Cambridge University Press
978-0-521-83940-2 — Matrix Analysis
Roger A. Horn , Charles R. Johnson
Excerpt
More Information

CHAPTER 0

Review and Miscellanea

0.0 Introduction

In this initial chapter we summarize many useful concepts and facts, some of which provide a foundation for the material in the rest of the book. Some of this material is included in a typical elementary course in linear algebra, but we also include additional useful items, even though they do not arise in our subsequent exposition. The reader may use this chapter as a review before beginning the main part of the book in Chapter 1; subsequently, it can serve as a convenient reference for notation and definitions that are encountered in later chapters. We assume that the reader is familiar with the basic concepts of linear algebra and with mechanical aspects of matrix manipulations, such as matrix multiplication and addition.

0.1 Vector spaces

A finite dimensional vector space is the fundamental setting for matrix analysis.

0.1.1 Scalar field. Underlying a vector space is its *field*, or set of scalars. For our purposes, that underlying field is typically the real numbers **R** or the complex numbers **C** (see Appendix A), but it could be the rational numbers, the integers modulo a specified prime number, or some other field. When the field is unspecified, we denote it by the symbol **F**. To qualify as a field, a set must be closed under two binary operations: "addition" and "multiplication." Both operations must be associative and commutative, and each must have an identity element in the set; inverses must exist in the set for all elements under addition and for all elements except the additive identity under multiplication; multiplication must be distributive over addition.

0.1.2 Vector spaces. A *vector space* V over a field **F** is a set V of objects (called *vectors*) that is closed under a binary operation ("addition") that is associative and commutative and has an identity (the *zero vector*, denoted by 0) and additive inverses in the set. The set is also closed under an operation of "scalar multiplication" of the vectors by elements of the scalar field **F**, with the following properties for all $a, b \in \mathbf{F}$ and all $x, y \in V$: $a(x + y) = ax + ay$, $(a + b)x = ax + bx$, $a(bx) = (ab)x$, and $ex = x$ for the multiplicative identity $e \in \mathbf{F}$.

For a given field **F** and a given positive integer n, the set $\mathbf{F}^n$ of n-tuples with entries from **F** forms a vector space over **F** under entrywise addition in $\mathbf{F}^n$. Our convention is that *elements of* $\mathbf{F}^n$ *are always presented as column vectors*; we often call them *n-vectors*. The special cases $\mathbf{R}^n$ and $\mathbf{C}^n$ are the basic vector spaces of this book; $\mathbf{R}^n$ is a real vector space (that is, a vector space over the real field), while $\mathbf{C}^n$ is both a real vector space and a complex vector space (a vector space over the complex field). The set of polynomials with real or with complex coefficients (of no more than a specified degree *or* of arbitrary degree) and the set of real-valued or complex-valued functions on subsets of **R** or **C** (all with the usual notions of addition of functions and multiplication of a function by a scalar) are also examples of real or complex vector spaces.

0.1.3 Subspaces, span, and linear combinations. A *subspace* of a vector space V over a field **F** is a subset of V that is, by itself, a vector space over **F** using the same operations of vector addition and scalar multiplication as in V. A subset of V is a subspace precisely when it is closed under these two operations. For example, $\{[a, b, 0]^T : a, b \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$; see (0.2.5) for the transpose notation. An intersection of subspaces is always a subspace; a union of subspaces need not be a subspace. The subsets $\{0\}$ and V are always subspaces of V, so they are often called *trivial subspaces*; a subspace of V is said to be *nontrivial* if it is different from both $\{0\}$ and V. A subspace of V is said to be a *proper subspace* if it is not equal to V. We call $\{0\}$ the *zero vector space*. Since a vector space always contains the zero vector, a subspace cannot be empty.

If S is a subset of a vector space V over a field **F**, span S is the intersection of all subspaces of V that contain S. If S is nonempty, then span $S = \{a_1v_1 + \cdots + a_kv_k : v_1, \ldots, v_k \in S, a_1, \ldots, a_k \in \mathbf{F}, \text{and } k = 1, 2, \ldots\}$. If S is empty, it is contained in every subspace of V; since the intersection of every subspace of V is the subspace $\{0\}$, the definition ensures that span $S = \{0\}$. Notice that span S is always a subspace even if S is not a subspace; S is said to *span* V if span $S = V$.

A *linear combination* of vectors in a vector space V over a field **F** is any expression of the form $a_1v_1 + \cdots + a_kv_k$ in which k is a positive integer, $a_1, \ldots, a_k \in \mathbf{F}$, and $v_1, \ldots, v_k \in V$. Thus, the span of a nonempty subset S of V consists of all linear combinations of finitely many vectors in S. A linear combination $a_1v_1 + \cdots + a_kv_k$ is *trivial* if $a_1 = \cdots = a_k = 0$; otherwise, it is *nontrivial*. A linear combination is by definition a sum of *finitely many* elements of a vector space.

Let S_1 and S_2 be subspaces of a vector space over a field **F**. The *sum* of S_1 and S_2 is the subspace

$$S_1 + S_2 = \text{span}\,\{S_1 \cup S_2\} = \{x + y : x \in S_1, y \in S_2\}$$

If $S_1 \cap S_2 = \{0\}$, we say that the sum of S_1 and S_2 is a *direct sum* and write it as $S_1 \oplus S_2$; every $z \in S_1 \oplus S_2$ can be written as $z = x + y$ with $x \in S_1$ and $y \in S_2$ in one and only one way.

0.1.4 Linear dependence and linear independence. We say that a finite list of vectors $v_1, \dots, v_k$ in a vector space V over a field $\mathbf{F}$ is *linearly dependent* if and only if there are scalars $a_1, \dots, a_k \in \mathbf{F}$, not all zero, such that $a_1 x_1 + \cdots + a_k x_k = 0$. Thus, a list of vectors $v_1, \dots, v_k$ is linearly dependent if and only if some nontrivial linear combination of $v_1, \dots, v_k$ is the zero vector. It is often convenient to say that "$v_1, \dots, v_k$ are linearly dependent" instead of the more formal statement "the list of vectors $v_1, \dots, v_k$ is linearly dependent." A list of vectors $v_1, \dots, v_k$ is said to have *length* k. A list of two or more vectors is linearly dependent if one of the vectors is a linear combination of some of the others; in particular, a list of two or more vectors in which two of the vectors in the list are identical is linearly dependent. Two vectors are linearly dependent if and only if one of the vectors is a scalar multiple of the other. A list consisting only of the zero vector is linearly dependent since $a_1 0 = 0$ for $a_1 = 1$.

A finite list of vectors $v_1, \dots, v_k$ in a vector space V over a field $\mathbf{F}$ is *linearly independent* if it is not linearly dependent. Again, it can be convenient to say that "$v_1, \dots, v_k$ are linearly independent" instead of "the list of vectors $v_1, \dots, v_k$ is linearly independent."

Sometimes one encounters natural lists of vectors that have infinitely many elements, for example, the monomials $1, t, t^2, t^3, \dots$ in the vector space of all polynomials with real coefficients or the complex exponentials $1, e^{it}, e^{2it}, e^{3it}, \dots$ in the vector space of complex-valued continuous functions that are periodic on $[0, 2\pi]$.

If certain vectors in a list (finite or infinite) are deleted, the resulting list is a *sublist* of the original list. An infinite list of vectors is said to be linearly dependent if some finite sublist is linearly dependent; it is said to be linearly independent if every finite sublist is linearly independent. Any sublist of a linearly independent list of vectors is linearly independent; any list of vectors that has a linearly dependent sublist is linearly dependent. Since a list consisting only of the zero vector is linearly dependent, any list of vectors that contains the zero vector is linearly dependent. A list of vectors can be linearly dependent, while any proper sublist is linearly independent; see (1.4.P12). An empty list of vectors is not linearly dependent, so it is linearly independent.

The *cardinality* of a finite set is the number of its (necessarily distinct) elements. For a given list of vectors $v_1, \dots, v_k$ in a vector space V, the cardinality of the *set* $\{v_1, \dots, v_k\}$ is less than k if and only if two or more vectors in the list are identical; if $v_1, \dots, v_k$ are linearly independent, then the cardinality of the set $\{v_1, \dots, v_k\}$ is k. The *span* of a list of vectors (finite or not) is the span of the set of elements of the list; a list of vectors *spans* V if V is the span of the list.

A set S of vectors is said to be linearly independent if every finite list of distinct vectors in S is linearly independent; S is said to be linearly dependent if some finite list of distinct vectors in S is linearly dependent.

0.1.5 Basis. A linearly independent list of vectors in a vector space V whose span is V is a *basis* for V. Each element of V can be represented as a linear combination of vectors in a basis in one and only one way; this is no longer true if any element whatsoever is appended to or deleted from the basis. A linearly independent list of vectors in V is a basis of V if and only if no list of vectors that properly contains it is linearly independent. A list of vectors that spans V is a basis for V if and only if no proper sublist of it spans V. The empty list is a basis for the zero vector space.

0.1.6 Extension to a basis. Any linearly independent list of vectors in a vector space V may be extended, perhaps in more than one way, to a basis of V. A vector space can have a basis that is not finite; for example, the infinite list of monomials $1, t, t^2, t^3, \ldots$ is a basis for the real vector space of all polynomials with real coefficients; each polynomial is a unique linear combination of (finitely many) elements in the basis.

0.1.7 Dimension. If there is a positive integer n such that a basis of the vector space V contains exactly n vectors, then every basis of V consists of exactly n vectors; this common cardinality of bases is the *dimension* of the vector space V and is denoted by $\dim V$. In this event, V is *finite-dimensional*; otherwise V is *infinite-dimensional*. In the infinite-dimensional case, there is a one-to-one correspondence between the elements of any two bases. The real vector space $\mathbf{R}^n$ has dimension n. The vector space $\mathbf{C}^n$ has dimension n over the field $\mathbf{C}$ but dimension $2n$ over the field $\mathbf{R}$. The basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$ in which each n-vector e_i has a 1 as its ith entry and 0s elsewhere is called the *standard basis*.

It is convenient to say "V is an n-dimensional vector space" as a shorthand for "V is a finite-dimensional vector space whose dimension is n." Any subspace of an n-dimensional vector space is finite-dimensional; its dimension is strictly less than n if it is a proper subspace.

Let V be a finite-dimensional vector space and let S_1 and S_2 be two given subspaces of V. The *subspace intersection theorem* is

$$\dim(S_1 \cap S_2) + \dim(S_1 + S_2) = \dim S_1 + \dim S_2 \tag{0.1.7.1}$$

Rewriting this identity as

$$\begin{aligned}\dim(S_1 \cap S_2) &= \dim S_1 + \dim S_2 - \dim(S_1 + S_2)\\ &\geq \dim S_1 + \dim S_2 - \dim V\end{aligned} \tag{0.1.7.2}$$

reveals the useful fact that if $\delta = \dim S_1 + \dim S_2 - \dim V \geq 1$, then the subspace $S_1 \cap S_2$ has dimension at least δ, and hence it contains δ linearly independent vectors, namely, any δ elements of a basis of $S_1 \cap S_2$. In particular, $S_1 \cap S_2$ contains a nonzero vector. An induction argument shows that if $S_1, \ldots, S_k$ are subspaces of V, and if $\delta = \dim S_1 + \cdots + \dim S_k - (k-1)\dim V \geq 1$, then

$$\dim(S_1 \cap \cdots \cap S_k) \geq \delta \tag{0.1.7.3}$$

and hence $S_1 \cap \cdots \cap S_k$ contains δ linearly independent vectors; in particular, it contains a nonzero vector.

0.1.8 Isomorphism. If U and V are vector spaces over the same scalar field $\mathbf{F}$, and if $f : U \to V$ is an invertible function such that $f(ax + by) = af(x) + bf(y)$ for all $x, y \in U$ and all $a, b \in \mathbf{F}$, then f is said to be an *isomorphism* and U and V are said to be isomorphic ("same structure"). Two finite-dimensional vector spaces over the same field are isomorphic if and only if they have the same dimension; thus, any n-dimensional vector space over $\mathbf{F}$ is isomorphic to $\mathbf{F}^n$. Any n-dimensional real vector space is, therefore, isomorphic to $\mathbf{R}^n$, and any n-dimensional complex vector space is isomorphic to $\mathbf{C}^n$. Specifically, if V is an n-dimensional vector space over a field $\mathbf{F}$

with specified basis $\mathcal{B} = \{x_1, \ldots, x_n\}$, then, since any element $x \in V$ may be written uniquely as $x = a_1x_1 + \cdots + a_nx_n$ in which each $a_i \in \mathbf{F}$, we may identify x with the n-vector $[x]_{\mathcal{B}} = [a_1 \ \ldots \ a_n]^T$. For any basis $\mathcal{B}$, the mapping $x \to [x]_{\mathcal{B}}$ is an isomorphism between V and $\mathbf{F}^n$.

0.2 Matrices

The fundamental object of study here may be thought of in two important ways: as a rectangular array of scalars and as a linear transformation between two vector spaces, given specified bases for each space.

0.2.1 Rectangular arrays. A *matrix* is an m-by-n array of scalars from a field $\mathbf{F}$. If $m = n$, the matrix is said to be *square*. The set of all m-by-n matrices over $\mathbf{F}$ is denoted by $M_{m,n}(\mathbf{F})$, and $M_{n,n}(\mathbf{F})$ is often denoted by $M_n(\mathbf{F})$. The vector spaces $M_{n,1}(\mathbf{F})$ and $\mathbf{F}^n$ are identical. If $\mathbf{F} = \mathbf{C}$, then $M_n(\mathbf{C})$ is further abbreviated to M_n, and $M_{m,n}(\mathbf{C})$ to $M_{m,n}$. Matrices are typically denoted by capital letters, and their scalar entries are typically denoted by doubly subscripted lowercase letters. For example, if

$$A = \begin{bmatrix} 2 & -\frac{3}{2} & 0 \\ -1 & \pi & 4 \end{bmatrix} = [a_{ij}]$$

then $A \in M_{2,3}(\mathbf{R})$ has entries $a_{11} = 2, a_{12} = -3/2, a_{13} = 0, a_{21} = -1, a_{22} = \pi$, $a_{23} = 4$. A *submatrix* of a given matrix is a rectangular array lying in specified subsets of the rows and columns of a given matrix. For example, $[\pi \ 4]$ is a submatrix (lying in row 2 and columns 2 and 3) of A.

Suppose that $A = [a_{ij}] \in M_{n,m}(\mathbf{F})$. The *main diagonal* of A is the list of entries $a_{11}, a_{22}, \ldots, a_{qq}$, in which $q = \min\{n, m\}$. It is sometimes convenient to express the main diagonal of A as a vector $\operatorname{diag} A = [a_{ii}]_{i=1}^{q} \in \mathbf{F}^q$. The pth *superdiagonal* of A is the list $a_{1,p+1}, a_{2,p+2}, \ldots, a_{k,p+k}$, in which $k = \min\{n, m - p\}$, $p = 0, 1, 2, \ldots, m - 1$; the pth *subdiagonal* of A is the list $a_{p+1,1}, a_{p+2,2}, \ldots, a_{p+\ell,\ell}$, in which $\ell = \min\{n - p, m\}$, $p = 0, 1, 2, \ldots, n - 1$.

0.2.2 Linear transformations. Let U be an n-dimensional vector space and let V be an m-dimensional vector space, both over the same field $\mathbf{F}$; let $\mathcal{B}_U$ be a basis of U and let $\mathcal{B}_V$ be a basis of V. We may use the isomorphisms $x \to [x]_{\mathcal{B}_U}$ and $y \to [y]_{\mathcal{B}_V}$ to represent vectors in U and V as n-vectors and m-vectors over $\mathbf{F}$, respectively. A *linear transformation* is a function $T : U \to V$ such that $T(a_1x_1 + a_2x_2) = a_1T(x_1) + a_2T(x_2)$ for any scalars a_1, a_2 and vectors x_1, x_2. A matrix $A \in M_{m,n}(\mathbf{F})$ corresponds to a linear transformation $T : U \to V$ in the following way: $y = T(x)$ if and only if $[y]_{\mathcal{B}_V} = A[x]_{\mathcal{B}_U}$. The matrix A is said to *represent the linear transformation* T (relative to the bases $\mathcal{B}_U$ and $\mathcal{B}_V$); the representing matrix A depends on the bases chosen. When we study a matrix A, we realize that we are studying a linear transformation relative to a particular choice of bases, but explicit appeal to the bases is usually not necessary.

0.2.3 Vector spaces associated with a matrix or linear transformation. Any n-dimensional vector space over $\mathbf{F}$ may be identified with $\mathbf{F}^n$; we may think of

$A \in M_{m,n}(\mathbf{F})$ as a linear transformation $x \to Ax$ from $\mathbf{F}^n$ to $\mathbf{F}^m$ (and also as an array). The *domain* of this linear transformation is $\mathbf{F}^n$; its *range* is range $A = \{y \in \mathbf{F}^m : y = Ax\}$ for some $x \in \mathbf{F}^n$; its *null space* is nullspace $A = \{x \in \mathbf{F}^n : Ax = 0\}$. The range of A is a subspace of $\mathbf{F}^m$, and the null space of A is a subspace of $\mathbf{F}^n$. The dimension of nullspace A is denoted by nullity A; the dimension of range A is denoted by rank A. These numbers are related by the *rank-nullity theorem*

$$\dim(\text{range } A) + \dim(\text{nullspace } A) = \text{rank } A + \text{nullity } A = n \tag{0.2.3.1}$$

for $A \in M_{m,n}(\mathbf{F})$. The null space of A is a set of vectors in $\mathbf{F}^n$ whose entries satisfy m homogeneous linear equations.

0.2.4 Matrix operations. Matrix addition is defined entrywise for arrays of the same dimensions and is denoted by $+$ ("$A + B$"). It corresponds to addition of linear transformations (relative to the same basis), and it inherits commutativity and associativity from the scalar field. The *zero matrix* (all entries are zero) is the additive identity, and $M_{m,n}(\mathbf{F})$ is a vector space over $\mathbf{F}$. Matrix multiplication is denoted by juxtaposition ("AB") and corresponds to the composition of linear transformations. Therefore, it is defined only when $A \in M_{m,n}(\mathbf{F})$ and $B \in M_{n,q}(\mathbf{F})$. It is associative, but not always commutative. For example,

$$\begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 8 \end{bmatrix}$$

The *identity matrix*

$$I = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in M_n(\mathbf{F})$$

is the multiplicative identity in $M_n(\mathbf{F})$; its main diagonal entries are 1, and all other entries are 0. The identity matrix and any scalar multiple of it (a *scalar matrix*) commute with every matrix in $M_n(\mathbf{F})$; they are the only matrices that do so. Matrix multiplication is distributive over matrix addition.

The symbol 0 is used throughout the book to denote each of the following: the zero scalar of a field, the zero vector of a vector space, the zero n-vector in $\mathbf{F}^n$ (all entries equal to the zero scalar in $\mathbf{F}$), and the zero matrix in $M_{m,n}(\mathbf{F})$ (all entries equal to the zero scalar). The symbol I denotes the identity matrix of any size. If there is potential for confusion, we indicate the dimension of a zero or identity matrix with subscripts, for example, $0_{p,q}$, 0_k, or I_k.

0.2.5 The transpose, conjugate transpose, and trace. If $A = [a_{ij}] \in M_{m,n}(\mathbf{F})$, the *transpose* of A, denoted by A^T, is the matrix in $M_{n,m}(\mathbf{F})$ whose i, j entry is a_{ji}; that is, rows are exchanged for columns and vice versa. For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Of course, $(A^T)^T = A$. The *conjugate transpose* (sometimes called the *adjoint* or *Hermitian adjoint*) of $A \in M_{m,n}(\mathbf{C})$, is denoted by A^* and defined by $A^* = \bar{A}^T$, in which $\bar{A}$ is the entrywise conjugate. For example,

$$\begin{bmatrix} 1+i & 2-i \\ -3 & -2i \end{bmatrix}^* = \begin{bmatrix} 1-i & -3 \\ 2+i & 2i \end{bmatrix}$$

Both the transpose and the conjugate transpose obey the *reverse-order law*: $(AB)^T = B^T A^T$ and $(AB)^* = B^* A^*$. For the complex conjugate of a product, there is no reversing: $\overline{AB} = \bar{A}\bar{B}$. If x, y are real or complex vectors of the same size, then $y^* x$ is a scalar and its conjugate transpose and complex conjugate are the same: $(y^* x)^* = \overline{y^* x} = x^* y = y^T \bar{x}$.

Many important classes of matrices are defined by identities involving the transpose or conjugate transpose. For example, $A \in M_n(\mathbf{F})$ is said to be *symmetric* if $A^T = A$, *skew symmetric* if $A^T = -A$, and *orthogonal* if $A^T A = I$; $A \in M_n(\mathbf{C})$ is said to be *Hermitian* if $A^* = A$, *skew Hermitian* if $A^* = -A$, *essentially Hermitian* if $e^{i\theta} A$ is Hermitian for some $\theta \in \mathbf{R}$, *unitary* if $A^* A = I$, and *normal* if $A^* A = AA^*$.

Each $A \in M_n(\mathbf{F})$ can be written in exactly one way as $A = S(A) + C(A)$, in which $S(A)$ is symmetric and $C(A)$ is skew symmetric: $S(A) = \frac{1}{2}(A + A^T)$ is the *symmetric part* of A; $C(A) = \frac{1}{2}(A - A^T)$ is the *skew-symmetric part* of A.

Each $A \in M_{m,n}(\mathbf{C})$ can be written in exactly one way as $A = B + iC$, in which $B, C \in M_{m,n}(\mathbf{R})$: $B = \frac{1}{2}(A + \bar{A})$ is the *real part* of A; $C = \frac{1}{2i}(A - \bar{A})$ is the *imaginary part* of A.

Each $A \in M_n(\mathbf{C})$ can be written in exactly one way as $A = H(A) + iK(A)$, in which $H(A)$ and $K(A)$ are Hermitian: $H(A) = \frac{1}{2}(A + A^*)$ is the *Hermitian part* of A; $iK(A) = \frac{1}{2}(A - A^*)$ is the *skew-Hermitian part* of A. The representation $A = H(A) + iK(A)$ of a complex or real matrix is its *Toeplitz decomposition*.

The *trace* of $A = [a_{ij}] \in M_{m,n}(\mathbf{F})$ is the sum of its main diagonal entries: $\operatorname{tr} A = a_{11} + \cdots + a_{qq}$, in which $q = \min\{m, n\}$. For any $A = [a_{ij}] \in M_{m,n}(\mathbf{C})$, $\operatorname{tr} AA^* = \operatorname{tr} A^* A = \sum_{i,j} |a_{ij}|^2$, so

$$\operatorname{tr} AA^* = 0 \text{ if and only if } A = 0 \tag{0.2.5.1}$$

A vector $x \in \mathbf{F}^n$ is *isotropic* if $x^T x = 0$. For example, $[1\ i]^T \in \mathbf{C}^2$ is a nonzero isotropic vector. There are no nonzero isotropic vectors in $\mathbf{R}^n$.

0.2.6 Metamechanics of matrix multiplication. In addition to the conventional definition of matrix-vector and matrix-matrix multiplication, several alternative viewpoints can be useful.

1. If $A \in M_{m,n}(\mathbf{F})$, $x \in \mathbf{F}^n$, and $y \in \mathbf{F}^m$, then the (column) vector Ax is a linear combination of the columns of A; the coefficients of the linear combination are the entries of x. The row vector $y^T A$ is a linear combination of the rows of A; the coefficients of the linear combination are the entries of y.
2. If b_j is the jth column of B and a_i^T is the ith row of A, then the jth column of AB is Ab_j and the ith row of AB is $a_i^T B$.

To paraphrase, in the matrix product AB, *left multiplication by A multiplies the columns of B* and *right multiplication by B multiplies the rows of A*. See (0.9.1) for an important special case of this observation when one of the factors is a diagonal matrix.

Suppose that $A \in M_{m,p}(\mathbf{F})$ and $B \in M_{n,q}(\mathbf{F})$. Let a_k be the kth column of A and let b_k be the kth column of B. Then

3. If $m = n$, then $A^T B = \left[a_i^T b_j\right]$: the i, j entry of $A^T B$ is the scalar $a_i^T b_j$.
4. If $p = q$, then $AB^T = \sum_{k=1}^p a_k b_k^T$: each summand is an m-by-n matrix, the *outer product* of a_k and b_k.

0.2.7 Column space and row space of a matrix. The range of $A \in M_{m,n}(\mathbf{F})$ is also called its *column space* because Ax is a linear combination of the columns of A for any $x \in \mathbf{F}^n$ (the entries of x are the coefficients in the linear combination); range A is the span of the columns of A. Analogously, $\{y^T A : y \in \mathbf{F}^m\}$ is called the *row space* of A. If the column space of $A \in M_{m,n}(\mathbf{F})$ is contained in the column space of $B \in M_{m,k}(\mathbf{F})$, then there is some $X \in M_{k,n}(\mathbf{F})$ such that $A = BX$ (and conversely); the entries in column j of X tell how to express column j of A as a linear combination of the columns of B.

If $A \in M_{m,n}(\mathbf{F})$ and $B \in M_{m,q}(\mathbf{F})$, then

$$\text{range } A + \text{range } B = \text{range} \begin{bmatrix} A & B \end{bmatrix} \tag{0.2.7.1}$$

If $A \in M_{m,n}(\mathbf{F})$ and $B \in M_{p,n}(\mathbf{F})$, then

$$\text{nullspace } A \cap \text{nullspace } B = \text{nullspace} \begin{bmatrix} A \\ B \end{bmatrix} \tag{0.2.7.2}$$

0.2.8 The all-ones matrix and vector. In $\mathbf{F}^n$, every entry of the vector $e = e_1 + \cdots + e_n$ is 1. Every entry of the matrix $J_n = ee^T$ is 1.

0.3 Determinants

Often in mathematics, it is useful to summarize a multivariate phenomenon with a single number, and the determinant function is an example of this. Its domain is $M_n(\mathbf{F})$ (square matrices only), and it may be presented in several different ways. We denote the determinant of $A \in M_n(\mathbf{F})$ by det A.

0.3.1 Laplace expansion by minors along a row or column. The determinant may be defined inductively for $A = [a_{ij}] \in M_n(\mathbf{F})$ in the following way. Assume that the determinant is defined over $M_{n-1}(\mathbf{F})$ and let $A_{ij} \in M_{n-1}(\mathbf{F})$ denote the submatrix of $A \in M_n(\mathbf{F})$ obtained by deleting row i and column j of A. Then, for any $i, j \in \{1, \ldots, n\}$, we have

$$\det A = \sum_{k=1}^n (-1)^{i+k} a_{ik} \det A_{ik} = \sum_{k=1}^n (-1)^{k+j} a_{kj} \det A_{kj} \tag{0.3.1.1}$$

The first sum is the *Laplace expansion by minors along row i*; the second sum is the *Laplace expansion by minors along column j*. This inductive presentation begins by defining the determinant of a 1-by-1 matrix to be the value of the single entry. Thus,

$$\det[\,a_{11}\,] = a_{11}$$

$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

and so on. Notice that $\det A^T = \det A$, $\det A^* = \overline{\det A}$ if $A \in M_n(\mathbf{C})$, and $\det I = 1$.

0.3.2 Alternating sums and permutations. A *permutation* of $\{1, \ldots, n\}$ is a one-to-one function $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. The *identity permutation* satisfies $\sigma(i) = i$ for each $i = 1, \ldots, n$. There are $n!$ distinct permutations of $\{1, \ldots, n\}$, and the collection of all such permutations forms a group under composition of functions.

Consistent with the low-dimensional examples in (0.3.1), for $A = [a_{ij}] \in M_n(\mathbf{F})$ we have the alternative presentation

$$\det A = \sum_{\sigma} \left(\operatorname{sgn} \sigma \prod_{i=1}^{n} a_{i\sigma(i)} \right) \qquad (0.3.2.1)$$

in which the sum is over all $n!$ permutations of $\{1, \ldots, n\}$ and $\operatorname{sgn} \sigma$, the “sign” or “signum” of a permutation σ, is $+1$ or -1 according to whether the minimum number of transpositions (pairwise interchanges) necessary to achieve it starting from $\{1, \ldots, n\}$ is even or odd. We say that a permutation σ is *even* if $\operatorname{sgn} \sigma = +1$; σ is *odd* if $\operatorname{sgn} \sigma = -1$.

If $\operatorname{sgn} \sigma$ in (0.3.2.1) is replaced by certain other functions of σ, one obtains *generalized matrix functions* in place of $\det A$. For example, the *permanent* of A, denoted by $\operatorname{per} A$, is obtained by replacing $\operatorname{sgn} \sigma$ by the function that is identically $+1$.

0.3.3 Elementary row and column operations. Three simple and fundamental operations on rows or columns, called *elementary row and column operations*, can be used to transform a matrix (square or not) into a simple form that facilitates such tasks as solving linear equations, determining rank, and calculating determinants and inverses of square matrices. We focus on *row operations*, which are implemented by matrices that act on the left. *Column operations* are defined and used in a similar fashion; the matrices that implement them act on the right.

Type 1: Interchange of two rows.

For $i \neq j$, interchange of rows i and j of A results from left multiplication of A by

$$\begin{bmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 0 & & \cdots & & 1 & & \\ & & & 1 & & & & & \\ & & \vdots & & \ddots & & \vdots & & \\ & & & & & 1 & & & \\ & & 1 & \cdots & \cdots & \cdots & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{bmatrix}$$

The two off-diagonal 1s are in the i, j and j, i positions, the two diagonal 0s are in positions i, i and j, j, and all unspecified entries are 0.

Type 2: Multiplication of a row by a nonzero scalar.

Multiplication of row i of A by a nonzero scalar c results from left multiplication of A by

$$\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & c & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

The i, i entry is c, all other main diagonal entries are 1, and all unspecified entries are 0.

Type 3: Addition of a scalar multiple of one row to another row.

For $i \neq j$, addition of c times row i of A to row j of A results from left multiplication of A by

$$\begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 1 & \\ & & c & & \ddots & \\ & & & & & 1 \end{bmatrix}$$

The j, i entry is c, all main diagonal entries are 1, and all unspecified entries are 0. The displayed matrix illustrates the case in which $j > i$.

The matrices of each of the three elementary row (or column) operations are just the result of applying the respective operation to the identity matrix I (on the left for a row operation; on the right for a column operation). The effect of a type 1 operation on the determinant is to multiply it by -1; the effect of a type 2 operation is to multiply it by the nonzero scalar c; a type 3 operation does not change the determinant. The determinant of a square matrix with a zero row is zero. The determinant of a square matrix is zero if and only if some subset of the rows of the matrix is linearly dependent.