Chapter 2
What Are Matrices

2.1 Introduction

In real life, a *matrix* is a rectangular array with prescribed numbers n of rows and m of columns ($n \times m$ matrix). To make this array as clear as possible, one encloses it between delimiters; we choose parentheses in this book. The position at the intersection of the ith row and jth column is labeled by the pair (i,j). If the name of the matrix is M (respectively, A, X, etc.), the entry at the (i,j)th position is usually denoted m_{ij} (respectively, a_{ij}, x_{ij}). An entry can be anything provided it gives the reader information. Here is a the real-life example.

$$M = \begin{pmatrix} 11 & 27 & 83 \\ \text{blue} & \text{green} & \text{yellow} \\ \text{undefined} & \text{Republican} & \text{Democrat} \end{pmatrix}.$$

Perhaps this matrix gives the age, the preferred color, and the political tendency of three people. In the present book, however, we restrict to matrices whose entries are mathematical objects. In practice, they are elements of a ring A. In most cases, this ring is Abelian; if it is a field, then it is denoted k or K, unless it is one of the classical number fields $\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p$. When writing a matrix blockwise, it becomes a smaller matrix whose elements are themselves matrices, and thus belong to some spaces that are not even rings; having possibly different sizes, these submatrices may even belong to distinct sets.

In some circumstances (extraction of matrices or minors, e.g.) the rows and the columns can be numbered in a different way, using nonconsecutive numbers i and j. In general one needs only two finite sets I and J, one for indexing the rows and the other for indexing the columns. For instance, the following extraction from a 4×5 matrix M corresponds to the choice $I = (1,3)$, $J = (2,5,3)$.

$$M_I^J = \begin{pmatrix} m_{12} & m_{15} & m_{13} \\ m_{32} & m_{35} & m_{33} \end{pmatrix}.$$

D. Serre, *Matrices,* Graduate Texts in Mathematics 216,
DOI 10.1007/978-1-4419-7683-3_2, © Springer Science+Business Media, LLC 2010

Notice that the indices need not be taken in increasing order.

2.1.1 Addition of Matrices

The set of matrices of size $n \times m$ with entries in A is denoted by $\mathbf{M}_{n\times m}(A)$. It is an additive group, where $M + M'$ denotes the matrix M'' whose entries are given by $m''_{ij} = m_{ij} + m'_{ij}$.

2.1.2 Multiplication by a Scalar

One defines the multiplication by a scalar $a \in A$: $M' := aM$ by $m'_{ij} = am_{ij}$. One has the formulæ $a(bM) = (ab)M$, $a(M + M') = (aM) + (aM')$, and $(a+b)M = (aM) + (bM)$. Likewise we define $M'' = Ma$ by $m''_{ij} := m_{ij}a$ and we have similar properties, together with $(aM)b = a(Mb)$.

With these operations, $\mathbf{M}_{n\times m}(A)$ is a left and right A-*module*. If A is Abelian, then $aM = Ma$. When the set of scalars is a field K, $\mathbf{M}_{n\times m}(K)$ is a K-vector space. The zero matrix is denoted by 0, or 0_{nm} when one needs to avoid ambiguity:

$$0_{n\times m} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

When $m = n$, one writes simply $\mathbf{M}_n(K)$ instead of $\mathbf{M}_{n\times n}(K)$, and 0_n instead of 0_{nn}. The matrices of sizes $n \times n$ are called *square* matrices of size n. When A has a unit 1, one writes I_n for the *identity* matrix, a square matrix of order n defined by

$$m_{ij} = \delta_i^j = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

In other words,

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

2.1.3 Special Matrices

The identity matrix is a special case of a *permutation matrix*, which is a square matrix having exactly one nonzero entry in each row and each column, that entry being a 1. In other words, a permutation matrix M reads

$$m_{ij} = \delta_i^{\sigma(j)}$$

for some permutation $\sigma \in S_n$.

A square matrix for which $i < j$ implies $m_{ij} = 0$ is called a *lower-triangular* matrix. It is *upper*-triangular if $i > j$ implies $m_{ij} = 0$. It is *strictly* upper- (respectively, *lower*)-triangular if $i \geq j$ (respectively, $i \leq j$) implies $m_{ij} = 0$. It is *diagonal* if m_{ij} vanishes for every pair (i,j) such that $i \neq j$. When $d_1, \dots, d_n \in A$ are given, one denotes by $\operatorname{diag}(d_1, \dots, d_n)$ the diagonal matrix M whose diagonal term m_{ii} equals d_i for every index i. See below typical triangular and diagonal matrices.

$$L = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & * \end{pmatrix}, \qquad U = \begin{pmatrix} * & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix}, \qquad D = \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & * \end{pmatrix}.$$

When $m = 1$, a matrix M of size $n \times 1$ is called a *column vector*. One identifies it with the vector of A^n whose ith coordinate in the canonical basis is m_{i1}. This identification is an isomorphism between $\mathbf{M}_{n\times 1}(A)$ and A^n. Likewise, the matrices of size $1 \times m$ are called *row vectors*.

A matrix $M \in \mathbf{M}_{n\times m}(A)$ may be viewed as the ordered list of its columns $M^{(j)}$ $(1 \leq j \leq m)$. When the set of scalars is a field, the dimension of the linear subspace spanned by the $M^{(j)}$s in K^n is called the *rank* of M and denoted by $\operatorname{rk} M$.

Here are examples of row and column matrices, and of an $n \times m$ matrix written rowwise and columnwise:

$$R = \begin{pmatrix} * & \cdots & * \end{pmatrix}, \qquad C = \begin{pmatrix} * \\ \vdots \\ * \end{pmatrix}, \qquad M = \begin{pmatrix} R_1 \\ \text{---} \\ \vdots \\ \text{---} \\ R_m \end{pmatrix} = \begin{pmatrix} C_1 \mid \cdots \mid C_m \end{pmatrix}.$$

2.1.4 Transposition

Definition 2.1 *The transpose matrix of $M \in \mathbf{M}_{n\times m}(A)$ is the matrix $M^T \in \mathbf{M}_{m\times n}(A)$ defined as*

$$m_{ij}^T := m_{ji}, \qquad \forall 1 \leq i \leq m, 1 \leq j \leq n.$$

Mind that the numbers of rows and columns are exchanged.

For instance, the transpose of a column is a row, and conversely.

The following formulæ are obvious.

$$(aM+N)^T = aM^T + N^T, \qquad \left(M^T\right)^T = M.$$

When M is a square matrix, M and M^T have the same size and we can compare them. We thus say that $M \in \mathbf{M}_n(A)$ is *symmetric* if $M^T = M$, and *skew-symmetric* if $M^T = -M$ (notice that these two notions coincide when K has characteristic 2). We denote by $\mathbf{Sym}_n(K)$ the subset of symmetric matrices in $\mathbf{M}_n(K)$. It is a linear subspace of $\mathbf{M}_n(K)$.

2.1.5 Writing a Matrix Blockwise

The size $n \times m$ of a matrix can be quite large, and its entries may have nice patterns such as repetitions or lots of zeroes. It often helps to partition a matrix into blocks, in order to better understand its overall structure. For this purpose, we may write a matrix blockwise. The standard way to do so is to choose partitions of n and m:

$$n = n_1 + \cdots + n_r, \qquad m = m_1 + \cdots + m_s,$$

with $0 \le n_p, m_q$. For each $1 \le p \le r$ and $1 \le q \le s$, let us form the submatrix $M_{pq} \in \mathbf{M}_{n_p \times m_q}(A)$ whose entries are

$$m_{pq,ij} := m_{\nu_{p-1}+i,\mu_{q-1}+j}, \qquad \nu_{p-1} := n_1 + \cdots + n_{p-1}, \, \mu_{q-1} := m_1 + \cdots + m_{q-1}.$$

As usual, $\nu_0 = \mu_0 = 0$. Then M is nothing but an $r \times s$ matrix whose (p,q)-entry is M_{pq}. Mind that these entries belong to distinct rings, inasmuch as the numbers n_p (respectively, m_q) need not be equal. Here is an example with $r = s = 2$, where we have indicated the partitions

$$M = \left(\begin{array}{cccc|c} 0 & 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & 8 & 9 \\ 10 & 11 & 12 & 13 & 14 \\ 15 & 16 & 17 & 18 & 19 \end{array}\right).$$

In this example, M_{11} is a row, M_{22} a column, and M_{12} is just a 1×1 matrix, that is, a scalar!

Multiplication of a matrix by a scalar can be done blockwise: we have $(aM)_{pq} = aM_{pq}$. The same remark holds true for the addition of two $n \times m$ matrices M and N, provided we choose the same partitions for each: $(M+N)_{pq} = M_{pq} + N_{pq}$.

2.1.6 Writing Blockwise Square Matrices

When $n = m$, it is often useful to choose the same partition for columns and rows: $r = s$ and $m_p = n_p$ for every $1 \leq p \leq r$. We say that M is *blockwise upper-* (respectively, lower)*-triangular* if $p > q$ (respectively, $p < q$) implies $M_{pq} = 0_{n_p \times n_q}$. We also speak of *block-triangular* matrices. A block-triangular matrix need not be triangular; after all, it is not necessarily a square matrix. Likewise, M is *blockwise diagonal* if $p \neq q$ implies $M_{pq} = 0_{n_p \times n_q}$. Again, a blockwise diagonal (or *block-diagonal*) matrix need not be diagonal. If $n_p \times n_p$ matrices M_{pp} are given, we form the block-diagonal matrix $\operatorname{diag}(M_{11}, \dots, M_{rr})$.

2.2 Matrices as Linear Maps

2.2.1 Matrix of a Linear Map

Let K be a field and E, F be finite-dimensional vector spaces over K. Let us choose a basis $\mathscr{B}_E = \{\mathbf{e}^1, \dots, \mathbf{e}^m\}$ of E and a basis $\mathscr{B}_F = \{\mathbf{f}^1, \dots, \mathbf{f}^n\}$ of F. Thus $\dim E = m$ and $\dim F = n$.

A linear map $u \in \mathscr{L}(E;F)$ can be described by its action over $\mathscr{B}_E$: let m_{ij} be the coordinate of $u(\mathbf{e}^j)$ in the basis $\mathscr{B}_F$; that is,

$$u(\mathbf{e}^j) = \sum_{i=1}^{n} m_{ij} \mathbf{f}^i.$$

The numbers m_{ij} are the entries of an $n \times m$ matrix which we call M. By linearity, one finds the image of a general vector $x \in E$:

$$u\left(\sum_{j=1}^{m} x_j \mathbf{e}^j\right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{m} m_{ij} x_j\right) \mathbf{f}^i. \tag{2.1}$$

Conversely, given a matrix $M \in \mathbf{M}_{n\times m}(K)$, the formula (2.1) defines a linear map u. We therefore have a one-to-one correspondance $u \leftrightarrow M$ between $\mathscr{L}(E;F)$ and $\mathbf{M}_{n\times m}(K)$. We say that M is the *matrix of u in the bases $\mathscr{B}_E$ and $\mathscr{B}_F$*. We warn the reader that this bijection is by no means canonical, because it depends upon the choice of the bases. We sometimes employ the notation M_u for the matrix associated with u, and u_M for the linear map associated with M, but this is dangerous because the bases are not indicated explicitly; this notation is recommended only when it is clear for both the writer and the reader what the bases of the underlying spaces are.

The addition of matrices is nothing but the addition of the linear maps, and the same can be said for multiplication by a scalar:

$$M_u + M_v = M_{u+v}, \qquad u_M + u_{M'} = u_{M+M'}, \qquad M_{\lambda u} = \lambda M_u, \qquad u_{\lambda M} = \lambda u_M.$$

The bijection above is thus an isomorphism between the vector spaces $\mathscr{L}(E;F)$ and $\mathbf{M}_{n\times m}(K)$.

The jth column of M is the represention of $u_M(\mathbf{e}^j)$ in the basis $\mathscr{B}_F$. The space spanned by the $M^{(j)}$s is thus in one-to-one correspondence with the space spanned by the $u_M(\mathbf{e}^j)$s, which is nothing but the range of u_M. Thus the rank of M equals that of u_M.

2.2.1.1 Transposition versus Duality

Let $u \in \mathscr{L}(E;F)$ be given and let us choose bases $\mathscr{B}_E$ and $\mathscr{B}_F$. We recall that the dual basis of $\mathscr{B}_E$ is a basis of the dual space E'. Likewise, the dual basis of $\mathscr{B}_F$ is a basis of the dual space F'.

Proposition 2.1 *Let M be the matrix associated with u in the bases $\mathscr{B}_E$ and $\mathscr{B}_F$. Then the matrix of the adjoint u^* in the dual bases is M^T.*

Proof. Let v^j be the elements of $\mathscr{B}_E$, w^k those of $\mathscr{B}_F$, and α^j, β^k those of the dual bases. We have $\alpha^j(v^i) = \delta_i^j$ and $\beta^\ell(w^k) = \delta_\ell^k$.

Let M' be the matrix of u^*. We have

$$u^*(\beta^\ell)(v^j) = \beta^\ell(u(v^j)) = \beta^\ell\left(\sum_i m_{ij} w^i\right) = m_{\ell j}.$$

Therefore

$$u^*(\beta^\ell) = \sum_j m_{\ell j}\alpha^j,$$

showing that $m'_{j\ell} = m_{\ell j}$. □

2.2.2 Multiplication of Matrices

Let E, F, and G be three vector spaces over K, of respective dimensions p,m,n. Let $\mathscr{B}_E$, $\mathscr{B}_F$, and $\mathscr{B}_G$ be respective bases. Using the isomorphism above, we can define a product of matrices by using the composition of maps. If $M \in \mathbf{M}_{n\times m}(K)$ and $M' \in \mathbf{M}_{m\times p}(K)$, then we have two linear maps

$$u_M \in \mathscr{L}(F;G), \qquad u_{M'} \in \mathscr{L}(E;F).$$

We define MM' as the matrix of $u_M \circ u_{M'}$.

At first glance, this definition depends heavily on the choice of three bases. But the following calculation shows that it does not at all. Denote M'' the product MM'. Then

$$\begin{aligned}\sum_{i=1}^{n} m''_{ik}\mathbf{g}^i &= u_M \circ u_{M'}(\mathbf{e}^k) = u_M\left(\sum_{j=1}^{n} m'_{jk}\mathbf{f}^j\right)\\ &= \sum_{j=1}^{n} m'_{jk}u_M(\mathbf{f}^j) = \sum_{j=1}^{n} m'_{jk}\left(\sum_{i=1}^{n} m_{ij}\mathbf{g}^i\right)\\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} m_{ij}m'_{jk}\right)\mathbf{g}^i.\end{aligned}$$

Finally, the matrix $M'' = MM'$ is given by the formula

$$m''_{ij} = \sum_{k=1}^{m} m_{ik}m'_{kj}, \quad 1 \le i \le n, 1 \le j \le p, \tag{2.2}$$

which is clearly independent of the chosen bases.

We point out that a product of matrices MM' makes sense as long as the number of columns of M equals the number of rows of M', and only in this situation. If MN makes sense, then $N^T M^T$ does too, and we have the obvious formula

$$(MN)^T = N^T M^T,$$

where we warn the reader that the positions of M and N are flipped under transposition.

Thanks to the associativity of the composition, the product is associative:

$$(MP)Q = M(PQ),$$

whenever the sizes agree. Likewise, the product is distributive with respect to the addition, and associates with the scalar multiplication:

$$M(P+Q) = MP+MQ, \qquad (P+Q)M = PM+QM, \qquad (aM)M' = a(MM').$$

The following formula extends that for linear maps (see Exercise 2)

$$\mathrm{rk}(MM') \le \min\{\mathrm{rk}\, M,\ \mathrm{rk}\, M'\}.$$

In particular, we have the following.

Proposition 2.2 *The rank of a submatrix of M is not larger than that of M.*

Proof. Just remark that the submatrix M' formed by retaining only the rows of indices $i_1 < \cdots < i_r$ and the columns of indices $j_1 < \cdots < j_r$ is given by a formula $M' = PMQ$ where P is the matrix of projection from K^n over the space spanned by $\mathbf{f}^{i_1}, \ldots, \mathbf{f}^{i_r}$, and Q is the embedding matrix from the space spanned by $\mathbf{e}^{j_1}, \ldots, \mathbf{e}^{j_r}$ over K^m. Then

$$\mathrm{rk}(M') = \mathrm{rk}(PMQ) \le \mathrm{rk}(MQ) \le \mathrm{rk}(M).$$

□

2.2.2.1 Matrices with Entries in a Ring

When the scalar set is a ring A, the formula (2.2) still makes sense and lets us define a product MN when $M \in \mathbf{M}_{n\times m}(A)$ and $N \in \mathbf{M}_{m\times p}(A)$. Associativity is straightforward. In particular $\mathbf{M}_n(A)$ is itself a ring, although a noncommutative one, even if A is Abelian.

2.2.2.2 The Case of Square Matrices

Square matrices of a given size can be multiplied together, which makes $\mathbf{M}_n(K)$ an algebra. We cannot emphasize enough that *the multiplication of matrices is not commutative*: in general, MM' differs from $M'M$. This is reminiscent of the lack of commutativity of the composition of endomorphisms. It is an endless source of interesting questions regarding matrices. For instance,

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

We say that two matrices $M,N \in \mathbf{M}_n(K)$ *commute* to each other if $MN = NM$. To quantify the lack of commutativity, we define the *commutator* of square matrices M,N by

$$[M,N] := MN - NM.$$

Section 4.4 discusses the amount of noncommutativity in $\mathbf{M}_n(K)$.

In $\mathbf{M}_n(K)$, the unit matrix I_n is a neutral element for the multiplication:

$$I_nM = MI_n = M.$$

We thus have the standard notion of *inverse*: a matrix $M \in \mathbf{M}_n(K)$ is invertible if there exists $N \in \mathbf{M}_n(K)$ such that $NM = MN = I_n$. We say that N is the *inverse* of M and we denote it M^{-1}. We could as well define right-inverse and left-inverse, but we show (Proposition 3.5) that the three notions coincide. We say that M is *invertible* or *nonsingular*. A characterization of invertible matrices is given in Chapter 3. As in every algebra, the product of nonsingular matrices is nonsingular, and we have

$$(MN)^{-1} = N^{-1}M^{-1}, \qquad \left(M^{-1}\right)^{-1} = M.$$

The subset of nonsingular matrices of size n is a multiplicative group, which is denoted by $\mathbf{GL}_n(K)$.

Powers of a square matrix M are defined inductively by $M^2 = MM$, $M^3 = MM^2 = M^2M$ (from associativity), ..., $M^{k+1} = M^kM$. We complete this notation by $M^1 = M$ and $M^0 = I_n$, so that $M^jM^k = M^{j+k}$ for all $j,k \in \mathbb{N}$. The powers of a square matrix M commute pairwise. In particular, the set $K[M]$ formed by polynomials in M, which consists of matrices of the form

$$a_0I_n + a_1M + \cdots + a_rM^r, \quad a_0,\ldots,a_r \in K, \quad r \in \mathbb{N},$$

is a commutative algebra.

If M is nonsingular, we define $M^{-k} := \left(M^k\right)^{-1} = \left(M^{-1}\right)^k$, which yields $M^jM^k = M^{j+k}$ for all $j,k \in \mathbb{Z}$.

If $M^k = 0_n$ for some integer $k \in \mathbb{N}$, we say that M is *nilpotent*. We say that M is *idempotent* if $I_n - M$ is nilpotent.

A matrix $M \in \mathbf{M}_n(K)$ is *orthogonal* if $M^TM = MM^T = I_n$. It is equivalent to saying that M is nonsingular and M^{-1} is the transpose of M. The set of orthogonal matrices is a multiplicative group in $\mathbf{M}_n(K)$, called the *orthogonal* and denoted $\mathbf{O}_n(K)$.

2.2.2.3 Multiplication of a Vector and a Matrix

Another interesting case is that of multiplication with a column vector. If $M \in \mathbf{M}_{n\times m}(K)$ and $X \in K^m$, the product MX makes sense because X can be viewed as an $m \times 1$ matrix. The result is an $n \times 1$ matrix, that is, a vector Y in K^n, given by

$$y_i = \sum_{j=1}^{m} m_{ij}x_j.$$

In the terminology of Section 2.2.1, M induces a linear map $u_M \in \mathscr{L}(K^m;K^n)$, which refers to the choice of the canonical bases; this correspondence is thus canonical somehow. When $n = m$, $\mathbf{M}_n(K)$ operates over K^n and is canonically isomorphic to $\mathbf{End}(K^n)$.

The above action of a given matrix is the straightforward translation of that of its associated linear map: if x and y are the vectors associated with the columns X and Y, then $y = u_M(x)$. This leads us to extend several notions already encountered for linear maps, such as the kernel and the range:

$$\ker M = \{X \in K^m \,|\, MX = 0\}, \qquad R(M) = \{MX \,|\, X \in K^m\}.$$

The *rank* is the dimension of $R(M)$ and is denoted by $\mathrm{rk}M$. Theorem 1.2 becomes the following.

Proposition 2.3 *Let K be a field. If $M \in \mathbf{M}_{n\times m}(K)$, then*

$$m = \dim\ker M + \mathrm{rk}M.$$

2.2.2.4 Scalar Product as a Matrix Product

When ℓ is a row vector and y a column vector with the same number of entries, then ℓy is a 1×1 matrix, that is, a scalar. This can be interpreted simply in terms of linear algebra: y is the matrix of an element of K^n (which we still denote y) in the canonical basis, and ℓ is the matrix of a linear form f over K^n, in the dual basis. Then ℓy is nothing but $f(y)$. We notice that $\ell = 0$ if and only if $\ell y = 0$ for every y. Likewise, $y = 0$ if and only if $\ell y = 0$ for every ℓ.

When x and y are both in K^n, then x^T is a row vector. We can form their product $x^T y$, which is the *canonical scalar product* over K^n:

$$x^T y = x_1 y_1 + \cdots + x_n y_n.$$

We notice that $x = 0$ if and only if $x^T y = 0$ for every y. Thus the scalar product is a nondegenerate bilinear form over K^n. When $x^T y = 0$, we say that x and y are *orthogonal* (to each other) and we denote $x \perp y$. If $x \perp x$, we say that x is *isotropic*. We warn the reader that for many fields K, there are nonzero isotropic vectors, even though there is not if $K = \mathbb{R}$. For instance, if $K = \mathbb{C}$ the vector

$$x = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

is isotropic.

If $x \in K^n$, the map $y \mapsto \ell_x(y) := x^T y$ is a linear form over K^n. In addition, the map $x \mapsto \ell_x$ is one-to-one. Because $(K^n)'$ has the same dimension n, this map is an isomorphism. We thus identify K^n with its dual space in a natural way.

Two subsets A and B of K^n are *orthogonal* if every vector of A is orthogonal to every vector of B. The *orthogonal* of a subset A is the set of all vectors in K^n that are orthogonal to A; it is denoted $A^\perp$. Because of the linearity of the scalar product with respect to each argument, the orthogonal $A^\perp$ is a subspace of K^n. For the same reason, we have

$$(\mathrm{Span}(A))^\perp = A^\perp. \tag{2.3}$$

Obviously, $A \subset B$ implies $B^\perp \subset A^\perp$.

When identifiying K^n with its dual, the orthogonal of S identifies to the polar set S^0. We therefore rephrase the results obtained in Paragraph 1.2.2:

Proposition 2.4 *If E is a subspace of K^n, then*

$$\dim E^\perp + \dim E = n.$$

Proposition 2.5 *Given a subset A of K^n, its biorthogonal is the subspace spanned by A:*

$$\left(A^\perp\right)^\perp = \mathrm{Span}(A).$$

2.2.2.5 Range, Kernel, and Duality

Let $M \in \mathbf{M}_{n\times m}(K)$ and $x \in \ker M^T$. Then $x^T M = (M^T x)^T = 0$. If $y \in K^m$, there follows $x^T M y = 0$. In other words, x is orthogonal to the range of M.

Conversely, let x be orthogonal to $R(M)$. Then $(M^T x)^T y = x^T (My) = 0$ for every $y \in K^m$. This tells us that $M^T x = 0$, and proves that the orthogonal of $R(M)$

is $\ker M^T$. Applying Proposition 2.5, we find also that the orthogonal of $\ker M^T$ is $R(M)$. Exchanging the roles of M and M^T leads to the following.

Proposition 2.6 *The orthogonal of* $R(M)$ *is* $\ker M^T$ *and that of* $\ker M$ *is* $R(M^T)$.

The following consequence is sometimes called the *Fredholm principle*.

Corollary 2.1 *Let* $M \in \mathbf{M}_{n\times m}(K)$ *and* $b \in K^n$. *In order that the linear equation* $Mx = b$ *be solvable, it is necessary and sufficient that* $z^T b = 0$ *for every* $z \in \ker(M^T)$.

Assembling Propositions 2.3, 2.4, and 2.6, we obtain the following identities for a matrix $M \in \mathbf{M}_{n\times m}(K)$:

$$\begin{aligned} m &= \dim\ker M + \operatorname{rk} M, & n &= \dim\ker M^T + \operatorname{rk} M^T, \\ n &= \dim\ker M^T + \operatorname{rk} M, & m &= \dim\ker M + \operatorname{rk} M^T. \end{aligned}$$

Besides some redundancy, this list has an interesting consequence:

Proposition 2.7 *For every* $M \in \mathbf{M}_{n\times m}(K)$, *there holds*

$$\operatorname{rk} M^T = \operatorname{rk} M.$$

The kernels, however, do not have the same dimension if $m \neq n$. Only for square matrices, we deduce the following.

Proposition 2.8 *If* $M \in \mathbf{M}_n(K)$ *is a square matrix, then*

$$\dim\ker M^T = \dim\ker M.$$

Corollary 2.2 *If* $M \in \mathbf{M}_n(K)$, *then*

$$(M : K^n \to K^n \textit{ is bijective}) \iff \ker M = \{0\} \iff \operatorname{rk} M = n.$$

In particular, there is a well-defined notion of inverse in $\mathbf{M}_n(K)$: a left-inverse exists if and only if a right-inverse exists, and then they are equal to each other. In particular, this inverse is unique.

Going back to $n \times m$ matrices, we say that M is a *rank-one* matrix if $\operatorname{rk} M = 1$. A rank-one matrix decomposes as xy^T where $x \in K^n$ spans $R(M)$ and $y \in K^m$ spans $R(M^T)$ (remark that M^T is rank-one too, because of Proposition 2.7).

2.2.3 Change of Basis

Let E be a K-vector space, in which one chooses a basis $\beta = \{e_1, \dots, e_n\}$. Choosing another n-tuple $\beta' = \{e'_1, \dots, e'_n\}$ in E amounts to prescribing the coordinates of each e'_i in the basis β:

$$e'_i = \sum_{j=1}^{n} p_{ji} e_j.$$

It is thus equivalent to selecting a matrix $P \in \mathbf{M}_n(K)$. Whether β' is a basis of E depends on whether P is nonsingular: If P is nonsingular and $Q := P^{-1}$, then a straightforward calculation yields

$$e_j = \sum_{i=1}^{n} q_{ji} e'_i,$$

which shows that β' is generating, thus a basis, because of cardinality. Conversely, if β' is a basis, then the coordinates of each e_i in β' provide a matrix Q which is nothing but the inverse of P.

Proposition 2.9 *The matrix P above is nonsingular if and only if β' is another basis of E.*

Definition 2.2 *The matrix P above is the matrix of the change of basis from β to β'.*

The matrix M_u of a linear map $u \in \mathscr{L}(E;F)$ depends upon the choice of the bases of E and F. Therefore it must be modified when they are changed. The following formula describes this modification. Let β, β' be two bases of E, and γ, γ' two bases of F. Let M be the matrix of u associated with the bases (β, γ), and M' be that associated with (β', γ'). Finally, let P be the matrix of the change of basis $\beta \mapsto \beta'$ and Q that of $\gamma \mapsto \gamma'$. We have $P \in \mathbf{GL}_m(K)$ and $Q \in \mathbf{GL}_n(K)$.

With obvious notations, we have

$$f'_k = \sum_{i=1}^{n} q_{ik} f_i, \qquad e'_j = \sum_{\ell=1}^{m} p_{\ell j} e_\ell.$$

We have

$$u(e'_j) = \sum_{k=1}^{n} m'_{kj} f'_k = \sum_{i,k=1}^{n} m'_{kj} q_{ik} f_i.$$

On the other hand, we have

$$u(e'_j) = u\left(\sum_{\ell=1}^{m} p_{\ell j} e_\ell\right) = \sum_{\ell=1}^{m} p_{\ell j} \sum_{i=1}^{m} m_{i\ell} f_i.$$

Comparing the two formulæ, we obtain

$$\sum_{\ell=1}^{m} m_{i\ell} p_{\ell j} = \sum_{k=1}^{n} q_{ik} m'_{kj}, \qquad \forall 1 \le i \le n, \, 1 \le j \le m.$$

This exactly means the formula

$$MP = QM'. \tag{2.4}$$

Definition 2.3 *Two matrices $M, M' \in \mathbf{M}_{n\times m}(K)$ are* equivalent *if there exist two matrices $P \in \mathbf{GL}_m(K)$ and $Q \in \mathbf{GL}_n(K)$ such that equality* (2.4) *holds true.*

Thus equivalent matrices represent the same linear map in different bases.

2.2.3.1 The Situation for Square Matrices

When $F = E$ and thus $m = n$, it is natural to represent $u \in \mathbf{End}(E)$ by using only one basis, that is, choosing $\beta' = \beta$ with the notations above. In a change of basis, we have likewise $\gamma' = \gamma$, which means that $Q = P$. We now have

$$MP = PM',$$

or equivalently

$$M' = P^{-1}MP. \tag{2.5}$$

Definition 2.4 *Two matrices $M, M' \in \mathbf{M}_n(K)$ are* similar *if there exists a matrix $P \in \mathbf{GL}_n(K)$ such that equality* (2.5) *holds true.*

Thus similar matrices represent the same endomorphism in different bases.

The equivalence and the similarity of matrices both are equivalence relations. They are studied in detail in Chapter 9.

2.2.4 Multiplying Blockwise

Let $M \in \mathbf{M}_{n\times m}(K)$ and $M' \in \mathbf{M}_{m'\times p}(K)$ be given. We assume that partitions

$$\begin{aligned} n &= n_1 + \cdots + n_r, & m &= m_1 + \cdots + m_s, \\ m' &= m'_1 + \cdots + m'_s, & p &= p_1 + \cdots + p_t \end{aligned}$$

have been chosen, so that M and M' can be written blockwise with blocks $M_{\alpha\beta}$ and $M'_{\beta\gamma}$ with $\alpha = 1,\dots,r$, $\beta = 1,\dots,s$, $\gamma = 1,\dots,t$. We can make the product MM', which is an $n \times p$ matrix, provided that $m' = m$. On the other hand, we wish to use the block form to calculate this product more concisely. Let us write blockwise MM' by using the partitions

$$n = n_1 + \cdots + n_r, \qquad p = p_1 + \cdots + p_t.$$

We expect that the blocks $(MM')_{\alpha\gamma}$ obey a simple formula, say

$$(MM')_{\alpha\beta} = \sum_{\beta=1}^{s} M_{\alpha\beta}M'_{\beta\gamma}. \tag{2.6}$$

The block products $M_{\alpha\beta}M'_{\beta\gamma}$ make sense provided $m'_\beta = m_\beta$ for every $\beta = 1,\dots,s$ (which in turn necessitates $m' = m$). Once this requirement is fulfilled, it is easy to see that formula (2.6) is correct. We leave its verification to the reader as an exercise.

In conclusion, multiplication of matrices written blockwise follows the same rule as when the matrices are given entrywise. The multiplication is done in two stages: one level using block multiplication, the other one using multiplication in K. Actually, we may have as many levels as wished, by writing blocks blockwise (using subblocks), and so on. This recursive strategy is employed in Section 11.1.

2.3 Matrices and Bilinear Forms

Let E,F be two K-vector spaces. One chooses two respective bases $\beta = \{e_1,\dots,e_n\}$ and $\gamma = \{f_1,\dots,f_m\}$. If $B : E \times F \to K$ is a bilinear form, then

$$B(x,y) = \sum_{i,j} B(e_i,f_j)x_i y_j,$$

where the x_is are the coordinates of x in β and the y_js are those of y in γ. Let us define a matrix $M \in \mathbf{M}_{n\times m}(K)$ by $m_{ij} = B(e_i,f_j)$. Then B can be recovered from the formula

$$B(x,y) := x^T My = \sum_{i,j} m_{ij} x_i y_j. \tag{2.7}$$

Conversely, if $M \in \mathbf{M}_{n\times m}(K)$ is given, one can construct a bilinear form on $E \times F$ by applying (2.7). We say that M is the *matrix of the bilinear form* B, or that B is the bilinear form associated with M. We warn the reader that once again, the correspondence $B \leftrightarrow M$ depends upon the choice of the bases.

This correspondence is a (noncanonical) isomorphism between $\mathbf{Bil}(E,F)$ and $\mathbf{M}_{n\times m}(K)$. We point out that, opposite to the isomorphism with $\mathscr{L}(E;F)$, n is now the dimension of E and m that of F.

If M is associated with B, its transpose M^T is associated with the bilinear form B_T defined on $F \times E$ by

$$B_T(y,x) := B(x,y).$$

When $F = E$, it makes sense to assume that $\gamma = \beta$. Then M is symmetric if and only if B is symmetric: $B(x,y) = B(y,x)$. Likewise, one says that M is *alternate* if B itself is an alternate form. This is equivalent to saying that

$$m_{ij} + m_{ji} = 0, \qquad m_{ii} = 0, \qquad \forall 1 \le i,j \le n.$$

An alternate matrix is skew-symmetric, and the converse is true if $\mathrm{charc}(K) \neq 2$. If $\mathrm{charc}(K) = 2$, an alternate matrix is a skew-symmetric matrix whose diagonal vanishes.

2.3.1 Change of Bases

As for matrices associated with linear maps, we need a description of the effect of a change of bases for the matrix associated with a bilinear form.

Denoting again by P,Q the matrices of the changes of basis $\beta \mapsto \beta'$ and $\gamma \mapsto \gamma'$, and by M,M' the matrices of B in the bases (β,γ) or (β',γ'), respectively, one has

$$m'_{ij} = B(e'_i, f'_j) = \sum_{k,l} p_{ki} q_{lj} B(e_k, f_\ell) = \sum_{k,l} p_{ki} q_{lj} m_{kl}.$$

Therefore,

$$M' = P^T MQ.$$

The case $F = E$

When $F = E$ and $\gamma = \beta$, $\gamma' = \beta'$, the change of basis has the effect of replacing M by $M' = P^T MP$. We say that M and M' are *congruent*. If M is symmetric, then M' is too. This was expected, inasmuch as one expresses the symmetry of the underlying bilinear form B.

If the characteristic of K is distinct from 2, there is an isomorphism between $\mathbf{Sym}_n(K)$ and the set of quadratic forms on K^n. This isomorphism is given by the formula

$$Q(e_i + e_j) - Q(e_i) - Q(e_j) = 2m_{ij}.$$

In particular, $Q(e_i) = m_{ii}$.

Exercises

1. Let G be an $\mathbb{R}$-vector space. Verify that its complexification $G \otimes_{\mathbb{R}} \mathbb{C}$ is a $\mathbb{C}$-vector space and that $\dim_{\mathbb{C}} G \otimes_{\mathbb{R}} \mathbb{C} = \dim_{\mathbb{R}} G$.
2. Let $M \in \mathbf{M}_{n\times m}(K)$ and $M' \in \mathbf{M}_{m\times p}(K)$ be given. Show that

   $$\operatorname{rk}(MM') \le \min\{\operatorname{rk} M,\ \operatorname{rk} M'\}.$$

   First show that $\operatorname{rk}(MM') \le \operatorname{rk} M$, and then apply this result to the transpose matrix.
3. Let K be a field and let A,B,C be matrices with entries in K, of respective sizes $n \times m$, $m \times p$, and $p \times q$.

   a. Show that $\operatorname{rk} A + \operatorname{rk} B \le m + \operatorname{rk} AB$. It is sufficient to consider the case where B is onto, by considering the restriction of A to the range of B.

   b. Show that $\operatorname{rk} AB + \operatorname{rk} BC \le \operatorname{rk} B + \operatorname{rk} ABC$. One may use the vector spaces $K^p/\ker B$ and $R(B)$, and construct three homomorphisms u,v,w, with v being onto.
4. a. Let $n,n',m,m' \in \mathbb{N}^*$ and let K be a field. If $B \in \mathbf{M}_{n\times m}(K)$ and $C \in \mathbf{M}_{n'\times m'}(K)$, one defines a matrix $B \otimes C \in \mathbf{M}_{nn'\times mm'}(K)$, the tensor product, whose block form is

$$B\otimes C=\begin{pmatrix} b_{11}C & \cdots & b_{1m}C \\ \vdots & & \vdots \\ b_{n1}C & \cdots & b_{nm}C \end{pmatrix}.$$

Show that $(B,C)\mapsto B\otimes C$ is a bilinear map and prove that its range spans $\mathbf{M}_{nn'\times mm'}(K)$. Is this map onto?

b. If $p,p'\in\mathbb{N}^*$ and $D\in\mathbf{M}_{m\times p}(K)$, $E\in\mathbf{M}_{m'\times p'}(K)$, then compute $(B\otimes C)(D\otimes E)$.

c. Show that for every bilinear form $\phi:\mathbf{M}_{n\times m}(K)\times\mathbf{M}_{n'\times m'}(K)\to K$, there exists one and only one linear form $L:\mathbf{M}_{nn'\times mm'}(K)\to K$ such that $L(B\otimes C)=\phi(B,C)$.