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 $v = (v_1, v_2, \ldots, v_n)$ 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 $i = 1, 2, \ldots, n \Leftrightarrow v = 0$.
(N2) For $\lambda \in \mathbb{R}$ and $v \in \mathbb{R}^n$,

$$\|\lambda v\|_2 = \left(\sum_{i=1}^n |\lambda v_i|^2\right)^{\frac{1}{2}} = \left(\sum_{i=1}^n |\lambda|^2 |v_i|^2\right)^{\frac{1}{2}} = \left(|\lambda|^2 \sum_{i=1}^n |v_i|^2\right)^{\frac{1}{2}}$$
$$= |\lambda| \left(\sum_{i=1}^n |v_i|^2\right)^{\frac{1}{2}} = |\lambda| \, \|v\|_2$$

(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 $C[-4, 4]$.

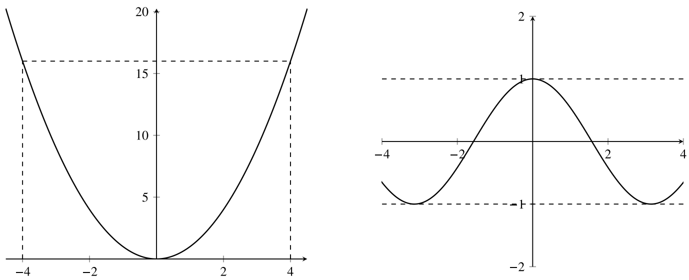


Fig. 5.1 Consider the functions $f(x) = x^2$ and $g(x) = \cos x$ in $C[-4, 4]$. Then $\|f\| = \max_{x\in[-4,4]} |x^2| = 16$ and $\|g\| = \max_{x\in[-4,4]} |\cos x| = 1$

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 $\lambda \neq 0 \in \mathbb{K}$,

$$\|\lambda v\| = \begin{cases} 1\,, & v \neq 0 \\ 0\,, & v = 0 \end{cases} \quad \neq |\lambda|\, \|v\| = \begin{cases} |\lambda|\,, & v \neq 0 \\ 0\,, & v = 0 \end{cases}$$

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.

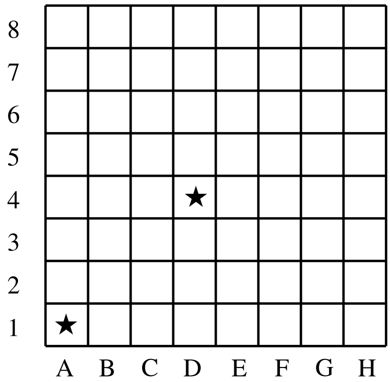


Fig. 5.2 Suppose that you have to move a chess piece from $A1$ to $D4$ in least number of moves. If the piece is a bishop we can move the piece directly from A_1 to D_4. If the piece is a rook, first we will have to move the piece either to $A4$ or $D1$ and then to $D4$. Now, if the piece is king, the 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

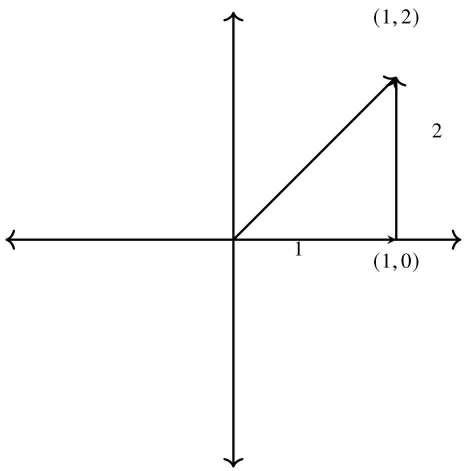


Fig. 5.3 Consider $\mathbb{R}^2$ with different norms defined on it. If we are using the 2-norm, the distance from the origin to the point $(1, 2)$ is $\sqrt{|1|^2 + |2|^2} = \sqrt{5}$ as it is length of the hypotenuse of a triangle with base 1 and height 2. If we are using 1-norm the distance will be 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

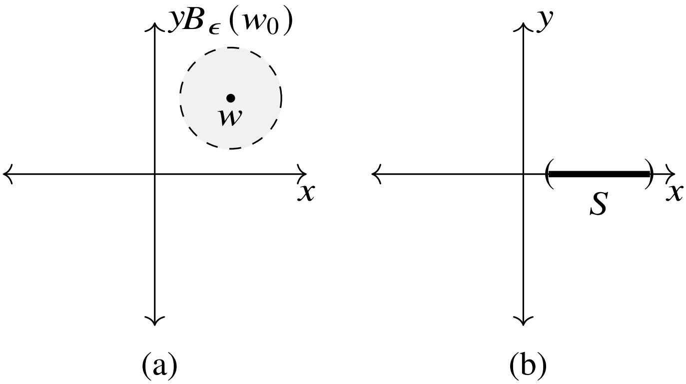
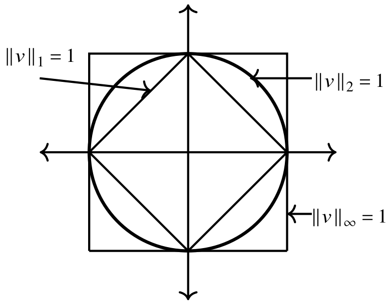


Fig. 5.4 The open balls in $(\mathbb{R}^2, \|.\|_2)$ are open circles as given in (a). Clearly, S is not an open ball in $(\mathbb{R}^2, \|.\|_2)$

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)

$$
B_1^\infty = \left\{(v_1, v_2) \in \mathbb{R}^2 \mid max\{|v_1|, |v_2|\} < 1\right\}
$$

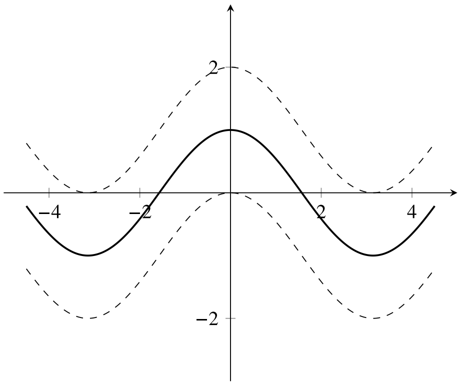


Fig. 5.6 Consider a function f in $C[-4, 4]$ with supremum norm. Continuous functions that lie between the dotted lines constitute $B_1(f) = \{g \in C[-4, 4] \mid \|f - g\| < 1\}$

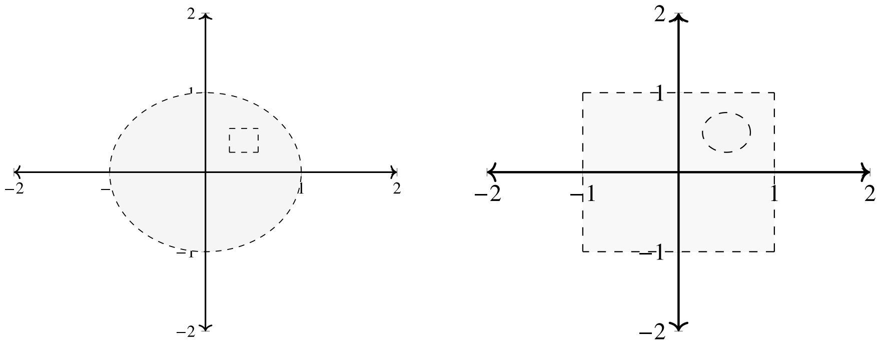


Fig. 5.7 Clearly, we can observe that every point in an open ball generated by the infinity norm is inside an open ball generated by 2-norm and vice versa

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.

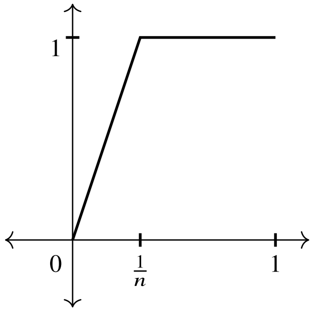


Fig. 5.8 Define $f_n(x) = \begin{cases} nx, \ x \in \left[0, \frac{1}{n}\right] \\ 1, \ x \in \left[\frac{1}{n}, 1\right] \end{cases}$ as shown in the figure. Clearly $f_n(x)$ belongs to $C[0, 1]$ for all n

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 $n > m$,

$$\begin{aligned} \|q_n(x) - q_m(x)\| &= \max_{x\in[0,1]} \left| \sum_{i=0}^{n} \frac{x^i}{2^i} - \sum_{i=0}^{m} \frac{x^m}{2^m} \right| \\ &= \max_{x\in[0,1]} \left| \sum_{i=m+1}^{n} \frac{x^i}{2^i} \right| \\ &\leq \frac{1}{2^m} \end{aligned}$$

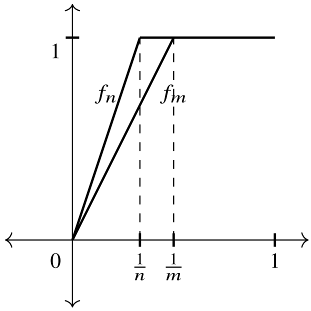


Fig. 5.9 As $\|f_n - f_m\|$ is the area of the triangle depicted in the figure, it is easy to observe that $\{f_n\}$ is Cauchy

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 $v_k = \left(v_1^k, v_2^k, \dots, v_n^k\right)$ 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 $i = 1, 2, \dots, n$ and $k, m \geq N$ and hence $|v_i^k - v_i^m| < \epsilon$ for each $i = 1, 2, \dots, n$ and $k, m \geq N$. Thus for a fixed i, the sequence $v_i^1, v_i^2, \dots$ 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 $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$. 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 $u = (u_1, u_2), v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2$,

$$\begin{aligned} u.(v + w) &= u_1(v_1 + w_1) + u_2(v_2 + w_2) \\ &= u_1 v_1 + u_2 v_2 + u_1 w_1 + u_2 w_2 \\ &= u.v + u.w \end{aligned}$$

(c) For $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$ and $\lambda \in \mathbb{K}$,

$$\begin{aligned} (\lambda u).v &= (\lambda u_1, \lambda u_2).(v_1, v_2) \\ &= \lambda u_1 v_1 + \lambda u_2 v_2 \\ &= \lambda(u_1 v_1 + u_2 v_2) = \lambda(u.v) \end{aligned}$$

(d) For $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{R}^2$,

$$u.v = u_1 v_1 + u_2 v_2 = v_1 u_1 + v_2 u_2 = v.u$$

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,

$$|w|^2 = |u|^2 + |v|^2 - 2|u||v| \cos\theta \tag{5.2}$$

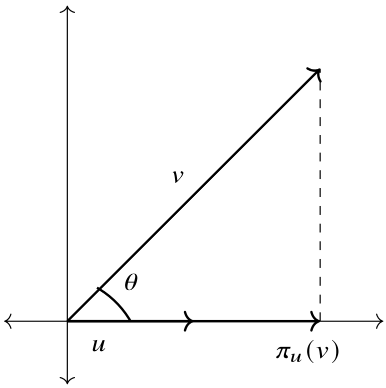


Fig. 5.10 Orthogonal projection of v on u

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, $\langle u, v\rangle = \sum_{i=1}^{n}u_iv_i$ 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 $v, w \in V$,*

$$|\langle v, w\rangle|^2 \le \langle v, v\rangle\langle w, w\rangle$$

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

$$\begin{aligned}0 \leq \langle u, u\rangle &= \langle \langle w, w\rangle v - \langle v, w\rangle w, \langle w, w\rangle v - \langle v, w\rangle w\rangle \\ &= |\langle w, w\rangle|^2\langle v, v\rangle - \langle w, w\rangle|\langle v, w\rangle|^2 - \langle w, w\rangle|\langle v, w\rangle|^2 + \langle w, w\rangle|\langle v, w\rangle|^2 \\ &= \langle w, w\rangle \left[\langle v, v\rangle\langle w, w\rangle - |\langle v, w\rangle|^2\right]\end{aligned}$$

Now suppose that $\langle w, w\rangle > 0$, then $\langle v, v\rangle\langle w, w\rangle - |\langle v, w\rangle|^2 \geq 0$, which implies that $|\langle v, w\rangle|^2 \leq \langle v, v\rangle\langle w, w\rangle$. 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

$$(u_1v_1 + u_2v_2 + \cdots + u_nv_n)^2 \leq (u_1 + u_2 + \cdots + u_n)^2(v_1 + v_2 + \cdots + v_n)^2$$

That is, $\left(\sum_{i=1}^n u_iv_i\right)^2 \leq \left(\sum_{i=1}^n u_i\right)^2\left(\sum_{i=1}^n v_i\right)^2$. 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

$$\left[\int_a^b f(x)g(x)dx\right]^2 \leq \int_a^b f^2(x)dx \int_a^b g^2(x)dx$$

That is, $|\langle f, g\rangle|^2 \leq \langle f, f\rangle\langle g, g\rangle$. 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,

$$|\langle f, g\rangle|^2 = \frac{121}{144} \leq \frac{56}{45} = \langle f, f\rangle\langle g, g\rangle$$

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 $\langle v, v\rangle \geq 0$, we have $\|v\| = \sqrt{\langle v, v\rangle} \geq 0$. 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 $u, v \in V$,

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&= \|u\|^2 + \|v\|^2 + 2Re(\langle u, v\rangle) \\
&\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle| \\
&\leq \|u\|^2 + \|v\|^2 + 2\,\|u\|\,\|v\| \;\; (Cauchy - Schwarz) \\
&= (\|u\| + \|v\|)^2
\end{aligned}$$

Hence $\|u+v\| \leq \|u\| + \|v\|$.
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 $u, v \in V$,*

$$\|u+v\|^2 + \|u-v\|^2 = 2\left(\|u\|^2 + \|v\|^2\right)$$

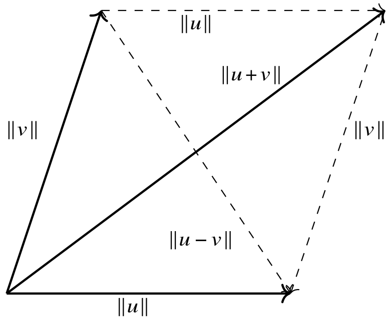


Fig. 5.11 Parallelogram law

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

$$\left|\left|\left(u+\frac{v}{2}\right)+\left(w+\frac{v}{2}\right)\right|\right|^2 + \left|\left|\left(u+\frac{v}{2}\right)-\left(w+\frac{v}{2}\right)\right|\right|^2 = 2\left|\left|u+\frac{v}{2}\right|\right|^2 + 2\left|\left|w+\frac{v}{2}\right|\right|^2$$

This implies

$$\|u+w+v\|^2 = 2\left|\left|u+\frac{v}{2}\right|\right|^2 + 2\left|\left|w+\frac{v}{2}\right|\right|^2 - \|u-w\|^2$$

Similarly,

$$\|u+w-v\|^2 = 2\left|\left|u-\frac{v}{2}\right|\right|^2 + 2\left|\left|w-\frac{v}{2}\right|\right|^2 - \|u-w\|^2$$

Then

$$\|u+w+v\|^2 - \|u+w-v\|^2 = 2\left[\left|\left|u+\frac{v}{2}\right|\right|^2 - \left|\left|u-\frac{v}{2}\right|\right|^2 + \left|\left|w+\frac{v}{2}\right|\right|^2 - \left|\left|w-\frac{v}{2}\right|\right|^2\right] \tag{5.4}$$

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*

$$\|v_1 + v_2 + \cdots + v_n\|^2 = \|v_1\|^2 + \|v_2\|^2 + \cdots + \|v_2\|^2$$

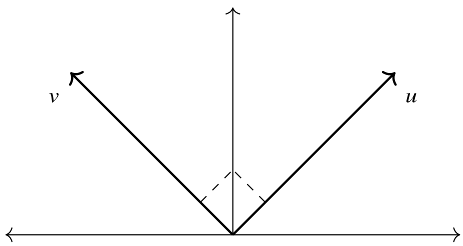


Fig. 5.12 Example for orthogonal vectors in $\mathbb{R}^2$

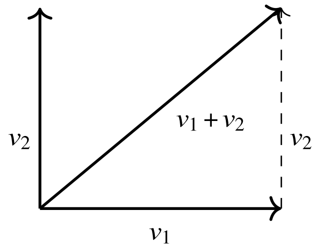


Fig. 5.13 Pythagoras theorem illustrated in $\mathbb{R}^2$

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 $S = \{e_1, e_2, \ldots, e_n\}$ 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 $i = 1, 2, \ldots, n$. 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 $\{v_1, v_2, \ldots, v_n\}$ 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 $\{v_1, v_2, \ldots, v_n\}$ 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 $S = \{v_1, v_2, \ldots, v_n\}$ 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 $w = \sum\limits_{i=1}^{n} \dfrac{\langle w, v_i \rangle}{\|v_i\|^2} v_i$. 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 $u_1 = \dfrac{w_1}{\|w_1\|}$. Then $\|u_1\| = \dfrac{\|w_1\|}{\|w_1\|} = 1$ and $span\{u_1\} = span\{v_1\}$ (Fig. 5.14).

For $0 \le i \le n-1$, define

$$w_i = v_i - \langle v_i, u_1\rangle u_1 - \cdots - \langle v_i, u_{i-1}\rangle u_{i-1},\quad u_i = \frac{w_i}{\|w_i\|}$$

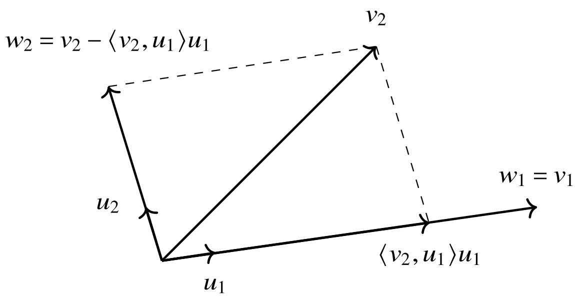


Fig. 5.14 Geometrical representation of first two steps of Gram–Schmidt process

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 $u_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{2}}(0, 1, 1, 0)$.

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 $u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{6}}(2, 1, -1, 0)$. Finally,

$$\begin{aligned} w_3 &= v_3 - \langle v_3, u_1 \rangle u_1 - \langle v_3, u_2 \rangle u_2 \\ &= (1, 0, 0, 1) - \langle (1, 0, 0, 1), \frac{1}{\sqrt{2}}(0, 1, 1, 0) \rangle \frac{1}{\sqrt{2}}(0, 1, 1, 0) \\ &\qquad - \langle (1, 0, 0, 1), \frac{1}{3}(2, 1, -1, 0) \rangle \frac{1}{3}(2, 1, -1, 0) \\ &= \frac{1}{3}(1, -1, 1, 3) \end{aligned}$$

and hence $u_3 = \frac{w_3}{\|w_3\|} = \frac{1}{2\sqrt{3}}(1, -1, 1, 3)$. 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

$$\begin{aligned} S_1^\perp &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid \langle (v_1, v_2, v_3), (1, 2, 3)\rangle = 0\} \\ &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid v_1 + 2v_2 + 3v_3 = 0\} \\ &= \text{plane passing through origin and perpendicular to the point } (1, 2, 3) \end{aligned}$$

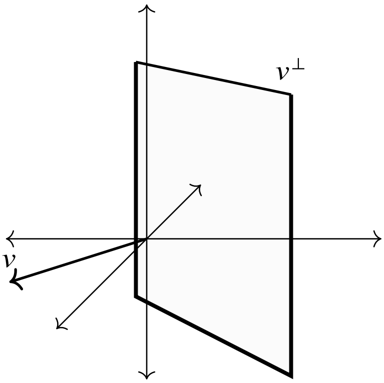


Fig. 5.15 Suppose v is a non-zero vector in $\mathbb{R}^3$. Then $v^\perp$ is the plane passing through origin O and perpendicular to the vector v

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 $w = \sum\limits_{i=1}^{k} \langle v, w_i\rangle w_i$.

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).

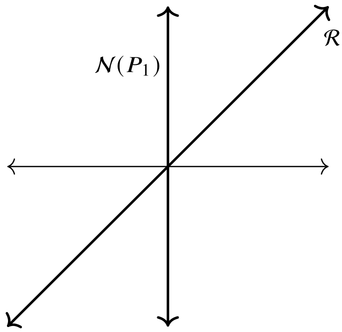


Fig. 5.16 Observe that $\mathcal{R}(P_1) \not\perp \mathcal{N}(P_1)$. Therefore P_1 is not an orthogonal projection

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

$$c = \{v = (v_1, v_2, \ldots) \in l^\infty \mid v_i \to \lambda \in \mathbb{K} \text{ as } i \to \infty\}$$
$$c_0 = \{v = (v_1, v_2, \ldots) \in l^\infty \mid v_i \to 0 \text{ as } i \to \infty\}$$
$$c_{00} = \{v = (v_1, v_2, \ldots) \in l^\infty \mid \text{all but finitely many } v_i's \text{ are equal to } 0\}$$

   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) $v = 1 + 2x + x^2$, $W = \{a_0 + a_1x + a_2x^2 \in \mathbb{P}_2[0, 1] \mid a_2 = 0\}$
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.