The Calculus of Functions of Several Variables

Section 1.6

Operations with Matrices

In the previous section we saw the important connection between linear functions and matrices. In this section we will discuss various operations on matrices which we will find useful in our later work with linear functions.

The algebra of matrices

If M is an $n \times m$ matrix with a_{ij} in the ith row and jth column, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$, then we will write $M = [a_{ij}]$. With this notation the definitions of addition, subtraction, and scalar multiplication for matrices are straightforward.

Definition Suppose $M = [a_{ij}]$ and $N = [b_{ij}]$ are $n \times m$ matrices and c is a real number. Then we define

$$M + N = [a_{ij} + b_{ij}], \tag{1.6.1}$$

$$M - N = [a_{ij} - b_{ij}], \tag{1.6.2}$$

and

$$cM = [ca_{ij}]. \tag{1.6.3}$$

In other words, we define addition, subtraction, and scalar multiplication for matrices by performing these operations on the individual elements of the matrices, in a manner similar to the way we perform these operations on vectors.

Example If

$$M = \begin{bmatrix} 1 & 2 & 3 \\ -5 & 3 & -1 \end{bmatrix}$$

and

$$N = \begin{bmatrix} 3 & 1 & 4 \\ 1 & -3 & 2 \end{bmatrix},$$

then, for example,

$$M + N = \begin{bmatrix} 1+3 & 2+1 & 3+4 \\ -5+1 & 3-3 & -1+2 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 7 \\ -4 & 0 & 1 \end{bmatrix},$$

$$M - N = \begin{bmatrix} 1-3 & 2-1 & 3-4 \\ -5-1 & 3+3 & -1-2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -1 \\ -6 & 6 & -3 \end{bmatrix},$$

and

$$3M = \begin{bmatrix} 3 & 6 & 9 \\ -15 & 9 & -2 \end{bmatrix}.$$

These operations have natural interpretations in terms of linear functions. Suppose $L : \mathbb{R}^m \to \mathbb{R}^n$ and $K : \mathbb{R}^m \to \mathbb{R}^n$ are linear with $L(\mathbf{x}) = M\mathbf{x}$ and $K(\mathbf{x}) = N\mathbf{x}$ for $n \times m$ matrices M and N. If we define $L + K : \mathbb{R}^n \to \mathbb{R}^m$ by

$$(L + K)(\mathbf{x}) = L(\mathbf{x}) + K(\mathbf{x}), \tag{1.6.4}$$

then

$$(L + K)(\mathbf{e}_j) = L(\mathbf{e}_j) + K(\mathbf{e}_j) \tag{1.6.5}$$

for $j = 1, 2, \ldots, m$. Hence the jth column of the matrix which represents $L + K$ is the sum of the jth columns of M and N. In other words,

$$(L + K)(\mathbf{x}) = (M + N)\mathbf{x} \tag{1.6.6}$$

for all $\mathbf{x}$ in $\mathbb{R}^m$. Similarly, if we define $L - K : \mathbb{R}^m \to \mathbb{R}^n$ by

$$(L - K)(\mathbf{x}) = L(\mathbf{x}) - K(\mathbf{x}), \tag{1.6.7}$$

then

$$(L - K)(\mathbf{x}) = (M - N)\mathbf{x}. \tag{1.6.8}$$

If, for any scalar c, we define $cL : \mathbb{R}^m \to \mathbb{R}^n$ by

$$cL(\mathbf{x}) = c(L(\mathbf{x})), \tag{1.6.9}$$

then

$$cL(\mathbf{e}_j) = c(L(\mathbf{e}_j)) \tag{1.6.10}$$

for $j = 1, 2, \ldots, m$. Hence the jth column of the matrix which represents cL is the scalar c times the jth column of M. That is,

$$cL(\mathbf{x}) = (cM)\mathbf{x} \tag{1.6.11}$$

for all $\mathbf{x}$ in $\mathbb{R}^m$. In short, the operations of addition, subtraction, and scalar multiplication for matrices corresponds in a natural way with the operations of addition, subtraction, and scalar multiplication for linear functions.

Now consider the case where $L : \mathbb{R}^m \to \mathbb{R}^p$ and $K : \mathbb{R}^p \to \mathbb{R}^n$ are linear functions. Let M be the $p \times m$ matrix such that $L(\mathbf{x}) = M\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$ and let N be the $n \times p$ matrix such that $K(\mathbf{x}) = N\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^p$. Since for any $\mathbf{x}$ in $\mathbb{R}^m$, $L(\mathbf{x})$ is in $\mathbb{R}^p$, we can form $K \circ L : \mathbb{R}^m \to \mathbb{R}^n$, the composition of K with L, defined by

$$K \circ L(\mathbf{x}) = K(L(\mathbf{x})). \tag{1.6.12}$$

Now

$$K(L(\mathbf{x})) = N(M\mathbf{x}), \tag{1.6.13}$$

so it would be natural to define NM, the product of the matrices N and M, to be the matrix of $K \circ L$, in which case we would have

$$N(M\mathbf{x}) = (NM)\mathbf{x}. \tag{1.6.14}$$

Thus we want the jth column of NM, $j = 1, 2, \ldots, m$, to be

$$K \circ L(\mathbf{e}_j) = N(L(\mathbf{e}_j)), \tag{1.6.15}$$

which is just the dot product of $L(\mathbf{e}_j)$ with the rows of N. But $L(\mathbf{e}_j)$ is the jth column of M, so the jth column of NM is formed by taking the dot product of the jth column of M with the rows of N. In other words, the entry in the ith row and jth column of NM is the dot product of the ith row of N with the jth column of M. We write this out explicitly in the following definition.

Definition If $N = [a_{ij}]$ is an $n \times p$ matrix and $M = [b_{ij}]$ is a $p \times m$ matrix, then we define the *product* of N and M to be the $n \times m$ matrix $NM = [c_{ij}]$, where

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \tag{1.6.16}$$

$i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

Note that NM is an $n \times m$ matrix since $K \circ L : \mathbb{R}^m \to \mathbb{R}^n$. Moreover, the product NM of two matrices N and M is defined only when the number of columns of N is equal to the number of rows of M.

Example If

$$N = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & -2 \end{bmatrix}$$

and

$$M = \begin{bmatrix} 2 & -2 & 1 & 3 \\ 1 & 2 & -1 & -2 \end{bmatrix},$$

then

$$\begin{aligned} NM &= \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & 3 \\ 1 & 2 & -1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 2+2 & -2+4 & 1-2 & 3-4 \\ -2+3 & 2+6 & -1-3 & -3-6 \\ 4-2 & -4-4 & 2+2 & 6+4 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 2 & -1 & -1 \\ 1 & 8 & -4 & -9 \\ 2 & -8 & 4 & 10 \end{bmatrix}. \end{aligned}$$

Note that N is 3×2, M is 2×4, and NM is 3×4. Also, note that it is not possible to form the product in the other order.

Example Let $L : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear function defined by

$$L(x, y) = (3x - 2y, x + y, 4y)$$

and let $K : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear function defined by

$$K(x, y, z) = (2x - y + z, x - y - z).$$

Then the matrix for L is

$$M = \begin{bmatrix} 3 & -2 \\ 1 & 1 \\ 0 & 4 \end{bmatrix},$$

the matrix for K is

$$N = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix},$$

and the matrix for $K \circ L : \mathbb{R}^2 \to \mathbb{R}^2$ is

$$NM = \begin{bmatrix} 2 & -1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ 1 & 1 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 6 - 1 + 0 & -4 - 1 + 4 \\ 3 - 1 + 0 & -2 - 1 - 4 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 2 & -7 \end{bmatrix}.$$

In other words,

$$K \circ L(x, y) = \begin{bmatrix} 5 & -1 \\ 2 & -7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5x - y \\ 2x - 7y \end{bmatrix}.$$

Note that it in this case it is possible to form the composition in the other order. The matrix for $L \circ K : \mathbb{R}^3 \to \mathbb{R}^3$ is

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

and so

$$L \circ K(x, y, z) = \begin{bmatrix} 4 & -1 & 5 \\ 3 & -2 & 0 \\ 4 & -4 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4x - y + 5z \\ 3x - 2y \\ 4x - 4y - 4z \end{bmatrix}.$$

In particular, note that not only is $NM \neq MN$, but in fact NM and MN are not even the same size.

Determinants

The notion of the determinant of a matrix is closely related to the idea of area and volume. To begin our definition, consider the 2×2 matrix

$$M = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$$

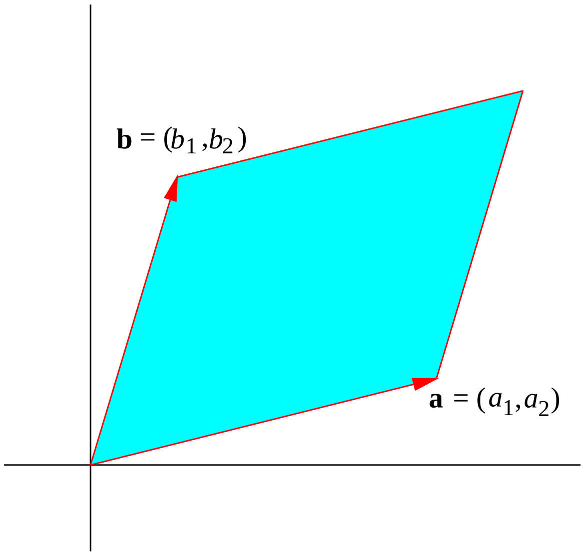

Figure 1.6.1 A parallelogram in $\mathbb{R}^2$ with adjacent sides **a** and **b**

and let $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$. If P is the parallelogram which has **a** and **b** for adjacent sides and A is the area of P (see Figure 1.6.1), then we saw in Section 1.3 that

$$A = \|(a_1, a_2, 0) \times (b_1, b_2, 0)\| = \|(0, 0, a_1 b_2 - a_2 b_1\| = |a_1 b_2 - a_2 b_1|. \tag{1.6.17}$$

This motivates the following definition.

Definition Given a 2×2 matrix

$$M = \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix},$$

the *determinant* of M, denoted $\det(M)$, is

$$\det(M) = a_1 b_2 - a_2 b_1. \tag{1.6.18}$$

Hence we have $A = |\det(M)|$. In words, for a 2×2 matrix M, the absolute value of the determinant of M equals the area of the parallelogram which has the rows of M for adjacent sides.

Example We have

$$\det \begin{bmatrix} 1 & 3 \\ -4 & 5 \end{bmatrix} = (1)(5) - (3)(-4) = 5 + 12 = 17.$$

Now consider a 3×3 matrix

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

and let $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$, and $\mathbf{c} = (c_1, c_2, c_3)$. If V is the volume of the parallelepiped P with adjacent edges $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, then, again from Section 1.3,

$$\begin{aligned} V &= |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})| \\ &= |a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)| \\ &= \left| a_1 \det \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} - a_2 \det \begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix} + a_3 \det \begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix} \right|. \end{aligned} \tag{1.6.19}$$

Definition Given a 3×3 matrix

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix},$$

the *determinant* of M, denoted $\det(M)$, is

$$\det(M) = a_1 \det \begin{bmatrix} b_2 & b_3 \\ c_2 & c_3 \end{bmatrix} - a_2 \det \begin{bmatrix} b_1 & b_3 \\ c_1 & c_3 \end{bmatrix} + a_3 \det \begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}. \tag{1.6.20}$$

Similar to the 2×2 case, we have $V = |\det(M)|$.

Example We have

$$\begin{aligned} \det \begin{bmatrix} 2 & 3 & 9 \\ 2 & 1 & -4 \\ 5 & 1 & -1 \end{bmatrix} &= 2 \det \begin{bmatrix} 1 & -4 \\ 1 & -1 \end{bmatrix} - 3 \det \begin{bmatrix} 2 & -4 \\ 5 & -1 \end{bmatrix} + 9 \det \begin{bmatrix} 2 & 1 \\ 5 & 1 \end{bmatrix} \\ &= 2(-1 + 4) - 3(-2 + 20) + 9(2 - 5) \\ &= 6 - 54 - 27 \\ &= -75. \end{aligned}$$

Given an $n \times n$ matrix $M = [a_{ij}]$, let M_{ij} be the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of M. If for $n = 1$ we first define $\det(M) = a_{11}$ (that is, the determinant of a 1×1 matrix is just the value of its single entry), then we could express, for $n = 2$, the definition of a the determinant of a 2×2 matrix given in (1.6.18) in the form

$$\det(M) = a_{11} \det(M_{11}) - a_{12} \det(M_{12}) = a_{11}a_{22} - a_{12}a_{21}. \tag{1.6.21}$$

Similarly, with $n = 3$, we could express the definition of the determinant of M given in (1.6.20) in the form

$$\det(M) = a_{11} \det(M_{11}) - a_{12} \det(M_{12}) + a_{13} \det(M_{13}). \tag{1.6.22}$$

Following this pattern, we may form a recursive definition for the determinant of an $n \times n$ matrix.

Definition Suppose $M = [a_{ij}]$ is an $n \times n$ matrix and let M_{ij} be the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of M, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$. For $n = 1$, we define the *determinant* of M, denoted $\det(M)$, by

$$\det(M) = a_{11}. \tag{1.6.23}$$

For $n > 1$, we define the *determinant* of M, denoted $\det(M)$, by

$$\begin{aligned} \det(M) &= a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + \cdots + (-1)^{1+n} a_{1n}\det(M_{1n}) \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(M_{1j}). \end{aligned} \tag{1.6.24}$$

We call the definition recursive because we have defined the determinant of an $n \times n$ matrix in terms of the determinants of $(n-1) \times (n-1)$ matrices, which in turn are defined in terms of the determinants of $(n-2) \times (n-2)$ matrices, and so on, until we have reduced the problem to computing the determinants of 1×1 matrices.

Example For an example of the determinant of a 4×4 matrix, we have

$$\begin{aligned} \det \begin{bmatrix} 2 & 1 & 3 & 2 \\ 2 & 1 & 4 & 1 \\ -2 & 3 & -1 & 2 \\ 1 & 2 & 1 & 1 \end{bmatrix} &= 2\det \begin{bmatrix} 1 & 4 & 1 \\ 3 & -1 & 2 \\ 2 & 1 & 1 \end{bmatrix} - \det \begin{bmatrix} 2 & 4 & 1 \\ -2 & -1 & 2 \\ 1 & 11 & 1 \end{bmatrix} \\ &\quad + 3\det \begin{bmatrix} 2 & 1 & 1 \\ -2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} - 2\det \begin{bmatrix} 2 & 1 & 4 \\ -2 & 3 & -1 \\ 1 & 2 & 1 \end{bmatrix} \\ &= 2((-1-2) - 4(3-4) + (3+2)) - (2(-1-2) \\ &\quad - 4(-2-2) + (-2+1)) + 3(2(3-4) - (-2-2) \\ &\quad + (-4-3)) - 2(2(3+2) - (-2+1) + 4(-4-3)) \\ &= 2(-3+4+5) - (-6+16-1) + 3(-2+4-7) \\ &\quad - 2(10+1-28) \\ &= 12 - 9 - 15 + 34 \\ &= 22. \end{aligned}$$

The next theorem states that there is nothing special about using the first row of the matrix in the expansion of the determinant specified in (1.6.24), nor is there anything special about expanding along a row instead of a column. The practical effect is that we may compute the determinant of a given matrix expanding along whichever row or column is most convenient. The proof of this theorem would take us too far afield at this point, so we will omit it (but you will be asked to verify the theorem for the special cases $n = 2$ and $n = 3$ in Problem 10).

Theorem Let $M = [a_{ij}]$ be an $n \times n$ matrix and let M_{ij} be the $(n-1) \times (n-1)$ matrix obtained by deleting the ith row and jth column of M. Then for any $i = 1, 2, \ldots, n$,

$$\det(M) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}), \tag{1.6.25}$$

and for any $j = 1, 2, \ldots, n$,

$$\det(M) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}), \tag{1.6.26}$$

Example The simplest way to compute the determinant of the matrix

$$M = \begin{bmatrix} 4 & 0 & 3 \\ 2 & 3 & 1 \\ -3 & 0 & -2 \end{bmatrix}$$

is to expand along the second column. Namely,

$$\begin{aligned} \det(M) = (-1)^{1+2}(0) \det \begin{bmatrix} 2 & 1 \\ -3 & -2 \end{bmatrix} + (-1)^{2+2}(3) \det \begin{bmatrix} 4 & 3 \\ -3 & -2 \end{bmatrix} \\ + (-1)^{3+2}(0) \det \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \\ = 3(-8+9) \\ = 3. \end{aligned}$$

You should verify that expanding along the first row, as we did in the definition of the determinant, gives the same result.

In order to return to the problem of computing volumes, we need to define a parallelepiped in $\mathbb{R}^n$. First note that if P is a parallelogram in $\mathbb{R}^2$ with adjacent sides given by the vectors $\mathbf{a}$ and $\mathbf{b}$, then

$$P = \{\mathbf{y} : \mathbf{y} = t\mathbf{a} + s\mathbf{b}, 0 \leq t \leq 1, 0 \leq s \leq 1\}. \tag{1.6.27}$$

That is, for $0 \leq t \leq 1$, $t\mathbf{a}$ is a point between $\mathbf{0}$ and $\mathbf{a}$, and for $0 \leq s \leq 1$, $s\mathbf{b}$ is a point between $\mathbf{0}$ and $\mathbf{b}$; hence $t\mathbf{a} + s\mathbf{b}$ is a point in the parallelogram P. Moreover, every point in P may be expressed in this form. See Figure 1.6.2. The following definition generalizes this characterization of parallelograms.

Definition Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ be linearly independent vectors in $\mathbb{R}^n$. We call

$$P = \{\mathbf{y} : y = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n, 0 \leq t_i \leq 1, i = 1, 2, \ldots, n\} \tag{1.6.28}$$

an *n-dimensional parallelepiped* with adjacent edges $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

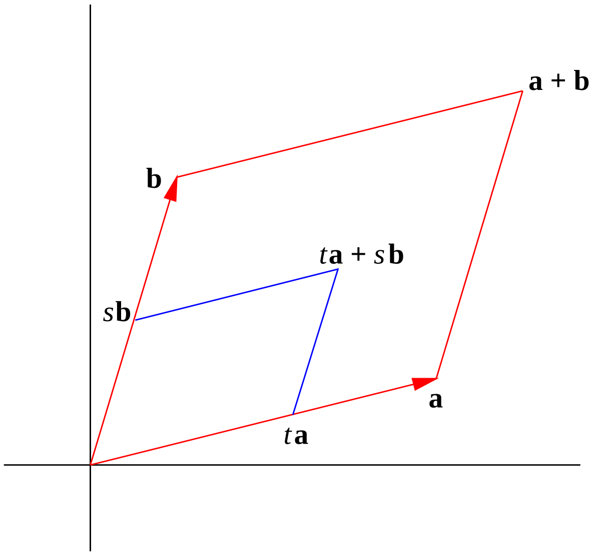

Figure 1.6.2 A parallelogram in $\mathbb{R}^2$ with adjacent sides $\mathbf{a}$ and $\mathbf{b}$

Definition Let P be an n-dimensional parallelepiped with adjacent edges $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ and let M be the $n \times n$ matrix which has $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ for its rows. Then the *volume* of P is defined to be $|\det(M)|$.

It may be shown, using (1.6.26) and induction, that if N is the matrix obtained by interchanging the rows and columns of an $n \times n$ matrix M, then $\det(N) = \det(M)$ (see Problem 12). Thus we could have defined M in the previous definition using $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ for columns rather than rows.

Now suppose $L : \mathbb{R}^n \to \mathbb{R}^n$ is linear and let M be the $n \times n$ matrix such that $L(\mathbf{x}) = M\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Let C be the n-dimensional parallelepiped with adjacent edges $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, the standard basis vectors for $\mathbb{R}^n$. Then C is a 1×1 square when $n = 2$ and a $1 \times 1 \times 1$ cube when $n = 3$. In general, we may think of C as an n-dimensional unit cube. Note that the volume of C is, by definition,

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

Suppose $L(\mathbf{e}_1), L(\mathbf{e}_2), \ldots, L(\mathbf{e}_n)$ are linearly independent and let P be the n-dimensional parallelepiped with adjacent edges $L(\mathbf{e}_1), L(\mathbf{e}_2), \ldots, L(\mathbf{e}_n)$. Note that if

$$\mathbf{x} = t_1\mathbf{e}_1 + t_2\mathbf{e}_2 + \cdots + t_n\mathbf{e}_n,$$

where $0 \leq t_k \leq 1$ for $k = 1, 2, \ldots, n$, is a point in C, then

$$L(\mathbf{x}) = t_1 L(\mathbf{e}_1) + t_2 L(\mathbf{e}_2) + \cdots + t_n L(\mathbf{e}_n)$$

is a point in P. In fact, L maps the n-dimensional unit cube C exactly onto the n-dimensional parallelepiped P. Since $L(\mathbf{e}_1), L(\mathbf{e}_2), \ldots, L(\mathbf{e}_n)$ are the columns of M, it follows that the volume of P equals $|\det(M)|$. In other words, $|\det(M)|$ measures how much L stretches or shrinks the volume of a unit cube.

Theorem Suppose $L : \mathbb{R}^n \to \mathbb{R}^n$ is linear and M is the $n \times n$ matrix such that $L(\mathbf{x}) = M\mathbf{x}$. If $L(\mathbf{e}_1), L(\mathbf{e}_2), \ldots, L(\mathbf{e}_n)$ are linear independent and P is the n-dimensional parallelepiped with adjacent edges $L(\mathbf{e}_1), L(\mathbf{e}_2), \ldots, L(\mathbf{e}_n)$, then the volume of P is equal to $|\det(M)|$.

Problems

1. Let $M = \begin{bmatrix} 2 & 3 \\ -2 & 1 \\ 4 & -1 \end{bmatrix}$ and $N = \begin{bmatrix} 3 & -2 \\ 1 & 0 \\ 2 & -5 \end{bmatrix}$. Evaluate the following.

   (a) $3M$ (b) $M - N$

   (c) $2M + N$ (d) $2N - 6M$

2. Evaluate the following matrix products.

   (a) $\begin{bmatrix} 3 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & -2 \end{bmatrix}$

   (c) $\begin{bmatrix} 2 & 1 & 3 \\ -3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & -1 \\ 0 & 2 & 4 \\ 2 & 1 & -2 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 2 & 3 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 1 \\ -2 & 4 \\ 0 & -4 \end{bmatrix}$

3. Suppose $L : \mathbb{R}^3 \to \mathbb{R}^3$ and $K : \mathbb{R}^3 \to \mathbb{R}^3$ are defined by
$$L(x, y, z) = (2x + 3y, y - x + 2z, x + 2y - z)$$
   and
$$K(x, y, z) = (2x + 4y - 3z, x + y + z, 3x - y + 4z).$$
   Find the matrices for the following linear functions.

   (a) $3L$ (b) $L + K$

   (c) $2L - K$ (d) $K + 2L$

   (e) $K \circ L$ (f) $L \circ K$

4. Let $\mathbb{R}_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear function which rotates a vector in $\mathbb{R}^2$ counterclockwise through an angle θ. In Section 1.5 we saw that
$$R_\theta(x, y) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$
   Show that the matrix for $R_\theta \circ R_\alpha$ is the same as the matrix for $R_{\theta+\alpha}$. In other words, show that $R_\theta \circ R_\alpha = R_{\theta+\alpha}$.

5. Compute the determinants of the following matrices.

(a) $\begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ (b) $\begin{bmatrix} -3 & -2 \\ 1 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 9 \\ 5 & -3 & -1 \end{bmatrix}$ (d) $\begin{bmatrix} -1 & 2 & -1 \\ 3 & 1 & 0 \\ 5 & -4 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 2 & -1 & 3 \\ 4 & 3 & -2 & 1 \\ 1 & 4 & -4 & 3 \\ 1 & 3 & 3 & 1 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 2 & -2 & 3 & 1 \\ 0 & 2 & 0 & 2 & 0 \\ -3 & 2 & 0 & 1 & 5 \\ 1 & 5 & -2 & 1 & 0 \\ 6 & -5 & 0 & 2 & -4 \end{bmatrix}$

6. Find the area of the parallelogram in $\mathbb{R}^2$ with vertices at $(1, -2)$, $(3, -1)$, $(4, 1)$, and $(2, 0)$.

7. Find the volume of the parallelepiped in $\mathbb{R}^3$ with bottom vertices at $(1, 1, 1)$, $(2, 3, 2)$, $(-1, 4, 3)$, and $(-2, 2, 2)$ and top vertices at $(1, 0, 5)$, $(2, 2, 6)$, $(-1, 3, 7)$, and $(-2, 1, 6)$.

8. Let P be the 4-dimensional parallelepiped with adjacent edges $\mathbf{a}_1 = (2, 1, 2, 1)$, $\mathbf{a}_2 = (-2, 0, 1, 1)$, $\mathbf{a}_3 = (1, 1, 3, 6)$, and $\mathbf{a}_4 = (-3, 1, 5, 0)$. Find the volume of P.

9. Find 2×2 matrices A and B for which $AB \neq BA$.

10. Verify that (1.6.25) and (1.6.26) hold for all 2×2 and 3×3 matrices.

11. An $n \times n$ matrix $M = [a_{ij}]$ is called a *diagonal matrix* if $a_{ij} = 0$ for all $i \neq j$. Show that if M is a diagonal matrix, then $\det(M) = a_{11}a_{22}\cdots a_{nn}$.

12. If M is an $n \times m$ matrix, then the $m \times n$ matrix M^T whose columns are the rows of M is called the *transpose* of M. For example, if

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

then

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

(a) Show that for a 2×2 matrix M, $\det(M^T) = \det(M)$.

(b) Show that for a 3×3 matrix M, $\det(M^T) = \det(M)$. (Hint: Using (1.6.26), expand $\det(M)$ along the first row and $\det(M^T)$ along the first column.)

(c) Use induction to show that for any $n \times n$ matrix M, $\det(M^T) = \det(M)$. (Hint: Note that $(M^T)_{ij} = (M_{ji})^T$.)

13. Let $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{y} = (y_1, y_2, y_3)$ be vectors in $\mathbb{R}^3$ and let $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ be the standard basis vectors for $\mathbb{R}^3$. Show that applying (1.6.20) to the array

$$\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}$$

yields $\mathbf{x} \times \mathbf{y}$. Discuss what is correct and what is incorrect about the statement

$$\mathbf{x} \times \mathbf{y} = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

14. Show that the set of all points $\mathbf{x} = (x, y, z)$ in $\mathbb{R}^3$ which satisfy the equation

$$\det \begin{bmatrix} x & y & z \\ 1 & 2 & -1 \\ 3 & 1 & 2 \end{bmatrix} = 0$$

is a plane passing through the points $(0, 0, 0)$, $(1, 2, -1)$, and $(3, 1, 2)$.

15. Verify directly that if $L : \mathbb{R}^m \to \mathbb{R}^p$ and $K : \mathbb{R}^p \to \mathbb{R}^n$ are linear functions, then $K \circ L : \mathbb{R}^m \to \mathbb{R}^n$ is also a linear function.