The matrix on the right of the vertical line is the required inverse. The matrix on the left of the line is the 3×3 identity matrix.

It is not necessary to know in advance whether A is nonsingular. If A is *singular* (not nonsingular), we can still apply the Gauss-Jordan method, but somewhere in the process one of the diagonal elements will become zero, and it will not be possible to transform A to the identity matrix.

A system of n linear equations in n unknowns, say

$$\sum_{k=1}^{n} a_{ik}x_k = c_i\,, \qquad i = 1, 2, \ldots, n\,,$$

can be written more simply as a matrix equation,

$$AX = C,$$

where $A = (a_{ij})$ is the coefficient matrix, and X and C are column matrices,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, \qquad C = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}.$$

If A is nonsingular there is a unique solution of the system given by $X = A^{-1}C$.

2.20 Exercises

Apply the Gauss-Jordan elimination process to each of the following systems. If a solution exists, determine the general solution.

1. $x + y + 3z = 5$
   $2x - y + 4z = 11$
   $-y + z = 3.$
2. $3x + 2y + z = 1$
   $5x + 3y + 3z = 2$
   $x + y - z = 1.$
3. $3x + 2y + z = 1$
   $5x + 3y + 3z = 2$
   $7x + 4y + 5z = 3.$
4. $3x + 2y + z = 1$
   $5x + 3y + 3z = 2$
   $7x + 4y + 5z = 3$
   $x + y - z = 0.$
5. $3x - 2y + 5z + u = 1$
   $x + y - 3z + 2u = 2$
   $6x + y - 4z + 3u = 7.$
6. $x + y - 3z + u = 5$
   $2x - y + z - 2u = 2$
   $7x + y - 7z + 3u = 3.$
7. $x + y + 2z + 3u + 4v = 0$
   $2x + 2y + 7z + 11u + 14v = 0$
   $3x + 3y + 6z + 10u + 15v = 0.$
8. $x - 2y + z + 2u = -2$
   $2x + 3y - z - 5u = 9$
   $4x - y + z - u = 5$
   $5x - 3y + 2z + u = 3.$
9. Prove that the system $x + y + 2z = 2$, $2x - y + 3z = 2$, $5x - y + az = 6$, has a unique solution if $a \neq 8$. Find all solutions when $a = 8$.
10. (a) Determine all solutions of the system

$$5x + 2y - 6z + 2u = -1$$
$$x - y + z - u = -2.$$

(b) Determine all solutions of the system

$$\begin{aligned} 5x + 2y - 6z + 2u &= -1 \\ x - y + z - u &= -2 \\ x + y + z \phantom{{}-u} &= 6. \end{aligned}$$

11. This exercise tells how to determine all nonsingular 2 x 2 matrices. Prove that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = (ad - bc)I.$$

Deduce that $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is nonsingular if and only if $ad - bc \neq 0$, in which case its inverse is

$$\frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Determine the inverse of each of the matrices in Exercises 12 through 16.

12. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{bmatrix}.$

13. $\begin{bmatrix} 1 & 2 & 2 \\ 2 & -1 & 1 \\ 1 & 3 & 2 \end{bmatrix}.$

14. $\begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -4 \\ 1 & -4 & 6 \end{bmatrix}.$

15. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$

16. $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{bmatrix}.$

2.21 Miscellaneous exercises on matrices

1. If a square matrix has a row of zeros or a column of zeros, prove that it is singular.
2. For each of the following statements about $n \times n$ matrices, give a proof or exhibit a counter example.
   (a) If $AB + BA = O$, then $A^2B^3 = B^3A^2$.
   (b) If A and B are nonsingular, then $A + B$ is nonsingular.
   (c) If A and B are nonsingular, then AB is nonsingular.
   (d) If A, B, and $A + B$ are nonsingular, then $A - B$ is nonsingular.
   (e) If $A^3 = O$, then $A - I$ is nonsingular.
   (f) If the product of k matrices $A_1 \cdots A_k$ is nonsingular, then each matrix A_i is nonsingular.

3. If $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$, find a nonsingular matrix P such that $P^{-1}AP = \begin{bmatrix} 6 & 0 \\ 0 & -1 \end{bmatrix}$.
4. The matrix $A = \begin{bmatrix} a & i \\ i & b \end{bmatrix}$, where $i^2 = -1$, $a = \frac{1}{2}(1 + \sqrt{5})$, and $b = \frac{1}{2}(1 - \sqrt{5})$, has the property that $A^2 = A$. Describe completely all 2 x 2 matrices A with complex entries such that $A^2 = A$.
5. If $A^2 = A$, prove that $(A + I)^k = Z + (2^k - l)A$.
6. The special theory of relativity makes use of a set of equations of the form $x' = a(x - vt)$, $y' = y$, $z' = z$, $t' = a(t - vx/c^2)$. Here v represents the velocity of a moving object, c the speed of light, and $a = c/\sqrt{c^2 - v^2}$, where $|v| < c$. The linear transformation which maps the two-dimensional vector (x, t) onto (x', t') is called a *Lorentz transformation.* Its matrix relative to the usual bases is denoted by $L(v)$ and is given by

$$L(v) = a\begin{bmatrix} 1 & -v \\ -vc^{-2} & 1 \end{bmatrix}.$$

Note that $L(v)$ is nonsingular and that $L(0) = I$. Prove that $L(v)L(u) = L(w)$, where $w = (u + v)c^2/(uv + c^2)$. In other words, the product of two Lorentz transformations is another Lorentz transformation.
7. If we interchange the rows and columns of a rectangular matrix A, the new matrix so obtained is called the **transpose** of A and is denoted by A^t. For example, if we have

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \qquad \text{then } A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Prove **that** transposes have the following properties :
(a) $(A^t)^t = A$. (b) $(A + B)^t = A^t + B^t$. (c) $(cA)^t = cA^t$.
(d) $(AB)^t = B^tA^t$. (e) $(A^t)^{-1} = (A^{-1})^t$ if A is nonsingular.
8. A square matrix A is called an orthogonal matrix if $AA^t = I$. Verify that the 2 x 2 matrix $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal for each real θ. If A is any $n \times n$ orthogonal matrix, prove that its rows, considered as vectors in V_n, form an orthonormal set.
9. For each of the following statements about $n \times n$ matrices, give a proof or else exhibit a counter example.
(a) If A and B are orthogonal, then $A + B$ is orthogonal.
(b) If A and B are orthogonal, then AB is orthogonal.
(c) If A and B are orthogonal, then B is orthogonal.
10. *Hadumard matrices,* named for Jacques Hadamard (1865–1963), are those $n \times n$ matrices with the following properties:
I. Each entry is 1 or -1 .
II. Each row, considered as a vector in V_n, has length $\sqrt{n}$.
III. The dot product of any two distinct rows is 0.
Hadamard matrices arise in certain problems in geometry and the theory of numbers, and they have been applied recently to the construction of optimum code words in space communication. In spite of their apparent simplicity, they present many unsolved problems. The

main unsolved problem at this time is to determine all n for which an $n \times n$ Hadamard matrix exists. This exercise outlines a partial solution.
(a) Determine all 2 x 2 Hadamard matrices (there are exactly 8).
(b) This part of the exercise outlines a simple proof of the following theorem: *If A is an $n \times n$ Hadamard matrix, where $n > 2$, then n is a multiple of 4.* The proof is based on two very simple lemmas concerning vectors in n-space. Prove each of these lemmas and apply them to the rows of Hadamard matrix to prove the theorem.

LEMMA 1. *If X, Y, Z are orthogonal vectors in V_n, then we have*

$$(X + Y) \cdot (X + Z) = \|X\|^2.$$

LEMMA 2. *Write $X = (x_1, \ldots, x_n)$, $Y = (y_1, \ldots, y_n)$, $Z = (z_1, \ldots, z_n)$. If each component x_i, y_i, z_i is either 1 or -1, then the product $(x_i + y_i)(x_i + z_i)$ is either 0 or 4.*

3

DETERMINANTS

3.1 Introduction

In many applications of linear algebra to geometry and analysis the concept of a determinant plays an important part. This chapter studies the basic properties of determinants and some of their applications.

Determinants of order two and three were intoduced in Volume I as a useful notation for expressing certain formulas in a compact form. We recall ‘that a determinant of order two was defined by the formula

$$(3.1) \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Despite similarity in notation, the determinant $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ (written with vertical bars) is conceptually distinct from the matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ (written with square brackets). The determinant is a *number* assigned to the matrix according to Formula (3.1). To emphasize this connection we also write

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

Determinants of order three were defined in Volume I in terms of second-order determinants by the formula

$$(3.2) \qquad \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

This chapter treats the more general case, the determinant of a square matrix of order n for any integer $n \geq 1$. Our point of view is to treat the: determinant as a function which assigns to each square matrix A a number called the determinant of A and denoted by $\det A$. It is possible to define this function by an explicit formula generalizing (3.1) and (3.2). This formula is a sum containing $n!$ products of entries of A. For large n the formula is unwieldy and is rarely used in practice. It seems preferable to study determinants from another point of view which emphasizes more clearly their essential properties. These properties, which are important in the applications, will be taken as *axioms* for a determinant function. Initially, our program will consist of three parts: (1) To motivate the choice of axioms. (2) To deduce further properties of determinants from the axioms. (3) To prove that there is one and only one function which satisfies the axioms.

3.2 Motivation for the choice of axioms for a determinant function

In Volume I we proved that the scalar triple product of three vectors A_1, A_2, A_3 in 3-space can be expressed as the determinant of a matrix whose rows are the given vectors. Thus we have

$$A_1 \times A_2 \cdot A_3 = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

where $A_1 = (a_{11}, a_{12}, a_{13})$, $A_2 = (a_{21}, a_{22}, a_{23})$, and $A_3 = (a_{31}, a_{32}, a_{33})$.

If the rows are linearly independent the scalar triple product is nonzero; the absolute value of the product is equal to the volume of the parallelepiped determined by the three vectors A_1, A_2, A_3. If the rows are dependent the scalar triple product is zero. In this case the vectors A_1, A_2, A_3 are coplanar and the parallelepiped degenerates to a plane figure of zero volume.

Some of the properties of the scalar triple product will serve as motivation for the choice of axioms for a determinant function in the higher-dimensional case. To state these properties in a form suitable for generalization, we consider the scalar triple product as a function of the three row-vectors A_1, A_2, A_3. We denote this function by d; thus,

$$d(A_1, A_2, A_3) = A_1 \times A_2 \cdot A_3 .$$

We focus our attention on the following properties:

(a) ***Homogeneity in each row.*** For example, homogeneity in the first row states that

$$d(tA_1, A_2, A_3) = t\, d(A_1, A_2, A_3) \qquad \text{for every scalar } t .$$

(b) ***Additivity in each row.*** For example, additivity in the second row states that

$$d(A_1, A_2 + C, A_3) = d(A_1, A_2, A_3) + d(A_1, C, A_3)$$

for every vector C.

(c) ***The scalar triple product is zero if two of the rows are equal.***

(d) ***Normalization:***

$$d(i, j, k) = 1, \qquad \text{where } i = (1, 0, 0),\ j = (0, 1, 0),\ k = (0, 0, 1).$$

Each of these properties can be easily verified from properties of the dot and cross product. Some of these properties are suggested by the geometric relation between the scalar triple product and the volume of the parallelepiped determined by the geometric vectors A_1, A_2, A_3. The geometric meaning of the additive property (b) in a special case is of particular interest. If we take C = *A,* in (b) the second term on the right is zero because of (c), and relation (b) becomes

$$d(A_1, A_2 + A_1, A_3) = d(A_1, A,, A_3). \tag{3.3}$$

This property is illustrated in Figure 3.1 which shows a parallelepiped determined by *A,, A,,* A,, and another parallelepiped determined by *A,, A,* + A_2, *A,*. Equation (3.3) merely states that these two parallelepipeds have equal volumes. This is evident geometrically because the parallelepipeds have equal altitudes and bases of equal area.

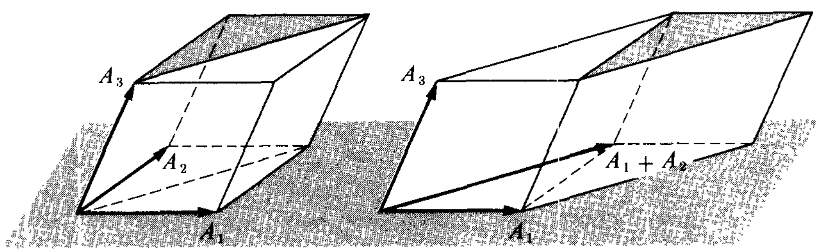

FIGURE 3.1

Geometric interpretation of the property $d(A_1$, A,, $A_3) = d(A_1$, A, + A,, $A_3)$. The two parallelepipeds have equal volumes.

3.3 A set of axioms for a determinant function

The properties of the scalar triple product mentioned in the foregoing section can be suitably generalized and used as axioms for determinants of order n. If $A = (a_{ij})$ is an n x n matrix with real or complex entries, we denote its rows by *A,, . . . , A,,* . Thus, the ith row of A is a vector in n-space given by

$$A_i = (a_{i1}, a_{i2}, \ldots, a_{in}).$$

We regard the determinant as a function of the n rows *A,, . . . , A,,* and denote its values by $d(A_1, \ldots,$ A,) or by det *A.*

AXIOMATIC DEFINITION OF A DETERMINANT FUNCTION. *A real- or complex-valuedfunction d, defined for each ordered n-tuple of vectors A,, . . . , A, in n-space, is called a determinant function of order n if it satisfies the following axioms for all choices of vectors A,, , . . , A, and C in n-space:*

AXIOM 1. HOMOGENEITY IN EACH ROW. *If the kth row A, is multiplied by a scalar t, then the determinant is also multiplied by t:*

$$d(\ldots, tA_k, \ldots) = td(\ldots, A,, \ldots).$$

AXIOM 2. ADDITIVITY IN EACH ROW. *For each k we have*

$$d(A_1, \ldots, A_k + C, \ldots, A_n) = d(A_1, \ldots, A, \ldots, A_n) + d(A_1, \ldots, C, \ldots, A,).$$

AXIOM 3. THE DETERMINANT VANISHES IF ANY TWO ROWS ARE EQUAL:

$$d(A_1, \ldots, A_n) = 0 \qquad \text{if } A_i = A_j \qquad \text{for some } i \text{ and } j \text{ with } i \neq j.$$

AXIOM 4. THE DETERMINANT OF THE IDENTITY MATRIX IS EQUAL TO 1:

$$d(I_1, \ldots, I_n) = 1, \qquad \text{where } I, \text{ is the kth unit coordinate vector.}$$

The first two axioms state that the determinant of a matrix is a linear function of each of its rows. This is often described by saying that the determinant is a *multilinear* function of its rows. By repeated application of linearity in the first row we can write

$$d\left(\sum_{k=1}^{p} t_k C_k, A_2, \ldots, A_n\right) = \sum_{k=1}^{p} t_k\, d(C_k, A_2, \ldots, A_n),$$

where $t_1, \ldots, t_p$ are scalars and $C_1, \ldots, C_p$ are any vectors in n-space.

Sometimes a weaker version of Axiom 3 is used:

AXIOM 3'. THE DETERMINANT VANISHES IF TWO ADJACENT ROWS ARE EQUAL:

$$d(A_1, \ldots, A,) = 0 \qquad \text{if } A_k = A_{k+1} \qquad \text{for some } k = 1, 2, \ldots, n-1.$$

It is a remarkable fact that for a given n there is one and only one function d which satisfies Axioms 1, 2, 3' and 4. The proof of this fact, one of the principal results of this chapter, will be given later. The next theorem gives properties of determinants deduced from Axioms 1, 2, and 3' alone. One of these properties is Axiom 3. It should be noted that Axiom 4 is not used in the proof of this theorem. This observation will be useful later when we prove uniqueness of the determinant function.

THEOREM 3.1. *A determinant function satisfying Axioms 1, 2, and 3' has the following further properties:*

(a) *The determinant vanishes if some row is 0:*

$$d(A_1, \ldots, A_n) = 0 \qquad \text{if } A_k = O \qquad \text{for some } k.$$

(b) *The determinant changes sign if two adjacent rows are interchanged:*

$$d(\ldots, A_k, A_{k+1}, \ldots) = -d(\ldots, A_{k+1}, A,, \ldots).$$

(c) The determinant changes sign if any two rows A_i and A_j with $i \neq j$ are interchanged.
(d) *The determinant vanishes if any two rows are equal:*

$$d(A_1, \ldots, A_n) = 0 \qquad if \ A_i = A_j \qquad for\ some\ i\ and\ j\ with\ i \neq j.$$

(e) The determinant vanishes if its rows are dependent.

Proof. To prove (a) we simply take $t = 0$ in Axiom 1. To prove (b), let B be a matrix having the same rows as A except for row k and row k + **11.** Let both rows B_k and B_{k+1} be equal to $A_k + A_{k+1}$. Then det $B = 0$ by Axiom 3'. Thus we may write

$$d(\ldots, A_k + A_{k+1}, A_k + A_{k+1}, \ldots) = 0.$$

Applying the additive property to row k and to row $k + 1$ we can rewrite this equation as follows :

$$d(\ldots A_k, A_k, \ldots) + d(\ldots, A_k, A_{k+1}, \ldots) + d(\ldots A_{k+1}, A_k, \ldots) \\ + d(\ldots A_{k+1}, A_{k+1}, \ldots) = 0.$$

The first and fourth terms are zero by Axiom 3'. Hence the second and third terms are negatives of each other, which proves (b).

To prove (c) we can assume that $i < j$. We can interchange rows A_i and A_j by performing an odd number of interchanges of *adjacent* rows. First we interchange row A, successively with the earlier adjacent rows $A_{j-1}, A_{j-2}, \ldots, A_i$. This requires $j - i$ interchanges. Then we interchange row A_i successively with the later adjacent rows $A_{i+1}, A_{i+2}, \ldots, A_{j-1}$. This requires $j - i -$ I further interchanges. Each interchange of adjacent rows reverses the sign of the determinant. Since there are $(j - i) + (j - i - 1) = 2(j - i) - 1$ interchanges altogether (an *odd* number), the determinant changes sign an odd number of times, which proves (c).

To prove (d), let B be the matrix obtained from A by interchanging rows A_i and A_j. Since $A_i = A_j$ we have $B = A$ and hence det B = det A. But by (c), det B = -det A. Therefore det $A = 0$.

To prove (e) suppose scalars $c_1, \ldots, c_n$ exist, not all zero, such that $\sum_{k=1}^n c_k A_k = 0$. Then any row A, with $c_k \neq 0$ can be expressed as a linear combination of the other rows. For simplicity, suppose that A, is a linear combination of the others, say $A, = \sum_{k=2}^n t_k A_k$. By linearity of the first row we have

$$d(A_1, A_2, \ldots, A_n) = d\left(\sum_{k=2}^n t_k A_k, A_2, \ldots, A_n\right) = \sum_{k=2}^n t_k\, d(A_k, A_2, \ldots, A_n).$$

But each term $d(A_k, A,, \ldots, A,)$ in the last sum is zero since A, is equal to at least one of $A_2, \ldots, A_n$. Hence the whole sum is zero. If row A, is a linear combination of the other rows we argue the same way, using linearity in the ith row. This proves (e).

3.4 Computation of determinants

At this stage it may be instructive to compute some determinants, using only the axioms and the properties in Theorem 3.1, assuming throughout that determinant functions exist. In each of the following examples we do not use Axiom 4 until the very last step in the computation.

EXAMPLE 1. *Determinant of a 2 x 2 matrix.* We shall prove that

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}. \tag{3.4}$$

Write the row vectors as linear combinations of the unit coordinate vectors $i = (1, 0)$ and $j = (0, 1)$:

$$A_1 = (a_{11}, a_{12}) = a_{11}i + a_{12}j, \qquad A_2 = (a_{21}, a_{22}) = a_{21}i + a_{22}j.$$

Using linearity in the first row we have

$$d(A_1, A_2) = d(a_{11}i + a_{12}j, A_2) = a_{11}\, d(i, A_2) + a_{12}\, d(j, A_2).$$

Now we use linearity in the second row to obtain

$$d(i, A_2) = d(i, a_{21}i + a_{22}j) = a_{21}\, d(i, i) + a_{22}\, d(i, j) = a_{22}\, d(i, j),$$

since $d(i, i) = 0$. Similarly we find

$$d(j, A_2) = d(j, a_{21}i + a_{22}j) = a_{21}\, d(j, i) = -a_{21}\, d(i, j).$$

Hence we obtain

$$d(A_1, A_2) = (a_{11}a_{22} - a_{12}a_{21})\, d(i, j).$$

But $d(i, j) = 1$ by Axiom 4, so $d(A_1, A_2) = a_{11}a_{22} - a_{12}a_{21}$, as asserted.

This argument shows that if a determinant function exists for 2 x 2 matrices, then it must necessarily have the form (3.4). Conversely, it is easy to verify that this formula does, indeed, define a determinant function of order 2. Therefore we have shown that there is one and only one determinant function of order 2.

EXAMPLE 2. *Determinant of a diagonal matrix.* A square matrix of the form

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

is called a *diagonal matrix.* Each entry a_{ij} off the main diagonal $(i \neq j)$ is zero. We shall prove that the determinant of A is equal to the product of its diagonal elements,

$$\det A = a_{11}a_{22} \cdots a_{nn}\,, \tag{3.5}$$

The kth row of .A is simply a scalar multiple of the kth unit 'coordinate vector, $A, = a_{kk}I_k$. Applying the homogeneity property repeatedly to factor out the scalars one at a time we get

$$\det A = d(A_1, \ldots, A_n) = d(a_{11}I_1, \ldots, a_{nn}I_n) = a_{11} \cdots a_{nn}\, d(I_1, \ldots, I_n).$$

This formula can be written in the form

$$\det A = a,, \cdots a,,,, \det I\,,$$

where Z is the identity matrix. Axiom 4 tells us that det Z = 1 so we obtain (3.5).

EXAMPLE 3. *Determinant of an upper triangular matrix.* A square matrix of the form

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

is called an upper triangular matrix. All the entries below the main diagonal are zero. We shall prove that the determinant of such a matrix is equal to the product of its diagonal elements,

$$\det U = u_{11}u_{22} \cdots u_{nn}\,.$$

First we prove that det $U = 0$ if some diagonal element $u_{ii} = 0$. If the last diagonal element u_{nn} is zero., then the last row is 0 and det $U = 0$ by Theorem 3.1 (a). Suppose, then, that some earlier diagonal element u_{ii} is zero. To be specific, say $u_{22} = 0$. Then each of the $n - 1$ row-vectors $U_2, \ldots, U_n$ has its first two components zero. Hence these vectors span a subspace of dimension at most $n - 2$. Therefore these $n - 1$ rows (and hence *all* the rows) are dependent. By Theorem 3.1(e), det $U = 0$. In the same way we find that det $U = 0$ if any diagonal element is zero.

Now we treat the general case. First we write the first row U_1 as a sum of two row-vectors,

$$U_1 = V_1 + V_1'\,,$$

where $V_1 = [u_{11}, 0, \ldots, 0]$ and $V_1' = [0, u_{12}, u_{13}, \ldots, u_{1n}]$. By linearity in the first row we have

$$\det. U = \det(V_1, U_2, \ldots, U_n) + \det(V_1', U_2, \ldots, U,).$$

But $\det(V_1', U_2, \ldots, U_n) = 0$ since this is the determinant of an upper triangular matrix with a diagonal element equal to 0. Hence we have

$$\det U = \det(V_1, U_2, \ldots, U_n). \tag{3.6}$$

Now we treat row-vector U_2 in a similar way, expressing it as a sum,

$$U_2 = V_2 + V_2',$$

where

$$V_2 = [0, u_{22}, 0, \ldots, 0] \quad \text{and} \quad V_2' = [0, 0, u_{23}, \ldots, u_{2n}].$$

We use this on the right of (3.6) and apply linearity in the second row to obtain

$$\det U = \det(V_1, V_2, U_3, \ldots, U_n), \tag{3.7}$$

since $\det(V_1, V_2', U_3, \ldots, U_n) = 0$. Repeating the argument on each of the succeeding rows in the right member of (3.7) we finally obtain

$$\det U = \det(V_1, V_2, \ldots, V_n),$$

where $(V_1, V_2, \ldots, V_n)$ is a diagonal matrix with the same diagonal elements as U. Therefore, by Example 2 we have

$$\det U = u_{11}u_{22} \cdots u_{nn},$$

as required.

EXAMPLE 4. *Computation by the Gauss-Jordan process.* The Gauss-Jordan elimination process for solving systems of linear equations is also one of the best methods for computing determinants. We recall that the method consists in applying three types of operations to the rows of a matrix:

(1) *Interchanging two rows.*

(2) *Multiplying all the elements in a row by a nonzero scalar.*

(3) *Adding to one row a scalar multiple of another.*

By performing these operations over and over again in a systematic fashion we can transform any square matrix A to an upper triangular matrix U whose determinant we now know how to compute. It is easy to determine the relation between $\det A$ and $\det U$. Each time operation (1) is performed the determinant changes sign. Each time (2) is performed with a scalar $c \neq 0$, the determinant is multiplied by c. Each time (3) is performed the determinant is unaltered. Therefore, if operation (1) is performed p times and if $c_1, \ldots, c_q$ are the nonzero scalar multipliers used in connection with operation (2), then we have

$$\det A = (-1)^p(c_1c_2 \cdots c_q)^{-1} \det U. \tag{3.8}$$

Again we note that this formula is a consequence of the first three axioms alone. Its proof does not depend on Axiom 4.

3.5 The uniqueness theorem

In Example 3 of the foregoing section we showed that Axioms 1, 2, and 3 imply the formula det $U == u_{11}u_{22} \cdots u_{nn}$ det I. Combining this with (3.8) we see that for every $n \times n$ matrix A there is a scalar c (depending on A) such that

$$d(A_1, \ldots, A,) = c\ d(I_1, \ldots, I_n). \tag{3.9}$$

Moreover, this formula is a consequence of Axioms 1, 2, and 3 alone. From this we can easily prove that there cannot be more than one determinant function.

THEOREM *3.2,* UNIQUENESS THEOREM FOR DETERMINANTS. *Let d be afunction satisfying all four axioms for a determinant function of order n, and let f be another function satisfying Axioms* 1, 2, *and* 3. *Then for every choice of vectors A,, . . . , A, in n-space we have*

$$f(A_1, \ldots, A_n) = d(A_1, \ldots, A_n) f(I_1, \ldots, I_n). \tag{3.10}$$

In particular, if f also satisfies Axiom 4 *we have f (A,, . . . , A,,)* $= d(A_1, \ldots, A,,)$.

Proof. Let $g(A_1, \ldots, A,) = f(A_1, \ldots, A,) - d(A_1, \ldots, A_n) f(I_1, \ldots, I_n)$. We will prove that $g(A_1, \ldots, A,,) = 0$ for every choice of $A,, \ldots, A,$. Since both d and f satisfy Axioms 1, 2, and 3 the same is true of g. Hence g also satisfies Equation (3.9) since this was deduced from the first three axioms alone. Therefore we can write

$$g(A_1, \ldots, A_n) = c\, g(I_1, \ldots, I_n), \tag{3.11}$$

where c is a scalar depending on A. Taking $A = I$ in the definition of g and noting that d satisfies Axiom 4 we find

$$g(I_1, \ldots, I_n) = f(I_1, \ldots, I_n) - f(I_1, \ldots, I_n) = 0.$$

Therefore Equation (3.11) becomes $g(A_1, \ldots, A,) = 0$. This completes the proof.

3.6 Exercises

In this set of exercises you may assume existence of a determinant function. Determinants of order 3 may be computed by Equation (3.2).

1. Compute each of the following determinants

(a) $\begin{vmatrix} 2 & 1 & 1 \\ 1 & 4 & -4 \\ 1 & 0 & 2 \end{vmatrix}$, (b) $\begin{vmatrix} 3 & 0 & 8 \\ 5 & 0 & 7 \\ -1 & 4 & 2 \end{vmatrix}$, (c) $\begin{vmatrix} a & 1 & 0 \\ 2 & a & 2 \\ 0 & 1 & a \end{vmatrix}$

2. If $\det \begin{bmatrix} x & y & z \\ 3 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} = 1$, compute the determinant of each of the following matrices:

$$\text{(a)} \begin{bmatrix} 2x & 2y & 2z \\ \frac{3}{2} & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} x & y & z \\ 3x+3 & 3y & 3z+2 \\ x+1 & y+1 & z+1 \end{bmatrix}, \quad \text{(c)} \begin{bmatrix} x-1 & y-1 & z-1 \\ 4 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix}.$$

3. (a) Prove that $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{vmatrix} = (b-a)(c-a)(c-b).$

(b) Find corresponding formulas for the determinants

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix}$$

4. Compute the determinant of each of the following matrices by transforming each of them to an upper triangular matrix.

$$\text{(a)} \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{bmatrix}, \quad \text{(c)} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^4 & b^4 & c^4 & d^4 \end{bmatrix},$$

$$\text{(d)} \begin{bmatrix} a & 1 & 0 & 0 & 0 \\ 4 & a & 2 & 0 & 0 \\ 0 & 3 & a & 3 & 0 \\ 0 & 0 & 2 & a & 4 \\ 0 & 0 & 0 & 1 & a \end{bmatrix}, \quad \text{(e)} \begin{bmatrix} & 1 & & & & \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}.$$

5. A lower triangular matrix $A = (a_{ij})$ is a square matrix with all entries above the main diagonal equal to 0; that is, $a_{ij} = 0$ whenever $i < j$. Prove that the determinant of such a matrix is equal to the product of its diagonal entries: $\det A = a_{11}a_{22} \cdots a_{nn}$.

6. Let f_1, f_2, g_1, g_2 be four functions differentiable on an interval (a, b). Define

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ g_1(x) & g_2(x) \end{vmatrix}$$

for each x in (a, b). Prove that

$$F'(x) = \begin{vmatrix} f_1'(x) & f_2'(x) \\ g_1(x) & g_2(x) \end{vmatrix} + \begin{vmatrix} f_1(x) & f_2(x) \\ g_1'(x) & g_2'(x) \end{vmatrix}.$$

7. State and prove a generalization of Exercise 6 for the determinant

$$F(x) = \begin{vmatrix} f_1(x) & f_2(x) & f_3(x) \\ g_1(x) & g_2(x) & g_3(x) \\ h_1(x) & h_2(x) & h_3(x) \end{vmatrix}.$$

8. (a) If $F(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1'(x) & f_2'(x) \end{vmatrix}$, prove that $F'(x) = \begin{vmatrix} f_1(x) & f_2(x) \\ f_1''(x) & f_2''(x) \end{vmatrix}$.

(b) State and prove a corresponding result for 3 x 3 determinants, assuming the validity of Equation (3.2).

9. Let U and V be two $n \times n$ upper triangular matrices.
(a) Prove that each of $U + V$ and UV is an upper triangular matrix.
(h) Prove that $\det(UV) = (\det U)(\det V)$.
(c) If $\det U \neq 0$ prove that there is an upper triangular matrix U^{-1} such that $UU^{-1} = I$, and deduce that $\det(U^{-1}) = 1/\det U$.

10. Calculate $\det A$, $\det(A^{-1})$, and A^{-1} for the following upper triangular matrix:

$$A = \begin{bmatrix} 2 & 3 & 4 & 5 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

3.7 The product formula for determinants

In this section we use the uniqueness theorem to prove that the determinant of a product of two square matrices is equal to the product of their determinants,

$$\det(AB) = (\det A)(\det B),$$

assuming that a determinant function exists.

We recall that the product AB of two matrices $A = (a_{ij})$ and $B = (b_{ij})$ is the matrix $C = (c_{ij})$ whose i, j entry is given by the formula

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}. \tag{3.12}$$

The product is defined only if the number of columns of the left-hand factor A is equal to the number of rows of the right-hand factor B. This is always the case if both A and B are square matrices of the same size.

The proof of the product formula will make use of a simple relation which holds between the rows of AB and the rows of A. We state this as a lemma. As usual, we let A_i denote the ith row of matrix A.