. Therefore they have the same trace (Corollary 4.4). That is,* $tr(B) = tr(-B) \Rightarrow tr(B) = 0$. *Similarly multiplying by* B^{-1} *gives* $tr(A) = 0$.

(100) Let A and B be $n \times n$ matrices over $\mathbb{C}$. Then,

(a) AB and BA always have the same set of eigenvalues.
(b) If AB and BA have the same set of eigenvalues then $AB = BA$.
(c) If A^{-1} exists then AB and BA are similar.
(d) The rank of AB is always the same as the rank of BA.

Ans. ***Options a and c***

(a) Let 0 be an eigenvalue of AB. *Then*

$$0 = det(AB) = det(A)det(B) = det(B)det(A) = det(BA)$$

This implies that 0 is also an eigenvalue of BA. *Now let* λ *be an eigenvalue of* AB, *then there exists* $v \neq 0$ *such that*

$$(AB)v = A(Bv) = \lambda v \Rightarrow B(AB)v = (BA)(Bv) = \lambda(Bv)$$

Clearly $Bv \neq 0$, *since* $Bv = 0$ *implies that either* $\lambda = 0$ *or* $v = 0$, *which are not possible. Therefore* λ *is an eigenvalue of* BA *with eigenvector* Bv *where* v *is an eigenvector of* AB *with respect to the eigenvalue* λ.

(b) Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. *Then* $AB = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ *and* $BA = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. *$AB$ and BA have the same set of eigenvalues (only 0). But $AB \neq BA$.*

(c) Since A is invertible

$$AB = (AB)I = (AB)(AA^{-1}) = A(BA)A^{-1}$$

AB and BA are similar.

(d) Consider A and B in option b. Then $Rank(AB) = 0$ and $Rank(BA) = 1$.

(101) The minimal polynomial associated with the matrix $\begin{bmatrix} 0 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ is

(a) $\lambda^3 - \lambda^2 - 2\lambda - 3$
(b) $\lambda^3 - \lambda^2 + 2\lambda - 3$
(c) $\lambda^3 - \lambda^2 - 3\lambda - 3$
(d) $\lambda^3 - \lambda^2 + 3\lambda - 3$.

Ans. ***Option a***
Since the characteristic polynomial of a 3×3 matrix A is given by $\lambda^3 - [tr(A)]\lambda^2 + [M_{11} + M_{22} + M_{33}]\lambda - det(A)$, the characteristic and minimal polynomial of the given matrix is $\lambda^3 - \lambda^2 - 2\lambda - 3$.

(102) The minimal polynomial of the matrix $\begin{bmatrix} 3 & 3 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$ is

(a) $\lambda(\lambda - 1)(\lambda - 6)$ (b) $\lambda(\lambda - 3)$ (c) $(\lambda - 3)(\lambda - 6)$ (d) $\lambda(\lambda - 6)$.

Ans. ***Option d***
Since the given matrix has determinant zero, zero is an eigenvalue. Since it has trace 12, the sum of the eigenvalues must be 12. From the given options, this is possible only when the minimal polynomial is $\lambda(\lambda - 6)$.

(103) The minimal polynomial of $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 5 \end{pmatrix}$ is

(a) $(\lambda - 2)(\lambda - 5)$ (b) $(\lambda - 2)^2(\lambda - 5)$ (c) $(\lambda - 2)^3(\lambda - 5)$ (d) $(\lambda - 2)^4$.

Ans. ***Option b***
Since the given matrix is a block diagonal matrix, the characteristic polynomial of the given matrix is $(\lambda - 2)^3(\lambda - 5)$ and the minimal polynomial is $(\lambda - 2)^2(\lambda - 5)$.

(104) Let A be a non-diagonal 2×2 matrix with complex entries such that $A = A^{-1}$. Write down its characteristic and minimal polynomials.

Ans. *As* $A = A^{-1}$*, we get* $A^2 = I$*. Since* A *is a* 2×2 *matrix, the characteristic polynomial of* A *is* $\lambda^2 - 1 = (\lambda + 1)(\lambda - 1)$*. Since each eigenvalue is a root of the minimal polynomial, minimal polynomial is same as the characteristic polynomial.*

(105) Let $f(\lambda)$ be the minimal polynomial of the 4×4 matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

then the rank of the 4×4 matrix $f(A)$ is
(a) 0 (b) 1 (c) 2 (d) 4.

Ans. ***Option a***
Since the minimal polynomial of a matrix A *is the least degree monic polynomial* $f(\lambda)$ *such that* $f(A) = 0$*, rank of* $f(A)$ *is 0.*

(106) If $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ (with $n \geq 2$) has rank 1, then show that the minimal polynomial of A has degree 2.

Ans. $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ *implies that there exists a linear transformation* $T : \mathbb{R}^n \to \mathbb{R}^n$*. Let* $\mathcal{R}(T)$ *denote the range space of* T *and* $\mathcal{N}(T)$ *denote the null space of* T*. Since*

$$Rank(A) = dim\,[\mathcal{R}(T)] = 1$$

$\mathcal{R}(T)$ *is spanned by some non-zero vector* $v \in \mathbb{R}^n$*. Since* $v \in \mathcal{R}(T)$ *there exists* $v_1 \in \mathbb{R}^n$ *such that* $Tv_1 = v$*. Since* $dim\,[\mathcal{N}(T)] = n - 1$*, null space of* T *is spanned by a set* $\{v_2, v_3, \ldots, v_n\}$*. Then* $\{v_1, v_2, v_3, \ldots, v_n\}$ *forms a basis for* $\mathbb{R}^n$*. Now we have* $T(v_1) = v$ *and* $T(v_i) = 0$ *for all* $i = 2, \ldots, n$ *and there are two possibilities for* $T(v_1)$*. Either* $T(v_1) \in \mathcal{N}(T)$ *or* $T(v_1) \notin \mathcal{N}(T)$*.*
Suppose that $T(v_1) \in \mathcal{N}(T)$*, then* $v = T(v_1) = c_2v_2 + \cdots + c_nv_n$ *since* $\mathcal{N}(T) = span\{v_1, v_2, v_3, \ldots, v_n\}$*. Then the matrix of* T *is given by*

$$\begin{bmatrix} 0 & 0 \cdots 0 \\ c_2 & 0 \cdots 0 \\ c_3 & 0 \cdots 0 \\ \vdots & \vdots \ddots \vdots \\ c_{n-1} & 0 \cdots 0 \end{bmatrix}$$

Then clearly $T^2 = 0$*. Now suppose that* $T(v_1) \notin \mathcal{N}(T)$*, then* $T(v_1) \in span\{v\}$*. That is,* $T(v_1) = c_1v$ *for some* $c_1 \neq 0 \in \mathbb{R}$*. In this case the matrix of* T *corresponding to the basis* $\{v, v_2, v_3, \ldots, v_n\}$ *is*

$$\begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Clearly the matrix of T is diagonalizable. In both cases, the minimal polynomial of A is of degree 2 (since the matrix corresponding to a linear transformation with respect to different basis is similar).

(107) Let A be a real 2×2 matrix such that $A^6 = I$. The total number of possibilities for the characteristic polynomial of A is ....

Ans. *Since A satisfies $A^6 = I$, $\lambda^6 - 1$ is an annihilating polynomial of A. Now*

$$\lambda^6 - 1 = (\lambda^3 - 1)(\lambda^3 + 1) = (\lambda + 1)(\lambda - 1)(\lambda^2 + \lambda + 1)(\lambda^2 - \lambda + 1)$$

So the possible minimal polynomials are

$$(\lambda + 1), (\lambda - 1), (\lambda + 1)(\lambda - 1), (\lambda^2 + \lambda + 1), (\lambda^2 - \lambda + 1)$$

and hence the possible characteristic polynomials are

$$(\lambda + 1)^2, (\lambda - 1)^2, (\lambda + 1)(\lambda - 1), (\lambda^2 + \lambda + 1), (\lambda^2 - \lambda + 1)$$

Therefore, the total number of possibilities for the characteristic polynomial of A is 5.

(108) Consider a matrix $A = \left(a_{ij}\right)_{n\times n}$ with integer entries such that $a_{ij} = 0$ for $i > j$ and $a_{ij} = 1$ for $i = 1, 2, \ldots, n$. Which of the following properties must be true?

(a) A^{-1} exists and it has integer entries.
(b) A^{-1} exists and it has some entries that are not integers.
(c) A^{-1} is a polynomial function of A with integer coefficients.
(d) A^{-1} is not a power of A unless A is the identity matrix.

Ans. ***Options a, c, and d***

The matrix is of the form $\begin{bmatrix} 1 & 1 & \ldots & 1 \\ 0 & 1 & \ldots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$. *Since the matrix is upper triangular, the characteristic polynomial of the given matrix is $(\lambda - 1)^n$. As the determinant is 1, A^{-1} exists and it has integer entries and by Cayley–Hamilton Theorem, A^{-1} is a polynomial function of A with integer coefficients. Now suppose that*

$A^{-1} = A^k$ for some positive integer k, then $A^{k+1} = I$. Thus, $\lambda^{k+1} - 1$ is an annihilating polynomial of A. Since the minimal polynomial divides the characteristic polynomial and every annihilating polynomial, the minimal polynomial is $\lambda - 1$.

(109) Consider the linear map $T : \mathbb{P}_3[a, b] \to \mathbb{P}_3[a, b]$ defined by

$$(Tp)(x) = p(x+1) + p(x-1)$$

Which of the following properties does the matrix of T (with respect to the standard basis $B = \{1, x, x^2, x^3\}$ of $\mathbb{P}_3[a, b]$) satisfy?

(a) $det(T) = 0$
(b) $(T - 2I)^4 = 0$ but $(T - 2I)^3 \neq 0$
(c) $(T - 2I)^3 = 0$ but $(T - 2I)^2 \neq 0$
(d) 2 is an eigenvalue with multiplicity 4.

Ans. ***Option d***
We have

$$T(1) = 2 = 2.1 + 0x + 0x^2 + 0x^3$$

$$T(x) = 2x = 0.1 + 2x + 0x^2 + 0x^3$$

$$T(x^2) = 2x^2 + 2 = 2.1 + 0x + 2x^2 + 0x^3$$

$$T(x^3) = 2x^3 + 6x = 0.1 + 6x + 0x^2 + 2x^3$$

Therefore the matrix of T is $A = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$. *The characteristic polynomial of A is $(\lambda - 2)^4$ and the minimal polynomial is $(\lambda - 2)^2$. Clearly $det(T) = 16$.*

(110) Let $A \in \mathbb{M}_{3\times 3}(\mathbb{R})$ be such that $A^8 = I$. Then

(a) minimal polynomial of A can only be of degree 2.
(b) minimal polynomial of A can only be of degree 3.
(c) either $A = I$ or $A = -I$.
(d) there are uncountably many A satisfying the above.

Ans. ***Option d***
Consider $A = I$. Then $A^8 = I$. But the minimal polynomial is of degree 1. Now consider the matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. *Then $A^8 = I$. But $A \neq I$ or $A \neq -I$. In fact any matrix with minimal polynomial $(\lambda - 1)(\lambda + 1) = 0$ satisfies $A^8 = I$.*

(111) Let V and W be finite dimensional vector spaces over $\mathbb{R}$, and let $T_1 : V \to V$ and $T_2 : W \to W$ be linear transformations whose minimal polynomials are given by

$$f_1(\lambda) = \lambda^3 + \lambda^2 + \lambda + 1 \text{ and } f_2(\lambda) = \lambda^4 - \lambda^2 - 2$$

Let $T : V \oplus W \to V \oplus W$ be the linear transformation defined by

$$T(v, w) = (T_1(v), T_2(w)) \ for \ (v, w) \in V \oplus W$$

and let $f(\lambda)$ be the minimal polynomial of T. Then
(a) $deg\,[f(\lambda)] = 7$ (b) $deg\,[f(\lambda)] = 5$
(c) $Nullity(T) = 1$ (d) $Nullity(T) = 0$.

Ans. ***Options b and d***
As $f_1(\lambda) = \lambda^3 + \lambda^2 + \lambda + 1 = (\lambda + 1)(\lambda^2 + 1)$ *and* $f_2(\lambda) = \lambda^4 - \lambda^2 - 2 = (\lambda^2 - 2)(\lambda^2 + 1)$, *we have*

$$\begin{aligned} f(\lambda) &= lcm\left(\lambda^3 + \lambda^2 + \lambda + 1, \lambda^4 - \lambda^2 - 2\right) \\ &= (\lambda + 1)(\lambda^2 + 1)(\lambda^2 - 2) \\ &= \lambda^5 + \lambda^4 - \lambda^3 - \lambda^2 - 2\lambda - 2 \end{aligned}$$

$deg\,[f(\lambda)] = 5$ *and since 0 is not an eigenvalue of* $f(\lambda)$, $Nullity(T) = 0$.

(112) Let A be an $n \times n$ matrix over $\mathbb{C}$ such that every non-zero vector of $\mathbb{C}^n$ is an eigenvector of A. Then

(a) All eigenvalues of A are equal.
(b) All eigenvalues of A are distinct.
(c) $A = \lambda I$ for some $\lambda \in \mathbb{C}$, where I is the $n \times n$ identity matrix.
(d) If χ_A and m_A denote the characteristic polynomial and the minimal polynomial respectively, then $\chi_A = m_A$.

Ans. ***Options a and c***
Suppose that λ_1 *and* λ_2 *are two distinct eigenvalues of A with eigenvectors* v_1 *and* v_2. *Since eigenvectors corresponding to distinct eigenvalues are linearly independent,* $\{v_1, v_2\}$ *is linearly independent. Now as every vector in* $\mathbb{C}^n$ *is an eigenvector of A, there exists* λ *such that* $A(v_1 + v_2) = \lambda(v_1 + v_2)$. *But* $A(v_1 + v_2) = Av_1 + Av_2 = \lambda_1 v_1 + \lambda_2 v_2$, *we get*

$$\lambda(v_1 + v_2) = \lambda_1 v_1 + \lambda_2 v_2 \Rightarrow (\lambda - \lambda_1)v_1 = (\lambda_2 - \lambda)v_2$$

Since $\lambda_1 \neq \lambda_2$ at least one of $(\lambda - \lambda_1)$ and $(\lambda_2 - \lambda)$ is non-zero. This implies $\{v_1, v_2\}$ is linearly dependent which is a contradiction. Therefore all eigenvalues of A are equal.
Since $Av = \lambda v$ for all $v \in \mathbb{C}^n$ and for some $\lambda \in \mathbb{C}$,

$$(A - \lambda I)v = 0 \;\forall\; v \in \mathbb{C}^n$$

This gives $(A - \lambda I)$ is the zero operator and hence $A = \lambda I$ for some $\lambda \in \mathbb{C}$. The characteristic polynomial is given by $\chi_A = (x - \lambda)^n$, and the minimal polynomial is given by $m_A = (x - \lambda)$.

(113) Given two $n \times n$ matrices A and B with entries in $\mathbb{C}$. Consider the following statements:
P : If A and B have the same minimal polynomial, then A is similar to B.
Q : If A has n distinct eigenvalues, then there exists $v \in \mathbb{C}^n$ such that $v, Av, \ldots, A^{n-1}v$ are linearly independent.
Which of the above statements hold TRUE?
(a) both P and Q (b) only P
(c) only Q (d) Neither P nor Q.

Ans. ***Option c***

(a) Consider the matrices $\begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}$ *and* $\begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{bmatrix}$. *They have the same minimal and characteristic polynomial. But they are not similar.*

(b) Since A has n distinct eigenvalues $\lambda_0, \lambda_1, \ldots, \lambda_{n-1}$, A has n linearly independent eigenvectors, say $v_0, v_1, v_2, \ldots v_{n-1}$. Consider the vector $v = v_0 + v_1 + \cdots + v_{n-1}$. Then

$$A^k v = A^k v_0 + A^k v_1 + \cdots + A^k v_{n-1} = \lambda_0^k v_0 + \lambda_1^k v_1 + \cdots + \lambda_{n-1}^k v_{n-1}$$

where $k = 1, 2, \ldots n - 1$. Now consider $c_0 v + c_1 Av + \cdots + c_{n-1} A^{n-1}v = 0$. This implies that $(c_0 + c_1\lambda_0 + c_2\lambda_0^2 + \cdots + c_{n-1}\lambda_0^{n-1})v_0 + (c_0 + c_1\lambda_1 + c_2\lambda_1^2 + \cdots + c_{n-1}\lambda_1^{n-1})v_1 + \cdots + (c_0 + c_1\lambda_{n-1} + c_2\lambda_{n-1}^2 + \cdots + c_{n-1}\lambda_{n-1}^{n-1})v_{n-1} = 0$. Since $v_0, v_1, v_2, \ldots v_{n-1}$ is linearly independent, this implies $c_0 + c_1\lambda_i + c_2\lambda_i^2 + \cdots + c_{n-1}\lambda_i^{n-1} = 0$ for all $i = 0, 1, \ldots, n-1$. This can be written in the matrix form $AX = B$ as

$$\begin{bmatrix} 1 & \lambda_0 & \lambda_0^2 & \ldots & \lambda_0^{n-1} \\ 1 & \lambda_1 & \lambda_1^2 & \ldots & \lambda_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{n-1} & \lambda_{n-1}^2 & \ldots & \lambda_{n-1}^{n-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Since the coefficient matrix is a Vandermonde matrix,[1] *with* $\lambda_i \neq \lambda_j \; \forall \, i \neq j$*, it is invertible and the given system has trivial solution only. Hence,* $c_0 = c_1 = \cdots = c_{n-1} = 0$*. Therefore the set* $v, Av, \ldots, A^{n-1}v$ *is linearly independent.*

(114) A non-zero matrix $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ is said to be nilpotent if $A^k = 0$ for some positive integer $k \geq 2$. If A is nilpotent, which of the following statements are true?

(a) Necessarily, $k \leq n$ for the smallest such k
(b) The matrix $I + A$ is invertible
(c) All eigenvalues of A are zero.

Ans. ***Options a, b, and c***

(a) Since A satisfies if $A^k = 0$ *for some positive integer* $k \geq 2$*,* λ^k *is an annihilating polynomial of A. Since the minimal polynomial divides annihilating polynomial, it is of the form* λ^r *for some* $r \leq n$ *as it divides the characteristic polynomial also.*
(b) Since the minimal polynomial is of the form λ^r *for some* $r \leq n$*, all eigenvalues of A are zero. Then the eigenvalues of* $I + A$ *are 1. Therefore* $I + A$ *is invertible.*
(c) All eigenvalues of A are zero.

(115) Let V be the vector space of polynomials of degree at most 3 in a variable x with coefficients in $\mathbb{R}$. Let $T = \frac{d}{dx}$ be the linear transformation of V into itself given by differentiation. Which of the following are correct?

(a) T is invertible.
(b) 0 is an eigenvalue of T.
(c) There is a basis with respect to which the matrix of T is nilpotent.
(d) The matrix of T with respect to the basis $\{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$ is diagonal.

Ans. ***Options b and c***
Consider the standard ordered basis $\{1, x, x^2, x^3\}$ *for* V*. Now*

[1] A Vandermonde matrix, named after the French Mathematician *Alexandre-Théophile Vandermonde (1735–1796)*, is a matrix of the form $A = \begin{bmatrix} 1 & x_0 & \ldots & x_0^n \\ 1 & x_1 & \ldots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & \ldots & x_m^n \end{bmatrix}$. The determinant of the Vandermonde matrix is given by $det(A) = \prod_{0\leq i<j\leq n} \left(x_j - x_i\right)$.

$$T(1) = 0 = 0 + 0x + 0x^2 + 0x^3$$

$$T(x) = 1 = 1 + 0x + 0x^2 + 0x^3$$

$$T(x^2) = 2x = 0 + 2x + 0x^2 + 0x^3$$

$$T(x^3) = 3x^2 = 0 + 0x + 3x^2 + 0x^3$$

Then the matrix of T *is* $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Clearly it is a nilpotent matrix as* $T^4 = 0$ *and 0 is an eigenvalue of* T. *Since the matrices of a linear transformation corresponding to different basis are similar, the matrix of T with respect to the basis* $\{1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3\}$ *is not diagonal.*

(116) Let $A = \begin{bmatrix} 2 & 0 & -3 \\ 3 & -1 & -3 \\ 0 & 0 & -1 \end{bmatrix}$. A matrix P such that $P^{-1}AP$ is a diagonal matrix is

(a) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} -1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & -1 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

Ans. ***Option a***

The characteristic polynomial of the given matrix A is

$$\lambda^3 - 3\lambda - 2 = (\lambda + 1)^2(\lambda - 2)$$

-1 *is an eigenvalue of the given matrix with eigenvectors* $(1, 0, 1)^T$, $(1, 1, 1)^T$ *and 2 is an eigenvalue with eigenvector* $(1, 1, 0)^T$.

(117) Which of the following matrices is NOT diagonalizable?

(a) $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

Ans. ***Option d***

- *(a) The characteristic equation of the given matrix is* $\lambda^2 - 3\lambda + 1 = 0$. *Since it has 2 distinct roots, the given matrix is diagonalizable.*
- *(b) The given matrix is a lower triangular matrix and hence the eigenvalues are its diagonal entries. Since it has 2 distinct eigenvalues, the given matrix is diagonalizable.*
- *(c) The characteristic equation of the given matrix is* $\lambda^2 + 1 = 0$. *Since it has two distinct complex roots, the given matrix is diagonalizable over* $\mathbb{C}$.
- *(d) The characteristic and minimal polynomial of the given matrix is* $(\lambda - 1)^2$. *Since the minimal polynomial of the given matrix is not of linear factors, the given matrix is not diagonalizable.*

(118) Which of the following matrices is not diagonalizable over $\mathbb{R}$?

(a) $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ (d) $\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$.

Ans. *(a) The given matrix is an upper triangular matrix and hence its diagonal entries are its eigenvalues. Hence its characteristic polynomial is* $(\lambda - 2)^2(\lambda - 3)$. *The minimal polynomial is the same as the characteristic polynomial. Since minimal polynomial does have non-linear factors, the matrix is not diagonalizable.*

(b) Since the given matrix is a 2×2 *matrix with distinct eigenvalues, it is diagonalizable. (Its characteristic polynomial is* $\lambda^2 - 2\lambda = \lambda(\lambda - 2)$.*)*

(c) The characteristic polynomial of the given matrix is $(\lambda - 2)^2(\lambda - 3)$ *and its minimal polynomial is* $(\lambda - 2)(\lambda - 3)$. *Since the minimal polynomial is of linear factors only, the matrix is diagonalizable.*

(d) Since the given matrix is a 2×2 *matrix with distinct eigenvalues, it is diagonalizable. (Its characteristic polynomial is* $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$.*)*

Option a

(119) Let $A = \begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & 4 \\ -2 & 1 & -4 \end{bmatrix}$. Given that 1 is an eigenvalue of A, which among the following are correct?

(a) The minimal polynomial of A is $(\lambda - 1)(\lambda + 4)$ (b) A is not diagonalizable
(c) The minimal polynomial of A is $(\lambda - 1)^2(\lambda + 4)$ (d) $A^{-1} = \frac{1}{4}(A + 3I)$.

Ans. ***Options b and c***

Let λ_1 *and* λ_2 *be the two eigenvalues of A. Since* $tr(A) = -2$ *and* $det(A) = -4$, $\lambda_1 + \lambda_2 = -3$ *and* $\lambda_1\lambda_2 = -4$. *Thus we get* $\lambda_1 = -4$ *and* $\lambda_2 = 1$. *Therefore the characteristic polynomial of A is* $(\lambda - 1)^2(\lambda + 4)$. *Since* $(A - I)(A + 4I) \neq 0$, *the minimal polynomial of A is* $(\lambda - 1)^2(\lambda + 4)$. *Since the minimal polynomial consists of non-linear factors, A is not diagonalizable.*

(120) Let A be a non-zero 2×2 matrix with real entries. Pick out the true statements:

(a) If $A^2 = A$, then A is diagonalizable.
(b) If $A^2 = 0$, then A is diagonalizable.
(c) If A is invertible, then $A = (tr(A))I - (det(A))A^{-1}$.

Ans. ***Options a and c***

(a) $A^2 = A$ *implies that A satisfies the polynomial equation*

$$\lambda^2 - \lambda = \lambda(\lambda - 1)$$

Then the possibilities of minimal polynomial are λ, $(\lambda - 1)$ *and* $\lambda(\lambda - 1)$. λ *is not the minimal polynomial since A is not zero. Therefore the possible minimal polynomials are* $(\lambda - 1)$ *and* $\lambda(\lambda - 1)$. *In either case, minimal polynomial consists of linear factors. Therefore A is diagonalizable.*

(b) *Consider the matrix* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. *Then* $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. *But A is not diagonalizable as the minimal polynomial of A is* λ^2.

(c) *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *then the characteristic equation of A is*

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0 \Rightarrow A^2 - (a+d)A + (ad - bc)I = 0$$
$$\Rightarrow A^2 = tr(A)A + (det(A))I$$
$$\Rightarrow A = (tr(A))I - (det(A))A^{-1}.$$

(121) Let A be an $n \times n$ matrix with real entries. Which of the following is correct?

(a) If $A^2 = I$, then A is diagonalizable over real numbers.
(b) If $A^2 = A$, then A is diagonalizable only over complex numbers.
(c) The only matrix of size n satisfying the characteristic polynomial of A is A.

Ans. ***Option b***

(a) *Since A satisfies* $A^2 = I$, $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$ *is an annihilating polynomial of A. Since the minimal polynomial of A divides annihilating polynomial, the possible minimal polynomials are* $\lambda - 1, \lambda + 1$, *and* $(\lambda - 1)(\lambda + 1)$. *Since each of them are of linear factors, A is diagonalizable over real numbers.*

(b) *Since A satisfies* $A^2 = A$, $\lambda^2 - \lambda = \lambda(\lambda - 1)$ *is an annihilating polynomial of A. The possible minimal polynomials are* $\lambda, \lambda - 1$, *and* $\lambda(\lambda - 1)$. *Since each of them are of linear factors, A is diagonalizable over real numbers and complex numbers.*

(c) *Every matrix similar to A has the same characteristic polynomial. For example,* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ *have the same characteristic polynomial.*

(122) An $n \times n$ complex matrix A satisfies $A^k = I$, where k is a positive integer greater than 1. Suppose 1 is not an eigenvalue of A. Then which of the following statements are necessarily true?
(a) A is diagonalizable.
(b) $A + A^2 + \cdots + A^{k-1} = 0$.
(c) $tr(A) + tr(A^2) + \cdots + tr(A^{k-1}) = -n$.
(d) $A^{-1} + A^{-2} + \cdots + A^{-(k-1)} = -I_n$.

Ans. ***Options a, c, and d***
Since A satisfies $A^k = I$, $\lambda^k - 1$ *is an annihilating polynomial of the given matrix. Since* 1 *is not an eigenvalue of A and as*

$$\lambda^k - 1 = (\lambda - 1)(1 + \lambda + \lambda^2 + \cdots + \lambda^{k-1})$$

$1 + \lambda + \lambda^2 + \cdots + \lambda^{k-1}$ *is an annihilating polynomial of A. Therefore*

$$A + A^2 + \cdots + A^{k-1} = -I$$

and hence

$$tr(A) + tr(A^2) + \cdots + tr(A^{k-1}) = -n$$

Since the annihilating polynomial can be linearly factorized using kth *roots of unity, A is diagonalizable over* $\mathbb{C}$. *Since* $I + A + A^2 + \cdots + A^{k-1} = 0$, *multiplying* $A^{-(k-1)}$ *on both sides*

$$A^{-1} + A^{-2} + \cdots + A^{-(k-1)} = -I_n$$

(123) Let A be a 2×2 matrix with real entries which is not a diagonal matrix and which satisfies $A^3 = I$. Pick out the true statements:

(a) $tr(A) = -1$.
(b) A is diagonalizable over $\mathbb{R}$.
(c) $\lambda = 1$ is an eigenvalue of A.

Ans. ***Option a***
Since A satisfies $A^3 = I$, $\lambda^3 - 1 = (\lambda - 1)(\lambda^2 + \lambda + 1)$ *is an annihilating polynomial of A. Therefore the possibilities of the minimal polynomial are* $\lambda - 1$ *and* $\lambda^2 + \lambda + 1$. *As A is not a diagonal matrix,* $\lambda - 1$ *cannot be the minimal polynomial of A. Thus the minimal polynomial of A is* $\lambda^2 + \lambda + 1$. *Since A is of order 2, the characteristic polynomial is the same as the minimal polynomial. A is not diagonalizable over* $\mathbb{R}$ *and* $\lambda = 1$ *is not an eigenvalue of A. Also* $tr(A) =$ *coefficient of* $\lambda = -1$.

(124) Let A be a non-zero 3×3 matrix with the property $A^2 = 0$. Which of the following is/are true?
(a) A is not similar to a diagonal matrix
(b) A is similar to a diagonal matrix
(c) A has one non-zero eigenvector
(d) A has 3 linearly independent eigenvectors.

Ans. ***Options a and c***
Since A is a non-zero 3×3 *matrix that satisfies* $A^2 = 0$, λ^2 *is the minimal polynomial of A and hence A is not similar to a diagonal matrix. 0 is the only eigenvalue of A. Since A has rank 1 (A is not of rank 3 as A is not invertible. If*

$Rank(A) = 2$, $Rank(A^2) \geq 1$), the solution space of $Ax = 0$ has dimension 2. Therefore A has one non-zero eigenvector.

(125) Let $A \in \mathbb{M}_n(\mathbb{R})$ with the property $A^n = 0$. Which of the following is/are true?
(a) A has n distinct eigenvalues (b) A has one eigenvalue of multiplicity n
(c) 0 is an eigenvalue of A (d) A is similar to a diagonal matrix.

Ans. ***Options b and c***
Clearly A is a nilpotent matrix. 0 is the only eigenvalue of A with multiplicity n. Consider the matrix $A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$. *Clearly A is not similar to a diagonal matrix. Since the characteristic polynomial of a nilpotent matrix is λ^n, its minimal polynomial is of the form λ^r where $r \leq n$. Therefore a nilpotent matrix is diagonalizable implies that the matrix is zero matrix.*

(126) Let A be an $n \times n (n \geq 2)$ non-zero real matrix with $A^2 = 0$, and let $\alpha \in \mathbb{R} \setminus \{0\}$. Then

(a) α is the only eigenvalue of $(A + \alpha I)$ and $(A - \alpha I)$
(b) α is the only eigenvalue of $(A + \alpha I)$ and $(\alpha I - A)$
(c) $-\alpha$ is the only eigenvalue of $(A + \alpha I)$ and $(A - \alpha I)$
(d) $-\alpha$ is the only eigenvalue of $(A + \alpha I)$ and $(\alpha I - A)$.

Ans. ***Option b***
As $A^2 = 0$, A is a nilpotent matrix and hence A has 0 as the only eigenvalue. Also if λ is an eigenvalue of A, $\lambda + k$ is an eigenvalue of $A + kI$. Therefore the only eigenvalue of $(A + \alpha I)$ and $(\alpha I - A)$ is α. Also $-\alpha$ is the only eigenvalue of $(A - \alpha I)$.

(127) Let A be a 3×3 real non-diagonal matrix with $A^{-1} = A$. Show that $tr(A) = det(A) = \pm 1$.

Ans. *A satisfies the polynomial equation $\lambda^2 - 1 = 0$. Therefore the possibilities for minimal polynomial of A are $\lambda - 1, \lambda + 1, \lambda^2 - 1$. Since A is a non-diagonal matrix, the minimal polynomial is $\lambda^2 - 1$. Thus the possible eigenvalues are $1, -1, 1$ or $1, -1, -1$. In the first case $tr(A) = 1$, $det(A) = -1$ and in the second case $tr(A) = -1$, $det(A) = 1$.*

(128) Let A be an $n \times n$ matrix with real entries such that $A^2 = A$. Show that

$$Rank(A - I) = Nullity(A)$$

Ans. *Since A satisfies $A^2 = A$, A is diagonalizable and the only possible eigenvalues of A are 0 and 1. If A is diagonalizable, then $A - I$ is also diagonalizable. For, A is diagonalizable implies that there exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. Now*

$$P^{-1}(A - I)P = P^{-1}AP - P^{-1}P = D - I$$

Thus the only possible eigenvalues of $A - I$ are -1 and 0. Also, as $A - I$ is diagonalizable

$$Rank(A - I) = AM(-1) = AM(0)(for\ A) = n - AM(1) = n - Rank(A) = Nullity(A)$$

(129) Let A be an $n \times n$ real matrix with $A^2 = A$. Then

(a) the eigenvalues of A are either 0 or 1
(b) A is a diagonal matrix with diagonal entries 0 or 1
(c) $Rank(A) = tr(A)$
(d) $Rank(I - A) = tr(I - A)$.

Ans. ***Options a, c, and d***
Since A satisfies $A^2 = A$, $\lambda^2 - \lambda$ is an annihilating polynomial of A and hence the only possible eigenvalues of A are 0 and 1. The possibilities for minimal polynomial are $\lambda, \lambda - 1, \lambda(\lambda - 1)$. In any case, the minimal polynomial has only linear factors. Hence A is diagonalizable. Therefore

$$Rank(A) = \text{ number of nonzero eigenvalues of } A = A.M(1) = tr(A)$$

The possible eigenvalues of $I - A$ are 0 and 1 and as above, $Rank(I - A) = tr(I - A)$. Now consider the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$. Then $A^2 = A$, but A is not a diagonal matrix.

(130) Let $A \neq I$ be an $n \times n$ matrix such that $A^2 = A$, where I is the identity matrix of order n. Which of the following statements is false?
(a) $(I - A)^2 = I - A$ (b) $tr(A) = Rank(A)$
(c) $Rank(A) + Rank(I - A) = n$ (d) The only eigenvalue of A is 1.

Ans. ***Option d***

(a) Since $A^2 = A$, we have

$$(I - A)^2 = I - 2A + A^2 = I - A$$

(b) Since $A^2 = A$, $\lambda^2 - \lambda$ is an annihilating polynomial of the given matrix. Therefore only possible eigenvalues of A are 1 and 0. As 1 is the only possible non-zero eigenvalue of A and A is diagonalizable,

$$tr(A) = algebraic\ multiplicity\ of\ 1 = Rank(A)$$

(c) The possible eigenvalues for $I - A$ are again 1 and 0. First, we will show that $I - A$ is diagonalizable. For, as A is diagonalizable, there exists an invertible matrix P such that $P^{-1}AP = D$, where D is a diagonal matrix. Then

$$I - A = P^{-1}P - P^{-1}DP = P^{-1}(I - D)P$$

and as $I - D$ is a diagonal matrix, $I - A$ is diagonalizable. Then,we have

$$Rank(I - A) = AM(1)\ for\ I - A = AM(0)\ for\ A$$

Therefore

$$Rank(A) + Rank(I - A) = AM(1)\ for\ A + AM(0)\ for\ A = n$$

(d) Consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $A^2 = A$, but A has 0 as an eigenvalue.

(131) Let $A \in \mathbb{M}_2(\mathbb{R})$ with $A \neq 0, I$ but $A^2 = A$. Which of the following statements are true?

(a) $\mathcal{N}(A)$ is the eigenspace of A corresponding to the eigenvalue 0.
(b) Let $v \neq 0 \in Col(A)$, then v is an eigenvector of A corresponding to the eigenvalue 1.
(c) Let $v \notin \mathcal{N}(A)$, then v is an eigenvector of A for the eigenvalue 1.
(d) $\mathbb{R}^n = Col(A) + \mathcal{N}(A)$.

Ans. ***Options a, b, and d***

(a) Let $v \in \mathcal{N}(A)$. Then $Av = 0 = 0.v$ implies that $\mathcal{N}(A)$ is the eigenspace of A corresponding to the eigenvalue 0.
(b) Let $v \neq 0 \in Col(A)$, then $v = Au$ for some $u \neq 0 \in \mathbb{R}^n$. Now,

$$Av = A(Au) = A^2u = Au = v$$

That is, v is an eigenvector of A corresponding to the eigenvalue 1.
(c) Consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Clearly, $A^2 = A$. Let $v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. As $Av = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $v \notin \mathcal{N}(A)$. Also, v is not an eigenvector of A corresponding to the eigenvalue 1.
(d) Let $v \in Col(A) \cap \mathcal{N}(A)$. Then $Av = 0$ and there exists $u \in \mathbb{R}^n$ such that $Au = v$. Now,

$$v = Au = A^2u = A(Au) = Av = 0$$

That is, $Col(A) \cap \mathcal{N}(A) = \{0\}$. Then by Rank-Nullity theorem, $\mathbb{R}^n = Col(A) + \mathcal{N}(A)$.

(132) Let A be a 100×100 matrix such that $a_{ij} = \begin{cases} i, \; if \; i+j = 101 \\ 0, \; otherwise \end{cases}$. Which of the following statements are true about A?

(a) A is similar to a diagonal matrix over $\mathbb{R}$.
(b) A is not similar to a diagonal matrix over $\mathbb{C}$.
(c) One of the eigenvalues of A is 10.
(d) None of the eigenvalues of A exceeds 51.

Ans. ***Options a, c, and d***

The matrix $A = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 2 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 99 & \dots & 0 & 0 \\ 100 & 0 & \dots & 0 & 0 \end{bmatrix}$. *Then,* $A^2 = \begin{bmatrix} 1.100 & 0 & \dots & 0 & 0 \\ 0 & 2.99 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 2.99 & 0 \\ 0 & 0 & \dots & 0 & 1.100 \end{bmatrix}$ *is a diagonal matrix. Thus, the eigenvalues of A^2 are $1.100, 2.99, \dots, 50.51$ each repeating twice. Therefore, the possible eigenvalues of A are $\pm\sqrt{1.100}$, $\pm\sqrt{2.99}, \dots, \pm\sqrt{50.51}$. As $trace(A) = 0$, $\pm\sqrt{1.100}, \pm\sqrt{2.99}, \dots, \pm\sqrt{50.51}$ are the eigenvalues of A and hence A is diagonalizable. Also, 10 is an eigenvalue of A and none of the eigenvalues of A exceeds 51.*

(133) Let $n \geq 1$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha \neq \beta$. Suppose $A_n(\alpha, \beta) = [a_{ij}]_{n \times n}$ be such that

$$a_{ij} = \begin{cases} \alpha, \; if \; i = j \\ 0, \; otherwise \end{cases}$$

Let D_n denote the determinant of $A_n(\alpha, \beta)$. Which of the following statements are true?

(a) $D_n = (\alpha - \beta)D_{n-1} + \beta$ for $n \geq 2$.
(b) $\frac{D_n}{(\alpha-\beta)^{n-1}} = \frac{D_{n-1}}{(\alpha-\beta)^{n-2}} + \beta$ for $n \geq 2$.
(c) $D_n = (\alpha + (n-1)\beta)^{n-1}(\alpha - \beta)$ for $n \geq 2$.
(d) $D_n = (\alpha + (n-1)\beta)(\alpha - \beta)^{n-1}$ for $n \geq 2$.

Ans. *The matrix*

$$A_n(\alpha, \beta) = \begin{bmatrix} \alpha & \beta & \dots & \beta \\ \beta & \alpha & \dots & \beta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \dots & \alpha \end{bmatrix} = \begin{bmatrix} \beta & \beta & \dots & \beta \\ \beta & \beta & \dots & \beta \\ \vdots & \vdots & \ddots & \vdots \\ \beta & \beta & \dots & \beta \end{bmatrix} + \begin{bmatrix} \alpha - \beta & 0 & \dots & 0 \\ 0 & \alpha - \beta & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha - \beta \end{bmatrix}.$$

That is, $A_n(\alpha, \beta) = B + (\alpha - \beta)I$, where B is the $n \times n$ matrix with all entries equal to β. As B is a symmetric matrix, it is diagonalizable. Hence, $Rank(B)$ is the number of non-zero eigenvalues of B. As the sum of each row of B is $n\beta$, the eigenvalues of B are $n\beta$ and 0 with multiplicities 1 and $n-1$, respectively. Therefore, the eigenvalues of $A_n(\alpha, \beta)$ are $\alpha + (n-1)\beta$ and $\alpha - \beta$ with multiplicities 1 and $n-1$, respectively. Thus, for $n \geq 2$

$$D_n = (\alpha + (n-1)\beta)(\alpha - \beta)^{n-1}$$

and

$$\frac{D_n}{(\alpha - \beta)^{n-1}} = \frac{D_{n-1}}{(\alpha - \beta)^{n-2}} + \beta$$

Options b and d

(134) Let $A \in \mathbb{M}_2(\mathbb{R})$. Which of the following statements are true?

(a) If $(tr(A))^2 > 4det(A)$, then A is diagonalizable over $\mathbb{R}$.
(b) If $(tr(A))^2 = 4det(A)$, then A is diagonalizable over $\mathbb{R}$.
(c) If $(tr(A))^2 < 4det(A)$, then A is diagonalizable over $\mathbb{R}$.

Ans. ***Option a***
The characteristic polynomial of a 2×2 matrix is $\lambda^2 - (tr(A))\lambda + det(A) = 0$. The given system has distinct roots when $(tr(A))^2 - 4det(A) > 0$. Therefore if $(tr(A))^2 > 4det(A)$, then A is diagonalizable over $\mathbb{R}$.

(135) If A and B are $n \times n$ matrices with real entries, then which of the following is/are TRUE?

(a) If $P^{-1}AP$ is diagonal for some real invertible matrix P, then there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors of A.
(b) If A is diagonal with distinct entries and $AB = BA$, then B is also diagonal.
(c) If A^2 is diagonal, then A is diagonal.
(d) If A is diagonal and $AB = BA$ for all B, then $AB = \lambda I$ for some $\lambda \in \mathbb{R}$.

Ans. ***Options a, b, and d***

(a) *$P^{-1}AP$ is diagonal implies that A is diagonalizable. Therefore $\mathbb{R}^n$ has a basis consisting of eigenvectors of A.*

(b) *Let $A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$ with $a_{11} \neq a_{22}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$. Then*

$$AB = BA \Rightarrow \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} \\ a_{22}b_{21} & a_{22}b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{22}b_{12} \\ a_{11}b_{21} & a_{22}b_{22} \end{bmatrix} \Rightarrow b_{12} = b_{21} = 0$$

since $a_{11} \neq a_{22}$. Thus, B is diagonal.

(c) *Let $A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$, then $A^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. But A is not diagonal.*

(d) *Let $A = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$. Since A commutes with every matrix, it commutes with $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Now*

$$\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \Rightarrow a_{11} = a_{22}$$

(136) Let $A \in \mathbb{M}_{n\times n}(\mathbb{C})$. Then $\begin{pmatrix} A & A \\ 0 & A \end{pmatrix}$ is diagonalizable if and only if

(a) $A = 0$ (b) $A = I$ (c) $n = 2$ (d) None of the above.

Ans. ***Option a***

Let $A = [1]$. Consider the matrix $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then B is not diagonalizable. So option (b) is false. Consider the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then clearly the matrix $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is not diagonalizable. The characteristic polynomial of the given matrix is the square of the characteristic polynomial of A, Then for the matrix to be diagonalizable, A must be equal to zero matrix for otherwise the minimal polynomial won't be of linear factors.

(137) Consider $\mathbb{M}_{n\times n}(\mathbb{R})$. Which of the following are true for every $n \geq 2$?

(a) there exists $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$ such that $AB - BA = I$.
(b) if $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$ and $AB = BA$, then A is diagonalizable over $\mathbb{R}$ if and only if B is diagonalizable over $\mathbb{R}$.
(c) if $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$, then AB and BA have the same minimal polynomial.
(d) if $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$, then AB and BA have the same eigenvalues in $\mathbb{R}$.

Ans. ***Option d***

(a) *Suppose there exists $A, B \in \mathbb{M}_{n\times n}(\mathbb{R})$ such that $AB - BA = I$. Taking trace on both sides,*

$$tr(AB - BA) = tr(I_n) = n$$

But $tr(AB - BA) = tr(AB) - tr(BA) = 0$. Thus our assumption is incorrect.

(b) *Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Then $AB = BA = B$. A is diagonalizable but B is not diagonalizable.*

(c) *Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Then $AB = A$ and $BA = 0$. The minimal polynomial for A is λ^2 and for AB the minimal polynomial is λ.*

(d) *Let λ be an eigenvalue of AB. Then there exists $v \neq 0$ such that $(AB)(v) = \lambda v$. Let $Bv = w$, then*

$$(BA)(w) = (BA)(Bv) = B[(AB)v] = \lambda Bv = \lambda w$$

Thus, λ is an eigenvalue of BA.

(138) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation whose matrix with respect to the standard basis $\{e_1, e_2, e_3\}$ of $\mathbb{R}^3$ is $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Then T

(a) maps the subspaces spanned by e_1 and e_3 into itself
(b) has distinct eigenvalues
(c) has eigenvectors that span $\mathbb{R}^3$
(d) has a non-zero null space.

Ans. ***Options a and c***

(a) *Since*

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$span\{e_1, e_3\}$ is mapped onto itself.

(b) *The characteristic polynomial of the given matrix is $(\lambda - 1)^2(\lambda + 1)$. So it does not have distinct eigenvalues.*

(c) *Since the minimal polynomial of the given matrix is $(\lambda - 1)(\lambda + 1)$ (linear factors only), the matrix is diagonalizable and hence it has eigenvectors that span $\mathbb{R}^3$.*

(d) *As the given matrix has rank 3 by Rank-Nullity Theorem, $Nullity(T) = 0$.*

(139) Let $T_1 : \mathbb{R}^n \to \mathbb{R}^n$ be given by $v \mapsto \alpha v$ for a fixed $\alpha \in \mathbb{R}, \alpha \neq 0$. Let $T_2 : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation such that $B = \{v_1, v_2, \ldots, v_n\}$ is a set of linearly independent eigenvectors of T. Then

(a) The matrix of T_2 with respect to B is diagonal.
(b) The matrix of $(T_2 - T_1)$ with respect to B is diagonal.
(c) The matrix of T_2 with respect to B is not necessarily diagonal, but upper triangular.
(d) The matrix of T_2 with respect to B is diagonal, but the matrix of $(T_2 - T_1)$ with respect to B is not diagonal.

Ans. ***Options a and b***
Since $T_1(v) = \alpha v$ for all $v \in \mathbb{R}^n$, the matrix of $T_1 = \alpha I$. Since T_2 has n linearly independent eigenvectors, the matrix of T_2 with respect to B is diagonal (Corollary 4.6). Also the matrix of $(T_2 - T_1)$ with respect to B is diagonal (see Question 134).

(140) Let $A \in M_{2\times 2}(\mathbb{R})$ be of trace 2 and determinant -3. Identifying $M_{2\times 2}(\mathbb{R})$ with $\mathbb{R}^4$, consider the linear transformation $T : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ defined by $T(B) = AB$. Then which of the following statements are true?
(a) T is diagonalizable. (b) 2 is an eigenvalue of T.
(c) T is invertible. (d) $T(B) = B$ for some $0 \neq B$ in $M_2(\mathbb{R})$.

Ans. ***Options a and c***
Take $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *Then $a + d = 2$ and $ad - bc = -3$. Now*

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & a \\ 0 & c \end{bmatrix} = 0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} b & 0 \\ d & 0 \end{bmatrix} = b\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} a & 0 \\ c & 0 \end{bmatrix} = 0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 0\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then the matrix of T is $\begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix}$.

(a) *As* $trace(A) = 2$ *and* $det(A) = -3$*, the eigenvalues of* A *are* -1 *and* 3*. Let* $V_1 = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ *and* $V_2 = \begin{bmatrix} v_3 \\ v_4 \end{bmatrix}$ *be the respective eigenvectors of* A*. That is,*

$$AV_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} av_1 + bv_2 \\ cv_1 + dv_2 \end{bmatrix} = (-1) \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

and

$$AV_2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} av_3 + bv_4 \\ cv_3 + dv_4 \end{bmatrix} = 3 \begin{bmatrix} v_3 \\ v_4 \end{bmatrix}$$

From this, we can see that $u_1 = \begin{bmatrix} v_1 \\ 0 \\ v_2 \\ 0 \end{bmatrix}$ *and* $u_2 = \begin{bmatrix} 0 \\ v_1 \\ 0 \\ v_2 \end{bmatrix}$ *are eigenvectors of matrix of* T *corresponding to the eigenvalue* -1*. Also,* $u_3 = \begin{bmatrix} v_3 \\ 0 \\ v_4 \\ 0 \end{bmatrix}$ *and* $u_4 = \begin{bmatrix} 0 \\ v_3 \\ 0 \\ v_4 \end{bmatrix}$ *are eigenvectors of matrix of* T *corresponding to the eigenvalue* 3*. As* V_1 *and* V_2 *are linearly independent eigenvectors of* A*, the set* $\{u_1, u_2, u_3, u_4\}$ *is linearly independent and hence forms a basis for* $\mathbb{R}^4$*. Therefore,* T *is diagonalizable.*

(b) *The eigenvalues of* T *are the eigenvalues of* A *repeated twice. Clearly,* 2 *is not an eigenvalue of* T*.*

(c) *As* 0 *is not an eigenvalue of* T*,* T *is invertible.*

(d) *Suppose there exists* B *such that* $T(B) = AB = B$*. Then* $det(AB) = det(B) \Rightarrow det(A) = 1$*, which is a contradiction.*
Observe that the properties of T *are exactly the same as that of the properties of* A*.*

(141) Suppose A is a real $n \times n$ matrix of rank r. Let V be the vector space of all $n \times n$ matrices X such that $AX = 0$. What is the dimension of V?
(a) r (b) nr (c) n^2r (d) $n^2 - nr$.

Ans. *Consider the map* $T : \mathbb{M}_{n\times n}(\mathbb{R}) \to \mathbb{M}_{n\times n}(\mathbb{R})$ *defined by* $T(B) = AB$*. Then* $V = \{X \in \mathbb{M}_n(\mathbb{R}) \mid AX = 0\}$ *is the null space of* T*. As* $Rank(A) = r$*, by Rank-Nullity Theorem,* $Nullity(A) = n - r$ *and hence* $Nullity(T) = n(n - r) = n^2 - nr$*.*
Option d

(142) Consider the linear transformation $T : \mathbb{R}^7 \to \mathbb{R}^7$ defined by

$$T(x_1, x_2, \ldots, x_6, x_7) = (x_7, x_6, \ldots, x_2, x_1)$$

Which of the following statements must be true?

(a) the determinant of T is 1.
(b) there is a basis of T with respect to which T is a diagonal matrix.
(c) $T^7 = I$.
(d) The smallest n such that $T^n = I$ is even.

Ans. ***Options b and d***
The matrix of the given linear transformation T is

$$\begin{bmatrix} 0&0&0&0&0&0&1 \\ 0&0&0&0&0&1&0 \\ 0&0&0&0&1&0&0 \\ 0&0&0&1&0&0&0 \\ 0&0&1&0&0&0&0 \\ 0&1&0&0&0&0&0 \\ 1&0&0&0&0&0&0 \end{bmatrix}$$

Since $T^2 = I$, $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$ is an annihilating polynomial of T. As the minimal polynomial divides annihilating polynomial, minimal polynomial of T is of linear factors. Therefore T is diagonalizable (Theorem 4.15). That is, there is a basis of T with respect to which T is a diagonal matrix. Clearly $det(T) = -1$ and $T^7 = T$.

(143) Let A be a 3×3 matrix with real entries. Identify the correct statements.

(a) A is necessarily diagonalizable over $\mathbb{R}$.
(b) if A has distinct real eigenvalues then it is diagonalizable over $\mathbb{R}$.
(c) if A has distinct eigenvalues then it is diagonalizable over $\mathbb{C}$.
(d) if all eigenvalues of A are non-zero then it is diagonalizable over $\mathbb{C}$.

Ans. ***Options b and c***
Consider the matrix $A = \begin{bmatrix} 1&1&0 \\ 0&1&1 \\ 0&0&1 \end{bmatrix}$. Then all the eigenvalues of A are non-zero. But it is not diagonalizable over $\mathbb{R}$ or $\mathbb{C}$ as its minimal polynomial is $(\lambda - 1)^3$. If A has distinct real eigenvalues then it is diagonalizable over both $\mathbb{R}$ and $\mathbb{C}$.

(144) Which of the following statements is correct for every linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $T^3 - T^2 - T + I = 0$?

(a) T is invertible as well as diagonalizable.

(b) T is invertible but not necessarily diagonalizable.
(c) T is diagonalizable, but not necessarily invertible.
(d) None of the other three statements.

Ans. ***Option b***
The given matrix satisfies the polynomial $\lambda^3 - \lambda^2 - \lambda + 1 = 0$. *Since the constant term is not zero, T is invertible. Consider the matrix* $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. *The matrix satisfies the given polynomial equation and is not diagonalizable.*

(145) Let A be a 4×4 matrix with $dim\,[\mathcal{N}(A - 2I)] = 2$, $dim\,[\mathcal{N}(A - 4I)] = 1$ and $Rank(A) = 3$. Then
(a) 0, 2, and 4 are eigenvalues of A. (b) $det(A) = 0$.
(c) A is not diagonalizable. (d) $tr(A) = 8$.

Ans. ***Options a, b, and d***
Since $Rank(A) = 3$, 0 is an eigenvalue of A. Also, A has three non-zero eigenvalues. As $dim\,[\mathcal{N}(A - 2I)] = 2$ and $dim\,[\mathcal{N}(A - 4I)] = 1$, 2 is an eigenvalue with geometric multiplicity 2 and 4 is an eigenvalue with geometric multiplicity 1. Thus in the Jordan canonical form there must be 4 Jordan blocks. Therefore A is diagonalizable.

(146) Let $A \in \mathbb{M}_{3\times3}(\mathbb{R})$ and let $X = \{P \in GL_3(\mathbb{R}) \mid PAP^{-1}\, is\, triangular\}$. Then

(a) $X \neq \phi$
(b) If $X = \phi$, then A is not diagonalizable over $\mathbb{C}$.
(c) If $X = \phi$, then A is diagonalizable over $\mathbb{C}$.
(d) If $X = \phi$, then A has no real eigenvalue.

Ans. ***Option c***
Since every odd degree matrix with real coefficient has at least one real root, A has at least one real eigenvalue and the complex eigenvalues occur as conjugate pairs. Thus X can be empty if A has complex eigenvalues and if $X = \phi$, then A is diagonalizable over $\mathbb{C}$.

(147) Which of the following matrices have Jordan canonical form equal to $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$?

(a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

Ans. ***Options a, b, and c***

The characteristic polynomial of the given matrix is λ^3 *and minimal polynomial is* λ^2*. Matrices* (a), (b)*, and* (c) *also have the same characteristic and minimal polynomials as that of the given matrix.* λ^3 *is both characteristic and minimal polynomial for matrix* (d).

(148) Let $T : \mathbb{R}^4 \to \mathbb{R}^4$ be a linear transformation with characteristic polynomial $(\lambda - 2)^4$ and minimal polynomial $(\lambda - 2)^2$. Jordan canonical form of T can be

$$(a) \begin{pmatrix} 2&0&0&0\\ 1&2&0&0\\ 0&0&2&0\\ 0&0&1&2 \end{pmatrix} \quad (b) \begin{pmatrix} 2&0&0&0\\ 0&2&0&0\\ 0&0&2&0\\ 0&0&1&2 \end{pmatrix} \quad (c) \begin{pmatrix} 2&0&0&0\\ 0&2&0&0\\ 0&0&2&0\\ 0&0&0&2 \end{pmatrix} \quad (d) \begin{pmatrix} 2&0&0&0\\ 1&2&0&0\\ 0&1&2&0\\ 0&0&0&2 \end{pmatrix}.$$

Ans. ***Options a and b***

Since the minimal polynomial has a factor of degree 2, it has at least one Jordan block of order 2.

(149) Let A be a 6×6 matrix over $\mathbb{R}$ with characteristic polynomial $= (\lambda - 3)^2(\lambda - 2)^4$ and minimal polynomial $= (\lambda - 3)(\lambda - 2)^2$. Then Jordan canonical form of A can be

$$(a) \begin{pmatrix} 3&0&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&1&0\\ 0&0&0&0&2&1\\ 0&0&0&0&0&2 \end{pmatrix} \quad (b) \begin{pmatrix} 3&0&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&2&0\\ 0&0&0&0&0&2 \end{pmatrix} \quad (c) \begin{pmatrix} 3&0&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&2&1\\ 0&0&0&0&0&2 \end{pmatrix} \quad (d) \begin{pmatrix} 3&1&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&2&1\\ 0&0&0&0&0&2 \end{pmatrix}.$$

Ans. ***Options b and c***

The given matrix has two eigenvalues 3 and 2. The algebraic multiplicity of $\lambda_1 = 3$ *is 2 and geometric multiplicity is 1, and algebraic multiplicity of* $\lambda_2 = 2$ *is 4 and geometric multiplicity is 2. Therefore there exists at least one Jordan block of order 2 for* $\lambda_2 = 2$ *and the possible Jordan forms are*

$$\begin{pmatrix} 3&0&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&2&0\\ 0&0&0&0&0&2 \end{pmatrix} \textit{ and } \begin{pmatrix} 3&0&0&0&0&0\\ 0&3&0&0&0&0\\ 0&0&2&1&0&0\\ 0&0&0&2&0&0\\ 0&0&0&0&2&1\\ 0&0&0&0&0&2 \end{pmatrix}$$

(150) Which of the following matrices is not diagonalizable over $\mathbb{R}$?

$$(a) \begin{bmatrix} 1&1&0\\ 0&2&0\\ 0&0&1 \end{bmatrix} \quad (b) \begin{bmatrix} 1&1&0\\ 0&2&1\\ 0&0&3 \end{bmatrix} \quad (c) \begin{bmatrix} 1&1&0\\ 0&1&0\\ 0&0&2 \end{bmatrix} \quad (d) \begin{bmatrix} 1&0&1\\ 0&2&0\\ 0&0&3 \end{bmatrix}.$$

Ans. ***Option c***

(a) The characteristic polynomial of the given matrix is $(\lambda-1)^2(\lambda-2)$ *and the minimal polynomial is* $(\lambda-1)(\lambda-2)$ *and hence it is diagonalizable.*
(b) The given matrix is an upper triangular matrix with distinct eigenvalues (diagonal elements are the eigenvalues) and hence it is diagonalizable.
(c) The given matrix contains a Jordan block of order 2 and hence is not diagonalizable.
(d) The given matrix is an upper triangular matrix with distinct eigenvalues (diagonal elements are the eigenvalues) and hence it is diagonalizable.

(151) Let $A = \begin{bmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{bmatrix}$. Then the Jordan Canonical form of A is

(a) $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$ (b) $\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$.

Ans. ***Option a***

Since determinant of the given matrix is 4 and trace is 0, $\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$ *is the required Jordan Canonical form.*

(152) Consider the matrices $A = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

(a) A and B are similar over the field of rational numbers $\mathbb{Q}$.
(b) A is diagonalizable over the field of rational numbers $\mathbb{Q}$.
(c) B is the Jordan canonical form of A.
(d) The minimal polynomial and the characteristic polynomial of A are the same.

Ans. ***Options a, c, and d***
The characteristic and minimal polynomial of A and B is $(\lambda-2)^2(\lambda-3)$. *Clearly it factors over* $\mathbb{Q}$. *Since they have the same minimal and characteristic polynomial, A and B are similar over the field of rational numbers* $\mathbb{Q}$. *A is not diagonalizable since the minimal polynomial has non-linear factors. Since the minimal polynomial is* $(\lambda-2)^2(\lambda-3)$, *B is the Jordan canonical form of A.*

(153) Let A be a 7×7 matrix such that $2A^2 - A^4 = I$. If A has two distinct eigenvalues and each eigenvalue has geometric multiplicity 3, then the total number of non-zero entries in the Jordan canonical form of A equals ....

Ans. *Since A satisfies* $A^4 - 2A^2 + I = 0$, $\lambda^4 - 2\lambda^2 + 1 = (\lambda^2 - 1)^2 = (\lambda + 1)^2(\lambda - 1)^2$ *is an annihilating polynomial of A. Therefore the two distinct eigenvalues of A are 1 and* -1. *Since geometric multiplicity of each eigenvalue is 3, each eigenvalue has 3 Jordan blocks and since A is of order 7, one of the Jordan blocks must be of order 2. Therefore the total number of non-zero entries in the Jordan canonical form of A is 8.*

(154) Let $D : \mathbb{P}_3[x] \to \mathbb{P}_3[x]$ be the linear operator given by differentiation with respect to x. Let A be the matrix representation of D with respect to some basis for $\mathbb{P}_3[x]$. Which of the following are true?

(a) A is a nilpotent matrix.
(b) A is a diagonalizable matrix.
(c) the rank of A is 2.
(d) the Jordan canonical form of A is $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

Ans. ***Options a and d***
Consider the standard ordered basis $\{1, x, x^2, x^3\}$ *for* V. *Now*

$$D(1) = 0 = 0 + 0x + 0x^2 + 0x^3$$

$$D(x) = 1 = 1 + 0x + 0x^2 + 0x^3$$

$$D(x^2) = 2x = 0 + 2x + 0x^2 + 0x^3$$

$$D(x^3) = 3x^2 = 0 + 0x + 3x^2 + 0x^3$$

Then $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. *Clearly A is a nilpotent matrix as* $A^4 = 0$. *The characteristic and minimal polynomial of A are* λ^4. *Therefore A is not diagonalizable. The rank of A is 3 and the Jordan canonical form of A is* $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(155) Let $T : \mathbb{P}_2[x] \to \mathbb{P}_2[x]$ be the linear transformation given by

$$T(p) = 2p + p' \; for \; p \in V$$

where p' is the derivative of p. Then the number of non-zero entries in the Jordan canonical form of a matrix of T equals ....

Ans. *Consider the standard ordered basis for $\mathbb{P}_2[x]$. Then, we have*

$$T(1) = 2 = 2.1 + 0x + 0x^2$$

$$T(x) = 2x + 1 = 1.1 + 2x + 0x^2$$

$$T(x^2) = 2x^2 + 2x = 0.1 + 2x + 2x^2$$

Therefore the matrix of T is given by $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}$. *Since the minimal polynomial is $(\lambda - 2)^3$, the number of non-zero entries in the Jordan canonical form of a matrix of T is 5.*

(156) Let A be a complex 3×3 matrix with $A^3 = -I$. which of the following statements are correct?
(a) A has three distinct eigenvalues. (b) A is diagonalizable over $\mathbb{C}$.
(c) A is triangularizable over $\mathbb{C}$. (d) A is non-singular.

Ans. ***Options b, c, and d***
Since A satisfies $A^3 = -I$, $\lambda^3 + 1$ is an annihilating polynomial of A. Also

$$\lambda^3 + 1 = (\lambda + 1)(\lambda^2 - \lambda + 1) = (\lambda + 1)\left(\lambda - \left(\frac{1 + i\sqrt{3}}{2}\right)\right)\left(\lambda - \left(\frac{1 - i\sqrt{3}}{2}\right)\right)$$

As 0 is not an eigenvalue of A, it is non-singular. Since the minimal polynomial divides annihilating polynomial, minimal polynomial has linear factors only. Thus A is diagonalizable over $\mathbb{C}$. Also A is triangularizable over $\mathbb{C}$. Take $A = -I$, then $A^3 = -I$, but -1 is the only eigenvalue.

(157) Let A be a real matrix with characteristic polynomial $(\lambda - 1)^3$. Pick the correct statements from below

(a) A is necessarily diagonalizable.
(b) If the minimal polynomial of A is $(\lambda - 1)^3$, then A is diagonalizable.
(c) Characteristic polynomial of A^2 is $(\lambda - 1)^3$.
(d) If A has exactly two Jordan blocks, then $(A - I)^2$ is diagonalizable.

Ans. ***Options c and d***
Consider the matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$. *Then the characteristic and minimal polynomial of A are $(\lambda - 1)^3$. Clearly A is not diagonalizable. Since 1 is the only eigenvalue of A, A^2 has only one eigenvalue 1 with algebraic multiplicity 3. Therefore characteristic polynomial of A^2 is $(\lambda - 1)^3$. Since A has two Jordan blocks and 1 is the only eigenvalue of A implying $(A - I)^2 = 0$, therefore $(A - I)^2$ is diagonalizable.*

(158) Let A be a 4×4 matrix over $\mathbb{C}$ such that $Rank(A) = 2$ and $A^3 = A^2 \neq 0$. Suppose that A is not diagonalizable. Then

(a) One of the Jordan blocks of the Jordan Canonical form of A is $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.
(b) $A^2 = A \neq 0$.
(c) There exists a vector v such that $Av \neq 0$ but $A^2 v = 0$.
(d) The characteristic polynomial of A is $\lambda^4 - \lambda^3$.

Ans. ***Options a, c, and d***
Since A satisfies $A^3 = A^2$, $\lambda^3 - \lambda^2 = \lambda^2(\lambda - 1)$ *is an annihilating polynomial of A. Therefore the possible minimal polynomials of A are* $\lambda, \lambda^2, \lambda - 1, \lambda(\lambda - 1)$, *and* $\lambda^2(\lambda - 1)$. *Since* $A^2 \neq 0$, x^2 *is not the minimal polynomial of A. As A is not diagonalizable* $\lambda, \lambda - 1$, *and* $\lambda(\lambda - 1)$ *cannot be the minimal polynomial of A. Therefore* $\lambda^2(\lambda - 1)$ *is its minimal polynomial and hence one of the Jordan blocks of the Jordan Canonical form of A is* $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *and A does not satisfy* $A^2 = A$. *For if* $A^2 = A$, $\lambda^2 - \lambda = \lambda(\lambda - 1)$ *is its annihilating polynomial and hence diagonalizable as it has linear factors only. Since* $A^3 = A^2$, $A^2(A - I)v = 0$ *for every v. But* $A(A - I)v \neq 0$ *for every v as* $A^2 \neq A$. *So there exists a vector v such that* $Av \neq 0$ *but* $A^2 v = 0$. *The possible characteristic polynomials are* $\lambda^2(\lambda - 1)^2$ *and* $\lambda^3(\lambda - 1)$. *But* $\lambda^2(\lambda - 1)^2$ *is not possible since* $Rank(A) = 2$ *and the minimal polynomial is* $\lambda^2(\lambda - 1)$. *Therefore the characteristic polynomial of A is* $\lambda^4 - \lambda^3$.

(159) For an $n \times n$ real matrix A, $\lambda \in \mathbb{R}$ and a non-zero vector $v \in \mathbb{R}^n$ suppose that $(A - \lambda I)^k v = 0$ for some positive integer k. Then which of the following is/are always true?
(a) $(A - \lambda I)^{k+r} v = 0, \ \forall r \in \mathbb{R}^+$. (b) $(A - \lambda I)^{k-1} v = 0$.
(c) $(A - \lambda I)$ is not injective. (d) λ is an eigenvalue of A.

Ans. ***Options a, c, and d***

(a) Since $(A - \lambda I)^k v = 0$, $(A - \lambda I)^{k+r} v = (A - \lambda I)^k (A - \lambda I)^r v = 0$ *for all positive integers r.*
(b) Consider $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. *Take* $\lambda = 1$. *Then* $(A - \lambda I)^2 v = 0$ *for every non-zero vector* $v \in \mathbb{R}^n$, *but* $(A - \lambda I)v \neq 0$ *for every non-zero vector* $v \in \mathbb{R}^n$.
(c) We know that there exists a non-zero vector $v \in \mathbb{R}^n$ *such that* $(A - \lambda I)^k v = 0$ *for some positive integer k. Now*

$$(A - \lambda I)^k v = 0 \Rightarrow det(A - \lambda I)^k = 0 \Rightarrow det(A - \lambda I) = 0$$

Therefore there exists a non-zero vector $v \in \mathbb{R}^n$ *such that* $(A - \lambda I)v = 0$ *and hence* $(A - \lambda I)$ *is not injective.*
(d) From option c, λ *is an eigenvalue of A.*

(160) Check whether the following statements are true or false.

(a) Let A be an $n \times n$ matrix whose row sums equal 1. Then for any positive integer m the row sums of the matrix A^m equal 1.

(b) Let A be a 2×2 matrix with complex entries. The number of 2×2 matrices A with complex entries satisfying the equation $A^3 = A$ is infinite.

(c) The matrix $\begin{bmatrix} 1 & \pi & 3 \\ 0 & 2 & 4 \\ 0 & 0 & 3 \end{bmatrix}$ is diagonalizable.

(d) If A and B are similar matrices then every eigenvector of A is an eigenvector of B.

(e) Let V be the subspace of the real vector space of real-valued functions on $\mathbb{R}$, spanned by $cos\ x$ and $sin\ x$. Let $D : V \to V$ be the linear map sending $f(x) \in V$ to $\frac{df(x)}{dx}$. Then D has a real eigenvalue.

(f) Any linear transformation $A : \mathbb{R}^4 \to \mathbb{R}^4$ has a proper non-zero invariant subspace.

(g) Let $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ be upper triangular with all diagonal entries 1 such that $A \neq I$. Then A is not diagonalizable.

(h) If A is a 2×2 complex matrix that is invertible and diagonalizable, and such that A and A^2 have the same characteristic polynomial, then A is the 2×2 identity matrix.

(i) The matrices $\begin{pmatrix} 0 & i & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 & 0 \\ -i & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$ are similar.

(j) For any matrix C with entries in $\mathbb{C}$, let $m(C)$ denote the minimal polynomial of C, and $p(C)$ its characteristic polynomial. Then for any $n \in \mathbb{N}$, two matrices $A, B \in \mathbb{M}_{n\times n}(\mathbb{C})$ are similar if and only if $m(A) = m(B)$ and $p(A) = p(B)$.

(k) Let $A, B \in \mathbb{M}_{3\times 3}(\mathbb{R})$. Then $det(AB - BA) = \dfrac{tr[(AB - BA)^3]}{3}$.

(l) For any $n \geq 2$, there exists $n \times n$ real matrix A such that the set $\{A^p \mid p \geq 1\}$ spans the $\mathbb{R}$- vector space $\mathbb{M}_{n\times n}(\mathbb{R})$.

(m) The matrices $\begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$ and $\begin{pmatrix} x & 1 \\ 0 & y \end{pmatrix}$, $x \neq y$ for any $x, y \in \mathbb{R}$ are conjugate/similar in $\mathbb{M}_{2\times 2}(\mathbb{R})$.

Ans. *(a) True. We will prove this by induction on m. Let A be an $n \times n$ matrix whose row sums equal 1. Then 1 is an eigenvalue of A with the eigenvector* $\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T_{n\times 1}$. *Now*

$$A \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} \Rightarrow A^2 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} = A \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1}$$

Thus the statement is true for $m = 2$. Now suppose that the statement is true for $m = k-1$. That is, the row sum of A^{k-1} is 1. Then as above,

$$A^{k-1}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times 1} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times 1} \Rightarrow A^{k}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times 1} = A\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times 1} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}_{n\times 1}$$

Thus the statement is true for every positive integer m.

(b) *True. Since A satisfies $A^3 = A$,*

$$\lambda^3 - \lambda = \lambda(\lambda-1)(\lambda+1)$$

is an annihilating polynomial of A. Then the possible minimal polynomials are

$$\lambda, \lambda-1, \lambda+1, \lambda(\lambda-1), \lambda(\lambda+1) \textit{ and } (\lambda-1)(\lambda+1)$$

If the minimal polynomial is λ, then $A = 0$. If the minimal polynomial is $\lambda - 1$, then $A = I$ and if the minimal polynomial is $\lambda + 1$, then $A = -I$. Now if the minimal polynomial is $\lambda(\lambda-1)$, any matrix with $trace = -1$ and $determinant = 0$ satisfies this. So there exist infinitely many matrices such that $A^3 = A$.

(c) *True. The eigenvalues of an upper triangular matrix are its diagonal entries. So the given matrix has three distinct eigenvalues. Therefore the matrix is diagonalizable.*

(d) *False. Consider the matrix $A = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$. Take $P = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\\ 1 & 1\end{bmatrix}$. Then*

$$PAP^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix}-1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\ -1 & -1\end{bmatrix} = \begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix} = B$$

Also $A\begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}1\\ 1\end{bmatrix}$.Therefore $\begin{bmatrix}1 & 1\end{bmatrix}^t$ is an eigenvector of A. But $\begin{bmatrix}1 & 1\end{bmatrix}^t$ is not an eigenvector of B as $B\begin{bmatrix}1\\ 1\end{bmatrix} = \begin{bmatrix}-1\\ 1\end{bmatrix}$.

(e) *False. We have*

$$D(sin\ x) = cos\ x = 0(sin\ x) + 1(cos\ x)$$

and

$$D(cos\ x) = -sin\ x = (-1)(sin\ x) + 0(cos\ x)$$

Then the matrix of T is given by $\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$ and it has no real eigenvalue.

(f) *True. If T has a real eigenvalue, then clearly T has a proper non-zero invariant subspace. Suppose T has complex eigenvalues only, then the complex eigenvalues occur as conjugate pairs. Then consider the real Jordan form of T. Then we get that any linear transformation $A : \mathbb{R}^4 \to \mathbb{R}^4$ has a proper non-zero invariant subspace.*

(g) *True. Since A is an upper triangular with all diagonal entries 1, its characteristic polynomial is given by $(\lambda - 1)^n$ and its minimal polynomial is of the form $(\lambda - 1)^r$ where $r \leq n$. Since $A \neq I$, $r \geq 2$. Therefore A is not diagonalizable.*

(h) *False. Consider the matrix $A = \begin{bmatrix} \omega & 0 \\ 0 & \omega^2 \end{bmatrix}$ where $\omega \in C_3$ where C_3 denotes the cube root of unity. Then A is invertible and diagonalizable. Also A and A^2 have the same characteristic polynomial.*

(i) *True. $n \times n$ matrices having the same characteristic and minimal polynomial are similar when $n \leq 3$.*

(j) *False. Let A and B be two matrices of order $n \leq 3$; A and B are similar if and only if they have the same minimal and characteristic polynomial. For $n \geq 3$, this is not true. For example, consider the matrices $\begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{bmatrix}$ and $\begin{bmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{bmatrix}$. They have the same minimal and characteristic polynomial. but they are not similar.*

(k) *True. Consider the characteristic equation of the matrix $(AB - BA)$. It will be of the form*

$$\lambda^3 - tr(AB - BA)\lambda^2 + c\lambda - det(AB - BA) = 0$$

where c is some real number. Since $tr(AB - BA) = tr(AB) - tr(BA) = 0$, this implies $\lambda^3 + c\lambda - det(AB - BA) = 0$. Then by Cayley–Hamilton Theorem,

$$det(AB - BA)I = (AB - BA)^3 + c(AB - BA)$$

Taking trace on both sides, we get

$$3det(AB - BA) = tr[(AB - BA)^3]$$

(l) *False. By Cayley–Hamilton theorem any matrix of power* $\geq n$ *can be written as a linear combination of* $\{I, A, A^2, \ldots A^{n-1}\}$. *So the dimension of span of* $\{A^p \mid p \geq 1\}$ *is less than or equal to n. But the vector space of all* $n \times n$ *real matrices has dimension* n^2. *So the set* $\{A^p \mid p \geq 1\}$ *does not span* $\mathbb{M}_{n\times n}(\mathbb{R})$.

(m) *False. Since the given matrices have the same characteristic and minimal polynomial, they are similar.*

Chapter 12
Solved Problems—Normed Spaces and Inner Product Spaces

(1) Consider the vector space V of real polynomials of degree less than or equal to n. Fix distinct real numbers $a_0, a_1, \ldots, a_k$. For $p \in V$, $max\{|p(a_j)| : 0 \leq j \leq k\}$ defines a norm on V
(a) only if $k < n$ (b) only if $k \geq n$ (c) if $k + 1 \leq n$ (d) if $k \geq n + 1$

Ans. ***Options b and d***
Given that $\|p\| = max\{|p(a_j)| : 0 \leq j \leq k\}$.
Clearly $\|p\| \geq 0$.

$$\|p\| = 0 \Rightarrow max\{|p(a_j)| : 0 \leq j \leq k\} = 0$$
$$\Rightarrow |p(a_j)| = 0,\ 0 \leq j \leq k$$

This implies $p = 0$ *only if* $k \geq n$ *or if* $k \geq n + 1$ *as* $\{a_0, a_1, \ldots, a_k\}$ *is a collection of* $k + 1$ *elements and by Fundamental Theorem of Algebra* $p = 0$ *if it has greater than n zeros.*

$$\|kp\| = max\{|kp(a_j)| : 0 \leq j \leq k\}$$
$$= |k| max\{|p(a_j)| : 0 \leq j \leq k\} = |k|\,\|p\|$$

Also

$$\|p + q\| = max\{|p(a_j) + q(a_j)| : 0 \leq j \leq k\}$$
$$\leq max\{|p(a_j)| : 0 \leq j \leq k\} + max\{|q(a_j)| : 0 \leq j \leq k\} = \|p\| + \|q\|$$

(2) Consider the real vector space V of polynomials of degree less than or equal to d. For $p \in V$ define $\|p\|_k = max\{|p(0)|, |p'(0)| \ldots, |p^{(k)}(0)|\}$ where $p^{(i)}(0)$ is the ith derivative of p evaluated at 0. Then $\|p\|_k$ defines a norm on V if and only if
(a) $k \geq d - 1$ (b) $k < d$ (c) $k \geq d$ (d) $k < d - 1$

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_12

Ans. ***Option c***
Clearly $\|p\|_k \geq 0$ *since* $|p^{(i)}(0)| \geq 0$ *for all* i.

$$\|p\|_k = 0 \Rightarrow |p^{(i)}(0)| = 0 \; \forall \; i = 1, 2, \ldots, k$$
$$\Rightarrow p = 0 \; only \; if \; k \geq d$$

This is because $p^{(i)}(0)$ *is the co-efficient of* x_i.

(3) Consider the following statements:

P: Let V be any normed space. Then

$$|\|u\| - \|v\|| \leq \|u - v\|, \; \forall \; u, v \in V$$

Q: For any $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$,

$$\|v\|_\infty \leq \|v\|_2 \leq \|v\|_1 \leq n \, \|v\|_\infty$$

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option a***
For $u, v \in V$, *we have*

$$\|v\| = \|v - u + u\| \leq \|v - u\| + \|u\|$$

This implies that $\|v\| - \|u\| \leq \|v - u\|$. *Similarly, we can show that* $\|u\| - \|v\| = -[\|v\| - \|u\|] \leq \|v - u\|$. *Therefore*

$$|\|u\| - \|v\|| \leq \|u - v\|, \; \forall \; u, v \in V$$

Now for any $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$,

$$\|v\|_\infty = \sup_i |v_i| = \sup_i \left(|v_i|^2\right)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} |v_i|^2\right)^{\frac{1}{2}} = \|v\|_2$$

Also,

$$\|v\|_2^2 = \sum_{i=1}^{n} |v_i|^2 \leq \left(\sum_{i=1}^{n} |v_i|\right)^2 \leq \|v\|_1^2$$

and

$$\|v\|_1 = \sum_{i=1}^{n} |v_i| \le \sum_{i=1}^{n} \sup_i |v_i| = n \sup_i |v_i| = n \, \|v\|_\infty$$

Therefore $\|v\|_\infty \le \|v\|_2 \le \|v\|_1 \le n \, \|v\|_\infty$.

(4) Which of the following statements about the spaces l^p and $L^p[0, 1]$ is true?
(a) $l^3 \subset l^7$ and $L^6[0, 1] \subset L^9[0, 1]$ (b) $l^3 \subset l^7$ and $L^9[0, 1] \subset L^6[0, 1]$
(c) $l^7 \subset l^3$ and $L^6[0, 1] \subset L^9[0, 1]$ (d) $l^7 \subset l^3$ and $L^9[0, 1] \subset L^6[0, 1]$

Ans. ***Option b***
For $1 \le p \le q$, *we have* $l^p \subseteq l^q$ *and* $L^q \subseteq L^p$.

(5) The space $C[0, 1]$ of continuous functions on $[0, 1]$ is complete with respect to the norm
(a) $\|f\|_\infty = sup\{|f(x)| : x \in [0, 1]\}$
(b) $\|f\|_1 = \int_0^1 |f(x)| dx$
(c) $\|f\|_\infty^{0,1} = \|f\|_\infty + |f(0)| + |f(1)|$
(d) $\|f\|_2 = \sqrt{\int_0^1 |f(x)|^2 dx}$

Ans. ***Option a and c***

(a) Let $\{f_n\}$ *be a Cauchy sequence in* $C[0, 1]$. *Then, for any* $\epsilon > 0$, *there is an* N *such that*

$$\|f_m - f_n\| = sup\{|f_m(x) - f_n(x)| : x \in [0, 1]\} < \epsilon, \ \forall \, m, n > N \tag{12.1}$$

Then for a fixed x, *say* $x_0 \in [0, 1]$, *we have* $|f_m(x_0) - f_n(x_0)| < \epsilon, \ \forall \, m, n > N$. *Therefore* $\{f_n(x_0)\}$ *is a Cauchy sequence of real numbers for each* $x_0 \in [0, 1]$. *Since* $\mathbb{R}$ *is complete,* $\{f_n(x_0)\}$ *converges to a real number, which is unique. Thus we can define a function* f *on* $[0, 1]$, *pointwise. From Eq.* (12.1), *as* $n \to \infty$, *we have*

$$sup\{|f_m(x) - f(x)| : x \in [0, 1]\} \le \epsilon, \ \forall \, m > N$$

Therefore $|f_m(x) - f(x)| \le \epsilon, \ \forall \ m > N$. *This implies that* $\{f_n(x)\}$ *converges to* $f(x)$ *uniformly on* $[0, 1]$. *Since the convergence is uniform* f *is continuous. Hence,* $f \in C[0, 1]$. *Thus* $C[0, 1]$ *with the norm* $\|.\|_\infty$ *is complete.*

(b) Consider the sequence $\{f_n; n \ge 3\} \in C[0, 1]$, *where*

$$f_n(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2} - \frac{1}{n} \\ nx - \frac{n}{2} + 1, & \frac{1}{2} - \frac{1}{n} \le x \le \frac{1}{2} \\ 1, & \frac{1}{2} \le x \le 1 \end{cases}$$

Then for $m \ge n \ge 3$,

$$\|f_n - f_m\|_1 = \int_{\frac{1}{2}-\frac{1}{n}}^{\frac{1}{2}} |f_n(x) - f_m(x)|dx \le \frac{1}{n}$$

This implies that, as $n \to \infty$, $\|f_n - f_m\|_1 \to 0$. *Therefore* $\{f_n\}$ *is a Cauchy sequence in* $C[0, 1]$.
Suppose that $f_n \to f$ *in* $C[0, 1]$. *Fix* $\alpha \in \left(0, \frac{1}{2}\right)$. *Take* n *such that* $\frac{1}{2} - \frac{1}{n} \ge \alpha$. *Then*

$$0 \le \int_0^\alpha |f(x)|dx = \int_0^\alpha |f(x) - f_n(x)|dx \le \|f_n - f\|_1 \to 0$$

as $n \to \infty$. *Therefore* $\int_0^\alpha |f(x)|dx = 0$. *Since* $f \in C[0, 1]$, *this implies that* $f(x) = 0$ *for* $x \in [0, \alpha]$ *for every* $\alpha < \frac{1}{2}$. *Therefore* $f\left(\frac{1}{2}\right) = 0$. *Also, we have*

$$0 \le \int_{\frac{1}{2}}^1 |f(x) - 1|dx = \int_{\frac{1}{2}}^1 |f(x) - f_n(x)|dx \le \|f_n - f\|_1 \to 0$$

as $n \to \infty$. *Therefore* $\int_{\frac{1}{2}}^1 |f(x) - 1|dx = 0$ *and by the continuity of* f, $f(x) = 1$ *for* $x \in [\frac{1}{2}, 1]$. *This is a contradiction.*

(c) Using (a), *we can prove that* $C[0, 1]$ *is complete with respect to the norm* $\|.\|_\infty^{0,1}$.
(d) Using the sequence in (b) *we can prove that* $C[0, 1]$ *is incomplete with respect to the norm* $\|.\|_2$.

(6) If V is the class of all polynomials on [0, 1], then
(a) V is complete when given the sup norm.
(b) V is complete when given the L^1 norm.
(c) V is not complete under any norm
(d) V is complete when given the L^1 norm.

Ans. ***Option c***
We know that the Hamel basis of an infinite-dimensional Banach space must be uncountable. As $\{1, x, x^2, \ldots\}$ *is a countable Hamel basis for* V, V *is not complete with respect to any norm.*

(7) Which of the following is not complete in any norm?
(a) c_{00}, the space of all sequences of real numbers having finitely many non-zero terms.
(b) l^∞, the space of all bounded sequences of real numbers.
(c) $C[0, 1]$, the space of all real valued continuous functions on [0, 1].
(d) $\mathbb{R}^n$, where $\mathbb{R}$ is the field of real numbers.

Ans. ***Option a***

(a) *Consider the sequence* $v_n = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, 0, \ldots\right) \in c_{00}$. *For* $m > n$, *we have*

$$\|v_m - v_n\| = \left\| \left(0, \ldots, 0, \frac{1}{n+1}, \ldots, \frac{1}{n}, 0, \ldots\right) \right\| = \frac{1}{n+1}$$

Now for any $\epsilon > 0$, *choose* $N \in \mathbb{N}$ *such that* $\frac{1}{N} < \epsilon$. *Then for all* $m > n > N$, *we have,* $\|v_m - v_n\| = \frac{1}{n+1} < \frac{1}{N+1} < \frac{1}{N} < \epsilon$. *Thus* $\{v_n\}$ *is a Cauchy sequence. But* $v_n \to v = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right) \notin c_{00}$. *Therefore* c_{00} *is not complete.*

(b) *Consider* $(l^\infty, \|.\|)$, *where* $\|.\|$ *is defined by* $\|v\| = sup_i |v_i|$ *for* $v = (v_1, v_2, \ldots) \in l^\infty$. *Let* $\{v_n\}$ *be a Cauchy sequence in* l^∞, *where* $v_n = \left(v_1^{(n)}, v_2^{(n)}, \ldots\right)$. *Then for any* $\epsilon > 0$, *there exists* N *such that*

$$\|v_m - v_n\| = sup_i |v_i^{(m)} - v_i^{(n)}| < \epsilon, \ \forall\, m, n > N$$

Thus, for every fixed i,

$$|v_i^{(m)} - v_i^{(n)}| < \epsilon, \ \forall\, m, n > N \tag{12.2}$$

Therefore the sequence $\left(v_i^{(1)}, v_i^{(2)}, \ldots\right)$ *is a Cauchy sequence of real numbers and hence is convergent, say* $v_i^{(n)} \to v_i$ *as* $n \to \infty$. *Consider the element* $v = (v_1, v_2, \ldots)$. *Letting* $n \to \infty$ *in* (12.2), *we get* $|v_i^{(m)} - v_i| < \epsilon, \ \forall\, m > N$. *As* $\{v_n\} \in l^\infty$, *there exists* $\lambda_n \in \mathbb{R}$ *such that* $|x_i^{(n)}| < \lambda_n$ *for all* i. *Therefore,*

$$|v_i| = |v_i - v_i^{(n)} + v_i^{(n)}| \leq |v_i - v_i^{(n)}| + |v_i^{(n)}| < \epsilon + \lambda_n, \ \forall\, n > N, \ \forall\, i$$

Thus $v_n \to v$ *and* $v \in l^\infty$.

(c) $C[0, 1]$ *is complete with respect to supremum norm.*

(d) *As every finite-dimensional space is complete,* $\mathbb{R}^n$ *is complete.*

(8) Let α be a primitive fifth root of unity. Define

$$A = \begin{bmatrix} \alpha^{-2} & 0 & 0 & 0 & 0 \\ 0 & \alpha^{-1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha^2 \end{bmatrix}$$

For a vector $v = (v_1, v_2, v_3, v_4, v_5) \in \mathbb{R}^5$, define $|\, v\,|_A = \sqrt{|\, vAv^T\,|}$. If $w = (1, -1, 1, 1, -1)$, then $|\, w\,|_A = \cdots\cdots$

(a) 0 (b) 1 (c) -1 (d) 2

Ans. ***Option a***

We have

$$
\begin{aligned}
wAw^T &= \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \alpha^{-2} & 0 & 0 & 0 & 0 \\ 0 & \alpha^{-1} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha^2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \\
&= \alpha^{-2} + \alpha^{-1} + 1 + \alpha + \alpha^2 \\
&= 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4 = 0
\end{aligned}
$$

(9) Let A be an $n \times n$ matrix with real entries. Define $\langle x, y\rangle_A = \langle Ax, Ay\rangle, x, y \in \mathbb{R}^n$. Then $\langle x, y\rangle_A$ defines an inner-product if and only if

(a) $Ker(A) = \{0\}$

(b) $Rank(A) = n$

(c) All eigenvalues of A are positive.

(d) All eigenvalues of A are non-negative.

Ans. ***Options a and b***

Clearly $\langle x, y\rangle_A = \langle Ax, Ay\rangle \geq 0$. *Also*

$$\langle x, y\rangle_A = 0 \Rightarrow \langle Ax, Ay\rangle = 0 \Rightarrow x = y \iff A \; is \; one - one$$

That is, if and only if $Ker(A) = \{0\}$. *Then by Rank-Nullity Theorem,* $Rank(A) = n$. *Since* A *can be any matrix, options* (c) *and* (d) *are false.*

(10) Which of the following is an inner product on the vector space V of $n \times n$ real symmetric matrices?

(a) $\langle A, B\rangle_1 = (tr(A))(tr(B))$ (b) $\langle A, B\rangle_2 = tr(AB)$

(c) $\langle A, B\rangle_3 = det(AB)$ (d) $\langle A, B\rangle_4 = tr(A) + tr(B)$

Ans. ***Option b***

As A *is symmetric,* A *is of the form* $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$.

(a) For $n = 2$, *consider* $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. *Then,*

$$\langle A, A\rangle_1 = (tr(A))^2 = 0$$

But $A \neq 0$. *As* $(IP1)$ *is violated,* $\langle , \rangle_1$ *is not an inner product on* V.

(b) (IP1) For any $A \in V$, *we have*

$$\langle A, A\rangle_2 = tr(A^2) = \sum_{i,j=1}^{n} a_{ij}^2 \geq 0$$

and

$$\langle A, A\rangle_2 = 0 \Leftrightarrow \sum_{i,j=1}^{n} a_{ij}^2 = 0 \Leftrightarrow a_{ij}^2 = 0,\ \forall i, j \Leftrightarrow A = 0$$

(IP2) For any $A, B, C \in V$, we have

$$\begin{aligned}\langle A + B, C\rangle_2 &= tr\,[(A + B)C] = tr(AC + BC) = tr(AC) + tr(BC)\\ &= \langle A, B\rangle_2 + \langle B, C\rangle_2\end{aligned}$$

(IP3) For any $A, B \in V$ and $\lambda \in \mathbb{R}$, we have

$$\langle \lambda A, B\rangle_2 = tr(\lambda AB) = \lambda tr(AB) = \lambda\langle A, B\rangle_2$$

(IP4) For any $A, B \in V$, we have

$$\langle A, B\rangle_2 = tr(AB) = tr(BA) = \langle B, A\rangle_2$$

Therefore, $\langle , \rangle_2$ defines an inner product on V.

(c) For $n = 2$, consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then

$$\langle A, A\rangle_3 = det(A^2) = 0$$

But $A \neq 0$. As $(IP1)$ is violated, $\langle , \rangle_3$ is not an inner product on V.

(d) For $n = 2$, consider $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Then,

$$\langle A, A\rangle_4 = 2.tr(A) = 0$$

But $A \neq 0$. As $(IP1)$ is violated, $\langle , \rangle_4$ is not an inner product on V.

(11) Which of the following are inner products on $\mathbb{R}^2$?
(a) $\langle (u_1, u_2), (v_1, v_2)\rangle_1 = u_1v_1 + 2u_1v_2 + 2u_2v_1 + u_2v_2$
(b) $\langle (u_1, u_2), (v_1, v_2)\rangle_2 = u_1v_1 + u_1v_2 + u_2v_1 + 2u_2v_2$
(c) $\langle (u_1, u_2), (v_1, v_2)\rangle_3 = u_1v_1 + u_1v_2 + u_2v_1 + u_2v_2$
(d) $\langle (u_1, u_2), (v_1, v_2)\rangle_4 = u_1v_1 - \frac{1}{2}u_1v_2 - \frac{1}{2}u_2v_1 + u_2v_2$

Ans. ***Option b and d***
We know that, for $u, v \in \mathbb{R}^2$ and a fixed matrix $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$, $\langle u, v\rangle = u^T Av$ defines an inner product on $\mathbb{R}^2$ if and only if A is positive definite.

(a) We have

$$\langle (u_1, u_2), (v_1, v_2)\rangle_1 = u_1v_1 + 2u_1v_2 + 2u_2v_1 + u_2v_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Here $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. *As the eigenvalues of A are 3 and* -1, *A is not positive definite and hence* $\langle , \rangle_1$ *does not define an inner product on* $\mathbb{R}^2$.

(b) We have

$$\langle (u_1, u_2), (v_1, v_2)\rangle_2 = u_1v_1 + u_1v_2 + u_2v_1 + 2u_2v_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Here $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. *As the eigenvalues of A are* $\frac{3 \pm \sqrt{5}}{2}$, *A is positive definite and hence* $\langle , \rangle_2$ *defines an inner product on* $\mathbb{R}^2$.

(c) We have

$$\langle (u_1, u_2), (v_1, v_2)\rangle_3 = u_1v_1 + u_1v_2 + u_2v_1 + u_2v_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Here $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. *As the eigenvalues of A are* 0 *and* 2, *A is not positive definite and hence* $\langle , \rangle_3$ *does not define an inner product on* $\mathbb{R}^2$.

(d) We have

$$\langle (u_1, u_2), (v_1, v_2)\rangle_4 = u_1v_1 - \frac{1}{2}u_1v_2 - \frac{1}{2}u_2v_1 + u_2v_2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

Here $A = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$. *As the eigenvalues of A are* $\frac{1}{2}$ *and* $\frac{3}{2}$, *A is positive definite and hence* $\langle , \rangle_4$ *defines an inner product on* $\mathbb{R}^2$.

(12) Define $\langle , \rangle : C[0, 1] \times C[0, 1] \to \mathbb{R}$ by

$$\langle f, g\rangle = \int_0^1 f(t)\,(g(t))^2\,dt$$

Then which of the following statements is true?
(a) $\langle , \rangle$ is an inner product on $C[0, 1]$.
(b) $\langle , \rangle$ is a bilinear form on $C[0, 1]$ but is not an inner product on $C[0, 1]$.
(c) $\langle , \rangle$ is not a bilinear form on $C[0, 1]$.
(d) $\langle f, f\rangle \geq 0$ for all $f \in C[0, 1]$.

Ans. ***Option c***
Take the constant function $f(t) = -1$ *for all* $t \in [0, 1]$. *Clearly* $f \in C[0, 1]$ *and*

$$\langle f, f\rangle = \int_0^1 (-1)^3 dt = -1 < 0$$

Thus $\langle , \rangle$ is not an inner product on $C[0, 1]$. $\langle , \rangle$ is a bilinear form on $C[0, 1]$ if it is linear on both first and second variable. Clearly $\langle , \rangle$ is not linear on the second variable. Thus, $\langle , \rangle$ is not a bilinear form on $C[0, 1]$.

(13) Which of the following normed spaces are inner product spaces?

(a) l^p, $1 \le p < \infty$ with $\|v\| = \left(\sum_{i=1}^{\infty} |v_i|^p\right)^{\frac{1}{p}}$, for $v = (v_1, v_2, \ldots) \in l^p$.

(b) $C[0, 1]$ with $\|f\| = \left(\int_0^1 |f(x)|^p dx\right)^{\frac{1}{p}}$, for $f \in C[0, 1]$.

(c) $L^p[0, 1]$, $1 \le p < \infty$ with $\|f\| = \left(\int_0^1 |f(x)|^p dx\right)^{\frac{1}{p}}$, for $f \in L^p[0, 1]$.

(d) None of the above.

Ans. ***Option d***

(a) Take $u = (1, 1, 0, \ldots)$, $v = (1, -1, 0, \ldots) \in l^p$. Then $u + v = (2, 0, 0, \ldots)$ and $u - v = (0, 2, 0, \ldots)$ with

$$\|u + v\| = \|u - v\| = 2$$

Observe that

$$\|u\| = \|v\| = \begin{cases} 2^{\frac{1}{p}}, & if\ 1 \le p < \infty \\ 1, & if\ p = \infty \end{cases}$$

Clearly parallelogram law is satisfied only when $p = 2$. Therefore, l^p, $p \neq 2$ are not inner product spaces.

(b) Consider $f(x) = 1$, $g(x) = 1 - x \in C[0, 1]$. Then $(f + g)(x) = 1$ and $(f - g)(x) = 2x - 1$ with $\|f + g\| = 1$ and

$$\|f - g\| = \int_0^1 |2x - 1|^p dx = \begin{cases} \frac{1}{(1+p)^{\frac{1}{p}}}, & if\ 1 \le p < \infty \\ 1 & , if\ p = \infty \end{cases}$$

Also

$$\|f\| = \|g\| = \int_0^1 |2x - 1|^p dx = \begin{cases} \frac{1}{(1+p)^{\frac{1}{p}}}, & if\ 1 \le p < \infty \\ 1 & , if\ p = \infty \end{cases}$$

Clearly parallelogram law is satisfied only when $p = 2$. Therefore, $C[0, 1]$ is not an inner product spaces.

(c) Consider the same example as above. L^p, $p \neq 2$ are not inner product spaces.

(14) The space l^p is a Hilbert space if and only if
(a) $p > 1$ (b) p is even (c) $p = \infty$ (d) $p = 2$

Ans. ***Option d***
The space l^p is a Hilbert space if and only if $p = 2$.

(15) Let V be an inner product space and $u, v \in V$ be such that

$$|\langle u, v\rangle| = \|u\| \, \|v\|$$

Then ......
(a) u and v are linearly independent. (b) u and v are orthogonal.
(c) u and v are linearly dependent. (d) None of these.

Ans. ***Option c***
By Cauchy-Schwartz inequality, $|\langle u, v\rangle| = \|u\| \, \|v\|$ if and only if u and v are linearly dependent.

(16) Let $c_0 = \{(v_n) : v_n \in \mathbb{R}, v_n \to 0\}$ and $M = \{(v_n) \in c_0 : v_1 + v_2 + \cdots + v_{10} = 0\}$. Then, $dim\ (c_0/M)$ is equal to ......

Ans. *Define $T : c_0 \to \mathbb{R}$ by*

$$T(v) = v_1 + v_2 + \cdots + v_{10}$$

where $v = (v_1, \ldots, v_n, \ldots) \in c_0$. Then, for $u = (u_1, \ldots, u_n, \ldots)$, $v = (v_1, \ldots, v_n, \ldots) \in c_0$ and $\lambda \in \mathbb{R}$,

$$\begin{aligned} T(\lambda u + v) &= (u_1 + \lambda v_1) + (u_2 + \lambda v_2) + \cdots + (u_{10} + \lambda v_{10}) \\ &= u_1 + u_2 + \cdots + u_{10} + \lambda(v_1 + v_2 + \cdots + v_{10}) \\ &= T(u) + \lambda T(v) \end{aligned}$$

Clearly, T is an onto homomorphism from c_0 to $\mathbb{R}$ with

$$Ker(T) = \{(v_n) \in c_0 : v_1 + v_2 + \cdots + v_{10} = 0\}$$

Therefore by First Isomorphism Theorem, $c_0/M \cong \mathbb{R}$. Therefore $dim(c_0/M)$ $= 1$.

(17) Let $\langle ., . \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be an inner product on the vector space $\mathbb{R}^n$ over $\mathbb{R}$. Consider the following statements:

P: $|\langle u, v\rangle| \leq \frac{1}{2}\left[\langle u, u\rangle + \langle v, v\rangle\right]$ for all $u, v \in \mathbb{R}^n$.
Q: If $\langle u, v\rangle = \langle 2u, -v\rangle$ for all $v \in \mathbb{R}^n$, then $u = 0$.

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option a***
For all $u, v \in \mathbb{R}^n$*, we have*

$$(\|u\| - \|v\|)^2 = \|u\|^2 + \|v\|^2 - 2\,\|u\|\,\|v\| \geq 0$$

which implies that

$$\|u\|\,\|v\| \leq \frac{1}{2}\left[\|u\|^2 + \|v\|^2\right] = \frac{1}{2}\left[\langle u, u\rangle + \langle v, v\rangle\right]$$

Then by Cauchy-Schwartz inequality, we have

$$|\langle u, v\rangle| \leq \|u\|\,\|v\| \leq \frac{1}{2}\left[\langle u, u\rangle + \langle v, v\rangle\right]$$

for all $u, v \in \mathbb{R}^n$*. Now* $\langle u, v\rangle = \langle 2u, -v\rangle$ *for all* $v \in \mathbb{R}^n$ *implies that* $\langle u, v\rangle = -2\langle u, v\rangle$*. That is,* $3\langle u, v\rangle = 0$ *for all* $v \in \mathbb{R}^n$*. Hence* $u = 0$*.*

(18) Let H be a complex Hilbert space. Let $u, v \in H$ be such that $\langle u, v\rangle = 2$. Then

$$\frac{1}{2\pi}\int_0^{2\pi} \left\|u + e^{it}v\right\|^2 e^{it}dt = \cdots\cdots\cdots$$

Ans. *We have*

$$\begin{aligned}\left\|u + e^{it}v\right\|^2 &= \langle u + e^{it}v, u + e^{it}v\rangle \\ &= \langle u, u\rangle + \langle u, e^{it}v\rangle + \langle e^{it}v, u\rangle + \langle e^{it}v, e^{it}v\rangle \\ &= \|u\|^2 + e^{it}\langle u, v\rangle + e^{-it}\langle v, u\rangle + \|v\|^2 \\ &= \|u\|^2 + 2e^{it} + 2e^{-it} + \|v\|^2\end{aligned}$$

Therefore

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} \left\|u + e^{it}v\right\|^2 e^{it}dt &= \frac{1}{2\pi}\int_0^{2\pi} \left[\|u\|^2 + 2e^{it} + 2e^{-it} + \|v\|^2\right] e^{it}dt \\ &= \frac{1}{2\pi}\left[\left(\|u\|^2\,\frac{e^{it}}{i}\right)_0^{2\pi} + 2\left(\frac{e^{i2t}}{2i}\right)_0^{2\pi} + 2(t)_0^{2\pi} + \left(\|v\|^2\,\frac{e^{it}}{i}\right)_0^{2\pi}\right] \\ &= \frac{1}{2\pi}\left[4\pi\right] \;\left(since\ e^{2\pi} = e^{4\pi} = 1\right) \\ &= 2\end{aligned}$$

(19) Let H be a Hilbert space. Consider the following statements:

P: If $\{e_i : i \in \mathbb{N}\}$ is an orthonormal set and $x \in H$, then the set $E = \{e_i : \langle x, e_i\rangle \neq 0\}$ is denumerable.

Q: If $\{e_i : i \in \mathbb{N}\}$ is an orthonormal set, then for any $x \in H$, $\langle x, e_n \rangle \to 0$ as $n \to \infty$.

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option a***
For $n \in \mathbb{N}$ and $x \in H$, define

$$E_n = \left\{e_i : |\langle x, e_i \rangle|^2 > \frac{1}{n} \|x\|^2\right\}$$

Suppose that E_n contains n or more than n elements, then for $e_i \in E_n$,

$$\sum |\langle x, e_i \rangle|^2 > n \left(\frac{1}{n} \|x\|^2\right) = \|x\|^2$$

But by Bessel's inequality, we must have $\sum |\langle x, e_i \rangle|^2 \leq \|x\|^2$ for any $x \in H$. Therefore E_n contains at most $n - 1$ elements. Thus E_n is finite for all $n \in \mathbb{N}$ Clearly, $E_n \subset E$ for all $n \in \mathbb{N}$. Now let $e_i \in E$. Then $\langle x, e_i \rangle \neq 0$. We can always choose an n, say n_0 such that $|\langle x, e_i \rangle|^2 > \frac{1}{n_0} \|x\|^2$. Then $e_i \in E_{n_0}$. Therefore $E = \cup_{n=1}^{\infty} E_n$. Since each E_n is finite, E is countable.
If $\{e_i : i \in \mathbb{N}\}$ is an orthonormal set in a Hilbert space H, by Bessel's inequality, we have $\sum |\langle x, e_i \rangle|^2 \leq \|x\|^2$. Therefore $|\langle x, e_i \rangle|^2$ converges as n tends to ∞. Thus $\langle x, e_n \rangle \to 0$ as $n \to \infty$ for any $x \in H$.

(20) Suppose $\{v_1, \ldots, v_n\}$ are unit vectors in $\mathbb{R}^n$ such that

$$\|v\|^2 = \sum_{i=1}^{n} |\langle v_i, v \rangle|^2, \ \forall\, v \in \mathbb{R}^n$$

Then decide the correct statements in the following
(a) $v_1, \ldots, v_n$ are mutually orthogonal.
(b) $\{v_1, \ldots, v_n\}$ is a basis for $\mathbb{R}^n$.
(c) $v_1, \ldots, v_n$ are not mutually orthogonal.
(d) At most $n - 1$ of the elements in the set $\{v_1, \ldots, v_n\}$ can be orthogonal.

Ans. ***Options a and b***
Take $v = v_j$. Then

$$\|v_j\|^2 = \sum_{i=1}^{n} |\langle v_i, v_j \rangle|^2 = |\langle v_j, v_j \rangle|^2 \Rightarrow |\langle v_i, v_j \rangle| = 0 \ \forall\, i \neq j$$

Therefore $v_1, \ldots, v_n$ are mutually orthogonal and hence $\{v_1, \ldots, v_n\}$ is a basis for $\mathbb{R}^n$.

(21) Let $(V, \langle ., . \rangle)$ be an inner product space. Consider the following statements;

P: If $u_n \to u$ and $v_n \to v$ in V, then $\langle u_n, v_n \rangle \to \langle u, v \rangle$ in V.
Q: If $v_n \to v$ and $u \perp v_n, \ \forall\, n$, then $u \perp v$.

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option a***
We have

$$\begin{aligned}|\langle u_n, v_n \rangle - \langle u, v \rangle| &= |\langle u_n, v_n \rangle - \langle u_n, v \rangle + \langle u_n, v \rangle - \langle u, v \rangle| \\ &\leq |\langle u_n, v_n \rangle - \langle u_n, v \rangle| + |\langle u_n, v \rangle - \langle u, v \rangle| \\ &\leq \|u_n\| \, \|v_n - v\| + \|u_n - u\| \, \|v\| \\ &\to 0 \textit{ as } u_n \to u \textit{ and } v_n \to v\end{aligned}$$

Since $v_n \to v$, from above we have $\langle u, v_n \rangle \to \langle u, v \rangle$. As $u \perp v_n, \ \forall\, n$, we have $\langle u, v_n \rangle = 0, \ \forall\, n$. Therefore $\langle u, v \rangle = lim_{n\to\infty} \langle u, v_n \rangle = 0$. Hence $u \perp v$.

(22) Let $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and $V = \{Ax^T : x \in \mathbb{R}^{1\times 3}\}$. Then an orthonormal basis for V.

(a) $\left\{(1, 0, 0)^T, \left(0, \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)^T, \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)^T\right\}$ (b) $\left\{(1, 0, 0)^T, \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T\right\}$

(c) $\left\{(1, 0, 0)^T, \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)^T, \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)^T\right\}$ (d) $\left\{(1, 0, 0)^T, (0, 0, 1)^T\right\}$

Ans. ***Option d***

Here V is clearly the column space of A. Since $A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ (Column reduced form), $\left\{(1, 0, 0)^T, (0, 0, 1)^T\right\}$ gives an orthonormal basis.

(23) Consider the subspaces W_1 and W_2 of $\mathbb{R}^3$ given by

$$W_1 = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$$

$$W_2 = \{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\}$$

If W is a subspace of $\mathbb{R}^3$ such that

(i) $W \cap W_2 = span\{(0, 1, 1)\}$
(ii) $W \cap W_1$ is orthogonal to $W \cap W_2$ with respect to the usual inner product of $\mathbb{R}^3$

then
(a) $W = \{(0, 1, -1), (0, 1, 1)\}$ (b) $W = \{(1, 0, -1), (0, 1, -1)\}$
(c) $W = \{(1, 0, -1), (0, 1, 1)\}$ (d) $W = \{(1, 0, -1), (1, 0, 1)\}$

Ans. ***Option a***
We have

$$\begin{aligned} W_1 &= \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\} \\ &= span\{(1, 0, -1), (0, 1, -1)\} \end{aligned}$$

and

$$\begin{aligned} W_2 &= \{(x, y, z) \in \mathbb{R}^3 : x - y + z = 0\} \\ &= span\{(1, 0, -1), (0, 1, 1)\} \end{aligned}$$

Since $W \cap W_2 = span\{(0, 1, 1)\}$ *and* $W \cap W_1$ *is orthogonal to* $W \cap W_2$, $W = \{(0, 1, -1), (0, 1, 1)\}$.

(24) Let U be an orthonormal set in a Hilbert space H and let $v \in H$ be such that $\|v\| = 2$. Consider the set

$$E = \{u \in U : |\langle v, u\rangle| \geq \frac{1}{4}\}$$

Then the maximum possible number of elements in E is ......

Ans. *Suppose that* E *contains* n *distinct elements, say* $u_1, u_2, \ldots, u_n$. *As* $|\langle v, u\rangle| \geq \frac{1}{4}$ *for each* $u \in E$, *we have* $\sum_{i=1}^{n} |\langle v, u_i\rangle|^2 \geq \frac{n}{16}$. *Also by Bessel's inequality, we get* $\sum_{i=1}^{n} |\langle v, u_i\rangle|^2 \leq \|v\|^2 = 4$. *Therefore, the maximum possible number of elements in* E *is* 64.

(25) The application of Gram-Schmidt process of orthonormalization to

$$u_1 = (1, 1, 0),\ u_2 = (1, 0, 0),\ u_3 = (1, 1, 1)$$

yields
(a) $\left\{\frac{1}{\sqrt{2}}(1, 1, 0), (1, 0, 0), (0, 0, 1)\right\}$ (b) $\left\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{2}}(1, -1, 0), \frac{1}{\sqrt{2}}(1, 1, 1)\right\}$
(c) $\{(0, 1, 0), (1, 0, 0), (0, 0, 1)\}$ (d) $\left\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{2}}(1, -1, 0), (0, 0, 1)\right\}$

Ans. ***Option d***
Let $v_1 = u_1 = (1, 1, 0)$. *Then*

$$v_2 = u_2 - \frac{\langle u_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (1, 0, 0) - \frac{1}{2}(1, 1, 0) = \frac{1}{2}(1, -1, 0)$$

Take $v_2 = (1, -1, 0)$. *Then*

$$v_3 = u_3 - \frac{\langle u_3, v_1\rangle}{\langle v_1, v_1\rangle} v_1 - \frac{\langle u_3, v_2\rangle}{\langle v_2, v_2\rangle} v_2 = (1, 1, 1) - (1, 1, 0) = (0, 0, 1)$$

Therefore $\left\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{2}}(1, -1, 0), (0, 0, 1)\right\}$ *is the required set.*

(26) Consider $\mathbb{R}^3$ with the standard inner product. Let $S = \{(1, 1, 1), (2, -1, 2), (1, -2, 1)\}$. For a subset W of $\mathbb{R}^3$, let $L(W)$ denote the linear span of W in $\mathbb{R}^3$. Then an orthonormal set T with $L(S) = L(T)$ is
(a) $\{\frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{6}}(1, -2, 1)\}$ (b) $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$
(c) $\{\frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{2}}(1, -1, 0)\}$ (d) $\{\frac{1}{\sqrt{3}}(1, 1, 1), \frac{1}{\sqrt{2}}(0, 1, -1)\}$

Ans. ***Option a***
Since $(1, -2, 1) = -(1, 1, 1) + (2, -1, 2)$,

$$L(S) = span\{(1, 1, 1), (2, -1, 2)\}$$

Now we construct an orthogonal basis using Gram-Schmidt Orthogonalisation process from the set $\{v_1 = (1, 1, 1), v_2 = (2, -1, 2)\}$.
Take $u_1 = v_1$, *then*

$$u_2 = v_2 - \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = (1, -2, 1)$$

and

$$\left\{\frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{3}}(1, 1, 1), \frac{u_2}{\|u_2\|} = \frac{1}{\sqrt{6}}(1, -2, 1)\right\}$$

is an orthonormal set T *with* $L(S) = L(T)$.

(27) Let V be the inner product space consisting of linear polynomials, $p : [0, 1] \to \mathbb{R}$($V$ consists of polynomials p of the form $p(x) = ax + b,\ a, b \in \mathbb{R}$), with the inner product defined by

$$\langle p, q\rangle = \int_0^1 p(x)q(x)dx \ for \ p, q \in V$$

An orthonormal basis of V is
(a) $\{1, x\}$ (b) $\{1, x\sqrt{3}\}$ (c) $\{1, (2x - 1)\sqrt{3}\}$ (d) $\{1, x - \frac{1}{2}\}$

Ans. ***Option c***
Consider the standard ordered basis $\{1, x\}$. *Take* $u_1 = 1$ *and*

$$u_2 = v_2 - \frac{\langle v_2, v_1\rangle}{\langle v_1, v_1\rangle} v_1 = x - \frac{\langle x, 1\rangle}{\langle 1, 1\rangle} = x - \frac{1}{2}$$

Then $\left\{1, x - \frac{1}{2}\right\}$ *forms an orthogonal basis. Now*

$$\frac{u_2}{\|u_2\|} = (2x - 1)\sqrt{3}$$

as $\|u_2\| = (\langle u_2, u_2\rangle)^{\frac{1}{2}} = 2\sqrt{3}$.

(28) Consider $\mathbb{R}^3$ with standard inner product. Let $W = span\{(1, 0, -1)\}$. Then, which of the following is a basis for the orthogonal complement of W?
(a) $\{(1, 0, 1), (0, 1, 0)\}$ (b) $\{(1, 2, 1), (0, 1, 1)\}$
(c) $\{(2, 1, 2), (4, 2, 4)\}$ (d) $\{(2, -1, 2), (1, 3, 1), (-1, -1, -1)\}$

Ans. ***Option a***
As $\mathbb{R}^3$ *can be written as a direct sum of W and its orthogonal complement,* $W^\perp$ *is a two-dimensional subspace of* $\mathbb{R}^3$.

(a) Observe that $\langle(1, 0, 1), (1, 0, -1)\rangle = 0$ *and* $\langle(1, 0, -1), (0, 1, 0)\rangle = 0$. *Also, the set* $\{(1, 0, 1), (0, 1, 0)\}$ *is linearly independent and hence is a basis of* $W^\perp$.
(b) We have $\langle(0, 1, 1), (1, 0, -1)\rangle = -1 \neq 0$. *Therefore* $(0, 1, 1) \notin W^\perp$.
(c) Clearly, $\langle(2, 1, 2), (1, 0, -1)\rangle = 0$ *and* $\langle(4, 2, 4), (1, 0, -1)\rangle = 0$. *But, the set* $\{(2, 1, 2), (4, 2, 4)\}$ *is linearly dependent and hence is not a basis of* $W^\perp$.
(d) As $W^\perp$ *is a two-dimensional subspace of* $\mathbb{R}^3$, $\{(2, -1, 2), (1, 3, 1), (-1, -1, -1)\}$ *cannot be a basis of* $W^\perp$.

(29) For $f, g \in \mathbb{P}_2[x]$, define

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt$$

Let $W = \{1 - t^2, 1 + t^2\}$. Which of the following conditions is satisfied for all $h \in W^\perp$?
(a) h is an even function.
(b) h is an odd function.
(c) $h(t) = 0$ has a real solution.
(a) $h(0)$.

Ans. ***Option c***
We have,

$$W^\perp = \left\{h \in \mathbb{P}_2[x] \mid \int_0^1 f(t)h(t)dt = 0,\ \forall f \in W\right\}$$
$$= \left\{h \in \mathbb{P}_2[x] \mid \langle(1 - t^2), h\rangle = 0 \textit{ and } \langle(1 + t^2), h\rangle = 0\right\}$$

Take $h(t) = a + bt + ct^2 \in \mathbb{P}_2[x]$. *Then,*

$$\begin{aligned}\langle(1-t^2),h\rangle = 0 &\Rightarrow \int_0^1 \left(1-t^2\right)\left(a+bt+ct^2\right)dt = 0\\ &\Rightarrow \int_0^1 \left(a+bt+(c-a)t^2-bt^3-ct^4\right)dt = 0\\ &\Rightarrow \frac{2a}{3}+\frac{b}{4}+\frac{2c}{15} = 0\end{aligned}$$

and

$$\begin{aligned}\langle(1+t^2),h\rangle = 0 &\Rightarrow \int_0^1 \left(1+t^2\right)\left(a+bt+ct^2\right)dt = 0\\ &\Rightarrow \int_0^1 \left(a+bt+(c+a)t^2+bt^3+ct^4\right)dt = 0\\ &\Rightarrow \frac{4a}{3}+\frac{3b}{4}+\frac{8c}{15} = 0\end{aligned}$$

Solving these equations, we get $a = \frac{3c}{15}$ *and* $b = \frac{16c}{15}$. *Therefore,*

$$W^{\perp} = span\left\{3+16t+15t^2\right\}$$

Any element $h \in W^{\perp}$ *is of the form* $3k + 16kt + 15kt^2$ *for some* $k \in \mathbb{R}$. *Clearly,* h *is neither an odd function nor an even function. Also* $h(0) \neq 0$. *As,*

$$(16k)^2 - 4(3k)(15k) = 256k^2 - 180k^2 = 76k^2 > 0$$

for all $k \in \mathbb{R}$, $h(t) = 0$ *has a real solution for any* $h \in W^{\perp}$.

(30) Consider $L^2[0, 2\pi]$ with inner product

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}dx$$

Which of the following is an orthogonal sequence in $L^2[0, 2\pi]$?
(a) $\{x^n : n \in \mathbb{N}\}$ (b) $\{cos\ nx : n \in \mathbb{N}\}$
(c) $\{e^{inx} : n \in \mathbb{N}\}$ (d) $\{sin\ nx : n \in \mathbb{N}\}$

Ans. ***Options b, c and d***

(a) We have

$$\langle x, x^2\rangle = \int_0^{2\pi} x^3 dx = \left[\frac{x^4}{4}\right]_0^{2\pi} \neq 0$$

Therefore $\{x^n : n \in \mathbb{N}\}$ *is not an orthogonal sequence in* $L^2[0, 2\pi]$.

(b) For $n_1 \neq n_2$, *we have*

$$\langle cos\ n_1x, cos\ n_2x\rangle = \int_0^{2\pi} cos\ n_1x\ cos\ n_2xdx$$
$$= \left[\frac{sin\ (n_1+n_2)\,x}{2\,(n_1+n_2)} + \frac{sin\ (n_1-n_2)\,x}{2\,(n_1-n_2)}\right]_0^{2\pi} = 0$$

Therefore $\{cos\ nx : n \in \mathbb{N}\}$ *is an orthogonal sequence in* $L^2[0, 2\pi]$.

(c) For $n_1 \neq n_2$,

$$\langle e^{in_1x}, e^{in_2x}\rangle = \int_0^{2\pi} e^{i(n_1-n_2)x}dx = \left[\frac{1}{i\,(n_1-n_2)}e^{i(n_1-n_2)x}\right]_0^{2\pi} = 0$$

Therefore $\{e^{inx} : n \in \mathbb{N}\}$ *is an orthogonal sequence in* $L^2[0, 2\pi]$.

(d) For $n_1 \neq n_2$, *we have*

$$\langle sin\ n_1x, sin\ n_2x\rangle = \int_0^{2\pi} sin\ n_1x\ sin\ n_2xdx$$
$$= \left[\frac{sin\ (n_1-n_2)\,x}{2\,(n_1-n_2)} - \frac{sin\ (n_1+n_2)\,x}{2\,(n_1+n_2)}\right]_0^{2\pi} = 0$$

Therefore $\{sin\ nx : n \in \mathbb{N}\}$ *is an orthogonal sequence in* $L^2[0, 2\pi]$.

(31) Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis of $\mathbb{C}^n$ as column vectors. Let $M = \left[v_1\ v_2\ \ldots\ v_k\right]$, $N = \left[v_{k+1}\ v_{k+2}\ \ldots\ v_n\right]$ and P be the diagonal $k \times k$ matrix with diagonal entries $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$. Then which of the following are true?
(a) $Rank(MPM^*) = k$ whenever $\alpha_i \neq \alpha_j,\ 1 \leq i, j \leq k$.
(b) $tr(MPM^*) = \sum_{i=1}^{k} \alpha_i$
(c) $Rank(M^*N) = min(k, n-k)$
(d) $Rank(MM^* + NN^*) < n$

Ans. ***Option b***
If $Rank(P) < k$, *then by Sylvester's inequality* $Rank(MPM^*) < k$. *Also*

$$tr(MPM^*) = tr(MM^*P) = tr(P) = \sum_{i=1}^{k} \alpha_i$$

Take $M = \left[e_1\ e_2\right]$ *and* $N = \left[e_3\ e_4\right]$. *Then* $M^*N = 0$. *Also* $MM^* + NN^* = I_4$.

(32) Which of the following statements are true?
(a) There exists $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ which is orthogonal and has 2 as an eigenvalue.
(b) There exists $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ which is orthogonal and has i as an eigenvalue.
(c) If $A \in \mathbb{M}_{2\times 2}(\mathbb{R})$ is orthogonal, then $\|Av\| = \|v\|$ for every $v \in \mathbb{M}_{2\times 2}(\mathbb{R})$, where $\|.\|$ denotes the usual euclidean norm on $\mathbb{R}^2$.

Ans. ***Options b and c***

(a) Let λ be an eigenvalue of A. Then there exists $v \neq 0 \in \mathbb{R}^2$ such that $Av = \lambda v$. We have

$$\|Av\|^2 = (Av)^T(Av) = v^T A^T Av = v^T v = \|v\|^2$$

Also

$$\|Av\|^2 = \langle \lambda v, \lambda v \rangle = |\lambda|^2 \|v\|^2 \Rightarrow |\lambda|^2 = 1 \Rightarrow |\lambda| = 1$$

Therefore 2 is not an eigenvalue of A.

(b) Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then

$$AA^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Also the eigenvalues of A are i and $-i$.

(c) From (a) we get $\|Ax\|^2 = \|x\|^2$. Since $\|Ax\|$ and $\|x\|$ are positive, $\|Ax\| = \|x\|$.

(33) Let A be an orthogonal 3×3 matrix with real entries. Pick out the true statements:
(a) The determinant of A is a rational number.
(b) $d(Au, Av) = d(u, v)$ for any two vectors u and $v \in \mathbb{R}^3$, where $d(u, v)$ denotes the usual Euclidean distance vectors u and $v \in \mathbb{R}^3$
(c) All the entries of A are positive.
(d) All the eigenvalues of A are real.

Ans. ***Options a and b***
A is orthogonal $\Rightarrow AA^T = I \Rightarrow det\ (AA^T) = 1 \Rightarrow (det\ A) = \pm 1$. Since A is orthogonal, geometrically it is either a rotation or reflection. Hence $d(Au, Av) = d(u, v)$ for any two vectors u and $v \in \mathbb{R}^3$. Consider the matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Clearly A is orthogonal. But all the entries are not positive and it has complex eigenvalues.

(34) Let S be the set of all 3×3 matrices A with integer entries such that the product AA^T is the identity matrix. Then $|S| = \ldots$
(a) 12 (b) 24 (c) 48 (d) 60

Ans. ***Option c***
Note that S consists of all matrices having orthonormal rows. Since A has only integer entries, a row of length one has only one non-zero entry which is equal to -1 or 1. Therefore there are 6 choices for the first row. The second row has to be perpendicular to the first. Therefore there are only 4 choices for the second row.

(i.e, if first row is $\begin{bmatrix}1 & 0 & 0\end{bmatrix}$*, the second row cannot be* $\begin{bmatrix}1 & 0 & 0\end{bmatrix}$ *or* $\begin{bmatrix}-1 & 0 & 0\end{bmatrix}$*). Finally for the last row there are only two choices. Therefore* $|S| = 6 \times 4 \times 2 = 48$ *choices.*

(35) Let A be a real $n \times n$ orthogonal matrix is, $A^T A = AA^T = I$. Which of the following statements are necessarily true?
(a) $\langle Au, Av\rangle = \langle u, v\rangle \; \forall\, u, v \in \mathbb{R}^n$
(b) All eigenvalues of A are either $+1$ or -1
(c) The rows of A form an orthonormal basis of $\mathbb{R}^n$.
(d) A is diagonalizable over $\mathbb{R}$.

Ans. ***Options a and c***
As $A^T A = AA^T = I$*, we have*

$$\langle Au, Av\rangle = \langle u, A^T Av\rangle = \langle u, v\rangle \; \forall\, u, v \in \mathbb{R}^n$$

Consider the matrix $A = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$. *Then* $A^T A = AA^T = I_n$. *The eigenvalues of* A *are either* i *or* $-i$. A *is not diagonalizable over* $\mathbb{R}$. *The* ith *diagonal element of* AA^T *is the length of the* ith *row of* A. *Therefore the rows of* A *form an orthonormal basis of* $\mathbb{R}^n$.

(36) Let n be an integer ≥ 2 and $B \in \mathbb{M}_{n\times n}(\mathbb{R})$ be an orthogonal matrix. Consider $W_B = \{B^T AB : A \in \mathbb{M}_{n\times n}(\mathbb{R})\}$. Which of the following are necessarily true?
(a) W_B is a subspace of $\mathbb{M}_{n\times n}(\mathbb{R})$ and $dim\ W_B \leq Rank(B)$.
(b) W_B is a subspace of $\mathbb{M}_{n\times n}(\mathbb{R})$ and $dim\ W_B = Rank(B)Rank(B^t)$.
(c) $W_B = \mathbb{M}_{n\times n}(\mathbb{R})$.
(d) W_B is not a subspace of $\mathbb{M}_{n\times n}(\mathbb{R})$.

Ans. ***Options b and c***
Let $C_1, C_2 \in W_B$. *Then there exists* $A_1, A_2 \in \mathbb{M}_{n\times n}(\mathbb{R})$ *such that* $C_1 = B^T A_1 B$ *and* $C_2 = B^T A_2 B$. *Then* $\lambda C_1 + C_2 = B^T(\lambda A_1 + A_2)B \in W_B$. *Therefore* W_B *is a subspace of* $\mathbb{M}_{n\times n}(\mathbb{R})$ *and* $dim\ W_B = n^2 = Rank(B)Rank(B^T)$. *Define* $T_B : M_n(R) \to M_n(R)$ *as* $T_B(A) = B^T ABT_B(A) = 0$ *if and only if* $A = 0$. *Hence* $\mathcal{N}(T) = 0$ *and by Rank-Nullity theorem,* $\mathcal{R}(T) = M_n(R)$.

(37) Let $A = \begin{bmatrix}3 & 0 & 0\\ 0 & 6 & 2\\ 0 & 2 & 6\end{bmatrix}$ and let $\lambda_1 \geq \lambda_2 \geq \lambda_3$ be the eigenvalues of A.

(a) The triple $(\lambda_1, \lambda_2, \lambda_3)$ equals
(i) $(9, 4, 2)$ (ii) $(8, 4, 3)$ (iii) $(9, 3, 3)$ (iv) $(7, 5, 3)$

(b) The matrix P such that $P^T AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$

(i) $\begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$ (ii) $\begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \end{bmatrix}$

(iii) $\begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{bmatrix}$ (iv) $\begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \end{bmatrix}$

Ans. *(a)* ***Option ii is correct.***
Since $det(A) = 96$*, the eigenvalues of A are 8, 4 and 3.*

(b) ***Option iii is correct.***
Since $8, 4, 3$ *are eigenvalues of A with eigenvectors* $(0, 1, 1)^T$, $(0, 1, -1)^T$, $(1, 0, 0)^T$, $P = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{bmatrix}$

(38) Let

$$A = \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(a) Then A is
i. non-invertible ii. skew-symmetric
iii. symmetric iv. orthogonal

(b) If B is any 3×3 real matrix, then $tr(ABA^t)$ is equal to
i. $[tr(A)]^2 tr(B)$ ii. $2tr(A) + tr(B)$
iii. $tr(B)$ iv. $[tr(A)]^2 + tr(B)$

Ans. *(a)* ***Option iv is correct.***
Since $det(A) \neq 0$*, N is invertible. Also*

$$A^T = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ \frac{-4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix} \neq A \neq -A$$

But $AA^T = \begin{bmatrix} \frac{3}{5} & \frac{-4}{5} & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{3}{5} & \frac{4}{5} & 0 \\ \frac{-4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix} = I = A^T A.$

Therefore A is orthogonal.

(b) ***Option iii is correct.***
Since $tr(ABA^T) = tr(A^TAB) = tr(B)$.

(39) Let A be an $n \times n$ non-zero skew-symmetric matrix. The matrix $(I - A)(I + A)^{-1}$ is always
(a) singular (b) symmetric (c) orthogonal (d) idempotent

Ans. ***Option c***
Consider $A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$. *Then*

$$(I - A)(I + A)^{-1} = \frac{1}{5}\begin{bmatrix} -3 & 4 \\ -4 & -3 \end{bmatrix}$$

is non-singular, not symmetric and is not idempotent. Now

$$\begin{aligned} \left[(I - A)(I + A)^{-1}\right]^T &= \left[(I + A)^T\right]^{-1}(I - A)^T \\ &= (I + A^T)^{-1}(I - A^T) \\ &= (I - A)^{-1}(I + A) \end{aligned}$$

Since

$$(I - A)(I + A) = I - A^2 = (I + A)(I - A)$$

we get

$$[(I - A)(I + A)]^{-1} = [(I + A)(I - A)]^{-1}$$

Therefore

$$\left[(I - A)(I + A)^{-1}\right]\left[(I - A)(I + A)^{-1}\right]^T = I$$

(40) If $\begin{bmatrix} \frac{\sqrt{5}}{3} & \frac{-2}{3} & \gamma \\ \frac{2}{3} & \frac{\sqrt{5}}{3} & \delta \\ \alpha & \beta & 1 \end{bmatrix}$ is a real orthogonal matrix, then $\alpha^2 + \beta^2 + \gamma^2 + \delta^2$ equals ......

Ans. *We have*

$$\begin{bmatrix} \frac{\sqrt{5}}{3} & \frac{-2}{3} & \gamma \\ \frac{-2}{3} & \frac{\sqrt{5}}{3} & \delta \\ \alpha & \beta & 1 \end{bmatrix}\begin{bmatrix} \frac{\sqrt{5}}{3} & \frac{2}{3} & \alpha \\ \frac{-2}{3} & \frac{\sqrt{5}}{3} & \beta \\ \gamma & \delta & 1 \end{bmatrix} = \begin{bmatrix} 1 + \gamma^2 & & \\ & 1 + \delta^2 & \\ & & \alpha^2 + \beta^2 + 1 \end{bmatrix} = I$$

$$Trace = \alpha^2 + \beta^2 + 1 + 1 + \delta^2 + 1 + \gamma^2 = 3 \Rightarrow \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 0$$

(41) For real constants α and β, let

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \alpha & \beta \end{bmatrix}$$

be an orthogonal matrix. Then which of the following statements is/are always true?
(a) $\alpha + \beta = 0$ (b) $\beta = \sqrt{1-\alpha^2}$ (c) $\alpha\beta = \frac{1}{2}$ (d) $A^2 = I$

Ans. ***Options a and d***
We have

$$AA^t = \begin{bmatrix} 1 & \frac{\alpha+\beta}{\sqrt{2}} \\ \frac{\alpha+\beta}{\sqrt{2}} & \alpha^2+\beta^2 \end{bmatrix} = I \Rightarrow \alpha + \beta = 0 \text{ and } \alpha^2 + \beta^2 = 1$$

$$\Rightarrow \alpha = \frac{1}{\sqrt{2}},\ \beta = -\frac{1}{\sqrt{2}} \text{ or } \alpha = -\frac{1}{\sqrt{2}},\ \beta = \frac{1}{\sqrt{2}}$$

Therefore $\beta = -\sqrt{1-\alpha^2}$ *and* $\alpha\beta = -\frac{1}{2}$. *Also* $A^2 = I$.

(42) Let $P = \dfrac{vv^T}{v^Tv}$ be an $n \times n (n > 1)$ matrix, where v is a non zero column vector. Then which one of the following statements is FALSE?
(a) P is idempotent (b) P is orthogonal
(c) P is symmetric (d) Rank of P is one

Ans. ***Option b***

$$P^2 = \frac{vv^T}{v^Tv}\frac{vv^T}{v^Tv} = \frac{v(v^Tv)v^T}{(v^Tv)^2} = \frac{vv^T}{v^Tv} = P \Rightarrow P \text{ is idempotent.}$$

$$P^T = \left(\frac{vv^T}{v^Tv}\right)^T = \frac{(v^T)^Tv^T}{v^Tv} = \frac{vv^T}{v^Tv} = P \Rightarrow P \text{ is symmetric.}$$

By Sylvester's inequality $Rank(P) = 1$. P *is not orthogonal.*

(43) Find an orthogonal matrix P such that $PAP^{-1} = B$, where

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Ans. *A is nilpotent matrix since* $A^3 = 0$. *Then,*

$$Av = 0 \Rightarrow v_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ is an eigenvector of } A.$$

$$Ax = v_1 \Rightarrow v_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ is generalized eigenvector of } A.$$

$Ax = v_2 \Rightarrow v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ *is generalized eigenvector of A.*

Now take $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Then* $P^{-1} = P$ *and* $PAP^{-1} = B$.

(44) Let $N \geq 2$. Let $v \in \mathbb{R}^N$, $v \neq 0$, be a column vector. Find the condition on v such that the matrix $I - 2vv^T$ is orthogonal.

Ans.

$$(I - 2vv^T)(I - 2vv^T) = I \Rightarrow 4vv^T(1 - v^T v) = 0$$

$v^T v$ *is a scalar and* $vv^T \neq 0$. *Hence the required condition is* $v^T v = 1$.

(45) Let $A : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. If W is a subspace of $\mathbb{R}^n$, define $W^\perp = \{v \in \mathbb{R}^n : \langle u, v \rangle = 0 \; \forall \; u \in W\}$.Which of the following statements are true?
(a) $\mathcal{R}(A) \subset \left[\mathcal{N}(A^T)\right]^\perp$.
(b) $\mathcal{R}(A) = \left[\mathcal{N}(A^T)\right]^\perp$.
(c) Neither of the above statements need be necessarily true.

Ans. ***Option a and b***
Let $v \in \mathcal{N}(A^T)$. *Then since* $A^T v = 0$ *if and only if* $\langle v, Au \rangle = \langle A^T v, u \rangle = 0$ *for all* $u \in \mathbb{R}^m$. *Therefore* $\mathcal{R}(A) = \left[\mathcal{N}(A^T)\right]^\perp$.

(46) Let V be the column space of the matrix $A = \begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 1 & -1 \end{pmatrix}$. Then the orthogonal projection of $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ on V is
(a) $(0, 1, 0)^T$ (b) $(0, 0, 1)^T$ (c) $(1, 1, 0)^T$ (d) $(1, 0, 1)^T$

Ans. ***Option a***
Since the columns v_1, v_2 *of A are orthogonal, the projection of the given vector* v *on* V *is* $= proj_{v_1}(v) + proj_{v_2}(v)$. *Now,*

$$proj_{v_1}(v) = \frac{v^T v_1}{v_1^T v_1} v_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

and

$$proj_{v_2}(v) = \frac{v^T v_2}{v_2^T v_2} v_2 = \frac{1}{3} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

Therefore $proj_V(v) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

(47) Let $W = span\left\{\frac{1}{\sqrt{2}}(0, 0, 1, 1), \frac{1}{\sqrt{2}}(1, -1, 0, 0)\right\}$ be a subspace of the Euclidean space $\mathbb{R}^4$. Then the square of the distance from the point $(1, 1, 1, 1)$ to the subspace W is equal to ......

Ans. *The projection of the given vector v on W is* $= proj_{w_1}(v) + proj_{w_2}(v)$. *Now*

$proj_{w_1}(v) = \dfrac{v^T w_1}{w_1^T w_1} w_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ *and* $proj_{w_2}(v) = \dfrac{v^T w_2}{w_2^T w_2} w_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. *Therefore*

$proj_W(v) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ *Now the square of the distance from the point* $(1, 1, 1, 1)$ *to the subspace W is the square of the length of the vector. i.e, 2.*

(48) Let $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ be a non-zero singular matrix. Consider the following problem: find $X \in \mathbb{M}_{n\times n}(\mathbb{R})$ such that

$$(i)\ AXA = A\ (ii)\ XAX = X \text{ and } (iii)\ AX = XA$$

Which of the following statements are true?
(a) If a solution to the above problem exists, then A is not nilpotent.
(b) If A represents a projection, then the above problem admits a solution.
(c) If $n = 2$ and if a solution to the above problem exists, then A is diagonalizable over $\mathbb{R}$.

Ans. ***Options a, b and c***

(a) $A = AXA = AAX = A^2X = A^2XAX = A^3X^2 = A^5X^4 = \cdots = A^{2^n+1} X^{2n}$ *for every* $n \in \mathbb{N}$. *Suppose that A is nilpotent, then* $A^m = 0$ *for some* $m \in \mathbb{N}$. *Then by the above equation, this implies that* $A = 0$, *which is a contradiction. Therefore A is not nilpotent.*
(b) If A is a projection $A^2 = A$. *Then take* $X = A$.
(c) When $n = 2$, *Since A is singular and A is not nilpotent, clearly A is diagonalizable over* $\mathbb{R}$.

(49) Let $A \in \mathbb{M}_{n\times n}(\mathbb{R})$ be a non-zero singular matrix that $tr(A) \neq 0$. Which of the following statements are true?
(a) For every such matrix A, the problem stated in the preceding exercise need not have a solution.
(b) For every such matrix A, the problem stated in the preceding exercise has

a solution given by $X = \frac{1}{tr(A)} A$.
(c) For every such matrix A, the problem stated in the preceding exercise has a solution given by $X = \frac{1}{(tr(A))^2} A$.

Ans. ***Option c***
The characteristic polynomial of A is $\lambda^2 - tr(A)\lambda = 0$. That is, A satisfies $A^2 = tr(A)A$ which implies $X = \frac{1}{(tr(A))^2} A$ is a solution for the given equations with such A.

(50) Let A be an $n \times n$ matrix with rank k. Consider the following statements:

(i) If A has real entries, then AA^T necessarily has rank k.
(ii) If A has complex entries, then AA^T necessarily has rank k.

Then
(a) (i) and (ii) are true. (b) (i) and (ii) are false.
(c) (i) is true and (ii) is false. (d) (i) is false and (ii) is true.

Ans. ***Option c***
First we prove that $\mathcal{N}(A^T) = \mathcal{N}(AA^T)$. Let $v \in \mathcal{N}(A^T)$.

$$\begin{aligned} v \in \mathcal{N}(A^T) &\Rightarrow A^T v = 0 \\ &\Rightarrow AA^T v = 0 \\ &\Rightarrow v \in \mathcal{N}(AA^T) \end{aligned}$$

Therefore $\mathcal{N}(A^T) \subset \mathcal{N}(AA^T)$. Now

$$\begin{aligned} v \in \mathcal{N}(AA^T) &\Rightarrow AA^T v = 0 \\ &\Rightarrow v^T AA^T v = 0 \\ &\Rightarrow (A^T v)^T A^T v = 0 \\ &\Rightarrow \|A^T v\| = 0 \; if \;\; A \; has \; real \; entries \\ &\Rightarrow A^T v = 0 \\ &\Rightarrow v \in \mathcal{N}(A^T) \end{aligned}$$

Therefore $\mathcal{N}(A^T) \supset \mathcal{N}(AA^T)$. Thus $\mathcal{N}(A^T) = \mathcal{N}(AA^T)$. As $Rank(A) = Rank(A^T)$, by Rank-Nullity theorem, If A has real entries, then AA^T necessarily has rank k.
But if A has complex entries, then AA^T need not have rank k. For example, consider the matrix $\begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$, then $AA^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

(51) Let A be an $n \times n$ real matrix. Pick the correct answer(s) from the following
(a) A has at least one real eigenvalue
(b) For all non-zero vectors $v, w \in \mathbb{R}^n$, $(Aw)^T(Av) > 0$
(c) Every eigenvalue of A^TA is a non-negative real number.
(d) $I + A^TA$ is invertible

Ans. ***Options c and d***
Consider the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. *The given matrix has no real eigenvalues. Take* $w = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and* $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. *Then* $(Aw)^T(Av) = 0$. *Let* λ *be an eigenvalue of* A^TA *with eigenvector* u, *then*

$$\lambda \|u\|^2 = \langle \lambda u, u \rangle = \langle A^TAu, u \rangle = \langle Au, Au \rangle \geq 0$$

Therefore every eigenvalue of A^TA *is a non-negative real number. Hence every eigenvalue of* $I + A^TA$ *is a non-negative real number.*

(52) Let T_1, T_2 be two linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $\{v_1, v_2, \ldots, v_n\}$ be a basis of $\mathbb{R}^n$. Suppose that $T_i v_i \neq 0$ for every $i = 1, 2, \ldots, n$ and $v_i \perp Ker(T_2)$ for every $i = 1, 2, \ldots, n$. Which of the following is/are necessarily true?
(a) T_1 is invertible. (b) T_2 is invertible.
(c) Both T_1, T_2 are invertible. (d) Neither T_1 nor T_2 is invertible.

Ans. ***Option b***
Since $\{v_1, v_2, \ldots, v_n\}$ *is a basis for* $\mathbb{R}^n$ *and* $v_i \perp Ker(T_2)$ *for every* $i = 1, 2, \ldots, n$, $Ker(T_2) = \phi \Rightarrow T_2$ *is one-one. Since* T_2 *is from* $\mathbb{R}^n$ *to* $\mathbb{R}^n$, T_2 *is invertible. Consider the linear transformation* T *by* $T(v_1) = \cdots = T(v_n) = v_1$. *Then* $T(v_i) \neq 0$ *for every* $i = 1, 2, \ldots, n$, *but* T *is not invertible.*

(53) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$T(x, y, z) = (x + y, y + z, z - x)$$

Then an orthonormal basis for the range of T is
(a) $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)\right\}$ (b) $\left\{\left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right)\right\}$
(c) $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right)\right\}$ (d) $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}\right)\right\}$

Ans. ***Option c***
We have

$$T(1, 0, 0) = (1, 0, -1) = 1(1, 0, 0) + 0(0, 1, 0) + (-1)(0, 0, 1)$$

$$T(0, 1, 0) = (1, 1, 0) = 1(1, 0, 0) + 1(0, 1, 0) + 0(0, 0, 1)$$

$$T(0,0,1) = (0,1,1) = 0(1,0,0) + 1(0,1,0) + 1(0,0,1)$$

Therefore the matrix of T is given by $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}$. *Since the third column is the linear combination of first and second column,*

$$\mathcal{R}(T) = span\{(1,1,0), (1,0,-1)\}$$

As $\{(1,1,0), (1,0,-1)\}$ *is not orthonormal we have to use Gram-Schmidt Orthonormalization. Take* $u_1 = (1,1,0)$ *and then*

$$u_2 = (1,0,-1) - \frac{\langle(1,1,0), (1,0,-1)\rangle}{\langle(1,1,0), (1,1,0)\rangle}(1,1,0) = \left(\frac{1}{2}, \frac{-1}{2}, -1\right)$$

Therefore $(1,-1,-2)$ *is an orthogonal vector to* $(1,1,0)$ *and*

$$span\{(1,1,0), (1,0,-1)\} = span\{(1,1,0), (1,-1,-2)\}$$

Therefore $\left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{-2}{\sqrt{6}}\right)\right\}$ *is an orthonormal basis for the range space of* T.

(54) Let A be an $m \times n$ matrix of rank m with $n > m$. If for some non-zero real number α, we have $u^T AA^T u = \alpha u^T u$ for all $u \in \mathbb{R}^m$ then $A^T A$ has
(a) exactly two distinct eigenvalues
(b) 0 as an eigenvalue with multiplicity $n - m$
(c) α as a non-zero eigenvalue
(d) exactly two non-zero distinct eigenvalues

Ans. ***Options a, b and c***
Since A is a matrix of rank m, by Sylvester's inequality $A^T A$ *is a matrix of rank m and hence 0 as an eigenvalue with multiplicity* $n - m$. *Let* λ *be a non-zero eigenvalue of* $A^T A$. *Then there exists* $v \neq 0$ *such that* $A^T Av = \lambda v$. *Take* $Av = u$. *Then* $A^T u = A^T Au = \lambda v$ *and* $AA^T u = \lambda Av = \lambda u$. *Multiplying with* u^T, *we get*

$$u^T AA^T u = \alpha u^T u = \lambda u^T u \Rightarrow \alpha \|u\| = \lambda \|u\|$$

Since $\|u\| \neq 0$ *this gives* $\alpha = \lambda$. *Therefore* α *is the only non-zero eigenvalue of* $A^T A$.

(55) Suppose V is a finite-dimensional non-zero vector space over $\mathbb{C}$ and $T : V \to V$ is a linear transformation such that $\mathcal{R}(T) = \mathcal{N}(T)$. Then which of the following statements is false?
(a) The dimension of V is even.
(b) 0 is the only eigenvalue of T.

(c) Both 0 and 1 are eigenvalues of T.
(d) $T^2 = 0$

Ans. ***Option c***
We have $\mathcal{N}(T) \subset \mathcal{N}(T^2)$. Now let $v \in (\mathcal{N}(T))^{\perp}$. Then $0 \neq T(v) \in \mathcal{R}(T)$. Since $\mathcal{R}(T) = \mathcal{N}(T)$, $T^2(v) = T(Tv) = 0$. That is, $v \in \mathcal{N}(T^2)$. As $V = \mathcal{N}(T) \oplus \mathcal{N}(T^2)$, $T^2 = 0$. Therefore 0 is the only eigenvalue of T. Also Since $\mathcal{R}(T) = \mathcal{N}(T)$, By Rank-Nullity Theorem, the dimension of V is even.

(56) Let V be a closed subspace of $L^2[0, 1]$ and let $f, g \in L^2[0, 1]$ be given by $f(x) = x$ and $g(x) = x^2$. If $V^{\perp} = span\{f\}$ and Pg is the orthogonal projection of g on V, then $(g - Pg)(x), x \in [0, 1]$ is
(a) $\frac{3}{4}x$ (b) $\frac{1}{4}x$ (c) $\frac{3}{4}x^2$ (d) $\frac{1}{4}x^2$

Ans. ***Option a***
We know that if P is a projection on V, $I - P$ is a projection on $V^{\perp}$. We have $\|f\|^2 = \int_0^1 x^2 dx = \frac{1}{3}$. Then

$$(g - Pg)(x) = \frac{\langle g, f\rangle}{\|f\|^2} f = 3\langle x^2, x\rangle x = \frac{3}{4}x$$

Chapter 13
Solved Problems—Bounded Linear Maps

(1) If $f : l^\infty \to \mathbb{R}$ be defined by $f(\{v\}) = v_2$, then the norm of f is
(a) 1 (b) 0 (c) 2 (d) $\frac{1}{2}$

Ans. ***Option a***
For $v = (v_1, v_2, \ldots) \in l^\infty$, *we have*

$$|f(v)| = |v_2| \le \sup_i |v_i| = \|v\|$$

Therefore $\|f\| \le 1$. *Also, for* $e_2 = (0, 1, \ldots) \in l^\infty$, *we have* $f(e_2) = 1$. *Hence* $\|f\| = 1$.

(2) Let $T : (C[0, 1], \|.\|_\infty) \to \mathbb{R}$ be defined by $T(f) = \int_0^1 2xf(x)dx$ for all $f \in C[0, 1]$. Then $\|T\|$ equals ......

Ans. *We have*

$$\|T\| = \sup_{\|f\|=1} \left| \int_0^1 2xf(x)dx \right| \le 2 \sup_{\|f\|=1} \int_0^1 |x||f(x)|dx \le 2 \int_0^1 xdx = 1$$

Thus $\|T\| \le 1$. *Now for* $f(x) = 1$, *we have* $\|f\| = 1$ *and* $\|T(f)\| = \int_0^1 2xdx = 1$. *Therefore* $\|T\| = 1$.

(3) Let $C[0, 1]$ be the real vector space of all continuous real valued functions on $[0, 1]$, and let T be a linear operator on $C[0, 1]$ given by

$$(Tf)(x) = \int_0^1 sin(x + y)f(y)dy, \ x \in [0, 1]$$

Then the dimension of range space of T equals ......

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4_13

Ans. *We have* $sin(x+y) = sin\ x\ cos\ y + cos\ x\ sin\ y$. *Then*

$$(Tf)(x) = sin\ x \int_0^1 cos\ yf(y)dy + cos\ x \int_0^1 sin\ yf(y)dy$$

Therefore, $T(f) \in span\{sin\ x, cos\ x\}$. *As* $\{sin\ x, cos\ x\}$ *is linearly independent the dimension of range space of* T *equals 2.*

(4) Let V be a real normed linear space of all real sequences with finitely many non zero terms, with supremum norm and $T : V \to V$ be a one-one and onto linear operator defined by

$$T(v_1, v_2, v_3, \ldots) = \left(v_1, \frac{v_2}{2^2}, \frac{v_3}{3^2}, \ldots\right)$$

Then which of the following is TRUE?
(a) T is bounded but T^{-1} is not bounded.
(b) T is not bounded but T^{-1} is bounded.
(c) Both T and T^{-1} are bounded.
(d) Neither T nor T^{-1} is not bounded.

Ans. ***Option a***
We have,

$$\|T(v_1, v_2, v_3, \ldots)\| = sup\left\{|v_1|, \left|\frac{v_2}{2^2}\right|, \ldots\right\} \leq sup\{|v_1|, |v_2|, \ldots\} = \|v\|$$

Thus T *is bounded. Now consider the map* $T_1 : V \to V$ *defined by*

$$T_1(v_1, v_2, v_3, \ldots) = \left(v_1, 2^2 v_2, 3^2 v_3, \ldots\right)$$

Observe that $T_1 T = T T_1 = I$. *That is,* $T_1 = T^{-1}$. *Also,*

$$\|T_1(v_1, v_2, v_3, \ldots)\| = sup\left\{|v_1|, 2^2|v_2|, 3^2|v_3|, \ldots\right\}$$

Clearly, T_1 *is not bounded.*

(5) Let $T : l^2 \to l^2$ be defined by

$$T((v_1, v_2, \ldots, v_n, \ldots)) = (v_2 - v_1, v_3 - v_2, \ldots, v_{n+1} - v_n, \ldots)$$

Then
(a) $\|T\| = 1$ (b) $\|T\| > 2$ but bounded.
(c) $1 < \|T\| \leq 2$ (d) $\|T\|$ is unbounded.

Ans. ***Option c***
Since $(a-b)^2 = a^2 + b^2 - 2ab \leq 2(a^2+b^2)$ *for* $a, b \in \mathbb{R}$, *for any* $v = (v_1, v_2, \ldots, v_n, \ldots) \in l^2$ *with* $\|v\| = 1$, *we have*

$$\|T(v)\|^2 = \sum_{n=1}^{\infty}|v_{n+1} - v_n|^2 \leq 2\left(\sum_{n=1}^{\infty}|v_{n+1}|^2 + \sum_{n=1}^{\infty}|v_n|^2\right) \leq 4$$

Therefore $\|T\| \leq 2$. *Also for* $v = (2, 0, \ldots, 0, \ldots)$, $\|T(v)\| = 2$. *Therefore* $\|T\| = 2$.

(6) Let $\{(e_n) : n = 1, 2, 3, \ldots\}$ be an orthonormal basis of a complex Hilbert space H. Consider the following statements:

P: There exists a bounded linear functional $f : H \to \mathbb{C}$ such that $f(e_n) = \frac{1}{n}$ for $n = 1, 2, 3, \ldots$

Q: There exists a bounded linear functional $g : H \to \mathbb{C}$ such that $g(e_n) = \frac{1}{\sqrt{n}}$ for $n = 1, 2, 3, \ldots$

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option b***
For any $x \in H$, *we have* $x = \sum_n \langle x, e_n\rangle e_n$. *Then*

$$f(x) = \sum_n \langle x, e_n\rangle f(e_n) = \sum_n \langle x, e_n\rangle \frac{1}{n}$$

and by Holder's inequality,

$$|f(x)|^2 \leq \left(\sum_n |\langle x, e_n\rangle|^2\right)\left(\sum_n \frac{1}{n^2}\right) \leq \frac{\pi^2}{6}\|x\|^2$$

Therefore f *is a bounded linear functional on* H. *Now consider* $H = l^2$ *with the inner product* $\langle v, w\rangle = \sum_n v_n\overline{w_n}$, *where* $v = (v_1, v_2, \ldots), w = (w_1, w_2, \ldots) \in l^2$. *Suppose that there exists a bounded linear functional* $g : l^2 \to \mathbb{C}$ *such that* $g(e_n) = \frac{1}{\sqrt{n}}$ *for* $n = 1, 2, 3, \ldots$. *Then by Riesz representation theorem, there exists* $v = (v_1, v_2, \ldots) \in l^2$ *such that* $g(x) = \langle x, v\rangle$. *Then, we have*

$$g(e_n) = \langle e_n, v\rangle = v_n = \frac{1}{\sqrt{n}}\ \forall\, n = 1, 2, 3, \ldots$$

But $v = \left(\frac{1}{\sqrt{n}}\right) \notin l^2$ *as* $\sum_n |v_n|^2 = \sum_n \frac{1}{n}$ *diverges. Therefore such a bounded linear functional* g *need not exist.*

(7) Let $C[0, 1]$ be the Banach space of real valued continuous functions on $[0, 1]$ equipped with supremum norm. Define $T : C[0, 1] \to C[0, 1]$ by

$$(Tf)(x) = \int_0^x xf(t)dt$$

Let $\mathcal{R}(T)$ be the range space of T. Consider the following statements:

P: T is a bounded linear operator.
Q: $T^{-1} : \mathcal{R}(T) \to C[0, 1]$ exists and is bounded.

Then
(a) both P and Q are TRUE. (b) P is TRUE and Q is FALSE.
(c) P is FALSE and Q is TRUE. (d) both P and Q are FALSE.

Ans. ***Option b***
For all $f \in C[0, 1]$, *we have*

$$\|T(f)\| = \sup_{x\in[0,1]} \left| \int_0^x xf(t)dt \right| \le \sup_{x\in[0,1]} \int_0^x |xf(t)|dt = \sup_{x\in[0,1]} |x| \int_o^x |f(t)|dt \le \|f\|$$

Consider $f_n(x) = x^n \in C[0, 1]$. *Clearly,* $\|f_n\| = 1$. *Now*

$$\|T(f_n)\| = \sup_{x\in[0,1]} \left| \int_0^x xt^n dt \right| = \sup_{x\in[0,1]} \left| \frac{x^{n+2}}{n+1} \right| = \frac{1}{n+1}$$

Thus there does not exist $\lambda \in \mathbb{K}$ *such that* $\|T(f)\| \ge \lambda \|f\|$ *for all* $f \in C[0, 1]$. *Therefore* T *is not injective.*

(8) Let

$$L^2[0, 10] = \{f : [0, 10] \to \mathbb{R} : f \text{ is Lebesgue measurable and } \int_0^{10} f^2 dx < \infty\}$$

equipped with the norm $\|f\| = \left(\int_0^{10} f^2 dx\right)^{\frac{1}{2}}$ and let T be a linear functional on $L^2[0, 10]$ given by

$$T(f) = \int_0^2 f(x)dx - \int_3^{10} f(x)dx$$

Then $\|T\|$ equals ......

Ans. *Since* $L^2[0, 10]$ *is a Hilbert space, by Riesz representation theorem, there exists* $g \in L^2[0, 10]$ *such that* $T(f) = \langle f, g\rangle = \int_0^{10} f(x)g(x)dx$ *and* $\|T\| = \|g\|$. *Take* $f_1(x) = \begin{cases} 1, & if\ 0 \le x \le 2 \\ 0, & 2 < x \le 10 \end{cases}$. *Then*

$$T(f_1) = \int_0^2 1dx = 2 = \int_0^2 g(x)dx$$

Now take $f_2(x) = \begin{cases} 0, & if\ 0 \le x < 3 \\ -1, & 3 \le x \le 10 \end{cases}$. *Then*

$$T(f_2) = \int_3^{10} 1dx = 7 = -\int_3^{10} g(x)dx$$

Take $g(x) = \begin{cases} 1, \ if \ 0 \le x \le 2 \\ 0, \ 2 < x < 3 \\ -1, \ 3 \le x \le 10 \end{cases}$ *. Then, clearly*

$$T(f) = \int_0^2 f(x)dx - \int_3^{10} f(x)dx = \langle f, g \rangle$$

As

$$\|g\|^2 = \int_0^{10} | \ g(x) \ |^2 \ dx = \int_0^2 1.dx + 0 + \int_3^{10} 1.dx = 9$$

we get, $\|T\| = \|g\| = 3.$

(9) Consider $C[-1, 1]$ equipped with supremum norm given by

$$\|f\|_\infty = \sup\{|f(x)| : x \in [-1, 1]\}$$

for $f \in C[-1, 1]$. Define a linear functional T on $C[0, 1]$ by

$$T(f) = \int_{-1}^0 f(x)dx - \int_0^1 f(x)dx$$

for all $f \in C[-1, 1]$. Then the value of $\|T\|$ is ......

Ans. *We have*

$$\begin{aligned}
\|T(f)\| &= \sup_{x\in[-1,1]} \left| \int_{-1}^0 f(x)dx - \int_0^1 f(x)dx \right| \\
&\le \sup_{x\in[-1,1]} \left| \int_{-1}^0 f(x)dx + \int_0^1 f(x)dx \right| \\
&= \sup_{x\in[-1,1]} \left| \int_{-1}^1 f(x)dx \right| \\
&\le 2 \left(\int_{-1}^1 |f(x)|dx \right)^{\frac{1}{2}} = 2 \, \|f\|
\end{aligned}$$

Now consider $f \in C[-1, 1]$ *defined by* $f(x) = -2x$ *for all* $x \in [-1, 1]$*. Then*

$$T(f) = \int_{-1}^0 -2xdx + \int_0^1 2xdx = 2$$

and hence $\|T\| = 2.$

(10) Let c_{00} be the vector space of all complex sequences having finitely many non-zero terms. Equip c_{00} with the inner product $\langle u, v\rangle = \sum_{n=1}^{\infty} u_n\overline{v_n}$ for all $u = (u_n)$ and $v = (v_n)$ in c_{00}. Define $f : c_{00} \to \mathbb{C}$ by $f(v) = \sum_{n=1}^{\infty} \frac{v_n}{n}$. Let N be the kernel of f.

(I) Which of the following is FALSE?
   (a) f is a continuous linear functional.
   (b) $\|f\| \le \frac{\pi}{\sqrt{6}}$
   (c) There does not exist any $u \in c_{00}$ such that $f(v) = \langle v, u\rangle$ for all $v \in c_{00}$.
   (d) $N^{\perp} \neq \{0\}$

(II) Which of the following is FALSE?
   (a) $c_{00} \neq N$ (b) c_{00} is not a complete inner product space.
   (c) N is closed. (d) $c_{00} = N \oplus N^{\perp}$

Ans. *(I)* ***Option d***

For $u = (u_n)$ and $v = (v_n)$ in c_{00}, we have

$$f(u+v) = \sum_{n=1}^{\infty} \frac{u_n + v_n}{n} = \sum_{n=1}^{\infty} \frac{u_n}{n} + \sum_{n=1}^{\infty} \frac{v_n}{n} = f(u) + f(v)$$

Also by Holders inequality, we have

$$\|f(v)\| = \left\| \sum_{n=1}^{\infty} \frac{v_n}{n} \right\| \le \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} |v_n|^2 \right)^{\frac{1}{2}} = \frac{\pi}{\sqrt{6}} \|v\|$$

Therefore f is a continuous linear functional with $\|f\| \le \frac{\pi}{\sqrt{6}}$.
Suppose there exist $u \in c_{00}$ such that $f(v) = \sum_{n=1}^{\infty} \frac{v_n}{n} = \langle v, u\rangle = \sum_{n=1}^{\infty} v_n\overline{u_n}$ for all $v \in c_{00}$. Let e_n denote the vector with 1 as the n^{th} entry and all other entries zero. Clearly, $e_n \in c_{00}$. Then

$$\overline{u_n} = \langle e_n, u\rangle = f(e_n) = \frac{1}{n}$$

This is a contradiction, since $u = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right) \notin c_{00}$. Therefore there does not exist any $u \in c_{00}$ such that $f(v) = \langle v, u\rangle$ for all $v \in c_{00}$.
Now let $v = (v_n) \in N^{\perp}$. Since $v \in c_{00}$, there exists m such that $v_n = 0$ for all $n > m$. For a fixed $j \in \{1, 2, \ldots m\}$, consider the element $u_j = (u_n) \in c_{00}$ with

$$u_n = \begin{cases} \overline{v_j}, \ if \ n = j \\ \frac{-(m+1)}{j}\overline{v_j}, if \ n = m+1 \\ 0, \ otherwise \end{cases}$$

As $f(u_j) = 0$ for all $j = 1, 2, \dots, m$, $u_j \in N^{\perp}$ for all $j = 1, 2, \dots, m$. Then $\langle v, u \rangle = |v_j|^2 = 0$ which implies that $v_j = 0$ for all $j = 1, 2, \dots, m$. Therefore $N^{\perp} = \{0\}$.

(II) ***Option d***

As $f(e_1) \neq 0$, $c_{00} \neq N$. We know that kernel of a linear operator is closed and c_{00} is not a complete inner product space. As $N^{\perp} = \{0\}$, $N \oplus N^{\perp} = N \neq c_{00}$.

(11) Let V be the Banach space of all complex $n \times n$ matrices equipped with the norm $\|A\| = \max_{1 \le i, j \le n} |a_{ij}|$. If $f : V \to \mathbb{C}$ is defined by $f(A) = tr(A)$, then
(a) f is not linear.
(b) f is linear but not continuous.
(c) f is bounded linear functional with $\|f\| = 1$.
(d) f is bounded linear functional with $\|f\| = n$.

Ans. ***Option d***
Take $A, B \in V$ and $\lambda \in \mathbb{C}$, then

$$f(\lambda A + B) = tr(\lambda A + B) = \lambda tr(A) + tr(B) = \lambda f(A) + f(B)$$

Therefore f is linear. Now

$$f(A) = tr(A) = a_{11} + a_{22} + \cdots + a_{nn} \le n \left(\max_{1 \le i, j \le n} |a_{ij}| \right) = n \|A\|$$

Therefore f is bounded and $\|f\| \le n$. Also for $A = I_n$, we have $f(A) = n$. Therefore $\|f\| = n$.

(12) Let H be a Hilbert space and let $\{e_n : n \ge 1\}$ be an orthonormal basis of H. Suppose that $T : H \to H$ is a bounded linear operator. Which of the following cannot be true?
(a) $T(e_n) = e_1$ for all $n \ge 1$ (b) $T(e_n) = e_{n+1}$ for all $n \ge 1$
(c) $T(e_n) = \sqrt{\frac{n+1}{n}} e_n$ for all $n \ge 1$ (d) $T(e_n) = e_{n-1}$ for all $n \ge 2$ and $T(e_1) = 0$

Ans. ***Option a***
Since $\{e_n : n \ge 1\}$ is an orthonormal basis of H, any element $v \in H$ can be written as $v = \sum_{n=1}^{\infty} \lambda_n e_n$, where $\lambda_i \in \mathbb{K}, i = 1, 2, \ldots \in \mathbb{N}$. Also, we have $\sum_{n=1}^{\infty} \lambda_n e_n \in H$ if and only if $\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$.

(a) Suppose that $T(e_n) = e_1$ for all $n \ge 1$. Then

$$T(v) = T\left(\sum_{n=1}^{\infty} \lambda_n e_n \right) = \sum_{n=1}^{\infty} \lambda_n T(e_n) = \sum_{n=1}^{\infty} \lambda_n e_1$$

Take $\lambda_n = \frac{1}{n}$. Since $\sum_{n=1}^{\infty} \left|\frac{1}{n}\right|^2 < \infty$, we have $v_0 = \sum_{n=1}^{\infty} \frac{1}{n} e_n \in H$. But

$$T(v_0) = \sum_{n=1}^{\infty} \frac{1}{n} e_1 \notin H$$

as $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. Therefore a bounded linear operator with $T(e_n) = e_1$ for all $n \geq 1$ is not possible.

(b) Suppose that $T(e_n) = e_{n+1}$ for all $n \geq 1$. Then as above

$$T(v) = \sum_{n=1}^{\infty} \lambda_n e_{n+1} = \sum_{n=2}^{\infty} \lambda_{n-1} e_n$$

As $v = \sum_{n=1}^{\infty} \lambda_n e_n \in H$, we have $\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$. Hence $\sum_{n=2}^{\infty} |\lambda_{n-1}|^2 < \infty$. Thus a linear operator with $T(e_n) = e_{n+1}$ for all $n \geq 1$ exists. For boundedness, we have

$$\|T(v)\|^2 = \left\| \sum_{n=2}^{\infty} \lambda_{n-1} e_n \right\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 = \|v\|^2$$

(c) Suppose that $T(e_n) = \sqrt{\frac{n+1}{n}} e_n$ for all $n \geq 1$. Then

$$T(v) = \sum_{n=1}^{\infty} \lambda_n \sqrt{\frac{n+1}{n}} e_n$$

Since $\sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$, we have

$$\sum_{n=1}^{\infty} \left| \lambda_n \sqrt{\frac{n+1}{n}} \right|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \, \frac{n+1}{n} \leq 2 \sum_{n=1}^{\infty} |\lambda_n|^2 < \infty$$

Thus a linear operator with $T(e_n) = \sqrt{\frac{n+1}{n}} e_n$ for all $n \geq 1$ exists. For boundedness, we have

$$\|T(v)\|^2 = \left\| \sum_{n=2}^{\infty} \lambda_n \sqrt{\frac{n+1}{n}} e_n \right\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2 \frac{n+1}{n} \leq 2 \sum_{n=1}^{\infty} |\lambda_n|^2 = 2 \, \|v\|^2$$

(d) Suppose that $T(e_n) = e_{n-1}$ for all $n \geq 2$ and $T(e_1) = 0$. Then

$$T(v) = \sum_{n=2}^{\infty} \lambda_n e_{n-1} = \sum_{n=1}^{\infty} \lambda_{n+1} e_n$$

Since $\sum_{n=1}^{\infty}|\lambda_n|^2 < \infty$, we have $\sum_{n=1}^{\infty}|\lambda_{n+1}|^2 < \infty$. Thus a linear operator with $T(e_n) = e_{n-1}$ for all $n \geq 2$ and $T(e_1) = 0$ exists. For boundedness, we have

$$\|T(v)\|^2 = \left\|\sum_{n=1}^{\infty}\lambda_{n+1}e_n\right\|^2 = \sum_{n=2}^{\infty}|\lambda_n|^2 \leq \sum_{n=1}^{\infty}|\lambda_n|^2 = \|v\|^2$$

(13) Let $e_i = (0, \ldots, 0, 1, 0, \ldots)$($e_i$ is the sequence with 1 at the i^{th} place and 0 elsewhere) for $i = 1, 2, \ldots$. Consider the following statements:

P: $\{f(e_i)\}$ converges for every continuous linear functional on l^2.
Q: $\{e_i\}$ converges on l^2.

Then which of the following holds?
(a) Both P and Q are TRUE. (b) P is TRUE and Q is not TRUE.
(c) P is not TRUE and Q is TRUE. (d) Neither P nor Q is TRUE.

Ans. ***Option b***
Since l^2 is a Hilbert space, by Riesz representation theorem, there exists $u \in l^2$ with $f(v) = \langle v, u\rangle$ for all $v \in l^2$. Then $f(e_i) = \langle e_i, u\rangle$. Then by Bessel's inequality, we have $\sum_{n=1}^{\infty}|\langle e_n, u\rangle|^2 \leq \|u\|^2$. Therefore the series $\sum_{n=1}^{\infty}\langle e_n, u\rangle$ converges. Hence $\langle e_i, u\rangle = f(e_i) \to 0$ on l^2. As $\|e_i - e_j\| = 2$ for all $i \neq j$, $\{e_i\}$ does not converges on l^2.

(14) Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be defined by $T\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} z_1 - iz_2 \\ iz_1 + z_2 \\ z_1 + z_2 + iz_3 \end{pmatrix}$. Then adjoint T^* of T is given by $T^*\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} =$

(a) $\begin{pmatrix} z_1 + iz_2 \\ -iz_1 + z_2 \\ z_1 + z_2 - iz_3 \end{pmatrix}$ (b) $\begin{pmatrix} z_1 - iz_2 + z_3 \\ -iz_1 + z_2 + z_3 \\ iz_3 \end{pmatrix}$ (c) $\begin{pmatrix} z_1 - iz_2 + z_3 \\ iz_1 + z_2 + z_3 \\ -iz_3 \end{pmatrix}$ (d) $\begin{pmatrix} iz_1 + z_2 \\ z_1 - iz_2 \\ z_1 - z_2 - iz_3 \end{pmatrix}$

Ans. ***Option c***
We have

$$T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + i\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$T\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -i \\ 1 \\ 1 \end{bmatrix} = (-i)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$T\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\i\end{bmatrix}=0\begin{bmatrix}1\\0\\0\end{bmatrix}+0\begin{bmatrix}0\\1\\0\end{bmatrix}+i\begin{bmatrix}0\\0\\1\end{bmatrix}$$

Therefore the matrix of T is given by $\begin{bmatrix}1 & -i & 0\\ i & 1 & 0\\ 1 & 1 & i\end{bmatrix}$. *Hence the matrix of T^* is given by* $\begin{bmatrix}1 & -i & 1\\ i & 1 & 1\\ 0 & 0 & -i\end{bmatrix}$ *and* $T^*\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\begin{pmatrix}z_1-iz_2+z_3\\ iz_1+z_2+z_3\\ -iz_3\end{pmatrix}$.

(15) Consider $\mathbb{R}^n$ with standard inner product. For a non-zero $w\in\mathbb{R}^n$, define $T_w:\mathbb{R}^n\to\mathbb{R}^n$ by

$$T_w(v)=v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w$$

Which of the following are true?
(a) $det\,(T_w)=1$ (b) $T_w=T_w^{-1}$
(c) $T_{2w}=2T_w$ (d) $\langle T_w(v_1),T_w(v_2)\rangle=\langle v_1,v_2\rangle$ for all $v_1,v_2\in\mathbb{R}^n$

Ans. ***Option b and c***

(a) Take $n=2$. Then,

$$T_w(w)=w-\frac{2\langle w,w\rangle}{\langle w,w\rangle}w=w-2w=-w$$

Therefore -1 is an eigenvalue of T_w. Now, take any non-zero element from $u\in(span\{w\})^{\perp}$. Then, $\langle u,w\rangle=0$ and hence

$$T_w(u)=u-\frac{2\langle u,w\rangle}{\langle w,w\rangle}w=u$$

Thus -1 and 1 are the eigenvalues of T_w. Hence $det\,(T_w)=-1$, when $n=2$.

(b) We have,

$$\begin{aligned}T_w\left(T_w(v)\right)&=T_w\left(v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w\right)\\&=v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w-\frac{2\left\langle v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w,w\right\rangle}{\langle w,w\rangle}w\\&=v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w-\frac{2}{\langle w,w\rangle}\left[\langle v,w\rangle-\frac{2\langle v,w\rangle}{\langle w,w\rangle}\langle w,w\rangle\right]w\\&=v-\frac{2\langle v,w\rangle}{\langle w,w\rangle}w+\frac{2\langle v,w\rangle}{\langle w,w\rangle}w=v\end{aligned}$$

(c) For all $v \in \mathbb{R}^n$,

$$T_{2w}(v) = v - \frac{2\langle v, 2w\rangle}{\langle 2w, 2w\rangle}(2w) = v - \frac{8\langle v, w\rangle}{4\langle w, w\rangle}w = v - \frac{2\langle v, w\rangle}{\langle w, w\rangle}w = T_w(v)$$

(d) For $v_1, v_2 \in \mathbb{R}^n$,

$$\begin{aligned}
\langle T_w(v_1), T_w(v_2)\rangle &= \left\langle v_1 - \frac{2\langle v_1, w\rangle}{\langle w, w\rangle}w, v_2 - \frac{2\langle v_2, w\rangle}{\langle w, w\rangle}w\right\rangle \\
&= \langle v_1, v_2\rangle - \left\langle v_1, \frac{2\langle v_2, w\rangle}{\langle w, w\rangle}w\right\rangle - \left\langle \frac{2\langle v_1, w\rangle}{\langle w, w\rangle}w, v_2\right\rangle \\
&+ \left\langle \frac{2\langle v_1, w\rangle}{\langle w, w\rangle}w, \frac{2\langle v_2, w\rangle}{\langle w, w\rangle}w\right\rangle \\
&= \langle v_1, v_2\rangle - \frac{2\langle v_2, w\rangle}{\langle w, w\rangle}\langle v_1, w\rangle - \frac{2\langle v_1, w\rangle}{\langle w, w\rangle}\langle w, v_2\rangle \\
&+ \frac{4\langle v_1, w\rangle\langle v_2, w\rangle}{\langle w, w\rangle\langle w, w\rangle}\langle w, w\rangle \\
&= \langle v_1, v_2\rangle
\end{aligned}$$

(16) Let A be a 2×2 complex matrix such that A^*A is the identity matrix, where A^* is the conjugate transpose of A. Then the eigenvalues of A are
(a) real (b) complex conjugates of each other
(c) of modulus 1 (d) reciprocals of each other

Ans. ***Option c***

Consider the matrix $A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$. *Then* $AA^* = I$, *but the eigenvalues are not real, eigenvalues are not complex conjugates of each other and eigenvalues are not reciprocals of each other.*

Now let A be a 2×2 *complex matrix such that* $A^*A = I$. *Let* λ *be an eigenvalue of A. Then there exists* $v \neq 0$ *such that* $Av = \lambda v$.

$$Av = \lambda v \Rightarrow v^*A^*Av = v^*A^*\lambda v \Rightarrow v^*v = \lambda\overline{\lambda}v^*v \Rightarrow \|v\|^2 = |\lambda|^2\,\|v\|^2$$

Since $\|v\|^2 \neq 0$, *we get* $|\lambda| = 1$.

(17) Consider the following statements P and Q :
P : If A is an $n \times n$ complex matrix, then $\mathcal{R}(A) = [\mathcal{N}(A^*)]^{\perp}$.
Q : There exists unitary matrix with an eigenvalue λ such that $|\lambda| < 1$.
Which of the above statements hold TRUE?
(a) both P and Q (b) only P (c) only Q (d) Neither P nor Q

Ans. ***Option b***

P : Let H be a Hilbert space and $A \in BL(H)$. Then we have $\mathcal{R}(A) = [\mathcal{N}(A^*)]^{\perp}$.
Q : From the above problem, $|\lambda| = 1$.

(18) For the matrix

$$A = \begin{bmatrix} 2 & 3+2i & -4 \\ 3-2i & 5 & 6i \\ -4 & -6i & 3 \end{bmatrix}$$

which of the following statements are correct?
P : A is skew-Hermitian and iA is Hermitian.
Q : A is Hermitian and iA is skew-Hermitian.
R : eigenvalues of A are real.
S : eigenvalues of iA are real.
(a) P and R only. (b) Q and R only.
(c) P and S only. (d) Q and S only.

Ans. ***Option b***
We have $A^* = \begin{bmatrix} 2 & 3+2i & -4 \\ 3-2i & 5 & 6i \\ -4 & -6i & 3 \end{bmatrix} = A$, A is Hermitian. Since $iA = \begin{bmatrix} i2 & 3i-2 & -4i \\ 3i+2 & 5i & -6 \\ -4i & 6 & 3i \end{bmatrix}$, $(iA)^* = \begin{bmatrix} -i2 & -3i+2 & 4i \\ -3i-2 & -5i & 6 \\ 4i & -6 & -3i \end{bmatrix} = -iA$, iA is skew-Hermitian.The eigenvalues of a Hermitian matrix are always real.
Let λ be an eigenvalue of a Hermitian matrix A, then $Ax = \lambda x$. Now

$$Ax = \lambda x \Rightarrow x^*Ax = x^*\lambda x \Rightarrow x^*Ax = \lambda \|x\|^2$$

Now taking conjugate transpose on both sides,

$$x^*A^*x = \overline{\lambda} \|x\|^2 \Rightarrow x^*Ax = \overline{\lambda} \|x\|^2$$

Now, $\lambda \|x\|^2 = \overline{\lambda} \|x\|^2 \Rightarrow \lambda = \overline{\lambda}$. Hence, λ is real. Similarly we can prove that The eigenvalues of a skew-Hermitian matrix are always either zero or purely imaginary.

(19) Let A be an invertible Hermitian matrix and let $a, b \in \mathbb{R}$ be such that $a^2 < 4b$. Then
(a) both $A^2 + aA + bI$ and $A^2 - aA + bI$ are singular.
(b) $A^2 + aA + bI$ is singular but $A^2 - aA + bI$ is non-singular.
(c) $A^2 + aA + bI$ is non-singular but $A^2 - aA + bI$ is singular.
(d) both $A^2 + aA + bI$ and $A^2 - aA + bI$ are non-singular.

Ans. ***Option d***
Let λ be an eigenvalue of A, then since A is Hermitian, λ must be real. Also $\lambda^2 + a\lambda + b$ is an eigenvalue of $A^2 + aA + bI$. Now $A^2 + aA + bI$ is singular

if $\lambda^2 + a\lambda + b = 0$ *for some* $\lambda \in \mathbb{R}$. *Since* $a^2 < 4b$, $\lambda^2 + a\lambda + b = 0 \Rightarrow \lambda$ *must be complex, which is a contradiction. Similarly for* $A^2 - aA + bI$ *is singular if* $\lambda^2 - a\lambda + b = 0$ *for some* $\lambda \in \mathbb{R}$. *In this case also since* $a^2 < 4b$, $\lambda^2 + a\lambda + b = 0 \Rightarrow \lambda$ *must be complex, which is a contradiction. Therefore both* $A^2 + aA + bI$ *and* $A^2 - aA + bI$ *are non-singular.*

(20) The matrix $A = \begin{bmatrix} cos\ \alpha & sin\ \alpha \\ i\ sin\ \alpha & i\ cos\ \alpha \end{bmatrix}$ is a unitary matrix when α is
(a) $(2n+1)\frac{\pi}{2}, n \in \mathbb{Z}$ (b) $(3n+1)\frac{\pi}{3}, n \in \mathbb{Z}$
(c) $(4n+1)\frac{\pi}{4}, n \in \mathbb{Z}$ (d) $(5n+1)\frac{\pi}{5}, n \in \mathbb{Z}$

Ans. ***Option a***
A is unitary when $AA^* = I$. *Now,*

$$AA^* = I \Rightarrow \begin{bmatrix} cos\ \alpha & sin\ \alpha \\ i\ sin\ \alpha & i\ cos\ \alpha \end{bmatrix} \begin{bmatrix} cos\ \alpha & -i\ sin\ \alpha \\ sin\ \alpha & -i\ cos\ \alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 1 & -isin\ 2\alpha \\ isin\ 2\alpha & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow \alpha \in \left\{(2n+1)\frac{\pi}{2}, n \in \mathbb{Z}\right\}$$

(21) Pick out the true statements:
(a) The eigenvalues of a unitary matrix are all equal to ± 1.
(b) The determinant of real orthogonal matrix is always ± 1.

Ans. *A is orthogonal if* $AA^T = I$. *Taking determinant on both sides we get,*

$$det(AA^T) = det(I) \Rightarrow det(A)det(A^T) = 1 \Rightarrow (det(A))^2 = 1 \Rightarrow det(A) = \pm 1$$

Consider the matrix $A = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix}$, *then* $AA^* = I$. *Clearly, A is unitary but the eigenvalues are not* ± 1.

(22) Let A be an $n \times n$ matrix with real entries. Pick out the true statements:
(a) There exists a real symmetric $n \times n$ matrix B such that $B^2 = A^*A$.
(b) If A is symmetric, there exists a real symmetric $n \times n$ matrix B such that $B^2 = A$.
(c) If A is symmetric, there exists a real symmetric $n \times n$ matrix B such that $B^3 = A$.

Ans. ***Option a and c***

(a) Since A is a matrix with real entries, $A^* = A^T$ *and* A^TA *is a real symmetric matrix. Also* A^TA *is positive semi definite. As every real symmetric matrix is diagonalizable there exists an orthonormal matrix P such that* $PDP^{-1} = PDP^T = A^TA$, *where D is a diagonal matrix with eigenvalues of* A^TA *as diagonal entries. Now take* $B = PD^{\frac{1}{2}}P^T$. *B is well defined since the diagonal entries of D are non negative real numbers. Then*

$$B^2 = PD^{\frac{1}{2}}P^TPD^{\frac{1}{2}}P^T = PDP^T = A^TA$$

(b) *Consider* $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. *Then there does not exists a real symmetric* 2×2 *matrix* B *such that* $B^2 = A$. *For if* $B = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$. *Then* $B^2 = \begin{bmatrix} a^2+b^2 & ab+bd \\ ab+bd & b^2+d^2 \end{bmatrix}$. *As* $a, b \in \mathbb{R}$, $a^2 + b^2$ *is a non-negative real number.*

(c) *Since A is real symmetric matrix, there exists an orthonormal matrix P such that* $PDP^{-1} = PDP^T = A$, *where D is a diagonal matrix with eigenvalues of A as diagonal entries. Now take* $B = PD^{\frac{1}{3}}P^T$. *then B is well-defined and* $B^3 = A$.

(23) Let $S = \{\lambda_1, \dots \lambda_n\}$ be an ordered set of n real numbers, not all equal, but not all necessarily distinct. Pick out the true statements:
(a) There exists an $n \times n$ matrix with complex entries, which is not self-adjoint, whose set of eigenvalues is given by S.
(b) There exists an $n \times n$ self-adjoint, non-diagonal matrix with complex entries whose set of eigenvalues is given by S.
(c) There exists an $n \times n$ symmetric, non-diagonal matrix with real entries whose set of eigenvalues is given by S.

Ans. ***Options a, b and c***

(a) *Consider* $A = \begin{bmatrix} \lambda_1 & \alpha & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix}$, *where* $\alpha \in \mathbb{C}$. *Then A is an* $n \times n$ *matrix with complex entries, which is not self-adjoint, whose set of eigenvalues is given by S.*

(b) *Consider* $A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$. *Now take a matrix P such that its columns form an orthonormal basis for* $\mathbb{R}^n$ *and P does not commute with A. Now consider the matrix* $B = PAP^T$. *Clearly B is a self-adjoint matrix which is non-diagonal and the set of eigenvalues of B is S.*

(c) *Consider the matrix B as above.*

(24) Pick out the true statements:
(a) Let A be a hermitian $n \times n$ positive definite matrix. Then, there exists a hermitian positive definite $n \times n$ matrix B such that $B^2 = A$.
(b) Let B be a non-singular $n \times n$ matrix with real entries. Then $B^T B$ is a symmetric and positive definite matrix.

Ans. *(a) A hermitian matrix is unitary diagonalizable. That is, there exists a unitary matrix U such that* $UDU^* = A$. *Since A is positive definite, the* *diagonal entries must be positive. Now define the matrix* B *as* $B = UD^{\frac{1}{2}}U^*$. *Then* $B^2 = UD^{\frac{1}{2}}U^*UD^{\frac{1}{2}}U^* = UDU^* = A$.

(b) Since $(B^TB)^T = B^T(B^T)^T = B^TB$, B^TB *is symmetric. Also,*

$$v^TB^TBv = (Bv)^TBv = \|Bv\|^2 \geq 0$$

Therefore B^TB *is a positive definite matrix.*

(25) A 3×3 real symmetric matrix A admits $(1, 2, 3)^T$ and $(1, 1, -1)^T$ transpose as eigenvectors. The transpose of which of the following is surely an eigenvector for A? Choose all the correct options.
(a) $(1, -1, 0)$ (b) $(-5, 1, 1)$ (c) $(3, 2, 1)$ (d) none of the above

Ans. ***Option d is correct***
Since the eigenvectors corresponding different eigenvalues of a symmetric matrices are orthogonal, none of the above three vectors can be an eigenvector of A for sure.

(26) Let A be a 3×3 real symmetric matrix with eigenvalues 0,2 and a with the respective eigenvectors $u = (4, b, c)^T$, $v = (-1, 2, 0)^T$ and $w = (1, 1, 1)^T$. Consider the following statements:

I. $a + b - c = 10$
II. The vector $x = \left(0, \frac{3}{2}, \frac{1}{2}\right)^t$ satisfies $Ax = v + w$.
III. For any $d \in span\{u, v, w\}$, $Ax = d$ has a solution.
IV. The trace of the matrix $A^2 + 2A$ is 8.

Which of the following statements are TRUE?
(a) I, II and III only. (b) I and II only.
(c) II and IV only. (d) III and IV only.

Ans. ***Option b***

(I) Since the eigenvectors of a real symmetric matrix corresponding to distinct eigenvalues are orthogonal, $u.v = -4 + 2b = 0 \Rightarrow b = 2$. *As* $v.w = -1 + 2 = 1$, $a = 2$. *Now* $u.w = 6 + c = 0 \Rightarrow c = -6$. *Therefore* $a + b - c = 10$.

(II) Clearly $x = \left(0, \frac{3}{2}, \frac{1}{2}\right)^T = \frac{1}{2}(v + w)$. *Therefore*

$$Ax = \frac{1}{2}A(v + w) = \frac{1}{2}(Av + Aw) = \frac{1}{2}(2v + 2w) = v + w$$

($Av = 2v$ *and* $Aw = 2w$, *since v and w are eigenvectors of A with eigenvalue 2.)*

(III) Since $\{u, v, w\}$ *is linearly independent* $span\{u, v, w\} = \mathbb{R}^3$ *and as* 0 *is an eigenvalue of A,* $Rank(A) = 2$. *Therefore for any* $d \in span\{u, v, w\}$, $Ax = d$ *need not have a solution.*

(IV) *Since the eigenvalues of A are* $0, 2$ *and* 2, *the eigenvalues of* A^2 *are* $0, 4$ *and* 4. *Now*

$$tr(A^2 + 2A) = tr(A^2) + 2tr(A) = 8 + 8 = 16$$

(27) Let A be a real symmetric $n \times n$ matrix whose eigenvalues are 0 and 1. Let the dimension of the null space of $A - I$ be m. Pick out the true statements :
(a) The characteristic polynomial of A is $(\lambda - 1)^m \lambda^{m-n}$.
(b) $A^k = A^{k+1}$ for all positive integers k.
(c) The rank of A is m.

Ans. ***Options a and c***
Dimension of null space of $A - I$ *is* $m \Rightarrow 1$ *is an eigenvalue with multiplicity* m. *Consider* $A = [\frac{1}{2}]$. *Then Option* (b) *is incorrect. As A is real symmetric, it is diagonalizable. Therefore,* $Rank(A) =$ *number of non zero eigenvalues* $= m$.

(28) Let A be an $n \times n$ complex matrix. Assume that A is self-adjoint and let B denote the inverse of $A + iI_n$. Then all eigenvalues of $(A - iI_n)B$ are
(a) purely imaginary (b) real
(c) of modulus one (d) of modulus less than one

Ans. ***Option c***
Since A is self-adjoint, every eigenvalue of A is real. Let λ *be an eigenvalue of A, then* $\lambda + i$ *is an eigenvalue of* $A + iI_n$ *and* $\lambda - i$ *is an eigenvalue of* $(A - iI_n)$.*(As* $Av = \lambda v \Rightarrow (A + iI_n)v = \lambda v + iv = (\lambda + i)v$*) Also* $\frac{1}{\lambda - i}$ *is an eigenvalue of* $(A - iI_n)^{-1}$. *Now*

$$(A + iI_n)(A - iI_n)^{-1}v = (A + iI_n)\frac{1}{\lambda - i}v = \frac{\lambda + i}{\lambda - i}v$$

Hence, $\frac{\lambda+i}{\lambda-i}v$ *is an eigenvalue of* $(A + iI_n)(A - iI_n)^{-1}$ *and it has modulus one.*

(29) Let A be a real symmetric matrix and $B = I + iA$, where $i^2 = -1$. Then
(a) B is invertible if and only if A is invertible
(b) all eigenvalues of B are necessarily real
(c) $B - I$ is necessarily invertible
(d) B is necessarily invertible

Ans. ***Option d***
Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. *Then* $B = I + iA = \begin{bmatrix} 1 + i & 0 \\ 0 & 1 \end{bmatrix}$. *Then options a, b and c are are false. The eigenvalue of a real symmetric matrix are always real. Therefore the eigenvalues of* $B = I + iA$ *are of the form* $1 + i\lambda$ *where* $\lambda \in \mathbb{R}$ *is an eigenvalue of A. This cannot be zero. Therefore B is necessarily invertible.*

(30) Which of the following 3×3 matrices are diagonalizable over $\mathbb{R}$?

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 3 & 4 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Ans. ***Options a and c***

(a) Since the eigenvalues of an upper triangular matrices are its diagonal entries, the given matrix has distinct eigenvalues and hence is diagonalizable.

(b) The characteristic polynomial of the given matrix is

$$x^3 - x^2 + x - 1 = (x - 1)(x^2 + 1)$$

Since it has complex roots the given matrix is not diagonalizable over $\mathbb{R}$.

(c) Since a real symmetric matrix is always diagonalizable, the given matrix is diagonalizable.

(d) The characteristic polynomial of the given matrix is x^3 *as it is an upper triangular matrix with all diagonal entries zero. The minimal polynomial is also* x^3 *and hence the given matrix is not diagonalizable.*

(31) Let A be a real symmetric matrix. Then we can conclude that

(a) A does not have 0 as an eigenvalue

(b) All eigenvalues of A are real

(c) If A^{-1} exists, then A^{-1} is real and symmetric

(d) A has at least one positive eigenvalue

Ans. ***Options b and c***

Consider the matrix $A = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$. *Then A is real symmetric an the eigenvalues of A are 0 and* -1. *Eigenvalues of a real symmetric matrix A are always real. Also* $(A^{-1})^T = (A^T)^{-1} = A^{-1}$.

(32) The distinct eigenvalues of the matrix $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ are

(a) 0 and 1 (b) 1 and –1 (c) 1 and 2 (d) 0 and 2

Ans. ***Option d***

As the given matrix is symmetric, it is diagonalizable. Therefore it has only one non-zero eigenvalue, as the given matrix has rank 1. Also the matrix has trace 2. Therefore the distinct eigenvalues are 0 and 2.

(33) The number of distinct eigenvalues of the matrix $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

(a) 1 (b) 2 (c) 3 (d) 4

Ans. ***Option b***

As the given matrix is symmetric, it is diagonalizable. Since the matrix has rank 1*, the matrix has only one non zero eigenvalue. As the eigenvalues of a real symmetric matrix are real, the number of distinct eigenvalues is 2.*

(34) Let J denote a 101×101 matrix with all the entries equal to 1, and let I denote the identity matrix of order 101. Then the determinant of $J - I$ is
(a) 101 (b) 1 (c) 0 (d) 100

Ans. ***Option d***

From the above problem, the only eigenvalues of A are 0 and 1 with multiplicities 100 and 1 respectively. Therefore the eigenvalues of $J - I$ *are* -1 *and 100 with multiplicities 100 and 1 respectively . Therefore the determinant of* $J - I$ *is 100.*

(35) The possible set of eigenvalues of a 4×4 skew-symmetric orthogonal real matrix is
(a) $\{\pm i\}$ (b) $\{\pm i, \pm 1\}$ (c) $\{\pm 1\}$ (d) $\{0, \pm i\}$

Ans. ***Option a***

The eigenvalues of a skew-symmetric matrix are either zero or purely imaginary and the eigenvalues of an orthogonal matrix is of modulus 1. hence the possible set of eigenvalues of a skew-symmetric orthogonal real matrix is $\{\pm i\}$.

(36) Let $A_{n\times n} = (a_{ij}), n \geq 3$, where $a_{ij} = ({b_i}^2 - {b_j}^2)\ (i, j = 1, 2, 3, \ldots)$ for some distinct real numbers $b_1, b_2, b_3, \ldots, b_n$. Then $det(A)$ is
(a) $\Pi_{i<j}(b_i - b_j)$ (b) $\Pi_{i<j}(b_i + b_j)$ (c) 0 (d) 1

Ans. ***Option c***

The matrix is given by $A = \begin{bmatrix} 0 & {b_1}^2 - {b_2}^2 & \cdots & {b_1}^2 - {b_n}^2 \\ {b_2}^2 - {b_1}^2 & 0 & \cdots & {b_2}^2 - {b_n}^2 \\ \vdots & \vdots & \ddots & \vdots \\ {b_n}^2 - {b_1}^2 & {b_n}^2 - {b_2}^2 & \cdots & 0 \end{bmatrix}$. *Clearly A is a skew-symmetric matrix. Since the determinant of the odd order skew symmetric matrix is zero, from the given options, option* (c) *is correct.*

(37) Let A be a 3×3 non zero, skew-symmetric real matrix. Then
(a) A is invertible.
(b) the matrix $I + A$ is invertible.
(c) there exists a non-zero real number α such that $\alpha I + A$ is not invertible.
(d) all eigenvalues of A are real.

Ans. ***Option b***

A skew-symmetric matrix of odd order is always singular, hence not invertible.The characteristic polynomial of A is of degree 3. Since the eigenvalues of a skew-symmetric matrix are either purely imaginary or 0. Clearly, 0 is an eigenvalue of A. Also the eigenvalues of $I + A$ *is of the form* $1, 1 + ai, 1 - ai$ *where* $a \in \mathbb{R}$*. As their product is never zero* $I + A$ *is invertible.The eigenvalues* *of $\alpha I + A$ is of the form $\alpha, \alpha + ai, \alpha - ai$ where $a \in \mathbb{R}$. Their product is zero only if $\alpha = 0$.*

(38) Let $A = (a_{ij}) \in M_3(\mathbb{R})$ be such that $a_{ij} = -a_{ji}$ for all $1 \leq i, j \leq 3$. If $3i$ is an eigenvalue of A, find its other eigenvalues.

Ans. *The given matrix is an odd order skew symmetric matrix. Hence its determinant is zero which implies zero is an eigenvalue of the given matrix. Since the characteristic polynomial has real coefficients complex roots occur as conjugate pairs. Therefore its other eigenvalues are $-3i, 0$*

(39) Which of the following are non-singular?
(a) $I + A$ where $A \neq 0$ is a skew symmetric real $n \times n$ matrix, $n \geq 2$.
(b) Every skew symmetric non zero real 5×5 matrix.
(c) Every skew symmetric non zero real 2×2 matrix.
(d) All the above.

Ans. ***Options a and c***

(a) The eigenvalues of a skew symmetric matrix are either zero or purely imaginary. Since the matrix is real, complex eigenvalues occur in conjugate pair. So eigenvalues of $I + A$ are of the form $1 + ai, 1 - ai, 1$. Therefore $I + A$ is non singular.
(b) The characteristic polynomial is of odd degree with real coefficients. Hence it has atleast one real eigenvalue which will be 0. Therefore, it will be singular.
(c) Since the diagonal entries of a skew symmetric matrix must be zero, every skew symmetric non zero real 2×2 matrix will be of the form $A = \begin{bmatrix} 0 & a+bi \\ -a+bi & 0 \end{bmatrix}$. Clearly they are non singular.

(40) Let A be a 5×5 skew-symmetric matrix with entries in $\mathbb{R}$ and B be the 5×5 symmetric matrix whose (i, j)th entry is the binomial coefficient $\binom{i}{j}$ for $1 \leq i \leq j \leq 5$. Consider the 10×10 matrix, given in the block form by $C = \begin{pmatrix} A & A+B \\ 0 & B \end{pmatrix}$. Then
(a) $det(C) = 1$ or -1 (b) $det(C) = 0$ (c) $Tr(C) = 0$ (d) $Tr(C) = 5$

Ans. ***Option b and d***
Since the determinant of a odd order skew symmetric matrix is zero, $det(A) = 0$ and hence $det(C) = det(A)det(B) = 0$. Since the diagonal entries of a skew symmetric matrix are zero, $tr(A) = 0$ and the diagonal entries of B are all 1. Therefore $tr(C) = tr(A) + tr(B) = 5$.

(41) Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation whose matrix with respect to the standard basis of $\mathbb{R}^3$ is $\begin{bmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{bmatrix}$, where a, b, c are real numbers not all

zero. Then T
(a) is one-one (b) is onto
(c) has $rank$ 1 (d) does not map any line through origin to itself

Ans. ***Option d***
Since T is an odd order skew-symmetric matrix, it has determinant 0 and hence has $rank < 3$. Therefore T is not both one-one and onto. Consider the square sub matrices $\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}$. As not all a, b, c are zero, atleast one of these matrices have non-zero determinant and hence T can have $rank$ 2. Since the matrix is skew-symmetric the eigenvalues are zero or purely imaginary and any line in $\mathbb{R}^3$ is of the form $\{\lambda v : \lambda \in \mathbb{R}\}$ for some non-zero vector v. Therefore T does not map any line through origin to itself.

(42) Let $A = (a_{ij})$ be an $n \times n$ complex matrix and let A^* denote the conjugate transpose of A. Which of the following statements are true?
(a) If A is invertible, then $tr(A^*A) \neq 0$.
(b) If $tr(A^*A) \neq 0$, then A is invertible.
(c) If $|tr(A^*A)| < n^2$, then $|a_{ij}| < 1$ for some i, j.
(d) $tr(A^*A) = 0$, then A is the zero matrix.

Ans. ***Options a, c and d***
Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$. *Then* $A^* = \begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \dots & \overline{a_{n1}} \\ \overline{a_{12}} & \overline{a_{22}} & \dots & \overline{a_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \dots & \overline{a_{nn}} \end{bmatrix}$ *and hence* $tr(AA^*)$ $= \sum_{i,j=1}^{n} |a_{ij}|^2$.

(a) Suppose that A is invertible. If $tr(AA^*) = \sum_{i,j=1}^{n} |a_{ij}|^2 = 0$, then $|a_{ij}| = 0$ for all i and j and hence the matrix is the zero matrix and is not invertible. Therefore $tr(A^*A) \neq 0$, if A is invertible.
(b) Consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, then $A^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $AA^* = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Here $tr(A^*A) \neq 0$, but A is not invertible.
(c) $|tr(A^*A)| = \sum_{i,j=1}^{n} |a_{ij}|^2$. If $|a_{ij}| > 1$ for all i, j, then $tr(A^*A) > n^2$. Therefore if $|tr(A^*A)| < n^2$, then $|a_{ij}| < 1$ for some i, j.
(d) By option (a) if $tr(A^*A) = 0$, then A is the zero matrix.

(43) Let A be a real 3×4 matrix of rank 2. Then the rank of $A^T A$ is:
(a) exactly 2 (b) exactly 3
(c) exactly 4 (d) at most 2 but not necessarily 2

Ans. ***Option a***
Since $Rank(A) = Rank(A^T A)$, rank of $A^T A$ is 2.

(44) Let S be the set of 3×3 real matrices A with $A^T A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. Then the set S contains

(a) a nilpotent matrix. (b) a matrix of rank one.
(c) a matrix of rank two. (d) a non-zero skew-symmetric matrix.

Ans. ***Option a and b***

Consider the matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, *then* $A^T A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow A \in S$. *Clearly A is nilpotent and has rank one. Since* $Rank(A^T A) = Rank(A)$, *S does not contain a set with rank two. For a non-zero skew symmetric matrix B,* $B^T B = (-B)B = -B^2$ *and hence S does not contain a non-zero skew-symmetric matrix.*

(45) Let $A = \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$ and $Q(X) = X^T AX$ for $X \in \mathbb{R}^3$. Then

(a) A has exactly two positive eigenvalues
(b) all the eigenvalues of A are positive
(c) $Q(X) \geq 0$ for all $X \in \mathbb{R}^3$
(d) $Q(X) < 0$ for some $X \in \mathbb{R}^3$

Ans. ***Options a and d***

The characteristic polynomial of the given matrix is

$$x^3 - 6x^2 - 3x + 18 = (x - 6)(x^2 - 3)$$

Therefore the eigenvalues of A are $6, \pm\sqrt{3}$.

(46) The matrix $\begin{pmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$ is

(a) positive definite.
(b) non-negative definite but not positive definite.
(c) negative definite.
(d) neither negative definite nor positive definite.

Ans. ***Option a***

The characteristic polynomial of the given matrix is

$$x^3 - 8x^2 + 19x - 12 = (x - 1)(x - 3)(x - 4)$$

Since all the eigenvalues of the given matrix are positive, the given matrix is positive definite.

(47) Let $V=\{f : [0, 1] \to \mathbb{R} \mid f\, is\, a\, polynomial\, of\, degree\, less\, than\, or\, equal\, to\, n\}$. Let $f_j(x) = x^j$ for $0 \leq j \leq n$ and let A be the $(n+1) \times (n+1)$ matrix given by $a_{ij} = \int_0^1 f_i(x) f_j(x) dx$. Then which of the following is/are true?
(a) $dim(V) = n$ (b) $dim(V) > n$ (c) $det(A) > 0$ (d) A is non-negative definite

Ans. ***Options b, c and d***
$v^T Av = \sum_{i,j} \int_0^1 v_i f_i(x) v_j f_j(x) dx = \int_0^1 \left(\sum_i v_i f_i(x)\right)^2 dx > 0$ *for any* $v \neq 0$.

(48) Which of the following matrices are positive definite?
(a) $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$

Ans. ***Options a and c***

(a) The characteristic polynomial of the given matrix is

$$x^2 - 4x + 3 = (x-3)(x-1)$$

Clearly the matrix is positive definite.

(b) The characteristic polynomial of the given matrix is

$$x^2 - 2x - 3 = (x-3)(x+1)$$

The matrix is not positive definite.

(c) The characteristic polynomial of the given matrix is

$$x^2 - 8x + 15 = (x-5)(x-3)$$

The matrix is positive definite.

(d) The characteristic polynomial of the given matrix is

$$x^2 - 16 = (x+4)(x-4)$$

The matrix is not positive definite.

(49) Let A be a symmetric $n \times n$ matrix with real entries, which is positive semi-definite, i.e., $v^T Av \geq 0$ for every (column) vector v. Pick out the true statements:
(a) the eigenvalues of A are all non-negative;
(b) A is invertible.
(c) the principal minor of A (i.e., the determinant of the $k \times k$ matrix obtained from the first k rows and first k columns of A) is non-negative for each $1 \leq k \leq n$.

Ans. ***Option a and c***
Option (a) *and* (c) *are alternative definitions for positive semi-definiteness. Since eigenvalues of A can be zero, A need not be invertible.*

(50) Let J be the 3×3 matrix all of whose entries are 1. Then:
(a) 0 and 3 are the only eigenvalues of A.
(b) J is positive semi-definite. i.e., $\langle Jv, v\rangle \geq 0$ for all $v \in \mathbb{R}^3$.
(c) J is diagonalizable.
(d) J is positive definite. i.e., $\langle Jv, v\rangle > 0$ for all $v \in \mathbb{R}^3$ with $v \neq 0$.

Ans. ***Options a, b and c are true.***
The characteristic polynomial of the given matrix is $x^2(x-3)$ and the minimal polynomial is $x(x-3)$.

(51) Let a, b, c be positive real numbers such that $b^2 + c^2 < a < 1$. Consider the 3×3 matrix

$$A = \begin{bmatrix} 1 & b & c \\ b & a & 0 \\ c & 0 & 1 \end{bmatrix}$$

(a) All the eigenvalues of A are negative real numbers.
(b) All the eigenvalues of A are positive real numbers.
(c) A can have a positive as well as a negative eigenvalue.
(d) Eigenvalues of A can be non-real complex numbers.

Ans. ***Option b***
Clearly A is symmetric. Also the minors of the diagonal elements are a, $1-c^2$ and $a-b^2$, which are all greater than zero. Thus A is positive definite and hence all the eigenvalues of A are positive real numbers.

(52) Let $A \in \mathbb{M}_{m\times n}(\mathbb{R})$. Consider the following statements:

I : If $XAY = 0$ for all $X \in \mathbb{M}_{1\times m}(\mathbb{R})$ and $Y \in \mathbb{M}_{n\times 1}(\mathbb{R})$, then $A = 0$.
II. If $m = n$, A is symmetric and $A^2 = 0$, then $A = 0$.

Then
(a) both I and II are true. (b) I is true but II is false.
(c) I is false but II is true. (d) both I and II are false.

Ans. ***Option a***

(a) Let $A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$, *Take* $X_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $X_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}$, $Y_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $Y_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, *and* $Y_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. *Then considering* $X_iAY_j = 0$ *for all* $i = 1, 2$ *and* $j = 1, 2, 3$ *we get* $A = 0$. *We can use this idea for all m and n.*

(b) Let A be symmetric and $A^2 = 0$. The diagonal entries of A^2 are the length of the vectors in the corresponding row. This means each entry must be zero. Consider for example $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$. *Then* $A^2 = \begin{bmatrix} a^2+b^2 & ab+bc \\ ab+bc & b^2+c^2 \end{bmatrix} \Rightarrow a = b = c = 0$.

(53) For every 4×4 real symmetric non-singular matrix A, there exists a positive integer k such that
(a) $kI + A$ is positive definite (b) A^k is positive definite
(c) A^{-k} is positive definite (d) $exp(kA) - I$ is positive definite

Ans. ***Options a, b and c***
Since every eigenvalue of $kI + A$ is of the form $k + \lambda_i$ where λ_i is an eigenvalue of A, if we choose $k > max_i|\lambda_i|$, $kI + A$ is positive definite. As every eigenvalue of A^k is of the form $(\lambda_i)^k$ if we choose k as an even number, A^k is positive definite. Similarly since every eigenvalue of A^{-k} is of the form $\frac{1}{(\lambda_i)^k}$(as A is non-singular) if we choose k as an even number, A^{-k} is positive definite.

(54) Let A be a $n \times n$ real symmetric non-singular matrix. Suppose there exists $v \in \mathbb{R}^n$ such that $v^T Av < 0$. Then we can conclude that
(a) $B = -A$ is positive definite. (b) $det(A) < 0$
(c) there exists $u \in \mathbb{R}^n : u^T A^{-1}u < 0$ (d) $\forall\, u \in \mathbb{R}^n : u^T A^{-1}u < 0$

Ans. ***Option c***
Consider the matrix $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. *Then A is a real symmetric non-singular matrix with $det(A) = 1$ and $A^{-1} = A$. Clearly $B = -A$ is not positive definite. Now let $u = (u_1, u_2, u_3) \in \mathbb{R}^3$ be an arbitrary element. Then $u^T A^{-1}u = -u_1^2 - u_2^2 + u_3^2 < 0$ only when $u_3^2 < u_1^2 + u_2^2$.*

(55) Suppose A is a 3×3 symmetric matrix such that

$$\begin{bmatrix} x & y & 1 \end{bmatrix} A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = xy - 1$$

Let p be the number of positive eigenvalues of A and let $q = Rank(A) - p$. Then
(a) $p = 1$ (b) $p = 2$ (c) $q = 2$ (d) $q = 1$

Ans. ***Options b and d***
Take $A = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$. *Then,*

$$\begin{bmatrix} x & y & 1 \end{bmatrix} A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = xy - 1 \Rightarrow ax^2 + 2bxy + c2x + dy^2 + 2ey + f = xy - 1$$

$$\Rightarrow a = c = d = e = 0, b = \frac{1}{2}, f = -1$$

Therefore $A = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. *The characteristic polynomial of the given matrix is* $x^3 + x^2 - \frac{1}{4}x - \frac{1}{4} = \left(x^2 - \frac{1}{4}\right)(x+1)$. *Therefore* $p = 2$ *and* $q = 1$.

(56) Suppose A, B are $n \times n$ positive definite matrices and I be the $n \times n$ identity matrix. Then which of the following are positive definite.
(a) $A + B$ (b) ABA^* (c) $A^2 + I$ (d) AB

Ans. ***Options a, b and c***

(a) Suppose that A and B are positive definite, then $v^*Av > 0$ *and* $v^*Bv > 0$ *for all v. Therefore*

$$v^*(A+B)v = v^*Av + v^*Bv > 0 \; \forall \; v$$

Therefore $A + B$ *is positive definite.*

(b) Since B is positive definite, $v^*ABA^*v = (A^*v)^*B(A^*v) > 0$. *Therefore* ABA^* *is positive definite.*

(c) Since the eigenvalues of $A^2 + I$ *are of the form* $\lambda^2 + 1$ *where* λ *is an eigenvalue of A,* $A^2 + I$ *is positive definite.*

(d) Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$. *Then A and B are positive definite, but* $AB = \begin{bmatrix} 12 & 5 \\ 7 & 3 \end{bmatrix}$ *is not symmetric.*

(57) Consider a matrix $A = \left(a_{ij}\right)_{5\times 5}$, $1 \le i, j \le 5$ such that $a_{ij} = \dfrac{1}{n_i + n_j + 1}$, where $n_i, n_j \in \mathbb{N}$. Then in which of the following cases A is positive definite matrix?
(a) $n_i = i$ for all $i = 1, 2, 3, 4, 5$ (b) $n_1 < n_2 < n_3 < n_4 < n_5$
(c) $n_1 = n_2 = n_3 = n_4 = n_5$ (d) $n_1 > n_2 > n_3 > n_4 > n_5$

Ans. ***Options a, b and d***

A Cauchy matrix is an $m \times n$ *matrix with elements of the form* $a_{ij} = \dfrac{1}{x_i - y_j}$; $x_i - y_j \neq 0$, $1 \le i \le m$, $1 \le j \le n$ *where* (x_i) *and* (y_j) *have distinct elements. Every sub-matrix of a Cauchy matrix is itself a Cauchy matrix. The determinant of a Cauchy matrix is* $\dfrac{\prod_{i=2}^{n}\prod_{j=1}^{i-1}(x_i - x_j)(y_j - y_i)}{\prod_{i=2}^{n}\prod_{j=1}^{n}(x_i - y_j)}$. *The given matrix is a Cauchy matrix with* $x_i = n_i$ *and* $y_j = -(n_j + 1)$. *Then the determinant of each minor is strictly greater than 0 by the above formula for options a, b and d.*

(58) Which of the following statements are true?
(a) If A is a complex $n \times n$ matrix with $A^2 = A$, then $Rank(A) = tr(A)$.
(b) Let A be a 3×3 real symmetric matrix such that $A^6 = I$. Then $A^2 = I$.
(c) There exists $n \times n$ matrices A and B with real entries such that

$$(I - (AB - BA))^n = 0$$

(d) If A is a symmetric positive definite (All eigenvalues are positive) matrix, then

$$(tr\ (A))^n \geq n^n det\ (A)$$

Ans. ***Options a, b and d***

(a) *Since $A^2 = A$, A is diagonalizable.0 and 1 are the only possible eigenvalues. For a diagonalizable matrix, rank = number of non zero eigenvalues. Therefore $Rank(A) = tr(A)$.*

(b) *Suppose $A^6 = I$. Then the matrix satisfies the polynomial equation $x^6 - 1$. Since the matrix is real symmetric its eigenvalues are real. Now,*

$$x^6 - 1 = (x^3 - 1)(x^3 + 1) = (x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1)$$

Both $(x^2 + x + 1)$ and $(x^2 - x + 1)$ have complex roots. Therefore the only possible eigenvalues are 1 and -1 and the possible minimal polynomials are $(x - 1)$, $(x + 1)$ and $(x - 1)(x + 1)$. For all these cases $A^2 = I$.

(c) $(I - (AB - BA))^n = 0 \Rightarrow C = (I - (AB - BA))$ *is nilpotent which implies C must have trace$= 0$. But $tr\ (C) = tr\ (I) = n$.*

(d) *Let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be the eigenvalues of A.Since $AM \geq GM$ we have*

$$\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_n}{n} \geq \sqrt[n]{\lambda_1.\lambda_2.\cdots.\lambda_n}$$

which implies $(tr\ (A))^n \geq n^n det\ (A)$.

(59) Let P_1 and P_2 be two projection operators on a vector space. Then
(a) $P_1 + P_2$ is a projection if $P_1 P_2 = P_2 P_1 = 0$.
(b) $P_1 - P_2$ is a projection if $P_1 P_2 = P_2 P_1 = 0$.
(c) $P_1 + P_2$ is a projection.
(d) $P_1 - P_2$ is a projection.

Ans. ***Option a***

(a) *Since P_1 and P_2 are projection operators $P_1^2 = P_1$ and $P_2^2 = P_2$.*

$$(P_1 + P_2)^2 = P_1^2 + P_1 P_2 + P_2 P_1 + P_2^2 = P_1 + P_2\ if\ P_1 P_2 = P_2 P_1 = 0$$

Therefore $P_1 + P_2$ is a projection if $P_1 P_2 = P_2 P_1 = 0$.

(b) *As above*

$$(P_1 - P_2)^2 = P_1^2 - P_1 P_2 - P_2 P_1 + P_2^2 = P_1 + P_2\ if\ P_1 P_2 = P_2 P_1 = 0.$$

(c) *Let* $P_1(x, y) = P_2(x, y) = (x, 0)$*, then both* P_1 *and* P_2 *are projection operators. Observe that* $(P_1 + P_2)(x, y) = (2x, 0)$ *is not a projection as*

$$(P_1 + P_2)^2(x, y) = (P_1 + P_2)(2x, 0) = (4x, 0) \neq (P_1 + P_2)(x, y)$$

(d) *Let* $P_1(x, y) = (x, 0)$ *and* $P_2(x, y) = (0, y)$*, then both* P_1 *and* P_2 *are projection operators. Observe that* $(P_1 - P_2)(x, y) = (x, -y)$ *is not a projection as*

$$(P_1 - P_2)^2(x, y) = (P_1 - P_2)(x, -y) = (x, y)$$

Appendix

A.1 Determinants

Permutations

In Chap. 1, we got familiarized with the *symmetric group of n letters*. We have seen that a permutation on a set $S = \{1, 2, \ldots, n\}$ is a rearrangement of the members among themselves. In other words, a permutation σ is a one-one map from S onto itself. Such an element is represented in the form,

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & i & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(i) & \ldots & \sigma(n) \end{pmatrix}$$

We have seen that the set of all permutations on S forms a group with $n!$ elements under the operation function composition. For example consider S_3. The elements of S_3 can be listed as

$$\left\{\rho_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \rho_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \rho_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \mu_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \mu_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}\right\}$$

Then the Cayley table for S_3 is

$\circ$	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_0	ρ_0	ρ_1	ρ_2	μ_1	μ_2	μ_3
ρ_1	ρ_1	ρ_2	ρ_0	μ_3	μ_1	μ_2
ρ_2	ρ_2	ρ_0	ρ_1	μ_2	μ_3	μ_1
μ_1	μ_1	μ_2	μ_3	ρ_0	ρ_1	ρ_2
μ_2	μ_2	μ_3	μ_1	ρ_2	ρ_0	ρ_1
μ_3	μ_3	μ_1	μ_2	ρ_1	ρ_2	ρ_0

Now, take $\mu_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$. This permutation represents the following mapping (Fig. A.1).

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4

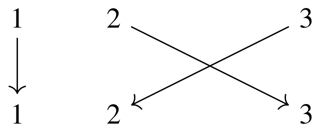


Fig. A.1 Schematic representation of μ_1

Then, we can can alternatively represent μ_1 as $\left(2\ 3\right) \circ \left(1\right)$ or $\left(2\ 3\right)$. This representation of μ_1 is called *cyclic representation* of μ_1, where $\left(2\ 3\right)$ represents the mapping $2 \longleftrightarrow 3$ and $\left(1\right)$ represents $1 \longleftrightarrow 1$. Then using the cyclic representation S_3 can be expressed as

$$S_3 \left\{e, \left(1\ 2\ 3\right), \left(1\ 3\ 2\right), \left(2\ 3\right), \left(1\ 3\right), \left(1\ 2\right)\right\}$$

Now, we will show that every permutation can be written as a cycle or as a product of disjoint cycles.

Theorem A.1 *Every permutation on the set S can be written as a cycle or as a product of disjoint cycles.*

Proof Let σ be a permutation on S_n. For $n = 1$, the proof is trivial. For $n > 1$, choose any member s_1 of S. Construct a sequence $s_1, s_2, s_3, \ldots$, where $s_i = \sigma^{i-1}(s_1)$. As S_n is finite, this sequence is finite and there exists m such that $\sigma^m(s_1) = s_1$. If all the members of the set have become a part of the sequence, we can write

$$\sigma = \left(s_1\ s_2\ \ldots\ s_m\right)$$

Otherwise,

$$\sigma = \left(s_1\ s_2\ \ldots\ s_m\right) \ldots$$

where three dots at the end suggest that we may not have exhausted set S during this process. Then choose another element r_1 in S which is not a member of the previous cycle and construct a new cycle as before. The new cycle will not contain any element from the previous cycle. If so, then $\sigma^i(s_1) = \sigma^j(r_1)$ and this would imply $r_1 = \sigma^{i-j}(s_1)$, which is not possible. As the set S is finite, continuing this process, every element in S will be a part of some cycle and

$$\sigma = \left(s_1\ s_2\ \ldots\ s_m\right)\left(r_1\ r_2\ \ldots\ r_k\right) \ldots \left(q_1\ q_2\ \ldots\ q_p\right)$$

Hence every permutation can be written as a cycle or as a product of disjoint cycles.

Theorem A.2 *Every permutation on S_n, $n > 1$ is a product of 2- cycles (transpositions).*

Proof From Theorem A.1, we have seen that any element $\sigma \in S_n$ can be written as

$$\sigma = \left(s_1\ s_2\ \ldots\ s_m\right)\left(r_1\ r_2\ \ldots\ r_k\right) \ldots \left(q_1\ q_2\ \ldots\ q_p\right)$$

Clearly, we can see that this representation is same as

$$\sigma = (s_1\ s_m)(s_1\ s_{m-1})\ldots(s_1\ s_2)(r_1\ r_k)(r_1\ r_{k-1})\ldots(r_1\ r_2)\ldots(q_1\ q_p)(q_1\ q_{p-1})\ldots(q_1\ q_2)$$

Hence the proof.

Now, we will prove that if the identity transformation(e) is written as a product of k transpositions, then k must be even. This idea will be later used to prove that whenever a permutation is written as a product of transpositions, the number of transpositions will be either always even or always odd.

Theorem A.3 *If $e = \sigma_1\sigma_2\ldots\sigma_k$, where, σ_i's are transpositions, then k is even.*

Proof Suppose on the contrary that k is odd. Clearly, we can say that $k = 1$ is not possible as the identity permutation must fix every element. Now suppose that $\sigma_k = (s_1\ s_2)$.

Case 1: $\sigma_{k-1} = (s_1\ s_2)$. Then,

$$e = \sigma_1\sigma_2\ldots\sigma_{k-1}\sigma_k = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_1\ s_2)(s_1\ s_2) = \sigma_1\sigma_2\ldots\sigma_{k-2}e = \sigma_1\sigma_2\ldots\sigma_{k-2}$$

That is, e can be written as a product of $k-2$ transpositions.

Case 2: $\sigma_{k-1} = (s_1\ s_3)$. Then

$$e = \sigma_1\sigma_2\ldots\sigma_{k-1}\sigma_k = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_1\ s_3)(s_1\ s_2) = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_1\ s_2)(s_2\ s_3)$$

Case 3: $\sigma_{k-1} = (s_3\ s_2)$. Then

$$e = \sigma_1\sigma_2\ldots\sigma_{k-1}\sigma_k = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_3\ s_2)(s_1\ s_2) = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_1\ s_3)(s_2\ s_3)$$

Case 4: $\sigma_{k-1} = (s_3\ s_4)$. Then

$$e = \sigma_1\sigma_2\ldots\sigma_{k-1}\sigma_k = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_3\ s_4)(s_1\ s_2) = \sigma_1\sigma_2\ldots\sigma_{k-2}(s_1\ s_2)(s_3\ s_4)$$

Consider the element s_1 in σ_k. Observe that in case 2, 3 and 4 we shifted it from σ_k to σ_{k-1}. We can continue this process till we must have a case 1. For, if that is not the case, we will reach a stage where,

$$e = (s_1\ s_5)\sigma_2\ldots\sigma_k$$

which is not possible as the permutation on the right hand side of the above equation does not fix s_1. Therefore we will eventually reach case 1.

We have now shown that if identity is represented as a product of odd number of transpositions, then we can delete 2 transpositions from this expression and still we will be having the identity permutation. That is, identity permutation is again expressed as a product of $m-2$ transpositions which is again an odd number. Again, repeating this elimination process, we will eventually have that identity equals a single transposition which is a contradiction. Hence k must be even.

Theorem A.4 *Let σ be a permutation in S_n. If σ can be represented as a product of even(odd) number of transpositions, then every decomposition of σ must also contain even(odd) number of transpositions.*

Proof Let $\sigma = \alpha_1\alpha_2\ldots\alpha_r$ and $\sigma = \beta_1\beta_2\ldots\beta_k$ be two different decomposition of σ. Then,

$$e = \alpha_1\alpha_2\ldots\alpha_r\beta_k^{-1}\ldots\beta_2^{-1}\beta_1^{-1}$$

As a transposition is its own inverse we have

$$e = \alpha_1\alpha_2\ldots\alpha_r\beta_k\ldots\beta_2\beta_1$$

Then, by Theorem A.3, $k+r$ is even. This implies that either both k and s must be even or both must be odd.

We can also prove that the collection of all even permutations on $S = \{1, 2, \ldots, n\}$ forms a group called *Alternating group of n letters* denoted by A_n. What about the collection of all odd permutations?

Definition A.1 (*Sign of a permutation*) Let σ be a permutation in S_n. Then the sign of σ is defined as

$$sgn(\sigma) = \begin{cases} 1, & if\ \sigma\ is\ even \\ -1, & if\ \sigma\ is\ odd \end{cases}$$

Example A.1 Consider the permutation $\sigma = \begin{pmatrix}2 & 3 & 5 & 1 & 4\end{pmatrix}$ in S_5. We have,

$$\sigma = \begin{pmatrix}2 & 3 & 5 & 1 & 4\end{pmatrix} = \begin{pmatrix}2 & 4\end{pmatrix}\begin{pmatrix}2 & 1\end{pmatrix}\begin{pmatrix}2 & 5\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix}$$

Thus σ is even and $sgn(\sigma) = 1$.

Example A.2 Consider $S_2 = \{e, \sigma = \begin{pmatrix}1 & 2\end{pmatrix}\}$. We have already shown that identity permutation is even. Hence $sgn(e) = 1$. Clearly, $sgn(\sigma) = 1$ as it is a transposition.

Example A.3 Consider $S_3 = \{\rho_0, \rho_1, \rho_2, \mu_1, \mu_2, \mu_3\}$. As ρ_0 is the identity element $sgn\ (\rho_0) = 1$. Verify that $sgn\ (\rho_1) = sgn\ (\rho_2) = 1$ and $sgn\ (\mu_1) = sgn\ (\mu_2) = sgn\ (\mu_3) = -1$.

By the definition of sign of a permutation, it is easy to verify the following properties. (Verify!)

Theorem A.5 *Let $\sigma, \tau \in S_n$. Then*

(a) $sgn(\sigma \circ \tau) = sgn(\sigma)sgn(\tau)$
(b) $sgn\left(\sigma^{-1}\right) = sgn(\sigma)$
(c) $sgn(e) = 1$
(d) $sgn(\sigma) = 1$*, where σ is any transposition.*
(e) $sgn(\tau) = (-1)^{k-1}$*, where τ is any cycle of length k.*

Now, we will define the determinant for a square matrix and discuss its properties.

Determinant of an n × n Matrix

Consider a 2×2 matrix, $A_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. We have,

$$det(A_1) = a_{11}a_{22} - a_{12}a_{21}$$

For a 3×3 matrix $A_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$,

$$det(A_2) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Keep in mind that just one element comes from each row and only one element comes from each column in each term of the expressions above. These are the only conceivable combinations of those, as well. If you observe the pattern of the column numbers in each term of the above two expressions, you can see that they are exactly S_2 and S_3. What about the *sign* of each terms in the expression? Does it have any relation with sign of the corresponding permutation? With these observations in our mind, we will define a function on $\mathbb{M}_{n\times n}(\mathbb{K})$ as follows.

$$f(A) = \sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} \tag{A.1}$$

Is this function well-defined? (Think!) Then, we can have the following theorem which tells us that the function f is nothing but the determinant function.

Theorem A.6 *Let $A = \left[a_{ij}\right]$ be an $n \times n$ matrix and M_{ij} denote the minor of A obtained by deleting its ith row and jth column. Then,*

$$f(A) = \sum_{j=1}^{N} (-1)^{i+j} a_{ij} M_{ij}$$

where, f is as defined in (A.1).

Now, we will prove the following properties of determinant function.

Theorem A.7 *Let $A = [a_{ij}]$ be an $n \times n$ matrix over the field $\mathbb{K}$.*

(a) $det(A) = det(A^T)$
(b) *If B is a matrix obtained from A by multiplying a row(column) by a scalar k, then $det(B) = kdet(A)$. Also, if $B = kA$, then $det(B) = k^n det(A)$.*
(c) *If two rows (columns) of A are equal, then $det(A) = 0$.*
(d) *If B is a matrix obtained from A by interchanging any two rows (columns) of A, then $det(B) = -det(A)$.*
(e) *If B is a matrix obtained from A by adding a multiple of one row(column) to another, then $det(B) = det(A)$.*

Proof (a) We have,

$$det(A) = \sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}$$

and

$$det(A^T) = \sum_{\sigma \in S_n} [sgn(\sigma)] a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}$$

If $\tau = \sigma^{-1}$, we have $a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n} = a_{1\tau(1)} a_{2\tau(2)} \dots a_{n\tau(n)}$ (Why?). Then,

$$det(A^T) = \sum_{\tau \in S_n} [sgn(\tau)] a_{1\tau(1)} a_{2\tau(2)} \dots a_{n\tau(n)}$$

As, S_N is a group and σ runs through all elements of S_N, τ also runs through all elements of S_n. Hence $det(A) = det(A^T)$.

(b) Each term in the expression (A.1) of $det(A)$ contains just one element from each row(column) of A. Therefore, multiplying a row(column) by a scalar k induces the factor k in each term of $det(A)$. Thus $det(B) = kdet(A)$. Similarly, $det(B) = k^n det(A)$.

(c) The matrix obtained from A by interchanging the two equal rows identical with A. However, by the previous result the sign of the determinant must change. This implies that $det(A) = 0$.

(d) Let $\tau = (i \ j)$ be the transposition that interchanges the ith and jth rows (columns) of A. This interchanging has the effect of replacing each permutation σ by $\sigma \circ \tau$. We have,

$$\sigma \circ \tau(i) = \sigma(\tau(i)) = \sigma(j)$$

$$\sigma \circ \tau(j) = \sigma(\tau(j)) = \sigma(i)$$

and for all $k \in \{1, 2, \dots, n\}$ except i and j, $\sigma \circ \tau(k) = \sigma(k)$. Also,

$$sgn(\sigma \circ \tau) = sgn(\sigma) sgn(\tau) = -sgn(\sigma)$$

Thus,

$$
\begin{aligned}
det(B) &= \sum_{\sigma \circ \tau \in S_n} [sgn(\sigma \circ \tau)] a_{1\sigma\circ\tau(1)} \dots a_{i\sigma\circ\tau(i)} \dots a_{j\sigma\circ\tau(j)} \dots a_{n\sigma\circ\tau(n)} \\
&= -\sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} \dots a_{(i-1)\sigma(i-1)} a_{i\sigma(j)} a_{(i+1)\sigma(i+1)} \cdots \\
&\quad \times a_{(j-1)\sigma(j-1)} a_{j\sigma(i)} a_{(j+1)\sigma(j+1)} \cdots a_{n\sigma(n)} \\
&= -det(A)
\end{aligned}
$$

(e) Let B be the matrix obtained from A by changing its ith row by a sum of ith row and a multiple of jth row. Then,

$$
\begin{aligned}
det(B) &= \sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} \dots \left(a_{i\sigma(i)} + k a_{j\sigma(j)}\right) \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} \\
&= \sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} \dots a_{i\sigma(i)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)} \\
&\quad + k \sum_{\sigma \in S_n} [sgn(\sigma)] a_{1\sigma(1)} \dots a_{j\sigma(j)} \dots a_{j\sigma(j)} \dots a_{n\sigma(n)}
\end{aligned}
$$

The second sum on the right side of this equation is zero, as it is the determinant of a matrix whose two rows are equal. Thus, $det(B) = det(A)$.

A.1.0.1 Elementary Operations and Elementary Matrices

Elementary matrices are essential linear algebra tools that play an important role in matrix operations and transformations. These are square matrices obtained by performing a single elementary row(column) operation on an identity matrix. Elementary row operations include scaling a row, swapping rows, or adding a multiple of one row to another. These basic matrices serve as the foundation for comprehending more advanced matrix operations including row reduction, matrix inverses, and solving systems of linear equations. They are an essential notion in linear algebra because they give a systematic approach to trace changes in a matrix caused by basic row or column operations.

Example A.4 Consider the matrices $E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Clearly E_1, E_2 and E_3 are elementary matrices as they are obtained from identity matrix by interchanging first and second row, multiplying the second row by 4 and adding -5 times the first row to the second row respectively.

Remark A.1 Elementary matrices obtained by interchanging the rows of identity matrix are also called *permutation matrices.*

Now, consider a matrix $A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. Observe the changes to A, when A is multiplied by E_1, E_2 and E_3. We have,

$$E_1 A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{bmatrix}, \; E_2 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ 4 & 0 & 4 \\ 1 & 0 & 0 \end{bmatrix}$$

and

$$E_3 A = \begin{bmatrix} 1 & 0 & 0 \\ -5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & -5 & 6 \\ 1 & 0 & 0 \end{bmatrix}$$

From this, we can clearly say that to conduct any of the three elementary row operations on a matrix A, take the product EA, where E is the elementary matrix generated by performing the required elementary row operation on the identity matrix. Now, consider the following result.

Theorem A.8 *Let A be an $n \times n$ non-singular matrix. Then A can be written as a product of elementary matrices.*

Proof As A is non-singular, atleast one element in the first column of A is non-zero. If $a_{11} = 0$, we can interchange rows to bring a non-zero element as a_{11}. Then, we multiply the first row by a_{11}^{-1}. Thus in the reduced matrix $a_{11} = 1$ and using this fact, we can add $-a_{i1}$ times the first row to the ith row to make every other element in the first row zero. Observe that, we have applied only the elementary row transformations. Hence the resulting matrix is still non-singular.

Now, we wish to do the same to the second row. At least one element in second column other than a_{12} must be non-zero. For, otherwise the first two columns will be linearly dependent. Thus, if we repeat the elementary operations as we have applied previously, we can have the $a_{22} = 1$ and $a_{i2} = 0$ for all $i \neq 2$. Continuing this process we will finally obtain the $n \times n$ identity matrix. Thus, if $E_1, E_2, \ldots, E_k$ are the elementary matrices representing the elementary operations applied on A successively, we have

$$I = E_k \ldots E_2 E_1 A$$

Hence

$$A = E_1^{-1} E_2^{-1} \ldots E_k^{-1}$$

Example A.5 Consider a matrix $A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.

Step I: Interchange the rows 1 and 3. That is, Multiply A by $E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Then,

$$E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Step II: Subtract the first row from the second row. That is, Multiply A by $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then, $E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$.

Step III: Interchange the rows 2 and 3. That is, Multiply A by $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. Then,

$$E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Step IV: Add the second row to the third row. That is, Multiply A by $E_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

Then, $E_4 E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

We can clearly observe that $A = E_1^{-1} E_2^{-1} E_3^{-1} E_4^{-1}$.

Theorem A.9 *Let A be an $n \times n$ matrix. If E is an elementary matrix, then*

$$det(EA) = det(E)det(A) = det(AE)$$

The above theorem is an immediate consequence of Theorems A.7 (b), (d) and (e). Now, by using Theorems A.8 and A.9, we can have the following result.

Theorem A.10 *Let A and B be two $n \times n$ matrices. Then,*

$$det(AB) = det(A)det(B) = det(BA)$$

A.2 Fourier Series

Consider the set of all functions in $[-\pi, \pi]$ that satisfy the condition

$$\int_{-\pi}^{\pi} |f(x)|^2 dx < \infty$$

Clearly, this collection will form a vector space under point-wise addition of functions and standard scalar multiplication (Verify!). Also,

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$$

defines an inner product on this space (Verify!). Consider the following set of functions;

$$\left\{\frac{1}{\sqrt{\pi}} sin\ nx \mid 1 \leq n < \infty\right\} \cup \left\{\frac{1}{\sqrt{\pi}} cos\ nx \mid 1 \leq n < \infty\right\}$$

We will show that this set is an orthonormal set. For $m \neq n$,

$$\begin{aligned}\langle cos\ nx, cos\ mx\rangle &= \int_{-\pi}^{\pi} cos\ nx cos\ mx dx \\ &= \frac{1}{2}\int_{-\pi}^{\pi} [cos(m+n)x + cos(m-n)x]\, dx \\ &= \frac{1}{2}\left[\frac{sin(m+n)x}{m+n} + \frac{sin(m-n)x}{m-n}\right]_{pi}^{\pi} = 0\end{aligned}$$

Similarly, for $m \neq n$

$$\langle sin\ nx, sin\ mx\rangle = 0$$

and for all m and n

$$\langle cos\ nx, sin\ mx\rangle = 0$$

Also,

$$\begin{aligned}\|sin\ nx\|^2 &= \int_{-\pi}^{\pi} [sin\ nx]^2\, dx = \frac{1}{2}\int_{-\pi}^{\pi} 2\,[sin\ nx]^2\, dx \\ &= \frac{1}{2}\int_{-\pi}^{\pi} (1 - cos\ 2nx)dx = \frac{1}{2}\left[x - \frac{sin\ 2nx}{2nx}\right]_{-\pi}^{\pi} = \pi\end{aligned}$$

and

$$\|cos\ nx\|^2 = \pi$$

Therefore, the set

$$\left\{\frac{1}{\sqrt{\pi}} sin\ nx \mid 1 \leq n < \infty\right\} \cup \left\{\frac{1}{\sqrt{\pi}} cos\ nx \mid 1 \leq n < \infty\right\}$$

is an orthonormal set. Our objective is to obtain a trigonometric series representation to every periodic function $f(x)$ defined on $[-\pi, \pi]$. Let us assume that

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n cos\ nx + b_n sin\ nx) \tag{A.2}$$

Integrating on both sides from $-\pi$ to π,

$$\int_{-\pi}^{\pi} f(x)dx = \int_{-\pi}^{\pi} \left[a_0 + \sum_{n=1}^{\infty} (a_n cos\ nx + b_n sin\ nx) \right] dx$$
$$= a_0 \int_{-\pi}^{\pi} dx + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} cos\ nxdx + b_n \int_{-\pi}^{\pi} sin\ nxdx \right) = 2\pi a_0$$

This gives,

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)dx \tag{A.3}$$

Now, consider $cos(mx)$, where m is a fixed integer. Multiply both sides of (A.6) by $cos(mx)$ and integrate from $-\pi$ to π,

$$\int_{-\pi}^{\pi} f(x)cos\ mxdx = \int_{-\pi}^{\pi} \left[a_0 + \sum_{n=1}^{\infty} (a_n cos\ nx + b_n sin\ nx) \right] cos\ mxdx$$
$$= a_0 \int_{-\pi}^{\pi} cos\ mxdx + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} cos\ nxcos\ mxdx + b_n \int_{-\pi}^{\pi} sin\ nxcos\ mxdx \right)$$
$$= a_m \int_{-\pi}^{\pi} [cos\ mx]^2 dx = a_m \pi$$

This gives,

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)cos\ mxdx,\ m = 1, 2, \ldots \tag{A.4}$$

Similarly, we get

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)sin\ mxdx,\ m = 1, 2, \ldots \tag{A.5}$$

The coefficients $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ are called the *Fourier coefficients* and the trigonometric series given in (A.6) is called the Fourier series expansion of $f(x)$ named after the famous French Mathematician *Jean–Baptiste Joseph Fourier (1786–1830)*. While Fourier series are well-known for their ability to approximate a wide range of functions, it's crucial to remember that they may not converge for all sorts of signals or functions. In such circumstances, rigorous analysis and consideration of convergence concerns are required to verify the approximation's accuracy. The following theorem gives sufficient conditions for the convergence of series (A.2) with coefficients given by Eqs. (A.3), (A.4) and (A.5).

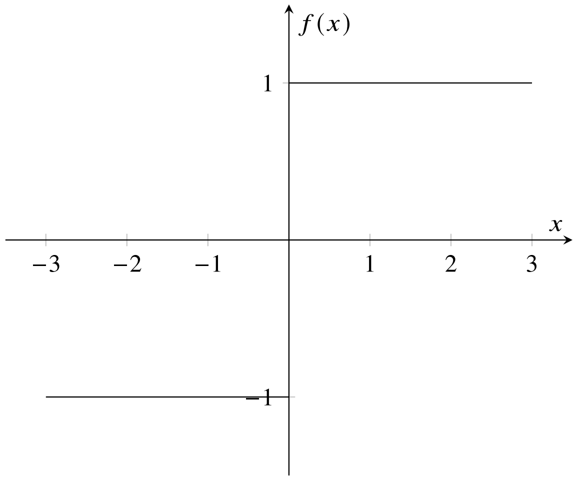


Fig. A.2 Graph of $f(x)$

Theorem A.11 *If a periodic function $f(x)$ with period 2π is piecewise continuous in the interval $-\pi \leq x \leq \pi$ and has a left-hand derivative and right-hand derivative at each point of that interval, then the series*

$$a_0 + \sum_{n=1}^{\infty} (a_n cos\ nx + b_n sin\ nx) \tag{A.6}$$

with coefficients

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)dx$$
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)cos\ nxdx,\ n = 1, 2, \ldots$$
$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)sin\ nxdx,\ n = 1, 2, \ldots$$

is convergent. The sum of the series (A.6) is $f(x)$, except at a point x_0 at which $f(x)$ is discontinuous. At x_0, sum of the series is the average of the left and right hand limits of $f(x)$ at x_0.

Example A.6 Consider the function (Fig. A.2)

$$f(x) = \begin{cases} -1, x \in [-\pi, 0) \\ 1, x \in (0, \pi] \end{cases} \tag{A.7}$$

Let us find the Fourier coefficients for $f(x)$. We have,

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{1}{2\pi}\left(-\int_{-\pi}^{0} dx + \int_{0}^{\pi} dx\right) = 0$$

Now,

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)cos\ nxdx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} -cos\ nxdx + \int_{0}^{\pi} cos\ nxdx\right] \\ &= \frac{1}{\pi}\left[\left(-\frac{sin\ nx}{n}\right)_{-\pi}^{0} + \left(\frac{sin\ nx}{n}\right)_{0}^{\pi}\right] = 0 \end{aligned}$$

and

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)sin\ nxdx \\ &= \frac{1}{\pi}\left[\int_{-\pi}^{0} -sin\ nxdx + \int_{0}^{\pi} sin\ nxdx\right] \\ &= \frac{1}{\pi}\left[\left(\frac{cos\ nx}{n}\right)_{-\pi}^{0} - \left(\frac{cos\ nx}{n}\right)_{0}^{\pi}\right] = \frac{2}{n\pi}(1 - cos\ n\pi) \end{aligned}$$

Then (Fig. A.3)

$$f(x) = \frac{4}{\pi}\left(sin\ x + \frac{1}{3}sin\ 3x + \frac{1}{5}sin\ 5x\right) + \ldots$$

Observe that $x = 0$ is a point of discontinuity of $f(x)$ all partial sums have the value zero, the arithmetic mean of values -1 and 1 of the given function.

Remark A.2 The Fourier series of an even function of period $2L$ is a *Fourier cosine series*,

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n cos\frac{n\pi}{L}x$$

with coefficients

$$a_0 = \frac{1}{L}\int_{0}^{L} f(x)dx, \quad a_n = \frac{2}{L} f(x)cos\frac{n\pi}{L}xdx,\ n = 1, 2, \ldots$$

Similarly, the Fourier series of an odd function of period $2L$ is a *Fourier sine series*,

$$f(x) = \sum_{n=1}^{\infty} b_n sin\frac{n\pi}{L}x$$

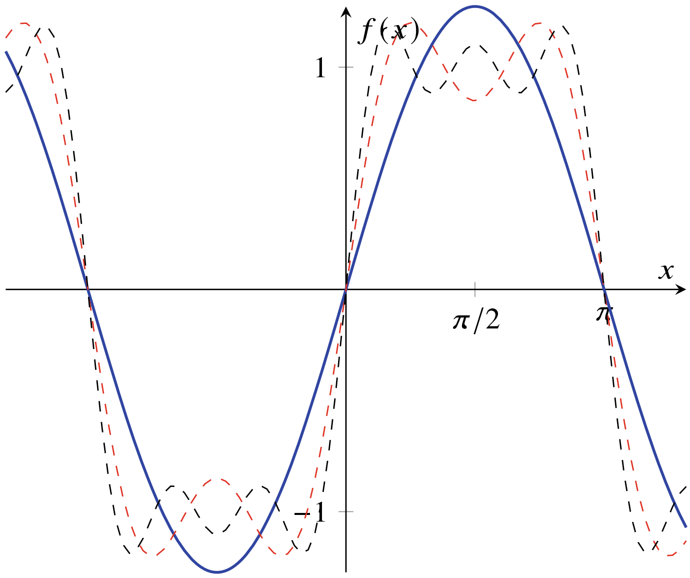


Fig. A.3 First three partial sums of the Fourier series for a square wave

with coefficients

$$b_n = \frac{2}{L}\int_0^L f(x) sin\frac{n\pi}{L}x$$

The Fourier series expansion is a powerful mathematical tool with numerous applications in a variety of domains. Signal processing relies on the Fourier series extensively. It is used to examine and manipulate signals, including electromagnetic waves, audio, and video. For instance, it is used in data compression methods like filtering to eliminate undesirable frequencies. Engineers investigate the vibrations and resonances of mechanical systems using Fourier analysis. This is essential for constructing machines, vehicles, and structures to stop or reduce vibrations. Fourier's law outlines how heat propagates through materials in the study of heat transfer. In a variety of engineering applications, the Fourier series can be used to address difficult heat conduction problems. Wave functions are frequently used in quantum physics and can be described as superpositions of multiple energy states using the Fourier series. This is critical for understanding particle behavior at the quantum level. These are just a few instances, and as technology progresses and our understanding of mathematical tools grows, so will the uses of Fourier series expansion. It is a useful and necessary tool in a wide range of scientific and engineering disciplines.

References

1. Adhikari, M.R.: Introduction to Linear Algebra with Application to Basic Cryptography. Asian Books Private Limited (2009)
2. Ajayakumar, A., George, R.K.: Controllability of networked systems with heterogeneous dynamics. Math. Control Signals Syst. **35**(2), 307–326 (2023)
3. Axler, S.: Linear Algebra Done Right. Springer, Cham (2016)
4. Bartle, R.G., Sherbert, D.R.: Introduction to Real Analysis. Wiley (2000)
5. Ben-Israel, A., Greville, T.N.: Generalized Inverses: Theory and Applications, vol. 15. Springer Science & Business Media (2003)
6. Brockett, R.W: Finite Dimensional Linear Systems. Society for Industrial and Applied Mathematics (2015)
7. Bronson, R., Costa, G.B.: Linear Algebra: An Introduction. Academic (2007)
8. Bronson, R., Costa, G.B.: Matrix Methods: Applied Linear Algebra. Academic (2008)
9. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill (1955)
10. Datta, B.N.: Numerical Linear Algebra and Applications. Society for Industrial and Applied Mathematics (2010)
11. Debnath, L., Mikusinski, P.: Introduction to Hilbert Spaces with Applications. Academic (2005)
12. Fraleigh, J.B.: A First Course in Abstract Algebra. Pearson Education India (2003)
13. Friedberg, S.H., Insel, A.J., Spence, L.E.: Linear Algebra, vol. 4. Pearson, Essex, NJ, USA (2014)
14. Gallian, J.: Contemporary Abstract Algebra. Chapman and Hall/CRC (2021)
15. Gantmacher, F.R.: The Theory of Matrices. American Mathematical Society (2000)
16. Goode, S.W., Annin, S.: Differential Equations and Linear Algebra. Prentice Hall, Upper Saddle River (2000)
17. Greenberg, M.D.: Advanced Engineering Mathematics, 2nd edn. Pearson (1998)
18. Gustafson, G.B., Wilcox, C.H.: Advanced Engineering Mathematics: Analytical and Computational Methods. Springer (1998)
19. Halmos, P.R.: Finite-Dimensional Vector Spaces. Springer, New York, NY (2011)
20. Hoffman, K., Kunze, R.: Linear Algebra. Prentice-Hall, Inc., New Jersey (1971)
21. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press (2012)
22. Joshi, M.C.: Ordinary Differential Equations: Modern Perspective. Alpha Science International (2006)
23. Joshi, M.C., George, R.K.: Controllability of nonlinear systems. Numer. Funct. Anal. Optim. **10**(1–2), 139–166 (1989)
24. Kreyszig, E.: Introductory Functional Analysis with Applications, vol. 17. Wiley (1991)
25. Kreyszig, E.: Advanced Engineering Mathematics, vol. 8. Wiley, India (1996)
26. Kumaresan, S.: Linear Algebra: A Geometric Approach. PHI Learning Pvt. Ltd (2000)
27. Lax, P.D.: Linear Algebra and Its Applications. Wiley (2007)
28. Lay, D.C.: Linear Algebra and Its Applications. Pearson Education India (2003)
29. Leontief, W.: Input-Output Economics. Oxford University Press (1986)
30. Limaye, B.V.: Functional Analysis. New Age International (1996)
31. Lipschutz, S.: Schaum's Outline of Theory and Problems of General Topology. McGraw-Hill Book Company (1965)
32. Lipschutz, S., Lipson, M.L.: Schaum's Outline of Linear Algebra. McGraw-Hill Education (2018)
33. Meyer, C.D., Stewart, I.: Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics (2023)
34. Moh, T.T.: Linear Algebra and Its Applications, vol. 10. World Scientific (2020)
35. Nair, M.T., Singh, A.: Linear Algebra. Springer (2018)
36. Nandakumaran, A.K., Datti, P.S., George, R.K.: Ordinary Differential Equations: Principles and Applications. Cambridge University Press (2017)

37. Nering, E.D.: Linear Algebra and Matrix Theory. Wiley (1991)
38. Penrose, R.: A generalized inverse for matrices. In: Mathematical Proceedings of the Cambridge Philosophical Society, vol. 51, No. 3, pp. 406–413. Cambridge University Press (1955)
39. Pokrovskii, V.N.: Econodynamics. The Theory of Social Production. Springer, Dordrecht (2011)
40. Roman, S., Axler, S., Gehring, F.W.: Advanced Linear Algebra, vol. 3. Springer, New York (2005)
41. Saikia, P.K.: Linear Algebra, 2nd edn. Pearson Education India (2014)
42. Serre, D.: Matrices Theory and Applications. Springer, New York, NY (2010)
43. Silverman, J.H., Pipher, J., Hoffstein, J.: An Introduction to Mathematical Cryptography, vol. 1. Springer, New York (2008)
44. Sontag, E.D.: Mathematical Control Theory: Deterministic Finite Dimensional Systems, vol. 6. Springer Science & Business Media (2013)
45. Strang, G.: Introduction to Linear Algebra. Wellesley-Cambridge Press (2022)
46. Sundarapandian, V.: Numerical Linear Algebra. PHI Learning Pvt. Ltd (2008)
47. Terrell, W.J.: Stability and Stabilization: An Introduction. Princeton University Press (2009)
48. Trefethen, L.N., Bau, D.: Numerical Linear Algebra. Siam (2022)
49. Zabczyk, J.: Mathematical Control Theory. Springer International Publishing (2020)
50. Zhang, F.: Linear Algebra: Challenging Problems for Students. JHU Press (1996)
51. Zhang, F.: Matrix Theory: Basic Results and Techniques. Springer, New York, NY (2011)
52. Zurada, J.: Introduction to Artificial Neural Systems. West Publishing Co. (1992)

Index

R. K. George and A. Ajayakumar, *A Course in Linear Algebra*, University Texts in the Mathematical Sciences, https://doi.org/10.1007/978-981-99-8680-4