© Pigprox/Shutterstock, Inc.

The Arc de Triomphe in Paris, France, was commissioned by Napoleon in 1806. He hired architect Jean-François Thérèse Chalgrin to determine the perfect location for the arch: the Place de l'Étoile. It took many years and subsequent architects to construct and was completed in 1836.

CHAPTER

Matrices and Linear Transformations

2

In this chapter we introduce operations of addition and multiplication for matrices. We discuss the algebraic properties of these operations. We define powers of matrices and inverses of matrices. These tools lead to further methods for solving linear systems and insights into their behavior. We lay the foundation for using matrices to define functions, called linear transformations, on vector spaces. These transformations include rotations, expansions, and reflections. Their implementation in computer graphics is discussed.

The reader will see how matrices are used in a wide range of applications. They are used in archaeology to determine the chronological order of artifacts, in cryptography to ensure security, and in demography to predict population movement. The inverse of a matrix is used in a model for analyzing the interdependence of economies. Wassily Leontief received a Nobel Prize for his work in this field. This model is now a standard tool for investigating economic structures ranging from cities and corporations to states and countries.

Throughout these discussions we shall be conscious of numerical implications. We shall be aware of the need for efficiency and accuracy in implementing matrix models.

2.1 Addition, Scalar Multiplication, and Multiplication of Matrices

A convenient notation has been developed for working with matrices. Matrices consist of rows and columns. Rows are labeled from the top of the matrix, columns from the left. The location of an element in a matrix is described by giving the row and column in which it lies. The element in row i, column j of the matrix A is denoted a_{ij}.

$$a_{ij}$$

1st subscript indicates row

2nd subscript indicates column

We refer to a_{ij} as the (i, j)th element of the matrix A. We can visualize an arbitrary $m \times n$ matrix A as in Figure 2.1.

If the number of rows m is equal to the number of columns n, A is said to be a **square matrix**. The elements of a square matrix A where the subscripts are equal, namely $a_{11}, a_{22}, \ldots, a_{nn}$, form the **main diagonal**. See Figure 2.2.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$m \times n$ matrix A

Figure 2.1

main diagonal

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

square $n \times n$ matrix A

Figure 2.2

For example, consider the matrix

$$A = \begin{bmatrix} 1 & -2 & -1 \\ 3 & -3 & 4 \\ 2 & 7 & 5 \end{bmatrix}$$

a_{12} is the element in row 1, column 2. Thus $a_{12} = -2$. We see that $a_{23} = 4$ and $a_{31} = 2$. A is a square matrix. The main diagonal of A consists of the elements $a_{11} = 1, a_{22} = -3, a_{33} = 5$.

We now begin our development of an algebraic theory of matrices. Some of this follows the same pattern as our development of vector algebra in the "Linear Equations, Vectors, and Matrices" chapter. We extend the idea of equality, and operations of addition and scalar multiplication to rectangular arrays.

DEFINITION Two matrices are *equal* if they are of the same size and if their corresponding elements are equal.

Thus $A = B$ if they are of the same size, and $a_{ij} = b_{ij}$ for all i and j.

This definition will enable us to introduce equations involving matrices. It immediately allows us to define an operation of addition for matrices.

Addition of Matrices

DEFINITION Let A and B be matrices of the same size. Their *sum* $A + B$ is the matrix obtained by adding together the corresponding elements of A and B. The matrix $A + B$ will be of the same size as A and B. If A and B are not of the same size, they cannot be added, and we say that *the sum does not exist.*

Thus if $C = A + B$, then $c_{ij} = a_{ij} + b_{ij}$.

For example, if $A = \begin{bmatrix} 1 & 4 & 7 \\ 0 & -2 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 5 & -6 \\ -3 & 1 & 8 \end{bmatrix}$, and $C = \begin{bmatrix} -5 & 4 \\ 2 & 7 \end{bmatrix}$, then

$$A + B = \begin{bmatrix} 1+2 & 4+5 & 7-6 \\ 0-3 & -2+1 & 3+8 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 1 \\ -3 & -1 & 11 \end{bmatrix}$$

A and C are not of the same size. $A + C$ does not exist.

Scalar Multiplication of Matrices

When working with matrices, it is customary to refer to numbers as *scalars*. We shall use uppercase letters to denote matrices and lowercase letters for scalars. The next step in the development of a theory of matrices is to introduce a rule for multiplying matrices by scalars.

DEFINITION Let A be a matrix and c be a scalar. The *scalar multiple* of A by c, denoted cA, is the matrix obtained by multiplying every element of A by c. The matrix cA will be the same size as A.

Thus if $B = cA$, then $b_{ij} = ca_{ij}$.

For example, if $A = \begin{bmatrix} 1 & -2 & 4 \\ 7 & -3 & 0 \end{bmatrix}$, then $3A = \begin{bmatrix} 3 & -6 & 12 \\ 21 & -9 & 0 \end{bmatrix}$.

Negation and Subtraction

The matrix $(-1)C$ is written $-C$ and is called the *negative of C*. Thus, for example, if

$$C = \begin{bmatrix} 1 & 0 & -7 \\ -3 & 6 & 2 \end{bmatrix}, \text{ then } -C = \begin{bmatrix} -1 & 0 & 7 \\ 3 & -6 & -2 \end{bmatrix}$$

We now define subtraction in terms of addition and scalar multiplication. Let

$$A - B = A + (-1)B$$

This definition implies that *subtraction is performed between matrices of the same size by subtracting corresponding elements*. Thus if $C = A - B$, then $c_{ij} = a_{ij} - b_{ij}$.

Suppose $A = \begin{bmatrix} 5 & 0 & -2 \\ 3 & 6 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 8 & -1 \\ 0 & 4 & 6 \end{bmatrix}$. Then

$$A - B = \begin{bmatrix} 5-2 & 0-8 & -2-(-1) \\ 3-0 & 6-4 & -5-6 \end{bmatrix} = \begin{bmatrix} 3 & -8 & -1 \\ 3 & 2 & -11 \end{bmatrix}$$

We now complete this discussion of matrix operations by defining matrix multiplication. We give a rule for computing the arbitrary element of a matrix product. This, in effect, defines the complete matrix product.

Matrix Multiplication

The most natural way of multiplying two matrices might seem to be to multiply corresponding elements when the matrices are of the same size, and to say that the product does not exist if they are of different size. However, mathematicians have introduced an alternative rule that is more useful. It involves multiplying the rows of the first matrix times the columns of the second matrix in a systematic manner.

DEFINITION Let the number of columns in a matrix A be the same as the number of rows in a matrix B. The product $C = AB$ then exists. The element in row i and column j of C is obtained by multiplying the corresponding elements of row i of A and column j of B and adding the products.

If the number of columns in A does not equal the number of rows in B, *the product does not exist*.

Let A have r columns and B have r rows so that AB exists. Then the elements in the ith row of A

are $a_{i1}, a_{i2}, \ldots, a_{ir}$, and in the jth column of B are $\begin{matrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{matrix}$. Thus, if $C = AB$,

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj}$$

EXAMPLE 1 Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 0 & 1 \\ 3 & -2 & 6 \end{bmatrix}$. Determine AB and BA if the products exist.

SOLUTION

A has two columns and B has two rows; thus AB exists. Interpret A in terms of its rows and B in terms of its columns, and multiply the rows by the columns in the following systematic manner.

$$AB = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 5 & 0 & 1 \\ 3 & -2 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} [1 \quad 3]\begin{bmatrix} 5 \\ 3 \end{bmatrix} & [1 \quad 3]\begin{bmatrix} 0 \\ -2 \end{bmatrix} & [1 \quad 3]\begin{bmatrix} 1 \\ 6 \end{bmatrix} \\ [2 \quad 0]\begin{bmatrix} 5 \\ 3 \end{bmatrix} & [2 \quad 0]\begin{bmatrix} 0 \\ -2 \end{bmatrix} & [2 \quad 0]\begin{bmatrix} 1 \\ 6 \end{bmatrix} \end{bmatrix}$$

Multiply 1st row times each column in turn.

Multiply 2nd row times each column in turn.

$$= \begin{bmatrix} (1 \times 5) + (3 \times 3) & (1 \times 0) + (3 \times -2) & (1 \times 1) + (3 \times 6) \\ (2 \times 5) + (0 \times 3) & (2 \times 0) + (0 \times -2) & (2 \times 1) + (0 \times 6) \end{bmatrix}$$

Multiply rows by columns.

$$= \begin{bmatrix} 14 & -6 & 19 \\ 10 & 0 & 2 \end{bmatrix}$$

Let us now look at the product BA.

$$BA = \begin{bmatrix} 5 & 0 & 1 \\ 3 & -2 & 6 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

When we try to compute the first element of this product, we get $(5 \times 1) + (0 \times 2) + (1 \times ?)$. The elements do not match. The same shortcoming applies to all the other elements of BA. We say that BA *does not exist*.

EXAMPLE 2 Let $C = \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix}$. Compute CD and DC. Comment on your answer.

SOLUTION

$$CD = \begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix}\begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} (1 \times 4) + (2 \times 2) & (1 \times -1) + (2 \times 3) \\ (-3 \times 4) + (0 \times 2) & (-3 \times -1) + (0 \times 3) \end{bmatrix}$$

$$= \begin{bmatrix} 8 & 5 \\ -12 & 3 \end{bmatrix}$$

$$DC = \begin{bmatrix} 4 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} (4 \times 1) + (-1 \times -3) & (4 \times 2) + (-1 \times 0) \\ (2 \times 1) + (3 \times -3) & (2 \times 2) + (3 \times 0) \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 8 \\ -7 & 4 \end{bmatrix}$$

Comment: Observe that CD and DC both exist but $CD \neq DC$. We see that the order in which two matrices are multiplied is important. Unlike the multiplication of real numbers, matrix multiplication is not commutative. Only rarely will CD and DC be equal. Note that Example 1 further illustrates this noncommutativity of matrix multiplication. In that example, $AB \neq BA$ because BA does not exist.

Matrix multiplication is not commutative.

The following example illustrates that we can use the definition of matrix multiplication to compute any desired element in a product matrix without computing the whole product.

EXAMPLE 3 Let $C = AB$ for the following matrices A and B. Determine the element c_{23} of C.

$$A = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} -7 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}$$

SOLUTION

c_{23} is the element in row 2, column 3 of C. It will be the product of row 2 of A and column 3 of B. We get

$$c_{23} = \begin{bmatrix} -3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-3 \times 2) + (4 \times 1) = -2$$

Size of a Product Matrix

Let us now discuss the size of a product matrix. Let A be an $m \times r$ matrix and B be an $r \times n$ matrix. A has r columns and B has r rows. AB thus exists. The first row of AB is obtained by multiplying the first row of A by each column of B in turn. Thus the number of columns in AB is equal to the number of columns in B. The first column of AB results from multiplying each row of A in turn with the first column of B. Thus the number of rows in AB is equal to the number of rows in A. AB will be an $m \times n$ matrix.

If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then AB will be an $m \times n$ matrix.

We can picture this result as follows:

$$\begin{array}{ccc} A & B & = \quad AB \\ m \times r & r \times n & \quad m \times n \end{array}$$

insides match

outsides give size of AB

For example, suppose A is a 5×6 matrix and B is a 6×7 matrix. Matrix A has six columns, whereas B has six rows. Thus AB exists. AB will be a 5×7 matrix.

Special Matrices

We now define three classes of matrices that play an important role in matrix theory.

DEFINITION *A zero matrix* is a matrix all of whose elements are zeros. *A diagonal matrix* is a square matrix in which all the elements not on the main diagonal are zeros. An *identity matrix* is a diagonal matrix in which every diagonal element is 1. See the following figure.

$$O_{mn} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \qquad I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Zero matrix O_{mn} Diagonal matrix A Identity matrix I_n

Zero matrices play a role in matrix theory similar to the role of the number 0 for real numbers, and identity matrices play a role similar to the number 1. These roles are described in the following theorem, which we illustrate by means of an example.

THEOREM 2.1

Let A be an $m \times n$ matrix and O_{mn} be the zero $m \times n$ matrix. Let B be an $n \times n$ square matrix, and O_n and I_n be the zero and identity $n \times n$ matrices. Then

$$A + O_{mn} = O_{mn} + A = A$$
$$BO_n = O_nB = O_n$$
$$BI_n = I_nB = B$$

EXAMPLE 4 Let $A = \begin{bmatrix} 2 & 1 & -3 \\ 4 & 5 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}$.

We see that

$$A + O_{23} = \begin{bmatrix} 2 & 1 & -3 \\ 4 & 5 & 8 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -3 \\ 4 & 5 & 8 \end{bmatrix} = A$$

$$BO_2 = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = O_2$$

$$BI_2 = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -3 & 4 \end{bmatrix} = B$$

Similarly, $O_{23} + A = A$, $O_2B = O_2$, $I_2B = B$.

We have introduced matrix multiplication by giving the rule for determining individual elements of the product. This element approach is the most useful for computing products of small matrices by hand. There are, however, numerous ways of looking at a product. We now introduce column approaches that are useful in theory and in applications.

Matrix Multiplication in Terms of Columns

(a) Consider the product AB where A is an $m \times n$ matrix and B is an $n \times r$ matrix (so that AB exists). Let the columns of B be the matrices $B_1, B_2, \ldots, B_r$. Write B as $[B_1 \quad B_2 \ldots B_r]$. Thus

$$AB = A[B_1 \quad B_2 \ldots B_r]$$

Matrix multiplication implies that the columns of the product are $AB_1, AB_2, \ldots, AB_r$. We can write

$$\boxed{AB = [AB_1 \quad AB_2 \ldots AB_r]}$$

For example, suppose $A = \begin{bmatrix} 2 & 0 \\ 1 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 1 & 3 \\ 0 & 2 & -1 \end{bmatrix}$. Then

$$AB = \left[\begin{bmatrix} 2 & 0 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 2 & 0 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} 3 \\ -1 \end{bmatrix} \right]$$
$$= \begin{bmatrix} 8 & 2 & 6 \\ 4 & 11 & -2 \end{bmatrix}$$

(b) The matrix product AB, where B is a column matrix, often occurs in practice. Consider the general case where A is an $m \times n$ matrix and B is an $n \times 1$ matrix. Write A in terms of its columns $[A_1 \quad A_2 \ldots A_n]$. Then

$$AB = [A_1 \quad A_2 \ldots A_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Matrix multiplication gives

$$\boxed{AB = b_1 A_1 + b_2 A_2 + \cdots + b_n A_n}$$

As for vectors, the expression $b_1 A_1 + b_2 A_2 + \cdots + b_n A_n$ is called a *linear combination* of $A_1, A_2, \ldots, A_n$. It is computed by performing the scalar multiples and then adding corresponding elements of the resulting matrices.

For example, suppose $A = \begin{bmatrix} 2 & 3 & 1 \\ -4 & 8 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}$. Then

$$AB = 3\begin{bmatrix} 2 \\ -4 \end{bmatrix} - 2\begin{bmatrix} 3 \\ 8 \end{bmatrix} + 5\begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ -12 \end{bmatrix} - \begin{bmatrix} 6 \\ 16 \end{bmatrix} + \begin{bmatrix} 5 \\ 25 \end{bmatrix}$$
$$= \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

This type of fluency in expressing the product in various ways is valuable when working with matrices. We shall, for example, use the first of these approaches in arriving at a way for finding the inverse of a matrix later in this chapter.

Partitioning of Matrices

In the previous discussion we subdivided matrices into column matrices. We now extend these ideas. A matrix can be subdivided into a number of *submatrices* (or *blocks*). For example, the following matrix A can be subdivided into the submatrices P, Q, R, and S.

$$A = \left[\begin{array}{c|cc} 0 & -1 & 2 \\ 3 & 1 & 4 \\ \hline -2 & 5 & -3 \end{array}\right] = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

where $P = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$, $Q = \begin{bmatrix} -1 & 2 \\ 1 & 4 \end{bmatrix}$, $R = [-2]$, and $S = [5 \quad 3]$.

Provided appropriate rules are followed, matrix addition and multiplication can be applied to submatrices as if they were elements of an ordinary matrix. Partitioning is used to reduce memory requirements and also to speed up computation on computers, especially when matrices have large blocks of zeros.[1]

Let us look at addition. Let A and B be matrices of the same kind. If A and B are partitioned in the same way, into $P, \ldots, U$ and $H, \ldots, M$, as follows, for example, their sum is the sum of the corresponding submatrices.

$$A + B = \begin{bmatrix} P & Q & R \\ S & T & U \end{bmatrix} + \begin{bmatrix} H & I & J \\ K & L & M \end{bmatrix} = \begin{bmatrix} P + H & Q + I & R + J \\ S + K & T + L & U + M \end{bmatrix}$$

In multiplication, any partition of the first matrix in a product determines the row partition of the second matrix. For example, let us consider the product of the following matrices AB,

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & -2 \\ 4 & -3 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 5 & 4 \end{bmatrix}$$

Let A be subdivided

$$A = \left[\begin{array}{cc|c} 1 & 2 & -1 \\ \hline 3 & 0 & -2 \\ 4 & -3 & 2 \end{array}\right] = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

where $P = [1 \quad 2]$, $Q = [-1]$, $R = \begin{bmatrix} 3 & 0 \\ 4 & -3 \end{bmatrix}$, and $S = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$.

A is interpreted as having two columns in this form. B must be subdivided into a suitable form having two rows for matrix multiplication to be possible. The following would be a suitable partition for B.

$$B = \left[\begin{array}{cc} -1 & 0 \\ 2 & 1 \\ \hline 5 & 4 \end{array}\right] = \begin{bmatrix} M \\ N \end{bmatrix}$$

where $M = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}$ and $N = [5 \quad 4]$.

Then we get, $AB = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}\begin{bmatrix} M \\ N \end{bmatrix} = \begin{bmatrix} PM + QN \\ RM + SN \end{bmatrix}$.

This method of multiplying partitioned matrices is called *block multiplication*.

EXAMPLE 5 Let $A = \begin{bmatrix} 1 & -1 \\ 3 & 0 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 1 \end{bmatrix}$. Consider the following partition of A.

$$A = \left[\begin{array}{c|c} 1 & -1 \\ 3 & 0 \\ \hline 2 & 4 \end{array}\right]$$

[1]We illustrate the ideas for small matrices. Results can be quickly checked using the standard element ways of adding and multiplying matrices.

Under this partition, A is interpreted as a 2×2 matrix. For the product AB to exist, B must be partitioned into a matrix having two rows.

One appropriate partition of B is

$$B = \left[\begin{array}{ccc} 1 & 2 & -1 \\ \hline 1 & 3 & 1 \end{array}\right]$$

B is interpreted as a 2×1 matrix.

Let us check these partitions to see that they do indeed work.

$$AB = \left[\begin{array}{c|c} 1 & -1 \\ 3 & 0 \\ \hline 2 & 4 \end{array}\right] \left[\begin{array}{ccc} 1 & 2 & -1 \\ \hline 1 & 3 & 1 \end{array}\right]$$

Multiply the submatrices,

$$\begin{aligned} \begin{bmatrix} 1 \\ 3 \end{bmatrix}\begin{bmatrix} 1 & 2 & -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 3 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \end{bmatrix} + \begin{bmatrix} -1 & -3 & -1 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 & -2 \\ 3 & 6 & -3 \end{bmatrix} \end{aligned}$$

and

$$[2][1 \quad 2 \quad -1] + [4][1 \quad 3 \quad 1] = [2 \quad 4 \quad -2] + [4 \quad 12 \quad 4] = [6 \quad 16 \quad 2]$$

Thus the product is

$$AB = \left[\begin{array}{ccc} 0 & -1 & -2 \\ 3 & 6 & -3 \\ \hline 6 & 16 & 2 \end{array}\right]$$

It can be verified using elementwise multiplication that this is indeed the product of A and B. There are three other possible partitions of B that can be used to compute AB for this partition of A, namely

$$\left[\begin{array}{c|c|c} 1 & 2 & -1 \\ \hline 1 & 3 & 1 \end{array}\right], \left[\begin{array}{c|cc} 1 & 2 & -1 \\ \hline 1 & 3 & 1 \end{array}\right], \text{ and } \left[\begin{array}{cc|c} 1 & 2 & -1 \\ \hline 1 & 3 & 1 \end{array}\right]$$

EXERCISE SET 2.1

Matrix Operations

1. Let $A = \begin{bmatrix} 5 & 4 \\ -1 & 7 \\ 9 & -3 \end{bmatrix}$, $B = \begin{bmatrix} -3 & 0 \\ 4 & 2 \\ 5 & -7 \end{bmatrix}$,

$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, and $D = \begin{bmatrix} 9 & -5 \\ 3 & 0 \end{bmatrix}$

Compute the following (if they exist).

(a) $A + B$ **(b)** $2B$ **(c)** $-D$
(d) $C + D$ **(e)** $A + D$ **(f)** $2A + B$
(g) $A - B$

2. Let $A = \begin{bmatrix} 9 \\ 2 \\ -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & -1 & 4 \\ 6 & -8 & 2 \\ -4 & 5 & 9 \end{bmatrix}$,

$C = \begin{bmatrix} 1 & 2 & -5 \\ -7 & 9 & 3 \\ 5 & -4 & 0 \end{bmatrix}$, and $D = \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix}$

Compute the following (if they exist).

(a) $A + B$ **(b)** $4B$ **(c)** $-3D$
(d) $B - 3C$ **(e)** $-A$ **(f)** $3A + 2D$
(g) $A + D$

3. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ -2 & 5 \end{bmatrix}$,
$C = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $D = \begin{bmatrix} -1 & 0 & 3 \\ 5 & 7 & 2 \end{bmatrix}$
Compute the following (if they exist).
(a) AB **(b)** BA **(c)** AC
(d) CA **(e)** AD **(f)** DC
(g) BD **(h)** A^2 [where $A^2 = AA$]

4. Let $A = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 & 5 \\ 3 & -7 & 8 \\ 2 & 3 & 1 \end{bmatrix}$,
$C = \begin{bmatrix} -2 & 0 & 5 \end{bmatrix}$, and $D = \begin{bmatrix} 9 & -5 \\ 3 & 0 \\ -4 & 2 \end{bmatrix}$
Compute the following (if they exist).
(a) BA **(b)** AB **(c)** CB
(d) CA **(e)** DA **(f)** DB
(g) AC **(h)** B^2

5. Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 3 \\ 5 & 6 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 0 \\ 3 & 5 \\ 2 & 6 \end{bmatrix}$,
$C = \begin{bmatrix} -4 & 0 \\ 3 & 2 \end{bmatrix}$, and $D = \begin{bmatrix} 5 & 0 \\ -2 & 1 \end{bmatrix}$
Compute the following (if they exist).
(a) $2A - 3(BC)$ **(b)** AB **(c)** $AC - BD$
(d) $CD - 2D$ **(e)** BA **(f)** $AD + 2(DC)$
(g) $C^3 + 2(D^2)$ (where $C^3 = CCC$).

6. Let $A = \begin{bmatrix} 1 & -8 & 4 \\ 5 & -6 & 3 \\ 2 & 0 & -1 \end{bmatrix}$, and $B = \begin{bmatrix} 0 & 2 & -3 \\ 5 & 6 & 7 \\ -1 & 0 & 4 \end{bmatrix}$.
Let O_3 and I_3 be the 3×3 zero and identity matrices. Show that

$$A + O_3 = O_3 + A = A,$$
$$BO_3 = O_3B = O_3,$$

and

$$BI_3 = I_3B = B.$$

7. **(a)** Let A be an $n \times n$ matrix and X be an $n \times 1$ column matrix of 1s. What can you say about the rows of A if $AX = X$? (We call such a matrix X an *eigenvector* of A. We shall study eigenvectors in the "Determinants and Eigenvectors" chapter.)
(b) Let A be an $n \times n$ matrix and X be a $1 \times n$ row matrix of 1s. What can you say about the columns of A if $XA = X$?

Sizes of Matrices

8. Let A be a 3×5 matrix, B a 5×2 matrix, C a 3×4 matrix, D a 4×2 matrix, and E a 4×5 matrix. Determine which of the following matrix expressions exist and give the size of the resulting matrices when they do exist.
(a) AB **(b)** EB
(c) AC **(d)** $AB + CD$
(e) $3(EB) + 4D$ **(f)** $CD - 2(CE)B$
(g) $2(EB) + DA$

9. Let A be a 2×2 matrix, B a 2×2 matrix, C a 2×3 matrix, D a 3×2 matrix, and E a 3×1 matrix. Determine which of the following matrix expressions exist and give the size of the resulting matrices when they do exist.
(a) AB **(b)** $(A^2)C$
(c) $B^3 + 3(CD)$ **(d)** $DC + BA$
(e) $DA - 2(DB)$ **(f)** $C + 3D$
(g) $3(BA)(CD) + (4A)(BC)$

Computing Certain Elements of Matrices

10. Let $C = AB$ and $D = BA$ for the following matrices A and B.

$$A = \begin{bmatrix} 0 & 3 & -5 \\ 2 & 6 & 3 \\ 1 & 0 & -2 \end{bmatrix} \text{ and } B = \begin{bmatrix} -1 & 2 & -3 \\ 5 & 7 & 2 \\ 0 & 1 & 6 \end{bmatrix}$$

Determine the following elements of C and D, without computing the complete matrices.
(a) c_{31} **(b)** c_{23} **(c)** d_{12} **(d)** d_{22}

11. Let $R = PQ$ and $S = QP$, where
$P = \begin{bmatrix} 1 & -2 \\ 4 & 6 \\ -1 & 3 \end{bmatrix}$ and $Q = \begin{bmatrix} 0 & 1 & 3 \\ 0 & -1 & 4 \end{bmatrix}$.
Determine the following elements (if they exist) of R and S, without computing the complete matrices.
(a) r_{21} **(b)** r_{33} **(c)** s_{11} **(d)** s_{23}

12. If $A = \begin{bmatrix} 1 & -3 \\ 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & -3 \\ 5 & 0 & -1 \end{bmatrix}$, and
$C = \begin{bmatrix} 2 & -4 & 5 \\ 7 & 1 & 0 \end{bmatrix}$, determine the following elements of $D = AB + 2C$, without computing the complete matrix.
(a) d_{12} **(b)** d_{23}

13. If $A = \begin{bmatrix} 1 & -3 & 0 \\ 4 & 5 & 1 \\ 3 & 8 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -2 \\ 3 & 0 & 4 \\ -1 & 3 & 2 \end{bmatrix}$, and
$C = \begin{bmatrix} 2 & 0 & -2 \\ 4 & 7 & -5 \\ 1 & 0 & -1 \end{bmatrix}$, determine the following elements of $D = 2(AB) + C^2$, without computing the complete matrix.
(a) d_{11} **(b)** d_{21} **(c)** d_{32}

Using Columns and Rows of Matrices

14. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$, $B = \begin{bmatrix} -2 & 3 \\ 4 & 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 0 & 5 \end{bmatrix}$

Compute the following products using the columns of B and C.

(a) AB (b) AC (c) BC

15. Let $A = \begin{bmatrix} 3 & -2 & 0 \\ 4 & 2 & 7 \\ 8 & -5 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 4 \\ 3 \\ -5 \end{bmatrix}$;

$P = \begin{bmatrix} 3 & 0 & 2 & 1 \\ 5 & 6 & 7 & 3 \end{bmatrix}$, $Q = \begin{bmatrix} -3 \\ 2 \\ 1 \\ 5 \end{bmatrix}$

(a) Express the product AB as a linear combination of the columns of A.

(b) Express PQ as a linear combination of the columns of P.

16. Let A and B be the following matrices. Compute row 2 of the matrix AB without computing the whole product.

$$A = \begin{bmatrix} 2 & -3 & 1 \\ 4 & 0 & 3 \\ 5 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 8 & 1 & 3 \\ 2 & 1 & 0 \\ 4 & 6 & 3 \end{bmatrix}$$

17. Let A be a matrix whose third row is all zeros. Let B be any matrix such that the product AB exists. Prove that the third row of AB is all zeros.

18. Let D be a matrix whose second column is all zeros. Let C be any matrix such that CD exists. Prove that the second column of CD is all zeros.

19. Let A be an $m \times r$ matrix, B an $r \times n$ matrix, and $C = AB$. Let the column submatrices of B be $B_1, B_2, \ldots, B_n$ and of C be $C_1, C_2, \ldots, C_n$. We can write B in the form $[B_1 \; B_2 \ldots B_n]$ and C as $[C_1 \; C_2 \ldots C_n]$. Prove that $C_j = AB_j$.

20. Let A and B be the following matrices. Use the result of Exercise 19 to compute column 3 of the matrix AB without computing the whole product.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 1 \\ 2 & 5 & 0 \end{bmatrix}, B = \begin{bmatrix} 2 & 1 & 4 \\ 6 & 0 & 1 \\ 2 & 3 & 5 \end{bmatrix}$$

Partitioning of Matrices

21. Use the given partitions of A and B below to compute AB.

(a) $A = \left[\begin{array}{c|c} 2 & 1 \\ -1 & 0 \\ \hline 3 & 1 \end{array}\right]$, $B = \left[\begin{array}{cc} 3 & 0 \\ \hline 2 & 1 \end{array}\right]$

(b) $A = \left[\begin{array}{cc|c} 1 & 2 & -1 \\ 3 & 0 & 1 \end{array}\right]$, $B = \left[\begin{array}{cc} 2 & 4 \\ 0 & -1 \\ \hline 1 & 3 \end{array}\right]$

(c) $A = \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 3 & -1 & 1 \\ \hline 4 & -2 & 0 \end{array}\right]$, $B = \left[\begin{array}{cc} 3 & -1 \\ 2 & 5 \\ \hline 0 & 1 \end{array}\right]$

22. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 0 & 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & -2 \\ 0 & 3 \\ 4 & 1 \end{bmatrix}$.

For each partition of A given below, find all the partitions of B that can be used to calculate AB.

(a) $A = \left[\begin{array}{c|cc} 1 & 2 & 3 \\ \hline -1 & 1 & 4 \\ 0 & 1 & 2 \end{array}\right]$ (b) $A = \left[\begin{array}{c|cc} 1 & 2 & 3 \\ -1 & 1 & 4 \\ 0 & 1 & 2 \end{array}\right]$

(c) $A = \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ -1 & 1 & 4 \\ \hline 0 & 1 & 2 \end{array}\right]$

23. Let $A = \begin{bmatrix} 2 & 0 & 3 & -1 \\ 6 & 2 & -5 & 9 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & -2 & 0 \\ 3 & 4 & 1 \\ 5 & 7 & 2 \\ 4 & 5 & -8 \end{bmatrix}$.

For each partition of B given below, find all the partitions of A that can be used to calculate AB.

(a) $B = \left[\begin{array}{c|cc} -1 & -2 & 0 \\ 3 & 4 & 1 \\ \hline 5 & 7 & 2 \\ 4 & 5 & -8 \end{array}\right]$ (b) $B = \left[\begin{array}{c|cc} -1 & -2 & 0 \\ 3 & 4 & 1 \\ 5 & 7 & 2 \\ 4 & 5 & -8 \end{array}\right]$

(c) $B = \left[\begin{array}{cc|c} -1 & -2 & 0 \\ 3 & 4 & 1 \\ 5 & 7 & 2 \\ \hline 4 & 5 & -8 \end{array}\right]$

24. Suggest suitable partitions involving zero and identity submatrices for computing the following products. Compute the products using these partitions (show your work). Check your answers using standard elementwise multiplication.

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

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

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

Miscellaneous Results

25. State (with a brief explanation) whether the following statements are true or false for matrices A, B, and C.

(a) If the sums $A + B$ and $B + C$ exist, then $A + C$ exists.

(b) If the products AB and BC exist, then AC exists.

(c) AB is never equal to BA.

(d) Let A be a column matrix and B a row matrix, both with the same number of elements. Then AB is a square matrix.

(e) If the element a_{ij} of a square matrix A lies below the main diagonal, then $i > j$.

2.2 Properties of Matrix Operations

We have defined operations of addition, scalar multiplication, and multiplication of matrices. We now list the most important properties of these operations. Since addition and scalar multiplication are defined the same way for matrices and vectors, these properties will be recognizable as ones that have their counterparts for vectors.

THEOREM 2.2

Let A, B, and C be matrices and r and s be scalars. Assume that the sizes of the matrices are such that the operations can be performed.

Properties of Matrix Addition and Scalar Multiplication

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $A + O = O + A = A$ *(where O is the appropriate zero matrix)*
4. $r(A + B) = rA + rB$
5. $(r + s)C = rC + sC$
6. $r(sC) = (rs)C$

Properties of Matrix Multiplication

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $AI = IA = A$ *(where I is the appropriate identity matrix)*
5. $r(AB) = (rA)B = A(rB)$

Note: $AB \neq BA$ *in general.* *Multiplication of matrices is not commutative*.

Each one of these results asserts an equality between matrices. We know that two matrices are equal if they are of the same size and their corresponding elements are equal. Each result is verified by showing this to be the case. We illustrate the method for the commutative property of addition. The reader is asked to use the same approach to prove some of the other results in the exercises that follow.

A + B = B + A By the rule of matrix addition we know that $A + B$ and $B + A$ are both matrices of the same size. It remains to show that their corresponding elements are equal. Consider the (i, j)th element of each matrix.

$$
\begin{aligned}
(i, j)\text{th element of } A + B &= a_{ij} + b_{ij} \\
(i, j)\text{th element of } B + A &= b_{ij} + a_{ij} \\
&= a_{ij} + b_{ij} \quad \text{(addition of real numbers is commutative)}
\end{aligned}
$$

The corresponding elements of $A + B$ and $B + A$ are equal. Thus $A + B = B + A$.

These properties enable us to extend addition, scalar multiplication, and multiplication to more than two matrices. We can write sums and products such as $aA + bB + cC$ and ABC without parentheses. The following examples illustrate these concepts. We remind you that we call an expression such as $aA + bB + cC$ a *linear combination* of the matrices A, B, and C.

EXAMPLE 1 Compute the linear combination $2A + 3B - 5C$ for the following three matrices.

$$A = \begin{bmatrix} 1 & 3 \\ -4 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & -7 \\ 2 & 1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 0 & 2 \\ 3 & -1 \end{bmatrix}$$

SOLUTION

We compute the scalar multiples first, then add or subtract corresponding elements in a natural way.

$$2\begin{bmatrix} 1 & 3 \\ -4 & 5 \end{bmatrix} + 3\begin{bmatrix} 3 & -7 \\ 2 & 1 \end{bmatrix} - 5\begin{bmatrix} 0 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 6 \\ -8 & 10 \end{bmatrix} + \begin{bmatrix} 9 & -21 \\ 6 & 3 \end{bmatrix} - \begin{bmatrix} 0 & 10 \\ 15 & -5 \end{bmatrix}$$
$$= \begin{bmatrix} 11 & -25 \\ -17 & 18 \end{bmatrix}$$

Certain products of matrices such as $ABCD$ will of course exist while other products will not. We can determine whether a product exists by comparing the numbers of rows and columns in adjacent matrices of the product, to see if they match.

If the product of a chain of matrices exists, the product matrix will have the same number of rows as the first matrix in the chain and the same number of columns as the last matrix.

EXAMPLE 2 Compute the product ABC of the following three matrices.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix}$$

SOLUTION

Let us check to see if the product ABC exists before we start spending time multiplying matrices. We get

$$\begin{array}{ccccc} A & B & C & = & ABC \\ 2 \times 2 & 2 \times 3 & 3 \times 1 & & 2 \times 1 \end{array}$$

match match

size of product is 2×1

The product exists and will be a 2×1 matrix. Since matrix multiplication is associative, the matrices in the product ABC can be grouped together in any manner for multiplying, as long as the order is maintained. Let us use the grouping $(AB)C$. This is probably the most natural. We get

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 3 \\ -1 & 0 & -2 \end{bmatrix} = \begin{bmatrix} -2 & 1 & -1 \\ 1 & 3 & 11 \end{bmatrix},$$

and

$$(AB)C = \begin{bmatrix} -2 & 1 & -1 \\ 1 & 3 & 11 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$$

Caution

In algebra we know that the following cancellation laws apply.

- If $ab = ac$ and $a \neq 0$, then $b = c$.
- If $pq = 0$, then $p = 0$ or $q = 0$.

However the corresponding results are not true for matrices.

- $AB = AC$ does not imply that $B = C$.
- $PQ = O$ does not imply that $P = O$ or $Q = O$.

We demonstrate these possibilities by means of examples.

Consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$, $C = \begin{bmatrix} -3 & 8 \\ 3 & -2 \end{bmatrix}$. Observe that $AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$, but $B \neq C$.

Consider the matrices $P = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$, $Q = \begin{bmatrix} 2 & -6 \\ 1 & -3 \end{bmatrix}$. Observe that $PQ = O$, but $P \neq O$ and $Q \neq O$.

Powers of Matrices

A similar notation is used for the powers of matrices as for powers of real numbers. If A is a square matrix, then A multiplied by itself k times is written A^k.

$$A^k = \underbrace{AA\ldots A}_{k \text{ times}}$$

Familiar rules of exponents of real numbers hold for matrices.

THEOREM 2.3

If A is an $n \times n$ square matrix and r and s are nonnegative integers, then
1. $A^rA^s = A^{r+s}$
2. $(A^r)^s = A^{rs}$
3. $A^0 = I_n$ (By definition.)

We verify the first rule. The proof of the second rule is similar.

$$A^rA^s = \underbrace{A\ldots A}_{r \text{ times}} \; \underbrace{A\ldots A}_{s \text{ times}} = \underbrace{A\ldots A}_{r+s \text{ times}} = A^{r+s}$$

EXAMPLE 3 If $A = \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix}$, compute A^4.

SOLUTION

This example illustrates how the preceding rules can be used to reduce the amount of computation involved in multiplying matrices. We know that $A^4 = AAAA$. We could perform three matrix multiplications to arrive at A^4. However we can apply rule 2 to write $A^4 = (A^2)^2$ and thus arrive at the result using two products. We get

$$A^2 = \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}.$$

$$A^4 = \begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 11 & -10 \\ -5 & 6 \end{bmatrix}.$$

The following example illustrates that the properties of matrix operations can be used to simplify matrix expressions in a similar way to the simplification of ordinary algebraic expressions.

EXAMPLE 4 Simplify the following matrix expression.

$$A(A + 2B) + 3B(2A - B) - A^2 + 7B^2 - 5AB$$

SOLUTION

Using the properties of matrix operations we get

$$\begin{aligned} A(A + 2B) + 3B(2A - B) - A^2 + 7B^2 - 5AB &= A^2 + 2AB + 6BA - 3B^2 - A^2 \\ &\quad + 7B^2 - 5AB \\ &= -3AB + 6BA + 4B^2 \end{aligned}$$

Resist the temptation to simplify $-3AB + 6BA$; matrix multiplication is not commutative!

We now introduce a valuable way of writing a system of linear equations as a single matrix equation. We will see how this technique is used in designing equipment for controlling currents and voltages in electrical circuits. Furthermore it opens up the possibility of using matrix algebra to further discuss the *behavior* of systems of linear equations.

Systems of Linear Equations

We can write a general system of m linear equations in n variables using matrix notation as follows:

$$\begin{matrix} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \vdots & & \cdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m \end{matrix}$$

Write each side of this equation as a column matrix,

$$\begin{bmatrix} a_{11}x_1 + & \cdots & + a_{1n}x_n \\ \vdots & \cdots & \vdots \\ a_{m1}x_1 + & \cdots & + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

The left matrix can be written as a product of the matrix of coefficients A and a column matrix of variables X. Let the column matrix of constants be B.

$$\overset{A}{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \cdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}} \overset{X}{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}} = \overset{B}{\begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}}$$

Thus we can write the system of equations in matrix form

$$\boxed{AX = B}$$

For example,

$$\begin{aligned} 3x_1 + 2x_2 - 5x_3 &= 7 \\ x_1 - 8x_2 + 4x_3 &= 9 \\ 2x_1 + 6x_2 - 7x_3 &= -2 \end{aligned} \qquad \text{can be written} \qquad \begin{bmatrix} 3 & 2 & -5 \\ 1 & -8 & 4 \\ 2 & 6 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \\ -2 \end{bmatrix}$$

We now use this notation and the properties of matrices to examine sums and scalar multiples of solutions to systems of linear equations.

Solutions to Systems of Linear Equations

In Section 4 of the "Linear Equations, Vectors, and Matrices" chapter, we found that the set of solutions to a specific system of homogeneous equations was closed under addition and scalar multiplication and was therefore a subspace. We are now in a position to show that this is true of all homogeneous systems of linear equations.

Consider a homogeneous system of linear equations $AX = 0$. Let X_1 and X_2 be solutions. Then

$$AX_1 = 0 \quad \text{and} \quad AX_2 = 0$$

Adding these equations, we get

$$AX_1 + AX_2 = 0, \text{ giving } A(X_1 + X_2) = 0$$

Thus $X_1 + X_2$ satisfies the equation $AX = 0$. This means that $X_1 + X_2$ is a solution. The set of solutions is closed under addition.

Furthermore, if c is a scalar, multiplying $AX_1 = 0$ by c,

$$cAX_1 = 0, \text{ giving } A(cX_1) = 0$$

Thus cX_1 is a solution. The set of solutions is closed under scalar multiplication.

> The set of solutions to a homogeneous system of linear equations is closed under addition and under scalar multiplication. It is a subspace.

We now find that the set of solutions to a nonhomogeneous system of linear equations on the other hand does not form a subspace. Let $AX = B$ $(B \neq 0)$ be such a nonhomogeneous system. Let X_1 and X_2 be solutions. Thus

$$AX_1 = B \text{ and } AX_2 = B$$

Adding these equations gives

$$AX_1 + AX_2 = 2B$$
$$A(X_1 + X_2) = 2B$$

Therefore $X_1 + X_2$ does not satisfy $AX = B$. It is not a solution. The set of solutions is not closed under addition. It is not a subspace. It can also be shown that the set of solutions is not closed under scalar multiplication. See Exercise 41.

The set of solutions to a nonhomogeneous system of linear equations is not closed under either addition or scalar multiplication. It is not a subspace.

Even though the sets of solutions to nonhomogeneous systems of linear equations do not form subspaces, subspaces are still at the heart of understanding such solutions. The following example illustrates the relationship between the set of solutions to a nonhomogeneous system of linear equations and the subspace of solutions to a corresponding homogeneous system.

EXAMPLE **5**

(a) Consider the following homogeneous system of linear equations:

$$x_1 + x_2 - 2x_3 = 0$$
$$x_1 - 2x_2 + 3x_3 = 0$$
$$3x_2 - 5x_3 = 0$$

It can be shown that there are many solutions, $x_1 = r$, $x_2 = 5r$, $x_3 = 3r$. The solutions are vectors in $\mathbf{R}^3$ of the form $(r, 5r, 3r)$, or $r(1, 5, 3)$. They make up a line through the origin defined by the vector $(1, 5, 3)$. The solutions form a subspace of $\mathbf{R}^3$ of dimension 1, with basis $\{(1, 5, 3)\}$. See Figure 2.3.

We discussed in Section 2 of the "Linear Equations, Vectors, and Matrices" chapter that $x_1 = 0, \ldots, x_n = 0$, is a solution to every homogeneous system of linear equations in n variables. *Geometrically this means that the space of solutions to every homogeneous system passes through the origin, as in the case of this line.*

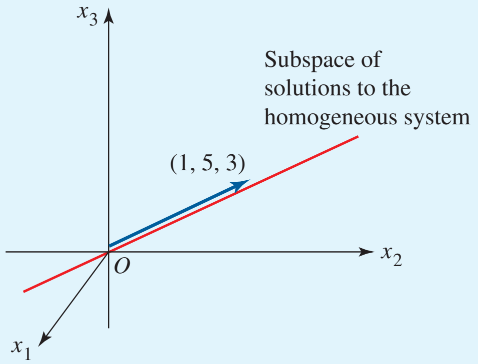


Figure 2.3

(b) Consider the following nonhomogeneous system of linear equations:

$$\begin{aligned} x_1 + x_2 - 2x_3 &= 4 \\ x_1 - 2x_2 + 3x_3 &= -5 \\ 3x_2 - 5x_3 &= 9 \end{aligned}$$

It can be shown that there are many solutions, $x_1 = r + 1, x_2 = 5r + 3, x_3 = 3r$. It can be shown that this set is not closed under addition or scalar multiplication. It is not a subspace.

Observe that this system differs from the homogeneous system of Part (a) only in the constant terms. Both systems have the same matrix of coefficients. We say that (a) is the *corresponding homogeneous system* of (b). There will be a corresponding homogeneous system for every nonhomogeneous system. The corresponding homogeneous system enables us to understand the solutions of the nonhomogeneous system. Express the above general solution in vector form, $(r + 1, 5r + 3, 3r)$. This can be written as the sum of the solutions to the corresponding homogeneous system and a constant vector.

$$(r + 1, 5r + 3, 3r) = r(1, 5, 3) + (1, 3, 0)$$

The implication is that the set of solutions to the nonhomogeneous system is the line defined by (1, 5, 3) slid in a manner described by the vector (1, 3, 0). This gives the set of solutions to be the line through the point (1, 3, 0) parallel to the line through the origin defined by the vector (1, 5, 3). See Figure 2.4. The solutions to all nonhomogeneous systems can be pictured in this way by relating the systems to their corresponding homogeneous systems. If the set of solutions to the homogeneous system is a line through the origin, the set of solutions to the nonhomogeneous system is a line parallel to this line. If the set of solutions to the homogeneous system is a plane through the origin, the set of solutions to the nonhomogeneous system is a plane parallel to this plane, and so on. We shall be able to give further insight into this geometrical way of looking at solutions to linear systems in Section 4 of the "General Vector Spaces" chapter when we have some very elegant mathematical tools involving transformations available.

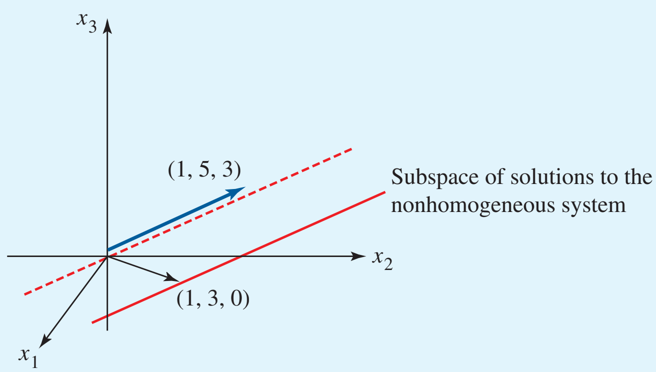


Figure 2.4

EXAMPLE 6 In this example we analyze a *two-port* in an electrical circuit.

Many networks are designed to accept signals at certain points and to deliver a modified version of the signals. The usual arrangement is illustrated in Figure 2.5. A current I_1 at voltage V_1 is delivered into a two-port and it in some way determines the output current I_2 at voltage V_2. In practice, the relationship between the input and output currents and voltages is usually linear—they are related by a matrix equation:

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

The matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is called the *transmission matrix* of the port. This matrix defines the two-port.

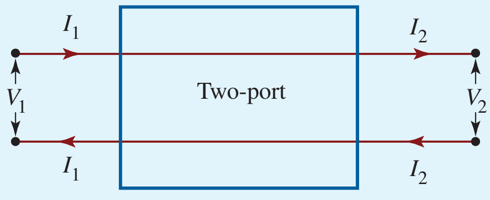


Figure 2.5

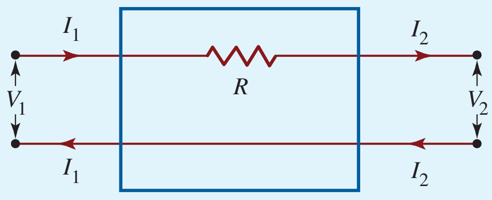


Figure 2.6

Figure 2.6 is an example of a two-port. The interior consists of a resistance R connected as shown. Let us show that the currents and voltages do indeed behave in a linear manner and determine the transmission matrix. Our approach will be to construct two equations, one expressing V_2 in terms of V_1 and I_1, and the other expressing I_2 in terms of V_1 and I_1 and then combine these two equations into a single matrix equation. We use the following law.

Ohm's Law
The voltage drop across a resistance is equal to the current times the resistance.

The voltage drop across the resistance R is $V_1 - V_2$. The current through the resistance is I_1. Thus Ohm's Law implies that $V_1 - V_2 = I_1 R$. The current I_1 passes through the resistance R unchanged and exits unchanged as I_1. Thus $I_2 = I_1$. Write these two equations in the following standard form.

$$\begin{aligned} V_2 &= V_1 - RI_1 \\ I_2 &= 0V_1 + I_1 \end{aligned}$$

Combine the two equations into a single matrix equation,

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & -R \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

The transmission matrix is $\begin{bmatrix} 1 & -R \\ 0 & 1 \end{bmatrix}$. Thus, for example, if R is 2 ohms and the input voltage and current are $V_1 = 5$ volts, $I_1 = 1$ amp, we get

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

The output voltage and current are 3 volts and 1 amp.

In practice a number of standard two-ports such as the one above are placed in series to provide a desired voltage and current change. Consider the three two-ports of Figure 2.7, with transmission matrices A, B, and C.

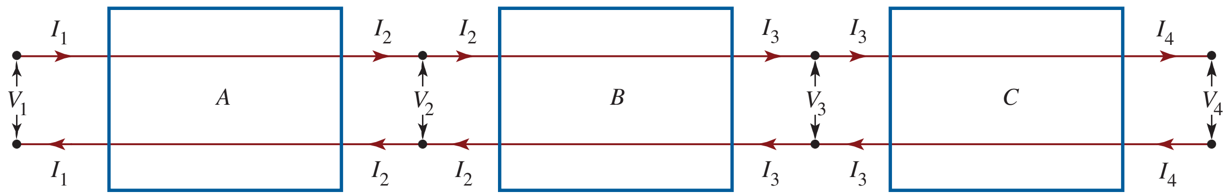


Figure 2.7

Considering each port separately, we have

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = A\begin{bmatrix} V_1 \\ I_1 \end{bmatrix}, \begin{bmatrix} V_3 \\ I_3 \end{bmatrix} = B\begin{bmatrix} V_2 \\ I_2 \end{bmatrix}, \begin{bmatrix} V_4 \\ I_4 \end{bmatrix} = C\begin{bmatrix} V_3 \\ I_3 \end{bmatrix}$$

Substituting for $\begin{bmatrix} V_2 \\ I_2 \end{bmatrix}$ from the first equation into the second gives $\begin{bmatrix} V_3 \\ I_3 \end{bmatrix} = BA\begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$. Substituting this $\begin{bmatrix} V_3 \\ I_3 \end{bmatrix}$ into the third equation gives

$$\begin{bmatrix} V_4 \\ I_4 \end{bmatrix} = CBA\begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

The three ports are thus equivalent to a single two-port. The transmission matrix of this two-port is the product CBA of the individual ports. Note that the placement of each port in the sequence is important since matrices are not commutative under multiplication.

EXERCISE SET 2.2

Computation

1. Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 5 & 4 \\ 2 & 1 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 3 \\ 6 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 2 & -2 \\ 1 & 3 \end{bmatrix}$. Calculate, if possible,

(a) AB and BA **(b)** AC and CA

(c) AD and DA

Observe that $AB \neq BA$ since BA does not exist, $AC \neq CA$ and $AD = DA$, illustrating the different possibilities when order is reversed in matrix multiplication.

2. Compute $A(BC)$ and $(AB)C$ for the matrices

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 4 \\ -2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Observe that these products are equal, illustrating the associative property of matrix multiplication.

3. Compute the product ABC for the following three matrices in two distinct ways.

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & 2 \\ 4 & 3 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 1 & 0 \end{bmatrix}$$

4. Compute each of the following linear combinations for $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} -2 & 0 \\ 3 & 4 \end{bmatrix}$.

(a) $2A + 3B$ **(b)** $A + 2B + 4C$

(c) $3A + B - 2C$

5. Compute each of the following expressions for $A = \begin{bmatrix} 2 & 0 \\ -1 & 5 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 2 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} 3 & 4 \\ 0 & 2 \end{bmatrix}$.

(a) $(AB)^2$ **(b)** $A - 3B^2$

(c) $A^2B + 2C^3$ **(d)** $2A^2 - 2A + 3I_2$

Sizes of Matrix Products

6. Given that A is a 4×2 matrix, B is 2×6, C is 3×4, and D is 6×3, determine the sizes of the following products, if they exist.

(a) ABC **(b)** ABD **(c)** CAB
(d) $DCAB$ **(e)** A^2BDC

7. If P is 3×2, Q is 2×1, R is 1×3, S is 3×1, and T is 3×3, determine the sizes of the following matrix expressions, if they exist.

(a) PQR **(b)** $PQ + TPQ$
(c) $5QR - 2TPR$ **(d)** $4SPQ + 3PQ$
(e) $QRSR + QR$

Matrix Operations

8. Let A be an $m \times n$ matrix. Prove that AB and BA both exist only if B is an $n \times m$ matrix.

9. Verify the following properties of matrix operations given in this section:

(a) the associative property of matrix addition $A + (B + C) = (A + B) + C$
(b) the distributive property $c(A + B) = cA + cB$
(c) $AI_n = I_nA = A$ if A is an $n \times n$ matrix

10. Let A be an $m \times n$ matrix. Show that $AI_n = A$.

11. Let A be any $m \times n$ matrix, O_{mn} be the $m \times n$ zero matrix, and c be a scalar. Show that if $cA = O_{mn}$ then either $c = 0$ or $A = O_{mn}$.

12. Simplify the following matrix expressions.

(a) $A(A - 4B) + 2B(A + B) - A^2 + 7B^2 + 3AB$
(b) $B(2I_n - BA) + B(4I_n + 5A)B - 3BAB + 7B^2A$
(c) $(A - B)(A + B) - (A + B)^2$

13. Simplify the following matrix expressions.

(a) $A(A + B) - B(A + B)$
(b) $A(A - B)B + B2AB - 3A^2$
(c) $(A + B)^3 - 2A^3 - 3ABA - A3B^2 - B^3$

14. Find all the matrices that commute with the following matrices.

(a) $\begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ **(c)** $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

15. What is incorrect about the following proof? Let $AX = B$ be a system of linear equations with solutions X_1 and X_2. Thus

$$AX_1 = B \quad \text{and} \quad AX_2 = B$$
$$AX_1 = AX_2$$
$$X_1 = X_2$$

Thus every system of linear equations has at most one solution.

Powers of Matrices

16. (a) Let A be an $n \times n$ matrix. Prove that A^2 is an $n \times n$ matrix.

(b) Let A be an $m \times n$ matrix, with $m \neq n$. Prove that A^2 does not exist.

Thus one can only talk about powers of square matrices.

17. If A and B are square matrices of the same size, prove that in general

$$(A + B)^2 \neq A^2 + 2AB + B^2$$

Under what condition does equality hold?

18. If A and B are square matrices of the same size such that $AB = BA$, prove that $(AB)^2 = A^2B^2$. By constructing an example, show that this result does not hold for all square matrices of the same size.

19. If n is a nonnegative integer and A and B are square matrices of the same size such that $AB = BA$, prove that $(AB)^n = A^nB^n$. By constructing an example, show that this identity does not hold in general for all square matrices of the same size.

20. Show that nonnegative integer powers of the same matrix commute. $(A^rA^s = A^sA^r)$

Diagonal Matrices

21. Let A and B be diagonal matrices of the same size and c a scalar. Prove that **(a)** $A + B$ is diagonal, **(b)** cA is diagonal, and **(c)** AB is diagonal.

22. If A and B are diagonal matrices of the same size, prove that $AB = BA$.

23. Prove that if a matrix A commutes with a diagonal matrix that has no two diagonal elements the same, then A is a diagonal matrix.

Idempotent and Nilpotent Matrices

A square matrix A is said to be **idempotent** if $A^2 = A$. A square matrix A is said to **nilpotent** if there is a positive integer p such that $A^P = 0$. The least integer such that $A^P = 0$ is called the **degree of nilpotency** of the matrix.

24. Determine whether the following matrices are idempotent.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **(d)** $\begin{bmatrix} 3 & -6 \\ 1 & -2 \end{bmatrix}$

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

25. Determine b, c, and d such that $\begin{bmatrix} 1 & b \\ c & d \end{bmatrix}$ is idempotent.

26. Determine a, c, and d such that $\begin{bmatrix} a & 0 \\ c & d \end{bmatrix}$ is idempotent.

27. Prove that if A and B are idempotent and $AB = BA$, then AB is idempotent.

28. Show that if A is idempotent, and if n is a positive integer, then $A^n = A$.

29. Show that the following matrices are nilpotent with degree of nilpotency 2.

(a) $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ **(b)** $\begin{bmatrix} -4 & 8 \\ -2 & 4 \end{bmatrix}$ **(c)** $\begin{bmatrix} 3 & -9 \\ 1 & -3 \end{bmatrix}$

30. Show that the following matrix is nilpotent with degree of nilpotency 3.

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Systems of Linear Equations

31. Write each of the following systems of linear equations as a single matrix equation $AX = B$.

(a) $\begin{aligned} 2x_1 + 3x_2 &= 4 \\ 3x_1 - 8x_2 &= -1 \end{aligned}$ **(b)** $\begin{aligned} 4x_1 + 7x_2 &= -2 \\ -2x_1 + 3x_2 &= -4 \end{aligned}$

(c) $\begin{aligned} -9x_1 - 3x_2 &= -4 \\ 6x_1 - 2x_2 &= 7 \end{aligned}$

32. Write each of the following systems of linear equations as a single matrix equation $AX = B$.

(a) $\begin{aligned} x_1 + 8x_2 - 2x_3 &= 3 \\ 4x_1 - 7x_2 + x_3 &= -3 \\ -2x_1 - 5x_2 - 2x_3 &= 1 \end{aligned}$

(b) $\begin{aligned} 5x_1 + 2x_2 &= 6 \\ 4x_1 - 3x_2 &= -2 \\ 3x_1 + x_2 &= 9 \end{aligned}$

(c) $\begin{aligned} x_1 - 3x_2 + 6x_3 &= 2 \\ 7x_1 + 5x_2 + x_3 &= -9 \end{aligned}$

(d) $\begin{aligned} 2x_1 + 5x_2 - 3x_3 + 4x_4 &= 4 \\ x_1 \qquad + 9x_3 + 5x_4 &= 12 \\ 3x_1 - 3x_2 - 8x_3 + 5x_4 &= -2 \end{aligned}$

33. Prove that if X_1 and X_2 are solutions to the homogeneous system of linear equations $AX = 0$, then the linear combination $aX_1 + bX_2$ is a solution for all scalars a and b.

34. Consider the following system of equations. You are given two solutions, X_1 and X_2. Generate four other solutions using the operations of addition and scalar multiplication. Use the result of Exercise 33 to find a solution for which $x_1 = 1$ and $x_2 = 0$.

$$\begin{aligned} x_1 + 2x_2 - x_3 - 2x_4 &= 0 \\ 2x_1 + 5x_2 \qquad - 2x_4 &= 0 \\ 4x_1 + 9x_2 - 2x_3 - 6x_4 &= 0 \\ x_1 + 3x_2 + x_3 \qquad &= 0 \end{aligned}$$

$$X_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 11 \\ -4 \\ 1 \\ 1 \end{bmatrix}$$

35. Consider the following system of four equations. You are given two solutions X_1 and X_2. Generate four other solutions using the operations of addition and scalar multiplication. Use the result of Exercise 33 to find a solution for which $x_1 = 6$ and $x_2 = 9$.

$$\begin{aligned} x_1 - x_2 + x_3 + 2x_4 &= 0 \\ 3x_1 - 2x_2 + x_3 + 3x_4 &= 0 \\ 5x_1 - 4x_2 + 3x_3 + 7x_4 &= 0 \\ 2x_1 - x_2 \qquad + x_4 &= 0 \end{aligned}$$

$$X_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 3 \\ 7 \\ 2 \\ 1 \end{bmatrix}$$

Consider the nonhomogeneous systems of linear equations in Exercises 36–39. For convenience, their general solutions are given. (a) Write down the corresponding homogeneous system and give its general solution. (b) Give a basis for this subspace of solutions to the homogeneous system and a written description of the subspace. (c) Give a written description of the subspace of solutions to the nonhomogeneous system.

36. $\begin{aligned} x_1 \qquad + x_3 &= 2 \\ x_1 + x_2 - 2x_3 &= 5 \\ 2x_1 + x_2 - x_3 &= 7 \end{aligned}$

General solution $(-r + 2, 3r + 3, r)$

37. $\begin{aligned} x_1 + x_2 + 5x_3 &= 5 \\ x_2 + 3x_3 &= 4 \\ x_1 + 2x_2 + 8x_3 &= 9 \end{aligned}$

General solution $(-2r + 1, -3r + 4, r)$

38. $\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ 2x_1 - x_2 - 4x_3 &= -5 \\ x_1 - x_2 - 3x_3 &= 4 \end{aligned}$

General solution $(r - 1, -2r + 3, r)$

39. $$\begin{aligned} x_1 + 2x_2 - x_3 &= 3 \\ 2x_1 + 4x_2 - 2x_3 &= 6 \\ -3x_1 - 6x_2 + 3x_3 &= -9 \end{aligned}$$
General solution $(-2r + s + 3, r, s)$

40. Solve the following nonhomogeneous system of linear equations. Show that the set of solutions is not a subspace. Give a geometrical description of the set of solutions.

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 5 \\ x_2 - x_3 &= 2 \\ 2x_1 + 5x_2 - 7x_3 &= 12 \end{aligned}$$

41. Show that a set of solutions to a system of nonhomogeneous linear equations is not closed under scalar multiplication.

42. Prove that a system of linear equations $AX = B$ has a solution if B is a linear combination of the columns of the matrix A.

Miscellaneous Results

43. State (with a brief explanation) whether the following statements are true or false for matrices A, B, C, and D.

(a) $A^2 - B^2 = (A + B)(A - B)$

(b) $A(B + C + D) = AB + AC + AD$

(c) If A is an $m \times n$ matrix, B is $n \times r$, and C is $r \times q$, then ABC has $m + q$ elements.

(d) If $A^2 = A$, then $A^r = A$ for all positive integer values of r.

(e) If X_1 is a solution of $AX = B_1$ and X_2 is a solution of $AX = B_2$, then $X_1 + X_2$ is a solution of $AX = B_1 + B_2$.

Two-Ports

44. Determine the transmission matrices of the two-ports in Figure 2.8.

Hints:

(a) $V_1 = V_2$ since terminals are connected directly. Current through resistance R is $I_1 - I_2$. Drop in voltage across R is V_1.

(b) Current through R_1 is $I_1 - I_2$. Drop in voltage across R_1 is V_1. Current through R_2 is I_2. Drop in voltage across R_2 is $V_1 - V_2$.

(c) Current through R_1 is I_1. Drop in voltage across R_1 is $V_1 - V_2$. Current through R_2 is $I_1 - I_2$. Drop in voltage across R_2 is V_2.

45. The two-port in Figure 2.9 consists of three two-ports placed in series. The transmission matrices are indicated. **(a)** What is the transmission matrix of the composite two-port? **(b)** If the input voltage is five volts and the current is two amps, determine the output voltage and current.

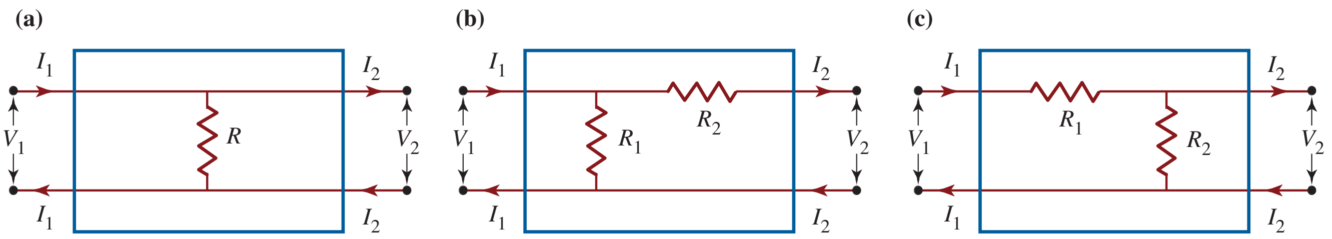


Figure 2.8

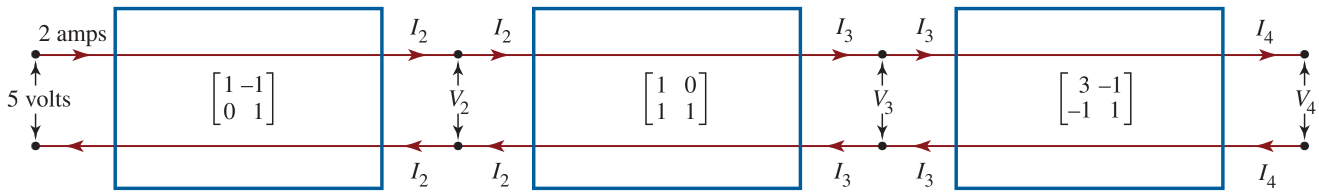


Figure 2.9

2.3 Symmetric Matrices and Seriation in Archaeology

In this section we continue the algebraic development of matrices and we see how the theory developed is used by archaeologists to determine the chronological order of graves and artifacts.

DEFINITION The **transpose** of a matrix A, denoted A^t, is the matrix whose columns are the rows of the given matrix A.

The first row of A becomes the first column of A^t, the second row of A becomes the second column of A^t, and so on. The (i, j)th element of A becomes the (j, i)th element of A^t. If A is an $m \times n$ matrix, then A^t is an $n \times m$ matrix.

EXAMPLE 1 Determine the transpose of each of the following matrices.

$$A = \begin{bmatrix} 2 & 7 \\ -8 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & -7 \\ 4 & 5 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} -1 & 3 & 4 \end{bmatrix}$$

SOLUTION

Writing rows as columns we get

$$A^t = \begin{bmatrix} 2 & -8 \\ 7 & 0 \end{bmatrix}, \quad B^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ -7 & 6 \end{bmatrix}, \quad C^t = \begin{bmatrix} -1 \\ 3 \\ 4 \end{bmatrix}$$

Observe that the $(1, 3)$ element of B namely -7, becomes the $(3, 1)$ element of B^t. Note also that B is a 2×3 matrix while B^t is 3×2, C is a 1×3 matrix while C^t is 3×1.

There are three operations that we have defined for matrices, namely addition, scalar multiplication, and multiplication. The following theorem tells us how transpose works in conjunction with these operations.

THEOREM 2.4

Properties of Transpose

Let A and B be matrices and c be a scalar. Assume that the sizes of the matrices are such that the operations can be performed.

1. $(A + B)^t = A^t + B^t$ *Transpose of a sum*
2. $(cA)^t = cA^t$ *Transpose of a scalar multiple*
3. $(AB)^t = B^tA^t$ *Transpose of a product*
4. $(A^t)^t = A$

We demonstrate the techniques that are used in verifying results involving transposes by verifying the third property. The reader is asked to derive the other results in the exercises that follow.

$(AB)^t = B^tA^t$ The expressions $(AB)^t$ and B^tA^t will be matrices upon carrying out all the products and transposes. We prove that these matrices are equal by showing that corresponding elements are equal.

$$\begin{aligned}(i, j)\text{th element of } (AB)^t &= (j, i)\text{th element of } AB \\ (i, j)\text{th element of } B^tA^t &= (\text{row } i \text{ of } B^t) \times (\text{column } j \text{ of } A^t) \\ &= (\text{column } i \text{ of } B) \times (\text{row } j \text{ of } A) \\ &= (j, i)\text{th element of } AB\end{aligned}$$

The corresponding elements of $(AB)^t$ and B^tA^t are equal, proving the result.

Remark: The results for the transpose of a sum and a product can be extended to any number of matrices. For example, for three matrices A, B, and C,

$$(A + B + C)^t = A^t + B^t + C^t \quad \text{and} \quad (ABC)^t = C^tB^tA^t$$

Note the reversal of the order of the matrices in the transpose of a product. (See the following exercises for the proofs.)

Comment: Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ be column vectors in $\mathbf{R}^n$, interpreted as $n \times 1$ matrices. The product $\mathbf{u}^t\mathbf{v}$ of a $1 \times n$ matrix and an $n \times 1$ matrix will be a 1×1 matrix. It is often useful to drop the matrix brackets and treat this product as a real number. We find, as follows, that $\mathbf{u}^t\mathbf{v}$ then results in the dot product.

$$\mathbf{u}^t\mathbf{v} = [u_1, u_1, \ldots, u_n]\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + \cdots + u_nv_n = \mathbf{u} \cdot \mathbf{v}.$$

$\boxed{\text{The dot product of two column vectors } \mathbf{u} \text{ and } \mathbf{v} \text{ in } \mathbf{R}^n \text{ can be written } \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^t\mathbf{v}}$

For example, if $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$ then $\mathbf{u}^t\mathbf{v} = [3, 4, -1]\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = (3 \times 0) + (4 \times 2) + (1 \times -1) = 7$, which is the dot product of $\mathbf{u}$ and $\mathbf{v}$.

This result leads to useful identities. For example, $\|\mathbf{u}\|^2 = \mathbf{u}^t\mathbf{u}$ (since $\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}} = \sqrt{\mathbf{u}^t \cdot \mathbf{u}}$). The vectors $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\mathbf{u}^t\mathbf{v} = 0$ (since $\mathbf{u}^t\mathbf{v} = \mathbf{u} \cdot \mathbf{v} = 0$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal). If $\mathbf{u}$ and $\mathbf{v}$ are viewed as row matrices, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}\mathbf{v}^t$. We shall find that this matrix way of looking at the dot product is important in theoretical work. See, for example, in the proof of Theorem 2.9, Section 2.5 on geometrical transformations.

We now introduce *symmetric matrices*. They are probably the single most important class of matrices. They are used in areas of mathematics such as geometry, and in fields such as theoretical physics, mechanical and electrical engineering, and sociology.

DEFINITION A **symmetric matrix** is a square matrix that is equal to its transpose.

The following are examples of symmetric matrices. Note the symmetry of these matrices about the main diagonal. All nondiagonal elements occur in pairs symmetrically located about the main diagonal.

match

$$\begin{bmatrix} 2 & 5 \\ 5 & -4 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & -4 \\ 1 & 7 & 8 \\ -4 & 8 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & -2 & 4 \\ 0 & 7 & 3 & 9 \\ -2 & 3 & 2 & -3 \\ 4 & 9 & -3 & 6 \end{bmatrix}$$

match

EXAMPLE 2 Consider the following matrix, which represents distances between various U.S. cities.

$$\begin{array}{l} \\ \text{Chicago} \\ \text{Los Angeles} \\ \text{Miami} \\ \text{New York} \end{array} \begin{array}{c} \begin{array}{cccc} \text{Chicago} & \text{LA} & \text{Miami} & \text{NY} \end{array} \\ \begin{bmatrix} 0 & 2092 & 1374 & 841 \\ 2092 & 0 & 2733 & 2797 \\ 1374 & 2733 & 0 & 1336 \\ 841 & 2797 & 1336 & 0 \end{bmatrix} \end{array}$$

Observe that the matrix is symmetric. All elements occur in pairs, symmetrically located about the main diagonal. There is a reason for this, namely that the distance from city X to city Y is the same as the distance from city Y to city X. For example, the distance from Chicago to Miami, which is 1374 miles (row 1, column 3) will be the same as the distance from Miami to Chicago (row 3, column 1). All such mileage matrices will be symmetric.

EXAMPLE 3 Let A and B be symmetric matrices of the same size. Let C be a linear combination of A and B. Prove that C is symmetric.

Proof Let $C = aA + bB$, where a and b are scalars. We now use the results of Theorem 2.4 to prove that C is symmetric.

$$\begin{aligned} C^t &= (aA + bB)^t \\ &= (aA)^t + (bB)^t && \textit{Transpose of sum} \\ &= aA^t + bB^t && \textit{Transpose of scalar multiple} \\ &= aA + bB && \textit{A and B are symmetric} \\ &= C \end{aligned}$$

Thus C is symmetric.

The Expression "If and Only If"

The expressions *"if and only if"* and *"necessary and sufficient"* (they mean the same thing) are frequently used in mathematics, and we shall use them periodically in this course. We have already used *if and only if* intuitively in this section. Let p and q be statements. Suppose that p implies q, written $p \Rightarrow q$, and that also $q \Rightarrow p$. The second implication is called the **converse** of the first. We say that "p if and only if q" or "p is necessary and sufficient for q." The next example states a result using this language.

EXAMPLE 4 Let A and B be symmetric matrices of the same size. Prove that the product AB is symmetric if and only if $AB = BA$.

Proof Every "if and only if" situation such as this consists of two parts. We have to show that (a) if AB is symmetric then $AB = BA$, and then conversely, (b) if $AB = BA$ then AB is symmetric.

(a) Let AB be symmetric. Then

$$\begin{aligned} AB &= (AB)^t && \textit{Definition of symmetric matrix} \\ &= B^tA^t && \textit{Transpose of a product} \\ &= BA && \textit{A and B are symmetric matrices} \end{aligned}$$

(b) Let $AB = BA$. Then

$$\begin{aligned} (AB)^t &= (BA)^t \\ &= A^tB^t && \textit{Transpose of a product} \\ &= AB && \textit{A and B are symmetric matrices} \end{aligned}$$

Therefore, AB is symmetric.

Thus, given two symmetric matrices of the same size, A and B, the product AB is symmetric if and only if $AB = BA$.

This result could also have been stated as follows: "Given two symmetric matrices of the same size, A and B, then a necessary and sufficient condition for the product AB to be symmetric is that $AB = BA$."

We now introduce a number that is associated with every square matrix, called the *trace of the matrix*.

DEFINITION Let A be a square matrix. The **trace** of A, denoted $tr(A)$ is the sum of the diagonal elements of A. Thus if A is an $n \times n$ matrix,

$$tr(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

EXAMPLE 5 Determine the trace of the matrix $A = \begin{bmatrix} 4 & 1 & -2 \\ 2 & -5 & 6 \\ 7 & 3 & 0 \end{bmatrix}$.

SOLUTION

We get,

$$tr(A) = 4 + (-5) + 0 = -1$$

The trace of a matrix plays an important role in matrix theory and matrix applications because of its properties and the ease with which it can be evaluated. It is important in fields such as statistical mechanics, general relativity, and quantum mechanics, where it has physical significance.

The following theorem tells us how the operation of trace interacts with the operations of matrix addition, scalar multiplication, multiplication, and transpose.

THEOREM 2.5

Properties of Trace

Let A and B be matrices and c be a scalar. Assume that the sizes of the matrices are such that the operations can be performed.

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(AB) = tr(BA)$
3. $tr(cA) = ctr(A)$
4. $tr(A^t) = tr(A)$

Proof We prove the first property, leaving the proofs of the other properties for the reader to complete in the exercises. Since the diagonal elements of $A + B$ are $(a_{11} + b_{11}), (a_{22} + b_{22}), \ldots, (a_{nn} + b_{nn})$, we get

$$\begin{aligned} tr(A + B) &= (a_{11} + b_{11}) + (a_{22} + b_{22}) + \cdots + (a_{nn} + b_{nn}) \\ &= (a_{11} + a_{22} + \cdots + a_{nn}) + (b_{11} + b_{22} + \cdots + b_{nn}) \\ &= tr(A) + tr(B) \end{aligned}$$

Matrices with Complex Elements (Optional)

The elements of a matrix may be complex numbers. A **complex number** is of the form

$$z = a + bi$$

where a and b are real numbers and $i = \sqrt{-1}$. a is called the **real part** and b the **imaginary part** of z.

The rules of arithmetic for complex numbers are as follows:

Let $z_1 = a + bi$, $z_2 = c + di$ be complex numbers.

Equality: $z_1 = z_2$ if $a = c$ and $b = d$

Addition: $z_1 + z_2 = (a + c) + (b + d)i$

Subtraction: $z_1 - z_2 = (a - c) + (b - d)i$

Multiplication:
$$\begin{aligned} z_1z_2 &= (a + bi)(c + di) = a(c + di) + bi(c + di) \\ &= ac + adi + bci + bdi^2 \\ &= ac + bdi^2 + (ad + bc)i = (ac - bd) + (ad + bc)i \end{aligned}$$

The **conjugate** of a complex number $z = a + bi$ is defined and written $\bar{z} = a - bi$.

EXAMPLE 6 Consider the complex numbers $z_1 = 2 + 3i$ and $z_2 = 1 - 2i$. Compute $z_1 + z_2$, z_1z_2, and $\bar{z}_1$.

SOLUTION

Using the above definitions we get

$$z_1 + z_2 = (2 + 3i) + (1 - 2i) = (2 + 1) + (3 - 2)i = 3 + i$$
$$z_1z_2 = (2 + 3i)(1 - 2i) = 2(1 - 2i) + 3i(1 - 2i) = 2 - 4i + 3i - 6i^2 = 8 - i$$
$$\bar{z}_1 = 2 - 3i$$

Matrices having complex elements are added, subtracted, and multiplied using the same rules as matrices having real elements.

EXAMPLE 7 Let $A = \begin{bmatrix} 2+i & 3-2i \\ 4 & 5i \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 2i \\ 1+i & 2+3i \end{bmatrix}$. Compute $A + B,\ 2A$, and AB.

SOLUTION

We get

$$A + B = \begin{bmatrix} 2+i & 3-2i \\ 4 & 5i \end{bmatrix} + \begin{bmatrix} 3 & 2i \\ 1+i & 2+3i \end{bmatrix}$$

$$= \begin{bmatrix} 2+i+3 & 3-2i+2i \\ 4+1+i & 5i+2+3i \end{bmatrix} = \begin{bmatrix} 5+i & 3 \\ 5+i & 2+8i \end{bmatrix}$$

$$2A = 2\begin{bmatrix} 2+i & 3-2i \\ 4 & 5i \end{bmatrix} = \begin{bmatrix} 4+2i & 6-4i \\ 8 & 10i \end{bmatrix}$$

$$AB = \begin{bmatrix} 2+i & 3-2i \\ 4 & 5i \end{bmatrix}\begin{bmatrix} 3 & 2i \\ 1+i & 2+3i \end{bmatrix}$$

$$= \begin{bmatrix} (2+i)3 + (3-2i)(1+i) & (2+i)(2i) + (3-2i)(2+3i) \\ (4)(3) + (5i)(1+i) & 4(2i) + (5i)(2+3i) \end{bmatrix}$$

$$= \begin{bmatrix} 11+4i & 10+9i \\ 7+5i & -15+8i \end{bmatrix}$$

The **conjugate** of a matrix A is denoted $\overline{A}$ and is obtained by taking the conjugate of each element of the matrix. The **conjugate transpose** of A is written and defined by $A^* = \overline{A}^{\,t}$.

For example if $A = \begin{bmatrix} 2+3i & 1-4i \\ 6 & 7i \end{bmatrix}$ then

$$\overline{A} = \begin{bmatrix} 2-3i & 1+4i \\ 6 & -7i \end{bmatrix} \quad \text{and} \quad A^* = \overline{A}^{\,t} = \begin{bmatrix} 2-3i & 6 \\ 1+4i & -7i \end{bmatrix}$$

A square matrix C is said to be **hermitian** if $C = C^*$.

Let us show that the matrix $C = \begin{bmatrix} 2 & 3-4i \\ 3+4i & 6 \end{bmatrix}$ is hermitian. We get

$$\overline{C} = \begin{bmatrix} 2 & 3+4i \\ 3-4i & 6 \end{bmatrix}; C^* = \overline{C}^t = \begin{bmatrix} 2 & 3-4i \\ 3+4i & 6 \end{bmatrix} = C$$

Hermitian matrices are more important than symmetric matrices for matrices having complex elements.

The properties of conjugate transpose are similar to those of transpose. We list these properties in the following theorem, leaving the proofs for the reader to do in the exercises that follow.

THEOREM 2.6

Properties of Conjugate Transpose

Let A and B be matrices with complex elements, and let z be a complex number.

1. $(A + B)^* = A^* + B^*$ *Conjugate transpose of a sum*
2. $(zA)^* = \bar{z}A^*$ *Conjugate transpose of a scalar multiple*
3. $(AB)^* = B^*A^*$ *Conjugate transpose of a product*
4. $(A^*)^* = A$

The time is now ripe for a good application of matrix algebra!

Seriation in Archaeology

A problem confronting archaeologists is that of placing sites and artifacts in proper chronological order. This branch of archaeology, called **sequence dating** or **seriation**, began with the work of Sir Flinders Petrie in the late nineteenth century. Petrie studied graves in the cemeteries of Nagada, Ballas, and Hu, all located in what was prehistoric Egypt. (Recent carbon dating shows that all the graves ranged from 6000 B.C. to 2500 B.C.) Petrie used the data from approximately 900 graves to order them and assign a time period to each type of pottery found.

Let us look at this general problem of seriation in terms of graves and varieties of pottery found in graves. An assumption usually made in archaeology is that two graves that have similar contents are more likely to lie close together in time than are two graves that have little in common. The mathematical model that we now construct leads to information concerning the common contents of graves and thus to the chronological order of the graves.

We construct a matrix A, all of whose elements are either 1 or 0, that describes the pottery content of the graves. Label the graves 1, 2, . . . , and the types of pottery 1, 2, ... Let the matrix A be defined by

$$a_{ij} = \begin{cases} 1 & \text{if grave } i \text{ contains pottery type } j \\ 0 & \text{if grave } i \text{ does not contain pottery type } j \end{cases}$$

The matrix A contains all the information about the pottery content of the various graves. The following result now tells us how information is extracted from A.

The element g_{ij} of the matrix $G = AA^t$ is equal to the number of types of pottery common to both grave i and grave j.

Thus, the larger g_{ij}, the closer grave i and grave j are in time. By examining the elements of G, the archaeologist can arrive at the chronological order of the graves. Let us verify this result.

$$\begin{aligned} g_{ij} &= \text{element in row } i\text{, column } j \text{ of } G \\ &= (\text{row } i \text{ of } A) \times (\text{column } j \text{ of } A^t) \\ &= [a_{i1}\ a_{i2} \ldots a_{in}] \begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{bmatrix} \\ &= a_{i1}a_{j1} + a_{i2}a_{j2} + \cdots + a_{in}a_{jn} \end{aligned}$$

Each term in this sum will be either 1 or 0. For example, the term $a_{i2}a_{j2}$ will be 1 if a_{i2} and a_{j2} are both 1; that is if pottery type 2 is common to graves i and j. It will be 0 if pottery type 2 is not common to graves i and j. Thus the number of 1's in this expression for g_{ij} (the actual value of g_{ij}), is the number of types of pottery common to graves i and j.

The matrix $P = A^tA$ leads in an analogous manner to information about the sequence dating of the pottery. The assumption is made that the larger the number of graves in which two types of pottery appear, the closer they are chronologically. The element p_{ij} of the matrix $P = A^tA$ gives the number of graves in which the ith and jth types of pottery both appear. Thus the larger p_{ij}, the closer pottery types i and j are in time. By examining the elements of P, we can arrive at the chronological order of the pottery (see Exercise 30).

It can be shown mathematically that the matrices $G(= AA^t)$ and $P(= A^tA)$ are symmetric matrices. Furthermore, it can be argued from the physical interpretation that G and P should be symmetric matrices (see Exercise 28). This illustrates the compatibility of the mathematics and the interpretation. The implication of this symmetry is that all the information is contained in the elements above the main diagonals of these matrices. The information is just duplicated in the elements below the main diagonals.

We now illustrate this method by means of an example. Let the following matrix A represent the three pottery contents of four graves.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Thus, for example, $a_{13} = 1$ implies that grave 1 contains pottery type 3; $a_{23} = 0$ implies that grave 2 does not contain pottery type 3. G is calculated:

$$G = AA^t = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Observe that G is indeed symmetric. The information contained in the elements above the main diagonal is duplicated in the elements below it. We systematically look at the elements above the main diagonal.

$g_{12} = 1$—graves 1 and 2 have one type of pottery in common.
$g_{13} = 1$—graves 1 and 3 have one type of pottery in common.
$g_{14} = 0$—graves 1 and 4 have no pottery in common.
$g_{23} = 0$—graves 2 and 3 have no pottery in common.
$g_{24} = 0$—graves 2 and 4 have no pottery in common.
$g_{34} = 1$—graves 3 and 4 have one type of pottery in common.

Graves 1 and 2 have pottery in common; they are close together in time. Let us start with graves 1 and 2 and construct a diagram.

$$1 - 2$$

Next add grave 3 to this diagram. $g_{13} = 1$ while $g_{23} = 0$. Thus grave 3 is close to grave 1 but not close to grave 2. We get

$$3 - 1 - 2$$

Finally, add grave 4 to the diagram. $g_{34} = 1$ while $g_{14} = 0$ and $g_{24} = 0$. Grave 4 is close to grave 3 but not close to grave 1 or to grave 2. We get

$$4 - 3 - 1 - 2$$

The mathematics does not tell us which way time flows in this diagram. There are two possibilities:

$$4 \to 3 \to 1 \to 2 \quad \text{and} \quad 4 \leftarrow 3 \leftarrow 1 \leftarrow 2$$

The archaeologist usually knows from other sources which of the two extreme graves (4 and 2 in our case) came first. Thus the chronological order of the graves is known.

The matrices G and P contain information about the chronological order of the graves and pottery, through the relative magnitudes of their elements. These matrices are, in practice, large and the information cannot be sorted out as easily or give results that are as unambiguous as in the above illustration. For example, Petrie examined 900 graves; his matrix G would be a 900×900 matrix. Special mathematical techniques have been developed for extracting information from these matrices; these methods are now being executed on computers. Readers who are interested in pursuing this topic further should consult "Some Problems and Methods in Statistical Archaeology" by David G. Kendall, *World Archaeology*, 1, 61–76, 1969.

In these sections we have developed the algebraic theory of matrices. This theory is extremely important in applications of mathematics. We shall, for example, use matrices in mathematical models of communication and of population movements. Albert Einstein used a matrix equation to describe the relationship between geometry and matter in his general theory of relativity:

$$T = R - \tfrac{1}{2}rG$$

T is a matrix that represents matter, R is a matrix that represents geometry, r is a scalar, and G is a matrix that describes gravity. All the matrices are symmetric 4×4 matrices. The theory involves solving this matrix equation to determine a gravitational field. The German physicist Werner Heisenberg made use of matrices in his development of quantum mechanics, for which he received a Nobel Prize. The theory of matrices is a cornerstone of two of the foremost physical theories of the twentieth century.

EXERCISE SET 2.3

Computation

1. Determine the transpose of each of the following matrices. Indicate whether or not the matrix is symmetric.

(a) $A = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix}$ **(b)** $B = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

(c) $C = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}$ **(d)** $D = \begin{bmatrix} 4 & 5 \\ -2 & 3 \\ 7 & 0 \end{bmatrix}$

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

(f) $F = \begin{bmatrix} 1 & -1 & 3 \\ -1 & 2 & 0 \\ 3 & 0 & 4 \end{bmatrix}$

(g) $G = \begin{bmatrix} -2 & 4 & 5 & 7 \\ 1 & 0 & 3 & -7 \end{bmatrix}$

(h) $H = \begin{bmatrix} 1 & -2 & 3 \\ 4 & 5 & 6 \\ -2 & 6 & 7 \end{bmatrix}$

(i) $K = \begin{bmatrix} 7 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{bmatrix}$

2. Each of the following matrices is to be symmetric. Determine the elements indicated with a •.

(a) $\begin{bmatrix} 1 & 2 & 4 \\ \bullet & 6 & \bullet \\ 4 & 5 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 3 & 5 & \bullet \\ \bullet & 8 & 4 \\ -3 & \bullet & 3 \end{bmatrix}$

(c) $\begin{bmatrix} -3 & \bullet & 8 & 9 \\ -4 & 7 & \bullet & 7 \\ \bullet & 2 & 6 & 4 \\ \bullet & 7 & \bullet & 9 \end{bmatrix}$

3. If A is 4×1, B is 2×3, C is 2×4, and D is 1×3, determine the sizes of the following matrices, if they exist.

(a) ADB^t **(b)** $C^tB - 5AD$

(c) $4CA - (CA)^2$ **(d)** $(ADB^tC)^2 + I_4$

(e) $(B^t C)^t - AD$

Transpose

4. Prove the following properties of transpose given in Theorem 2.4.
   (a) $(A + B)^t = A^t + B^t$ **(b)** $(cA)^t = cA^t$
   (c) $(A^t)^t = A$
5. Prove the following properties of transpose using the results of Theorem 2.4.
   (a) $(A + B + C)^t = A^t + B^t + C^t$
   (b) $(ABC)^t = C^t B^t A^t$
6. Let A be a diagonal matrix. Prove that $A = A^t$.
7. Let A be a square matrix. Prove that $(A^n)^t = (A^t)^n$.

Symmetric Matrices

8. Prove that a square matrix A is symmetric if and only if $a_{ij} = a_{ji}$ for all elements of the matrix.
9. Let A be a symmetric matrix. Prove that A^t is symmetric.
10. Prove that the sum of two symmetric matrices of the same size is symmetric. Prove that the scalar multiple of a symmetric matrix is symmetric. Thus a set of all symmetric matrices of the same size is closed under addition and under scalar multiplication.

Antisymmetric Matrices

11. A square matrix A is said to be **antisymmetric** if $A = -A^t$.
    (a) Give an example of an antisymmetric matrix.
    (b) Prove that the diagonal elements of an antisymmetric matrix are zero.
    (c) Prove that the sum of two antisymmetric matrices of the same size is an antisymmetric matrix.
    (d) Prove that the scalar multiple of an antisymmetric matrix is antisymmetric.
12. If A is a square matrix, prove that
    (a) $A + A^t$ is symmetric. **(b)** $A - A^t$ is antisymmetric.
13. Prove that any square matrix A can be decomposed into the sum of a symmetric matrix B and an antisymmetric matrix C: $A = B + C$.
14. **(a)** Prove that if A is idempotent, then A^t is also idempotent.
    (b) Prove that if A^t is idempotent, then A is idempotent.

Trace of a Matrix

15. Determine the trace of each of the following matrices.

(a) $\begin{bmatrix} 2 & 3 \\ -1 & -4 \end{bmatrix}$ **(b)** $\begin{bmatrix} 5 & 1 & 2 \\ 4 & -3 & 5 \\ -7 & 2 & 8 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & -1 & 2 & 3 \\ -4 & 5 & 3 & 2 \\ 1 & 6 & -7 & 2 \\ 3 & 9 & 2 & 1 \end{bmatrix}$

16. Prove the following properties of trace given in Theorem 2.5.
    (a) $tr(cA) = ctr(A)$ **(b)** $tr(AB) = tr(BA)$
    (c) $tr(A^t) = tr(A)$
17. Prove the following property of trace using the results of Theorem 2.5. $tr(A + B + C) = tr(A) + tr(B) + tr(C)$

If and Only If Condition

18. Consider two matrices A and B of the same size. Prove that $A = B$ if and only if $A^t = B^t$.
19. Prove that the matrix product AB exists if and only if the number of columns in A is equal to the number of rows in B.
20. Prove that $AB = O_n$ for all $n \times n$ matrices B if and only if $A = O_n$.

Complex Matrices

21. Compute $A + B$, AB, and BA for the matrices

$$A = \begin{bmatrix} 5 & 3 - i \\ 2 + 3i & -5i \end{bmatrix}, B = \begin{bmatrix} -2 + i & 5 + 2i \\ 3 - i & 4 + 3i \end{bmatrix}$$

22. Compute $A + B$, AB, and BA for the matrices

$$A = \begin{bmatrix} 4 + i & 2 - 3i \\ 6 + 2i & 1 - i \end{bmatrix}, \quad B = \begin{bmatrix} 2 + i & -3 \\ 2 & 4 - 5i \end{bmatrix}$$

23. Find the conjugate and conjugate transpose of each of the following matrices. Determine which matrices are hermitian.

$$A = \begin{bmatrix} 2 - 3i & 5i \\ 2 & 5 - 4i \end{bmatrix}, \quad B = \begin{bmatrix} 4 & 5 - i \\ 5 + i & 6 \end{bmatrix},$$
$$C = \begin{bmatrix} 7i & 4 - 3i \\ 6 + 8i & -9 \end{bmatrix}, \quad D = \begin{bmatrix} -2 & 3 - 5i \\ 3 + 5i & 9 \end{bmatrix}$$

24. Find the conjugate and conjugate transpose of each of the following matrices. Determine which matrices are hermitian.

$$A = \begin{bmatrix} 3 & 7 + 2i \\ 7 - 2i & 5 \end{bmatrix}, B = \begin{bmatrix} 3 + 5i & 1 - 2i \\ 1 + 2i & 5 + 6i \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \qquad D = \begin{bmatrix} 9 & -3i \\ 3i & 8 \end{bmatrix}$$

25. Prove the following four properties of conjugate transpose.
    (a) $(A + B)^* = A^* + B^*$ **(b)** $(zA)^* = \bar{z}A^*$
    (c) $(AB)^* = B^*A^*$ **(d)** $(A^*)^* = A$
26. Prove that the diagonal elements of a hermitian matrix are real numbers.

Applications

27. The following matrices describe the pottery contents of various graves. For each situation, determine possible chronological orderings of the graves and then the pottery types.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

28. Let $G = AA^t$ and $P = A^tA$, for an arbitrary matrix A.

(a) Prove that G and P are both symmetric matrices.

(b) G and P both have physical interpretation in the archaeological model. Use this physical interpretation to reason that G and P should be symmetric. The mathematical result and the physical interpretation are compatible.

29. Let A be an arbitrary matrix. What information does the ith diagonal element of the matrix AA^t normally give? Discuss.

30. Derive the result for analyzing the pottery in graves. Let A describe the pottery contents of various graves. Prove that:

The element p_{ij} of the matrix $P = A^tA$ gives the number of graves in which the ith and jth types of pottery both appear.

Thus the larger p_{ij}, the closer pottery types i and j are in time. By examining the elements of P the archaeologist can arrive at the chronological order of the pottery.

31. The model introduced here in archaeology is used in sociology to analyze relationships within a group of people. For example, consider the relationship of "friendship" within a group. Assume that all friendships are mutual. Label the people $1, \ldots, n$ and define a square matrix A as follows: $a_{ii} = 0$ for all i (diagonal elements of A are zero)

$$a_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are friends} \\ 0 & \text{if } i \text{ and } j \text{ are not friends} \end{cases}$$

(a) Prove that if $F = AA^t$, then f_{ij} is the number of friends that i and j have in common.

(b) Suppose that all friendships are not mutual. How does this affect the model?

2.4 The Inverse of a Matrix and Cryptography

In this section we introduce the concept of matrix inverse. We shall see how an inverse can be used to solve certain systems of linear equations, and we shall see the role it plays in implementing color on computer monitors, and in cryptography, the study of codes.

We motivate the idea of the inverse of a matrix by looking at the multiplicative inverse of a real number. If number b is the inverse of a, then

$$ab = 1 \quad \text{and} \quad ba = 1.$$

For example, $\frac{1}{4}$ is the inverse of 4 and we have

$$4(\tfrac{1}{4}) = (\tfrac{1}{4})4 = 1.$$

These are the ideas that we extend to matrices.

DEFINITION Let A be an $n \times n$ matrix. If a matrix B can be found such that $AB = BA = I_n$, then A is said to be **invertible** and B is called an **inverse** of A. If such a matrix B does not exist, then A has no inverse.

EXAMPLE 1 Prove that the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ has an inverse $B = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

SOLUTION

We have that

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

and

$$BA = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Thus $AB = BA = I_2$, proving that the matrix A has an inverse B.

We know that a real number can have at most one inverse. We now see that this is also the case for a matrix.

THEOREM 2.7

If a matrix has an inverse, that inverse is unique.

Proof Let B and C be inverses of A. Thus $AB = BA = I_n$ and $AC = CA = I_n$. Multiply both sides of the equation $AB = I_n$ by C and use the algebraic properties of matrices.

$$\begin{aligned} C(AB) &= CI_n \\ (CA)B &= C \\ I_nB &= C \\ B &= C \end{aligned}$$

Thus an invertible matrix has only one inverse.

Notation

The notation for the inverse of a matrix is similar to that used for the inverse of a real number. Let A be an invertible matrix. We denote its inverse A^{-1}. Thus

$$AA^{-1} = A^{-1}A = I_n$$

Let k be a positive integer. We define A^{-k} to be $(A^{-1})^k$. Therefore

$$A^{-k} = \underbrace{A^{-1}A^{-1}\ldots A^{-1}}_{k \text{ times}}$$

Determining the Inverse of a Matrix

We now derive a method for finding the inverse of a matrix. The method is based on the Gauss-Jordan algorithm. Let A be an invertible matrix. Then $AA^{-1} = I_n$. Let the columns of A^{-1} be $X_1, X_2, \ldots, X_n$, and the columns of I_n be $C_1, C_2, \ldots, C_n$. Express A^{-1} and I_n in terms of their columns,

$$A^{-1} = [X_1 \;\; X_2 \ldots X_n] \quad \text{and} \quad I_n = [C_1 \;\; C_2 \ldots C_n]$$

We shall find A^{-1} by finding $X_1, X_2, \ldots, X_n$. Write the equation $AA^{-1} = I_n$ in the form

$$A[X_1 \;\; X_2 \ldots X_n] = [C_1 \;\; C_2 \ldots C_n]$$

Using the column form of matrix multiplication,

$$[AX_1 \;\; AX_2 \ldots AX_n] = [C_1 \;\; C_2 \ldots C_n]$$

Thus

$$AX_1 = C_1, AX_2 = C_2, \ldots, AX_n = C_n$$

Therefore $X_1, X_2, \ldots, X_n$ are solutions to the systems $AX = C_1, AX = C_2, \ldots, AX = C_n$, all of which have the same matrix of coefficients A. Solve these systems by using Gauss-Jordan elimination on the large augmented matrix $[A\colon C_1\, C_2 \ldots C_n]$. Since the solutions $X_1, X_2, \ldots, X_n$ are unique (they are the columns of A^{-1}),

$$[A\colon C_1\ \ C_2 \ldots C_n] \approx \cdots \approx [I_n\colon X_1\ \ X_2 \ldots X_n]$$

Thus, when A^{-1} exists,

$$[A\colon I_n] \approx \cdots \approx [I_n\colon B] \text{ where } B = A^{-1}.$$

On the other hand, if the reduced echelon form of $[A\colon I_n]$ is computed and the first part is not of the form I_n, then A has no inverse.

We now summarize the results of this discussion.

Finding the Inverse of a Matrix Using Elimination

Let A be an $n \times n$ matrix.

1. Adjoin the identity $n \times n$ matrix I_n to A to form the matrix $[A\colon I_n]$.
2. Compute the reduced echelon form of $[A\colon I_n]$. If the reduced echelon form is of the type $[I_n\colon B]$, then B is the inverse of A. If the reduced echelon form is not of the type $[I_n\colon B]$, in that the first $n \times n$ submatrix is not I_n, then A has no inverse.

This discussion also leads to a result about the reduced echelon form of an invertible matrix. Suppose A is invertible. Then as $[A\colon I_n]$ is transformed to $[I_n\colon B]$, A is transformed to I_n. A is row equivalent to I_n. Conversely, if A is not row equivalent to I_n, then $[A\colon I_n]$ is not row equivalent to a matrix of the form $[I_n\colon B]$ and is not invertible. Thus

An $n \times n$ matrix A is invertible if and only if it is row equivalent to I_n.

The following example illustrates this method for finding the inverse of a matrix.

EXAMPLE 2 Determine the inverse of the matrix

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

SOLUTION

Applying the method of Gauss-Jordan elimination, we get

$$[A{:}I_n] = \begin{bmatrix} 1 & -1 & -2 & 1 & 0 & 0 \\ 2 & -3 & -5 & 0 & 1 & 0 \\ -1 & 3 & 5 & 0 & 0 & 1 \end{bmatrix} \underset{\substack{R2 + (-2)R1 \\ R3 + R1}}{\approx} \begin{bmatrix} 1 & -1 & -2 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 2 & 3 & 1 & 0 & 1 \end{bmatrix}$$

$$\underset{(-1)R2}{\approx} \begin{bmatrix} 1 & -1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 2 & 3 & 1 & 0 & 1 \end{bmatrix} \underset{\substack{R1 + R2 \\ R3 + (-2)R2}}{\approx} \begin{bmatrix} 1 & 0 & -1 & 3 & -1 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{bmatrix}$$

$$\underset{\substack{R1 + R3 \\ R2 + (-1)R3}}{\approx} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 5 & -3 & -1 \\ 0 & 0 & 1 & -3 & 2 & 1 \end{bmatrix}$$

Thus

$$A^{-1} = \begin{bmatrix} 0 & 1 & 1 \\ 5 & -3 & -1 \\ -3 & 2 & 1 \end{bmatrix}$$

The following example illustrates what happens when the method is used for a matrix that does not have an inverse.

EXAMPLE 3 Determine the inverse of the following matrix, if it exists.

$$A = \begin{bmatrix} 1 & 1 & 5 \\ 1 & 2 & 7 \\ 2 & -1 & 4 \end{bmatrix}$$

SOLUTION

Applying the method of Gauss-Jordan elimination, we get

$$[A\colon I_3] = \begin{bmatrix} 1 & 1 & 5 & 1 & 0 & 0 \\ 1 & 2 & 7 & 0 & 1 & 0 \\ 2 & -1 & 4 & 0 & 0 & 1 \end{bmatrix} \underset{\substack{R2 + (-1)R1 \\ R3 + (-2)R1}}{\approx} \begin{bmatrix} 1 & 1 & 5 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & -3 & -6 & -2 & 0 & 1 \end{bmatrix}$$

$$\underset{\substack{R1 + (-1)R2 \\ R3 + (3)R2}}{\approx} \begin{bmatrix} 1 & 0 & 3 & 2 & -1 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 0 & -5 & 3 & 1 \end{bmatrix}$$

There is no need to proceed further. The reduced echelon form cannot have a one in the $(3, 3)$ location. The reduced echelon form cannot be of the form $[I_n\colon B]$. Thus A^{-1} does not exist.

We now summarize some of the algebraic properties of matrix inverse.

Properties of Matrix Inverse
Let A and B be invertible matrices and c a nonzero scalar. Then

1. $(A^{-1})^{-1} = A$
2. $(cA)^{-1} = \frac{1}{c}A^{-1}$
3. $(AB)^{-1} = B^{-1}A^{-1}$
4. $(A^n)^{-1} = (A^{-1})^n$
5. $(A^t)^{-1} = (A^{-1})^t$

We verify results 1 and 3 to illustrate the techniques involved, leaving the remaining results for the reader to verify in the exercises that follow.

$\boldsymbol{(A^{-1})^{-1} = A}$ This result follows directly from the definition of inverse of a matrix. Since A^{-1} is the inverse of A we have

$$AA^{-1} = A^{-1}A = I_n$$

This statement also tells us that A is the inverse of A^{-1}. Thus $(A^{-1})^{-1} = A$.

$(AB)^{-1} = B^{-1}A^{-1}$ We want to show that the matrix $B^{-1}A^{-1}$ is the inverse of the matrix AB. We get, using the properties of matrices,

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AI_nA^{-1} \\ &= AA^{-1} \\ &= I_n\end{aligned}$$

Similarly, it can be shown that $(B^{-1}A^{-1})(AB) = I_n$. Thus $B^{-1}A^{-1}$ is the inverse of the matrix AB.

EXAMPLE 4 If $A = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}$, then it can be shown that $A^{-1} = \begin{bmatrix} 1 & -1 \\ -3 & 4 \end{bmatrix}$. Use this information to compute $(A^t)^{-1}$.

SOLUTION

Result 5 above tells us that if we know the inverse of a matrix we also know the inverse of its transpose. We get

$$(A^t)^{-1} = (A^{-1})^t = \begin{bmatrix} 1 & -1 \\ -3 & 4 \end{bmatrix}^t = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}$$

Systems of Linear Equations

We now see that matrix inverse enables us to conveniently express the solutions to certain systems of linear equations.

THEOREM 2.8

Let $AX = Y$ be a system of n linear equations in n variables. If A^{-1} exists, the solution is unique and is given by $X = A^{-1}Y$.

Proof We first prove that $X = A^{-1}Y$ is a solution.
Substitute $X = A^{-1}Y$ into the matrix equation. Using the properties of matrices we get

$$AX = A(A^{-1}Y) = (AA^{-1})Y = I_nY = Y$$

$X = A^{-1}Y$ satisfies the equation; thus it is a solution.

We now prove the uniqueness of the solution. Let X_1 be any solution. Thus $AX_1 = Y$. Multiplying both sides of this equation by A^{-1} gives

$$\begin{aligned}A^{-1}AX_1 &= A^{-1}Y \\ I_nX_1 &= A^{-1}Y \\ X_1 &= A^{-1}Y\end{aligned}$$

Thus there is a unique solution $X_1 = A^{-1}Y$.

EXAMPLE 5 Solve the following system of equations using the inverse of the matrix of coefficients.

$$\begin{aligned}x_1 - x_2 - 2x_3 &= 1 \\ 2x_1 - 3x_2 - 5x_3 &= 3 \\ -x_1 + 3x_2 + 5x_3 &= -2\end{aligned}$$

SOLUTION

This system can be written in the following matrix form,

$$\begin{bmatrix} 1 & -1 & -2 \\ 2 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

If the matrix of coefficients is invertible, the unique solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -2 \\ 2 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}^{-1}\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

This inverse has already been found in Example 2. Using that result we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 5 & -3 & -1 \\ -3 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

The unique solution is $x_1 = 1$, $x_2 = -2$, $x_3 = 1$.

Numerical Considerations

The knowledge that if A is invertible the solution to the system of equations $AX = Y$ is $X = A^{-1}Y$, is primarily of theoretical importance. It gives an algebraic expression for the solution. It is not usually used in practice to solve a specific system of equations. An elimination method such as Gauss-Jordan elimination or Gaussian elimination of Section 1 of the "Numerical Methods" chapter are more efficient. Most systems of equations are solved on a computer. As mentioned earlier, two factors that are important when using a computer are efficiency and accuracy. To solve a system of n equations, the matrix inverse method requires $n^3 + n^2$ multiplications and $n^3 - n^2$ additions, while Gauss-Jordan elimination requires $\frac{1}{2}n^3 + \frac{1}{2}n^2$ multiplications (half as many) and $\frac{1}{2}n^3 - \frac{1}{2}n$ additions. Thus, for example, for a system of ten equations in ten variables, the matrix inverse method would involve 1,100 multiplications and 900 additions, while Gauss-Jordan elimination would involve 550 multiplications and 495 additions. Furthermore, the more operations that are performed, the larger the possible round-off error. Thus Gauss-Jordan elimination is in general also more accurate than the matrix inverse method.

One may be tempted to assume that given a number of linear systems $AX = Y_1$, $AX = Y_2, \ldots, AX = Y_k$, all having the same invertible matrix of coefficients A, that it would be efficient to calculate A^{-1} and then compute the solutions using $X_1 = A^{-1}Y_1$, $X_2 = A^{-1}Y_2, \ldots, X_n = A^{-1}Y_k$. This approach would involve only one computation of A^{-1} and then a number of matrix multiplications. In general, however, it is more efficient and accurate to solve such systems using a large augmented matrix that represents all systems, and an elimination method such as Gauss-Jordan elimination, as discussed earlier.

In certain instances, the matrix inverse method is used to arrive at specific solutions. In Section 2.7 we illustrate such a situation in a model for analyzing the interdependence of industries. The elements of the matrix of coefficients in that example lend themselves to an efficient algorithm for computing the inverse.

Elementary Matrices

We now introduce a very useful class of matrices called *elementary matrices*. Row operations and their inverses can be performed using these matrices. This way of implementing row operations is particularly appropriate for computers.

An elementary matrix is one that can be obtained from the identity matrix I_n through a single elementary row operation.

Illustration Consider the following three row operations T_1, T_2, and T_3 on I_3 (one representing each kind of row operation). They lead to the three elementary matrices E_1, E_2, and E_3.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Elementary Row Operation	Corresponding Elementary Matrix
T_1: interchange rows 2 and 3 of I_3.	$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
T_2: multiply row 2 of I_3 by 5.	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
T_3: add 2 times row 1 of I_3 to row 2.	$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Suppose we want to perform a row operation T on an $m \times n$ matrix A. Let E be the elementary matrix obtained from I_n through the operation T. This row operation can be performed by multiplying A by E.

Illustration Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ be an arbitrary 3×3 matrix. Consider the three row operations above. Let us show that the corresponding elementary matrices can indeed be used to peform these operations.

Interchange rows 2 and 3 of A: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}$

Multiply row 2 by 5: $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 5d & 5e & 5f \\ g & h & i \end{bmatrix}$$

Add 2 row 1 to row 2:
$$\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d + 2a & e + 2b & f + 2c \\ g & h & i \end{bmatrix}$$

Each row operation has an inverse, namely the row operation that returns the original matrix. The elementary matrices of a row operation and its inverse operation are inverse matrices.

Each elementary matrix is square and invertible.

We now illustrate a way elementary matrices are used to arrive at theoretical results. Matrices that can be obtained from one another by a finite sequence of elementary row operations are said to be *row equivalent*.

If A and B are row equivalent matrices and A is invertible, then B is invertible.

Let us prove this result. Since A and B are row equivalent, there exists a sequence of row operations $T_1, \ldots, T_n$ such that $B = T_n \circ \cdots \circ T_1(A)$. Let the elementary matrices of these operations be $E_1, \ldots, E_n$. Thus

$$B = E_n \ldots E_1 A$$

The matrices $A, E_1, \ldots, E_n$ are all invertible. Repeatedly applying the property of matrix inverse of a product to the following expression we get

$$\begin{aligned} A^{-1}E_1^{-1}E_2^{-1}E_3^{-1}\ldots E_n^{-1} &= (E_1A)^{-1}E_2^{-1}E_3^{-1}\ldots E_n^{-1} \\ &= (E_2E_1A)^{-1}E_3^{-1}\ldots E_n^{-1} = \cdots \\ &= (E_n \ldots E_1A)^{-1} = B^{-1} \end{aligned}$$

Thus B is invertible and the inverse is given by

$$B^{-1} = A^{-1}E_1^{-1}E_2^{-1}\ldots E_n^{-1}$$

Elementary matrices are used in arriving at the so-called LU decomposition of certain square matrices. These are decompositions into products of lower (L) and upper (U) triangular matrices—matrices that have zeros above or below the main diagonal. The importance of this decomposition lies in the fact that once it is accomplished for a matrix A, the LU form provides a powerful starting point for performing many matrix tasks such as solving equations, computing matrix inverses, and finding determinants of matrices. (See Section 2 of the "Numerical Methods" chapter.) LU decomposition is, for example, used extensively in MATLAB® (discussed in the "MATLAB Manual" appendix). MATLAB is probably the most widely used matrix software package.

Color Models

A color model in the context of graphics is a method of implementing colors. There are numerous models that are used in practice, such as the RGB model (Red, Green, Blue) used in computer monitors, and the YIQ model used in television screens. An RGB computer signal can be converted to a YIQ television signal using what is known as an NTSC encoder. (NTSC stands for National Television System Committee.) The conversion is accomplished by using the following matrix transformation*

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.523 & .311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Let us look at the RGB model for Microsoft® Word®. The default text color is black. Let us find the RGB values for black, change the text color to a purple, and find the RGB values for this color. On the right of the tool bar of Microsoft Word, observe A▼. The bar under the A is black, indicating the current text color. Point the cursor at this bar. It shows "Font Color (RGB(0, 0, 0))." The RGB setting for black is (0, 0, 0). To change the color, select the sequence "▼→ More Colors → Custom." A spectrum of colors is displayed. Select a purple hue. The corresponding RGB values are seen to be R = 213, G = 77, B = 187. The bar under the A has now changed to purple and any text entered at the keyboard is in purple. The range of values for each of R, G, and B is 0 to 255, the set of numbers that can be represented by a byte on a computer (note that $2^8 = 256$). You are asked to use the matrix transformation to find the range of Y, I, and Q values in the exercises that follow.

If we enter the RGB values for black, namely R = 0, G = 0, B = 0, into the preceding transformation, we find that Y = 0, I = 0, Q = 0. Black has the same RGB and YIQ values. The RGB values R = 213, G = 77, B = 187 for purple become Y = 130.204, I = 45.746, Q = 63.042. These are the YIQ values that would be used to duplicate this purple color on a television screen.

A signal is converted from a television screen to a computer monitor using the inverse of the above matrix,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.523 & .311 \end{bmatrix}^{-1} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

That is,

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & .956 & .620 \\ 1 & -.272 & -.647 \\ 1 & -1.108 & 1.705 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

Cryptography

In the previous application, we talked about two different ways colors are coded. We now turn our attention to coding messages. Cryptography is the process of coding and

*Numbers in this field are usually written to three decimal places.

decoding messages. The word comes from the Greek "kryptos," meaning "hidden." The technique can be traced back to the ancient Greeks. Today governments use sophisticated methods of coding and decoding messages. One type of code that is extremely difficult to break makes use of a large invertible matrix to encode a message. The receiver of the message decodes it using the inverse of the matrix. This first matrix is called the *encoding matrix*, and its inverse is called the *decoding matrix*. We illustrate the method for a 3×3 matrix.

Let the message be

BUY IBM STOCK

and the encoding matrix be

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

We assign a number to each letter of the alphabet. For convenience, let us associate each letter with its position in the alphabet. A is 1, B is 2, and so on. Let a space between words be denoted by the number 27. The digital form of the message is

B	U	Y	–	I	B	M	–	S	T	O	C	K
2	21	25	27	9	2	13	27	19	20	15	3	11

Since we are going to use a 3×3 matrix to encode the message, we break the digital message up into a sequence of 3×1 column matrices as follows.

$$\begin{bmatrix} 2 \\ 21 \\ 25 \end{bmatrix}, \begin{bmatrix} 27 \\ 9 \\ 2 \end{bmatrix}, \begin{bmatrix} 13 \\ 27 \\ 19 \end{bmatrix}, \begin{bmatrix} 20 \\ 15 \\ 3 \end{bmatrix}, \begin{bmatrix} 11 \\ 27 \\ 27 \end{bmatrix}$$

Observe that it was necessary to add two spaces at the end of the message in order to complete the last matrix. We now put the message into code by multiplying each of the above column matrices by the encoding matrix. This can be conveniently done by writing the given column matrices as columns of a matrix and premultiplying that matrix by the encoding matrix. We get

$$\begin{bmatrix} -3 & -3 & -4 \\ 0 & 1 & 1 \\ 4 & 3 & 4 \end{bmatrix}\begin{bmatrix} 2 & 27 & 13 & 20 & 11 \\ 21 & 9 & 27 & 15 & 27 \\ 25 & 2 & 19 & 3 & 27 \end{bmatrix} = \begin{bmatrix} -169 & -116 & -196 & -117 & -222 \\ 46 & 11 & 46 & 18 & 54 \\ 171 & 143 & 209 & 137 & 233 \end{bmatrix}$$

The columns of this matrix give the encoded message. The message is transmitted in the following linear form.

$$-169, 46, 171, -116, 11, 143, -196, 46, 209, -117, 18, 137, -222, 54, 233$$

To decode the message, the receiver writes this string as a sequence of 3×1 column matrices and repeats the technique using the inverse of the encoding matrix. The inverse of this encoding matrix, the decoding matrix is

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

Thus, to decode the message

$$\begin{bmatrix} 1 & 0 & 1 \\ 4 & 4 & 3 \\ -4 & -3 & -3 \end{bmatrix}\begin{bmatrix} -169 & -116 & -196 & -117 & -222 \\ 46 & 11 & 46 & 18 & 54 \\ 171 & 143 & 209 & 137 & 233 \end{bmatrix} = \begin{bmatrix} 2 & 27 & 13 & 20 & 11 \\ 21 & 9 & 27 & 15 & 27 \\ 25 & 2 & 19 & 3 & 27 \end{bmatrix}$$

The columns of this matrix, written in linear form, give the original message.

2	21	25	27	9	2	13	27	19	20	15	3	11
B	U	Y	–	I	B	M	–	S	T	O	C	K

Readers who are interested in an introduction to cryptography are referred to *Coding Theory and Cryptography* edited by David Joyner, Springer-Verlag, 2000. This is an excellent collection of articles that contain historical, elementary, and advanced discussions.

EXERCISE SET 2.4

Checking for Matrix Inverse

1. Use the definition $AB = BA = I_2$ of inverse to check whether B is the inverse of A, for each of the following 2×2 matrices A and B.

(a) $A = \begin{bmatrix} 7 & -3 \\ 5 & -2 \end{bmatrix}$, $B = \begin{bmatrix} -2 & 3 \\ -5 & 7 \end{bmatrix}$

(b) $A = \begin{bmatrix} 3 & 4 \\ 5 & 7 \end{bmatrix}$, $B = \begin{bmatrix} 7 & -4 \\ -5 & 3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 & -4 \\ -5 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$

(d) $A = \begin{bmatrix} 7 & 6 \\ 8 & 7 \end{bmatrix}$, $B = \begin{bmatrix} 7 & -6 \\ -8 & 7 \end{bmatrix}$

2. Use the definition $AB = BA = I_3$ of inverse to check whether B is the inverse of A, for each of the following 3×3 matrices A and B.

(a) $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & -2 \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 & -1 \\ -3 & 2 & -1 \\ 3 & -3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 4 \\ 3 & 6 & 5 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & 1 & -1 \\ 2 & -2 & -1 \\ -1 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

Finding the Inverse of a Matrix

3. Determine the inverse of each of the following 2×2 matrices, if it exists, using the method of Gauss-Jordan elimination.

(a) $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 2 \\ 9 & 4 \end{bmatrix}$

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

(e) $\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$ **(f)** $\begin{bmatrix} 2 & -3 \\ 6 & -7 \end{bmatrix}$

4. Determine the inverse of each of the following 3×3 matrices, if it exists, using the method of Gauss-Jordan elimination.

(a) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 4 & 5 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 0 & 4 \\ -1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

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

5. Determine the inverse of each of the following 3×3 matrices, if it exists, using the method of Gauss-Jordan elimination.

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

(c) $\begin{bmatrix} 1 & -2 & -1 \\ -2 & 4 & 6 \\ 0 & 0 & 5 \end{bmatrix}$ (d) $\begin{bmatrix} 7 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$

6. Determine the inverse of each of the following 4×4 matrices, if it exists, using the method of Gauss-Jordan elimination.

(a) $\begin{bmatrix} -3 & -1 & 1 & -2 \\ -1 & 3 & 2 & 1 \\ 1 & 2 & 3 & -1 \\ -2 & 1 & -1 & -3 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

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

7. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, show that $A^{-1} = \dfrac{1}{(ad - bc)}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

This formula can be quicker than Gauss-Jordan elimination to compute the inverse of a 2×2 matrix. Compute the inverses of the following 2×2 matrices using both methods to see which you prefer.

(a) $\begin{bmatrix} 3 & 8 \\ 1 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}$

(c) $\begin{bmatrix} 5 & 6 \\ 3 & 4 \end{bmatrix}$ (d) $\begin{bmatrix} 4 & -6 \\ 2 & -2 \end{bmatrix}$

Systems of Linear Equations

8. Solve the following systems of two equations in two variables by determining the inverse of the matrix of coefficients and then using matrix multiplication.

(a) $\begin{aligned} x_1 + 2x_2 &= 2 \\ 3x_1 + 5x_2 &= 4 \end{aligned}$ (b) $\begin{aligned} x_1 + 5x_2 &= -1 \\ 2x_1 + 9x_2 &= 3 \end{aligned}$

(c) $\begin{aligned} x_1 + 3x_2 &= 5 \\ 2x_1 + x_2 &= 10 \end{aligned}$ (d) $\begin{aligned} 2x_1 + x_2 &= 4 \\ 4x_1 + 3x_2 &= 6 \end{aligned}$

(e) $\begin{aligned} 2x_1 + 4x_2 &= 6 \\ 3x_1 + 8x_2 &= 1 \end{aligned}$ (f) $\begin{aligned} 3x_1 + 9x_2 &= 9 \\ 2x_1 + 7x_2 &= 4 \end{aligned}$

9. Solve the following systems of three equations in three variables by determining the inverse of the matrix of coefficients and then using matrix multiplication.

(a) $\begin{aligned} x_1 + 2x_2 - x_3 &= 2 \\ x_1 + x_2 + 2x_3 &= 0 \\ x_1 - x_2 - x_3 &= 1 \end{aligned}$

(b) $\begin{aligned} x_1 - x_2 &= 1 \\ x_1 + x_2 + 2x_3 &= 2 \\ x_1 + 2x_2 + x_3 &= 0 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 + 8x_3 &= 15 \end{aligned}$

(d) $\begin{aligned} x_1 - 2x_2 + 2x_3 &= 3 \\ -x_1 + x_2 + 3x_3 &= 2 \\ x_1 - x_2 - 4x_3 &= -1 \end{aligned}$

(e) $\begin{aligned} -x_1 + x_2 &= 5 \\ -x_1 + x_3 &= -2 \\ 6x_1 - 2x_2 - 3x_3 &= 1 \end{aligned}$

10. Solve the following system of four equations in four variables by determining the inverse of the matrix of coefficients and then using matrix multiplication.

$$\begin{aligned} x_1 + x_2 + 2x_3 + x_4 &= 5 \\ 2x_1 + 2x_3 + x_4 &= 6 \\ x_2 + 3x_3 - x_4 &= 1 \\ 3x_1 + 2x_2 + 2x_4 &= 7 \end{aligned}$$

11. Solve the following systems of equations, all having the same matrix of coefficients, using the matrix inverse method,

$$\begin{aligned} x_1 + 2x_2 - x_3 &= b_1 \\ x_1 + x_2 + 2x_3 &= b_2 \quad \text{for} \\ x_1 - x_2 - x_3 &= b_3 \end{aligned}$$

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -3 \end{bmatrix} \text{ in turn.}$$

Miscellaneous Results

12. Prove the following properties of matrix inverse that were listed in this section.

(a) $(cA)^{-1} = \frac{1}{c}A^{-1}$

(b) $(A^n)^{-1} = (A^{-1})^n$
(c) $(A^t)^{-1} = (A^{-1})^t$

13. If $A^{-1} = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$, find A.

14. If $A^{-1} = \frac{1}{2}\begin{bmatrix} -3 & 2 \\ -10 & 6 \end{bmatrix}$, find A.

15. Consider the matrix $A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$, having inverse $\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$. Determine
(a) $(3A)^{-1}$ **(b)** $(A^2)^{-1}$ **(c)** A^{-2} **(d)** $(A^t)^{-1}$
[*Hint*: Use the algebraic properties of matrix inverse.]

16. If $A = \begin{bmatrix} 5 & 1 \\ 9 & 2 \end{bmatrix}$ then $A^{-1} = \begin{bmatrix} 2 & -1 \\ -9 & 5 \end{bmatrix}$. Use this information to determine
(a) $(2A^t)^{-1}$ **(b)** A^{-3} **(c)** $(AA^t)^{-1}$

17. Find x such that $\begin{bmatrix} 2x & 7 \\ 1 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}$.

18. Find x such that $2\begin{bmatrix} 2x & x \\ 5 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}$.

19. Find A such that $(4A^t)^{-1} = \begin{bmatrix} 2 & 3 \\ -4 & -4 \end{bmatrix}$.

20. Prove that $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.

21. Prove that $(A^tB^t)^{-1} = (A^{-1}B^{-1})^t$.

22. Prove that, in general, $(A + B)^{-1} \neq A^{-1} + B^{-1}$.

23. Prove that if A has no inverse, then A^t also has no inverse.

24. Prove that if A is an invertible matrix such that
(a) $AB = AC$, then $B = C$ **(b)** $AB = 0$, then $B = 0$

25. Prove that a matrix has no inverse if
(a) two rows are equal.
(b) two columns are equal. (*Hint*: Use the transpose.)
(c) it has a column of zeros.

26. Prove that a diagonal matrix is invertible if and only if all its diagonal elements are nonzero. Can you find a quick way for determining the inverse of an invertible diagonal matrix?

27. Prove that the set of 2×2 invertible matrices is not closed under either addition or scalar multiplication. (*Hint*: Give examples.)

28. State (with a brief explanation) whether the following statements are true or false for a square matrix A.
(a) If A is invertible, A^{-1} is invertible.
(b) If A is invertible, A^2 is invertible.
(c) If A has a zero on the main diagonal, it is not invertible.
(d) If A is not invertible, then AB is not invertible.
(e) A^{-1} is row equivalent to I_n.

Numerical Considerations

29. Let $AX = Y$ be a system of 25 linear equations in 25 variables, where A is invertible. Find the number of multiplications and additions needed to solve this system using **(a)** Gauss-Jordan elimination, **(b)** the matrix inverse method.

30. Let A be an invertible 2×2 matrix. Show that it takes the same amount of computation to find the solution to the system of equations $AX = Y$ using Gauss-Jordan elimination as it does to find the first column of the inverse of A. This exercise emphasizes the fact that it is more efficient to solve a system of equations by using Gauss-Jordan elimination than by using the inverse of the matrix of coefficients.

Elementary Matrices

31. Let A be a 3×3 matrix and T_1, T_2, T_3, be the following row operations. T_1: interchange rows 1 and 2; T_2: multiply row 3 by -2; T_3: add 4row 1 to row 3. Show that T_1, T_2, T_3 can be performed using the following elementary matrices E_1, E_2, E_3.

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix},$$
$$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{bmatrix}$$

32. Let T_1, T_2 be the following row operations. T_1: multiply row 1 by -3. T_2: add 3 times row 2 to row 1. Find the elementary 3×3 matrices of T_1, T_2.

33. Determine the row operation defined by each of the following elementary matrices. Find the inverse of that row operation and use it to find the inverse of the elementary matrix. Arrive at rules that will enable you to quickly write down the inverse of any given elementary matrix.

(a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, **(b)** $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$, **(d)** $\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$, **(f)** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$

Color Model

34. In the color model discussion we indicated that the range of each of the RGB values is 0 to 255. The interval for each of Y, I, and Q, however, is different. Use the encoding matrix equation to find the range of each of Y, I, and Q.

35. Consider the following YIQ values: $(176, -111, -33)$, $(184, 62, -18)$, $(171, 5, -19)$, and $(165, -103, -23)$. **(a)** Use the inverse matrix transformation to find the corresponding RGB values. **(b)** Use Microsoft® Word® to find the colors corresponding to these YIQ values; describe them in your own terms.

36. Black-and-white television monitors use only the Y signal. **(a)** Show that every YIQ signal of the form $(s, 0, 0)$ transforms into an RGB signal of the form (s, s, s). What is there about the matrix of the transformation from YIQ to RGB that makes this happen? **(b)** Find the RGB signals correponding to the YIQ signals $(255, 0, 0)$, $(200, 0, 0)$, $(150, 0, 0)$, $(100, 0, 0)$, $(0, 0, 0)$ of a black-and-white set. Use Microsoft Word to investigate these signals. Describe the effect of decreasing a in the television signal $(a, 0, 0)$ from 255 to zero.

Cryptography

In Exercises 37–42 associate each letter with its position in the alphabet.

37. Encode the message RETREAT using the matrix $\begin{bmatrix} 4 & -3 \\ 3 & -2 \end{bmatrix}$.

38. Encode the message THE BRITISH ARE COMING using the matrix

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

39. Decode the message 49, 38, -5, -3, -61, -39, which was encoded using the matrix of Exercise 37.

40. Decode the message 71, 100, -1, 28, 43, -5, 84, 122, -11, 63, 98, -27, 69, 102, -12, 88, 126, -3, which was encoded using the matrix of Exercise 38.

41. Intelligence sources give the information that the message BOSTON CAFE AT TWO was sent as 32, 47, 59, 79, 43, 57, 33, 36, 13, 19, 59, 86, 41, 61, 67, 87, 53, 68. **(a)** Find the encoding matrix. **(b)** Decode the message 43, 64, 49, 70, 59, 79, 39, 45, 45, 63, 59, 79.

42. Base station sends messages to an agent using the encoding matrix $A = \begin{bmatrix} 4 & -3 \\ 3 & -2 \end{bmatrix}$. The agent sends messages to an informer using the encoding matrix $B = \begin{bmatrix} 3 & 8 \\ 4 & 11 \end{bmatrix}$. Find the encoding matrix that is consistent with this communication circle, that enables base to send messages directly to the informer.

2.5 Matrix Transformations, Rotations, and Dilations

A function, or transformation, is a rule that assigns to each element of a set a unique element of another set. Transformations are used in many areas of mathematics and are important in applications for describing the dependency of one variable upon another. We shall be especially interested in linear transformations, which are transformations that preserve the mathematical structure of a vector space. In this section and the next, we shall see how linear transformations are used in computer graphics and in fractal geometry.

The reader will be familiar with functions such as $f(x) = 3x^2 + 4$. The set of allowable x values is called the *domain of the function*. The domain is often the set of real numbers, as here. When $x = 2$, for example, we see that $f(2) = 16$. We say that the image of 2 is 16. We extend these ideas to functions between vector spaces. We usually use the term *transformation* rather than *function* in linear algebra.

For example, consider the transformation T of $\mathbf{R}^3$ into $\mathbf{R}^2$ defined by

$$T(x, y, z) = (2x, y - z)$$

The *domain* of T is $\mathbf{R}^3$ and we say that the *codomain* is $\mathbf{R}^2$. The image of a vector such as $(1, 4, -2)$ can be found by letting $x = 1$, $y = 4$, and $z = -2$ in this equation for T. We get $T(1, 4, -2) = (2, 6)$. The *image* of $(1, 4, -2)$ is $(2, 6)$.

We shall often find it convenient to write vectors in column form when discussing transformations. The preceding transformation can also be written

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x \\ y - z \end{bmatrix}, \text{ and the image of } \begin{bmatrix} 1 \\ 4 \\ -2 \end{bmatrix} \text{ is } \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

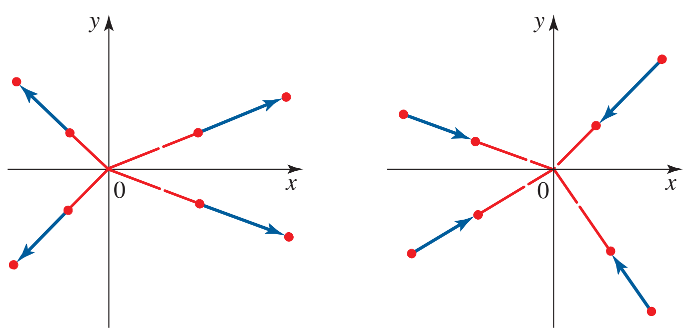
DEFINITION A *transformation* T of $\mathbf{R}^n$ into $\mathbf{R}^m$ is a rule that assigns to each vector $\mathbf{u}$ in $\mathbf{R}^n$ a unique vector $\mathbf{v}$ in $\mathbf{R}^m$. $\mathbf{R}^n$ is called the *domain* of T and $\mathbf{R}^m$ is the *codomain*. We write $T(\mathbf{u}) = \mathbf{v}$; $\mathbf{v}$ is the *image* of $\mathbf{u}$ under T. The term *mapping* is also used for a transformation.

We now introduce a number of useful geometric transformations and find that they can be described by matrices.

Dilation and Contraction

Consider the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = r\begin{bmatrix} x \\ y \end{bmatrix}$, where r is a positive scalar. T maps every point in $\mathbf{R}^2$ into a point r times as far from the origin. If $r > 1$, T moves points away from the origin and is called a *dilation of factor* r. If $0 < r < 1$, T moves points closer to the origin and is then called a *contraction of factor* r. See Figure 2.10. This equation can be written in the following useful matrix form.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

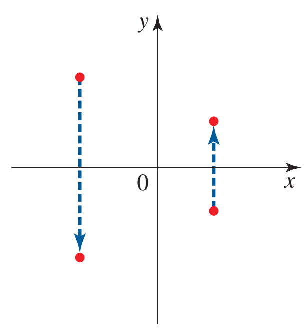


$r > 1$, dilation

$0 < r < 1$, contraction

Transformations defined by $\begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$

Figure 2.10



reflection in x-axis

Transformations defined by $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Figure 2.11

For example, when $r = 3$, we see in the following equation that an image is three times as far from the origin. When $r = 1/2$, an image is half the distance from the origin.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 3\begin{bmatrix} x \\ y \end{bmatrix}, \qquad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}$$

Reflection

Consider the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ -y \end{bmatrix}$. T maps every point in $\mathbf{R}^2$ into its mirror image in the x-axis. T is called a *reflection*. See Figure 2.11. This equation can be written in the following matrix form.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

For example, the image of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ under this reflection is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

We now find that a rotation about the origin is also a matrix transformation.

Rotation about the Origin

Consider a rotation T about the origin through an angle θ, as shown in Figure 2.12. T *maps* the point $A\begin{bmatrix} x \\ y \end{bmatrix}$ into the point $B\begin{bmatrix} x' \\ y' \end{bmatrix}$.

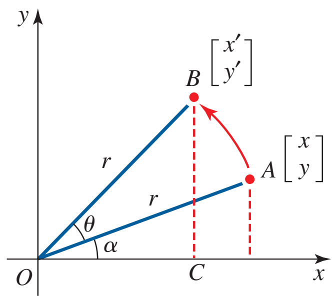


Figure 2.12

The distance OA is equal to OB; let it be r. Let the angle AOC be α. We get

$$\begin{aligned} x' = OC = r\cos(\alpha + \theta) &= r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ &= x\cos\theta - y\sin\theta \\ y' = BC = r\sin(\alpha + \theta) &= r\sin\alpha\cos\theta + r\cos\alpha\sin\theta \\ &= y\cos\theta + x\sin\theta \\ &= x\sin\theta + y\cos\theta \end{aligned}$$

These expressions for x' and y' can be combined into a single matrix equation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

We thus get the following result.

A rotation through an angle θ is described by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

Consider a rotation of $\pi/2$ about the origin. Since $\cos(\pi/2) = 0$ and $\sin(\pi/2) = 1$, the transformation is

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

The image of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, for example, is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.
Note that θ is positive for a counterclockwise rotation and negative for a clockwise rotation.

Matrix Transformations

In previous discussions we found that we could use matrices to define certain transformations. We now see that every matrix in fact defines a transformation. Let A be a matrix and $\mathbf{x}$ be a column vector such that $A\mathbf{x}$ exists. Then A defines the *matrix transformation* $T(\mathbf{x}) = A\mathbf{x}$. For example,

$$A = \begin{bmatrix} 5 & 3 & -2 \\ 0 & 4 & -1 \end{bmatrix} \text{ defines the transformation } T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 5 & 3 & -2 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

The image of a vector such as $\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ is $\begin{bmatrix} 6 \\ 8 \end{bmatrix}$. We write $\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \mapsto \begin{bmatrix} 6 \\ 8 \end{bmatrix}$. Similarly, for example, $\begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \mapsto \begin{bmatrix} 14 \\ 2 \end{bmatrix}$. We say that T *maps* $\mathbf{R}^3$ into $\mathbf{R}^2$ and write $T\colon \mathbf{R}^3 \to \mathbf{R}^2$. The *domain* of the transformation is $\mathbf{R}^3$, the *codomain* is $\mathbf{R}^2$. We can convey this information in a diagram, Figure 2.13.

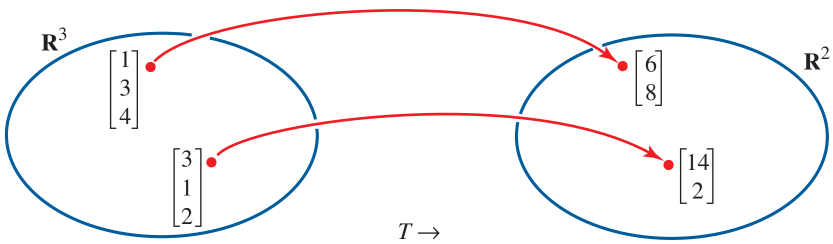


Figure 2.13

DEFINITION Let A be an $m \times n$ matrix. Let $\mathbf{x}$ be an element of $\mathbf{R}^n$ written in column matrix form. A defines a *matrix transformation* $T(\mathbf{x}) = A\mathbf{x}$ of $\mathbf{R}^n$ into $\mathbf{R}^m$. The vector $A\mathbf{x}$ is the *image* of $\mathbf{x}$. The *domain* of the transformation is $\mathbf{R}^n$ and the *codomain* is $\mathbf{R}^m$.

These transformations have the following geometrical properties (which we do not prove).

Matrix transformations map line segments into line segments (or points). If the matrix is invertible, the transformation also maps parallel lines into parallel lines.

The following example illustrates how a square is deformed.

EXAMPLE 1 Consider the transformation $T: \mathbf{R}^2 \to \mathbf{R}^2$ defined by the matrix $A = \begin{bmatrix} 4 & 2 \\ 2 & 3 \end{bmatrix}$. Determine the image of the unit square under this transformation.

SOLUTION

The unit square is the square whose vertices are the points

$$P\begin{bmatrix}1\\0\end{bmatrix}, Q\begin{bmatrix}1\\1\end{bmatrix}, R\begin{bmatrix}0\\1\end{bmatrix}, O\begin{bmatrix}0\\0\end{bmatrix}$$

See Figure 2.14(a). Let us compute the images of these points under the transformation. Multiplying each point by the matrix, we get

$$\overset{P}{\begin{bmatrix}1\\0\end{bmatrix}} \mapsto \overset{P'}{\begin{bmatrix}4\\2\end{bmatrix}}, \overset{Q}{\begin{bmatrix}1\\1\end{bmatrix}} \mapsto \overset{Q'}{\begin{bmatrix}6\\5\end{bmatrix}}, \overset{R}{\begin{bmatrix}0\\1\end{bmatrix}} \mapsto \overset{R'}{\begin{bmatrix}2\\3\end{bmatrix}}, \overset{O}{\begin{bmatrix}0\\0\end{bmatrix}} \mapsto \overset{O}{\begin{bmatrix}0\\0\end{bmatrix}}$$

The line segments are mapped as follows, Figure 2.14(b),

$$OP \mapsto OP', PQ \mapsto P'Q', QR \mapsto Q'R', OR \mapsto OR'$$

This matrix A is invertible. It thus maps parallel lines into parallel lines. The square $PQRO$ is mapped into the parallelogram $P'Q'R'O$.

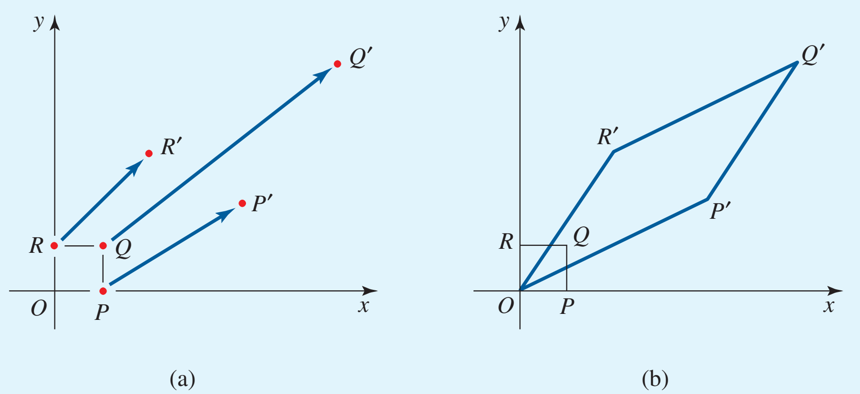


Figure 2.14

Solid bodies can be described geometrically. When loads are applied to bodies, changes in shape called *deformations* occur. For example, the square $PQRO$ in Figure 2.15 could represent a physical body that is deformed into the shape $P'Q'R'O$. Such deformations can be modeled and analyzed on computers using these mathematical techniques. The fields of science that investigate such problems are called *elasticity* and *plasticity*.

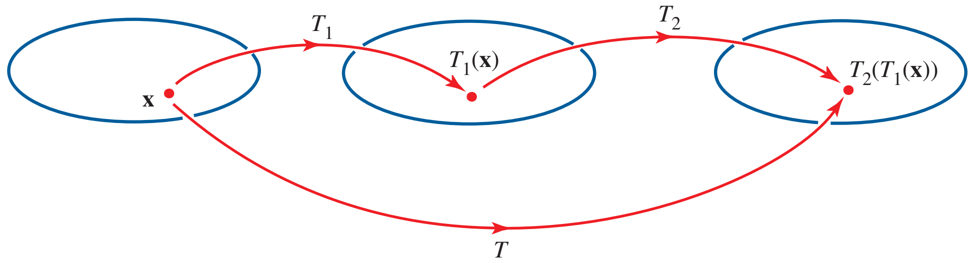
Composition of Transformations

The reader will be familiar with the concept of combining functions into composite functions, Figure 2.15. Matrix transformations can be combined in a very useful way. Consider the matrix transformations $T_1(\mathbf{x}) = A_1\mathbf{x}$ and $T_2(\mathbf{x}) = A_2\mathbf{x}$. The composite transformation $T = T_2 \circ T_1$ is given by

$$T(\mathbf{x}) = T_2(T_1(\mathbf{x})) = T_2(A_1\mathbf{x}) = A_2A_1\mathbf{x}$$

Thus T is defined by the matrix product A_2A_1.

$$T(\mathbf{x}) = A_2A_1\mathbf{x}$$



The composite transformation T of T_1 and T_2. $T(\mathbf{x}) = T_2(T_1(\mathbf{x}))$.

Figure 2.15

We can extend the results of this discussion in a natural way. Let $T_1, \ldots, T_n$, be a sequence of transformations defined by matrices $A_1, \ldots, A_n$. The composite transformation $T = T_n \circ \cdots \circ T_1$ is defined by the matrix product $A_n, \ldots, A_1$ (assuming this product exists).

The following example illustrates how matrix transformations such as rotations and dilations can be used as building blocks to construct more intricate transformations. (Movement in a video game, for example, is accomplished by using a sequence of such transformations.)

EXAMPLE 2 Determine the single matrix that describes a reflection in the x-axis, followed by a rotation through $\pi/2$ followed by a dilation of factor 3. Find the image of the point $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ under this sequence of mappings.

SOLUTION

The matrices that define the reflection, rotation, and dilation are

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} \cos(\frac{\pi}{2}) & -\sin(\frac{\pi}{2}) \\ \sin(\frac{\pi}{2}) & \cos(\frac{\pi}{2}) \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

A reflection F followed by a rotation R and then a dilation D is the composite transformation $D \circ R \circ F$. The matrix of this transformation is

$$\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$$

The image of the point $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$.

The reflection followed by the rotation and then dilation maps $A \mapsto B$, $B \mapsto C$, $C \mapsto D$ in Figure 2.16. The composite transformation described by the single matrix maps $A \mapsto D$ directly.

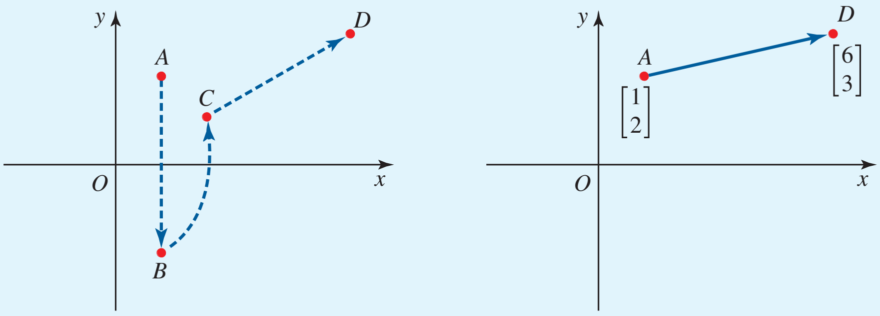


Figure 2.16

The next class of transformations are important in that they *preserve the geometry* of Euclidean Space. (Section 6 of the "Linear Equations, Vectors, and Matrices" chapter is a prerequisite.)

Orthogonal Transformation

An *orthogonal matrix* A is an invertible matrix that has the property

$$A^{-1} = A^t$$

An *orthogonal transformation* is a transformation $T(\mathbf{u}) = A\mathbf{u}$ where A is an orthogonal matrix.

An orthogonal transformation has the following geometrical properties.

THEOREM 2.9

Let T be an orthogonal transformation on $\mathbf{R}^n$. Let $\mathbf{u}$ and $\mathbf{v}$ be elements of $\mathbf{R}^n$. Let P and Q be the points in $\mathbf{R}^n$ defined by $\mathbf{u}$ and $\mathbf{v}$ and let R and S be their images under T. Then

$$\|\mathbf{u}\| = \|T(\mathbf{u})\|$$
$$\text{angle between } \mathbf{u} \text{ and } \mathbf{v} = \text{angle between } T(\mathbf{u}) \text{ and } T(\mathbf{v})$$
$$d(P, Q) = d(R, S)$$

Orthogonal transformations preserve norms, angles, and distances.

See Figure 2.17.

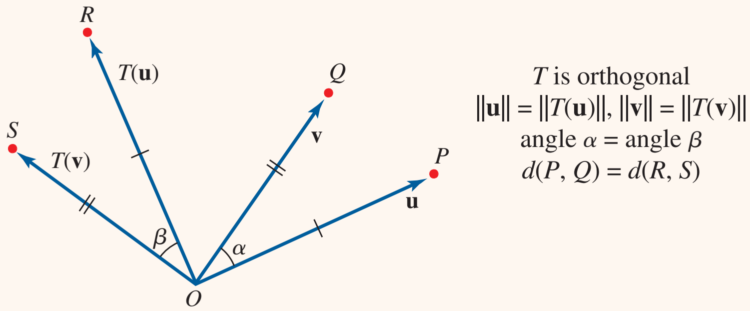


Figure 2.17

Orthogonal transformations preserve the shapes of rigid bodies and are often referred to as *rigid motions.*

Proof We first show that orthogonal transformations preserve dot products, $T(\mathbf{u})\cdot T(\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}$. Since norms, angles, and distances are defined in terms of dot products, this leads to their preservation. Let the orthogonal transformation T be defined by an orthogonal matrix A. We get

$$\begin{aligned} T(\mathbf{u})\cdot T(\mathbf{v}) &= (A\mathbf{u})\cdot(A\mathbf{v}) = (A\mathbf{u})^t(A\mathbf{v}) = (\mathbf{u}^tA^t)(A\mathbf{v}) \\ &= \mathbf{u}^tA^tA\mathbf{v} = \mathbf{u}^tI\mathbf{v} = \mathbf{u}^t\mathbf{v} = \mathbf{u}\cdot\mathbf{v} \end{aligned}$$

Thus $T(\mathbf{u})\cdot T(\mathbf{v}) = \mathbf{u}\cdot\mathbf{v}$; orthogonal transformations preserve dot products.

We now look at norms.

$$\begin{aligned} \|T(\mathbf{u})\| &= \|A\mathbf{u}\| = \sqrt{(A\mathbf{u})\cdot(A\mathbf{u})} \\ &= \sqrt{\mathbf{u}\cdot\mathbf{u}}, \text{ since dot product is preserved} \\ &= \|\mathbf{u}\| \end{aligned}$$

Thus norm is preserved.

We leave it for the reader to show that orthogonal transformations also preserve angles and distances in Exercise 21.

EXAMPLE 3 Let T be the orthogonal transformation defined by the following orthogonal matrix A. Show that T preserves norms, angles, and distances for the vectors $\mathbf{u}$ and $\mathbf{v}$.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

SOLUTION

We have that

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} \sqrt{2} \\ -\sqrt{2} \end{bmatrix} \text{ and } T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} \frac{7}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

It can be shown that $\|\mathbf{u}\| = \|T(\mathbf{u})\| = 2$ and $\|\mathbf{v}\| = \|T(\mathbf{v})\| = 5$. Norms of $\mathbf{u}$ and $\mathbf{v}$ are preserved. The angle between $\mathbf{u}$ and $\mathbf{v}$ and the angle between $T(\mathbf{u})$ and $T(\mathbf{v})$ are both found to be 53.13°. The angle is preserved.

Furthermore,

$$d\left(\begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}3\\4\end{bmatrix}\right) = d\left(\begin{bmatrix}\sqrt{2}\\-\sqrt{2}\end{bmatrix}, \begin{bmatrix}\frac{7}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix}\right) = \sqrt{17}.$$

Distance is preserved.

Observe that T defines a rotation of points in a plane through an angle of $\pi/4$ in a clockwise direction about the origin. Intuitively, we expect rotations to preserve norms, angles, and distances. Rotation matrices do in fact define orthogonal transformations (see Exercise 20).

We complete this section with a discussion of transformations that, even though they are not truly matrix transformations, are important in mathematics and in applications.

Translation

A *translation* is a transformation $T\colon \mathbf{R}^n \to \mathbf{R}^n$ defined by

$$T(\mathbf{u}) = \mathbf{u} + \mathbf{v},$$

where $\mathbf{v}$ is a fixed vector.

A translation slides points in a direction and distance defined by the vector $\mathbf{v}$. For example, consider the following translation on $\mathbf{R}^2$:

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}x\\y\end{bmatrix} + \begin{bmatrix}4\\2\end{bmatrix}$$

Let us determine the effect of T on the triangle PQR having vertices $\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}2\\8\end{bmatrix}, \begin{bmatrix}3\\2\end{bmatrix}$. We see that

$$\overset{P}{\begin{bmatrix}1\\2\end{bmatrix}} \mapsto \overset{P'}{\begin{bmatrix}5\\4\end{bmatrix}},\ \overset{Q}{\begin{bmatrix}2\\8\end{bmatrix}} \mapsto \overset{Q'}{\begin{bmatrix}6\\10\end{bmatrix}},\ \overset{R}{\begin{bmatrix}3\\2\end{bmatrix}} \mapsto \overset{R'}{\begin{bmatrix}7\\4\end{bmatrix}}$$

The triangle PQR is transformed into the triangle $P'Q'R'$ in Figure 2.18.

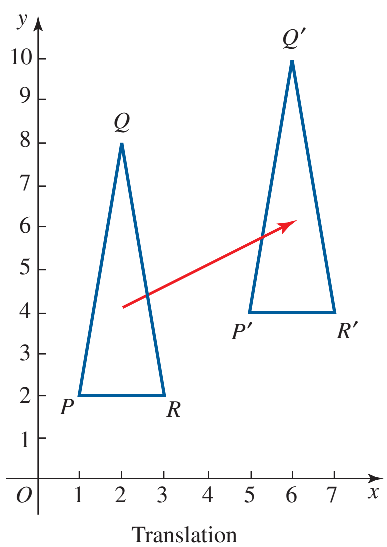


Figure 2.18

EXAMPLE 4 Find the equation of the image of the line $y = 2x + 3$ under the translation

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

SOLUTION

The equation $y = 2x + 3$ describes points on the line of slope 2 and y-intercept 3. T will slide this line into another line. We want to find the equation of this image line. We get

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ 2x + 3 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} x + 2 \\ 2x + 4 \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

We see that $y' = 2x'$ for the image point. Thus the equation of the image line is $y = 2x$.

Affine Transformation

An *affine transformation* is a transformation $T\colon \mathbf{R}^n \to \mathbf{R}^n$ defined by

$$T(\mathbf{u}) = A\mathbf{u} + \mathbf{v}$$

where A is a matrix and $\mathbf{v}$ is a fixed vector.

An affine transformation can be interpreted as a matrix transformation followed by a translation.

For example, consider the following affine transformation on $\mathbf{R}^2$.

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Let us find the image of the unit square in Figure 2.19. We get

$$\overset{P}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} \mapsto \overset{P'}{\begin{bmatrix} 3 \\ 3 \end{bmatrix}}, \overset{Q}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}} \mapsto \overset{Q'}{\begin{bmatrix} 4 \\ 4 \end{bmatrix}}, \overset{R}{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} \mapsto \overset{R'}{\begin{bmatrix} 2 \\ 3 \end{bmatrix}}, \overset{O}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}} \mapsto \overset{O'}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}$$

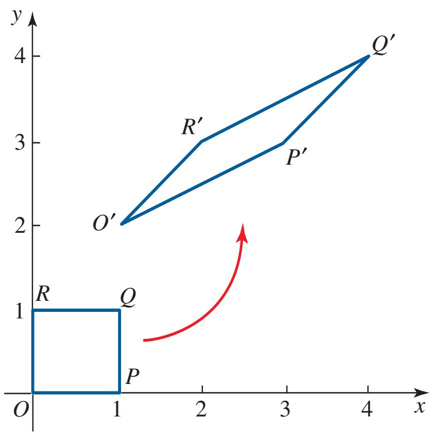


Figure 2.19

Line segments are mapped into line segments. We get

$$OP \to O'P',\ PQ \to P'Q',\ QR \to Q'R',\ OR \to O'R'$$

The square PQR is transformed into the parallelogram $P'Q'R'O'$.

We have introduced a number of fundamental transformations in this section. The reader will meet other transformations—namely *projection*, *scaling,* and *shear*—in the following section. These basic transformations are the building blocks for creating other transformations using composition.

EXERCISE SET 2.5

Matrix Transformations

1. Consider the transformation T defined by the following matrix A. Find $T(\mathbf{x})$, $T(\mathbf{y})$, and $T(\mathbf{z})$.

$$A = \begin{bmatrix} 0 & -4 \\ 1 & 2 \end{bmatrix};\ \mathbf{x} = \begin{bmatrix} 5 \\ -2 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} 4 \\ 0 \end{bmatrix},\ \mathbf{z} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

2. Consider the transformation T defined by the following matrix A. Determine the images of the vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$.

$$A = \begin{bmatrix} 3 & -2 & 0 \\ 4 & 2 & 6 \end{bmatrix};\ \mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix},\ \mathbf{z} = \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$$

3. The following matrix A defines a transformation T. Find the images of the vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$.

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \\ 1 & 2 \end{bmatrix};\ \mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix},\ \mathbf{z} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

Dilations, Reflections, Rotations

4. Determine the matrix that defines a reflection in the y-axis. Find the image of $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ under this transformation.

5. Find the matrix that defines a rotation of a plane about the origin through each of the following angles. Determine the image of the point $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ under each transformation.

(a) $\frac{\pi}{2}$ **(b)** $-\frac{\pi}{2}$ **(c)** $\frac{\pi}{4}$

(d) π **(e)** $-\frac{3\pi}{2}$ **(f)** $\frac{\pi}{6}$

(g) $-\frac{\pi}{3}$

6. Find the equation of the image of the unit circle, $x^2 + y^2 = 1$, under a dilation of factor 3.

7. Find the equation of the image of the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = 1$ under a rotation through an angle of $\pi/2$.

Geometry

8. Consider the transformations on $\mathbf{R}^2$ defined by each of the following matrices. Sketch the image of the unit square under each transformation.

(a) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

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

9. Sketch the image of the unit square under the transformation defined by each of the following transformations.

(a) $\begin{bmatrix} -2 & -3 \\ 0 & 4 \end{bmatrix}$ **(b)** $\begin{bmatrix} -2 & -4 \\ -4 & -1 \end{bmatrix}$

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

Composition of Transformations

10. Let $T_1(\mathbf{x}) = A_1\mathbf{x}$ and $T_2(\mathbf{x}) = A_2\mathbf{x}$ be defined by the following matrices A_1 and A_2. Let $T = T_2 \circ T_1$. Find the matrix that defines T and use it to determine the image of the vector $\mathbf{x}$ under T.

(a) $A_1 = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix},\ A_2 = \begin{bmatrix} -1 & 0 \\ 1 & 5 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$

(b) $A_1 = \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & -1 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & 2 \\ 1 & -1 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$

(c) $A_1 = \begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & 2 \\ 1 & -1 \\ 0 & 4 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$

11. Let $T_1(\mathbf{x}) = A_1\mathbf{x}$ and $T_2(\mathbf{x}) = A_2\mathbf{x}$ be defined by the following matrices A_1 and A_2. Let $T = T_2 \circ T_1$. Find the matrix that defines T and use it to determine the image of the vector $\mathbf{x}$ under T.

(a) $A_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & 3 \\ 0 & -4 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

(b) $A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 3 & -2 & 5 \end{bmatrix}$, $A_2 = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$

(c) $A_1 = \begin{bmatrix} 5 & -2 \\ 3 & 6 \end{bmatrix}$, $A_2 = \begin{bmatrix} 3 & 0 \\ 1 & -7 \\ 2 & 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$

12. Construct single 2×2 matrices that define the following transformations on $\mathbf{R}^2$. Find the image of the point $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ under each transformation.

(a) A rotation through $\pi/2$ counterclockwise, then a contraction of factor 0.5.

(b) A dilation of factor of 4, then a reflection in the x-axis.

(c) A reflection about the x-axis, a dilation of factor 3, then a rotation through $\pi/2$ in a clockwise direction.

13. Find a single matrix that defines a counterclockwise rotation of the plane through an angle of $\pi/2$ about the origin, while at the same time moves points to twice their original distance from the origin.

14. Determine a single matrix that defines both a rotation about the origin through an angle θ and a dilation of factor r.

15. Let A be the rotation matrix for $\pi/4$. Show that $A^8 = I$, the identity 2×2 matrix. Give a geometrical reason for expecting this result.

16. Show that $A^2BA^2 = I$ for the following matrices A and B. Give a geometrical reason for expecting this result.

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

17. Determine a rotation through an angle θ followed by a dilation of factor r. Show both algebraically and geometrically that this is equivalent to the dilation followed by the rotation. We say that the rotation and dilation transformations are *commutative*.

18. Let T_1 be a rotation and T_2 be a reflection in the x-axis. Are these transformations commutative? Discuss both algebraically and geometrically.

Orthogonal Transformations

19. Show that the following matrix A is orthogonal. Show that the transformation defined by A preserves the norms of the vectors $\mathbf{u}$ and $\mathbf{v}$, preserves the angle between these vectors, and also preserves the distance between the points defined by the vectors.

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

20. Prove that a rotation matrix is an orthogonal matrix.

21. Let A be an $n \times n$ orthogonal matrix. A defines a transformation of $\mathbf{R}^n$ into itself. Let $\mathbf{u}$ and $\mathbf{v}$ be elements of $\mathbf{R}^n$. Let P and Q be the points defined by $\mathbf{u}$ and $\mathbf{v}$, and let R and S be the points defined by $A\mathbf{u}$ and $A\mathbf{v}$. See Figure 2.18. Prove that

(a) the angle between $\mathbf{u}$ and $\mathbf{v}$ is equal to the angle between $A\mathbf{u}$ and $A\mathbf{v}$.

(b) the distance between P and Q is equal to the distance between R and S.

Thus the transformation defined by A preserves angles and distances.

Translations and Affine Transformations

22. Find the image of the triangle having vertices $(1, 2)$, $(3, 4)$, and $(4, 6)$ under the translation that takes the point $(1, 2)$ to $(2, -3)$.

23. Find the image of the line $y = 3x + 1$ under the translation

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} p \\ q \end{bmatrix}$$

where **(a)** $p = 2, q = 5$ **(b)** $p = -1, q = 1$.

24. Find and sketch the image of the unit square and the unit circle under the affine transformations $T(\mathbf{u}) = A\mathbf{u} + \mathbf{v}$ defined by the following matrices and vectors.

(a) $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$

(b) $A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

(c) $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

(d) $A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

2.6 Linear Transformations, Graphics, and Fractals

Let us now examine the *properties* of a matrix transformation T. We know that a vector space has two operations, namely addition and scalar multiplication. Let us look at how T interacts with these operations. Consider the matrix transformation $T(\mathbf{u}) = A(\mathbf{u})$. The matrix properties of A imply that

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \\ &= T(\mathbf{u}) + T(\mathbf{v}) \end{aligned}$$

and

$$\begin{aligned} T(c\mathbf{u}) &= A(c\mathbf{u}) = cA\mathbf{u} \\ &= cT(\mathbf{u}) \end{aligned}$$

The implication is that T maps the sum of two vectors into the sum of the images (preserves addition) and maps the scalar multiple of a vector into that same scalar multiple of the image (preserves scalar multiplication). We say that T *preserves vector space structure*. We call any transformation that has these properties a linear transformation.

DEFINITION Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbf{R}^n$ and let c be a scalar. A transformation $T\colon \mathbf{R}^n \to \mathbf{R}^m$ is said to be a *linear transformation* if

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T(\mathbf{u}) + T(\mathbf{v}) \quad &\text{(preserves addition)} \\ T(c\mathbf{u}) &= cT(\mathbf{u}) \quad &\text{(preserves scalar multiplication)} \end{aligned}$$

Every matrix transformation is linear. Since dilations, contractions, reflections, and rotations can all be described by matrices, these transformations are linear. Orthogonal transformations are also linear, but translations and affine transformations are not linear (see Exercise 39).

The structure-preserving ideas of a linear transformation are illustrated in Figure 2.20.

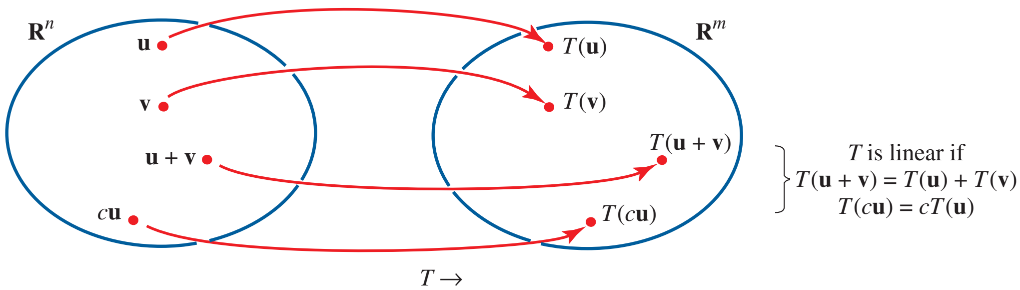


Figure 2.20

We now give examples to show how the linearity conditions are in general checked.

EXAMPLE 1 Prove that the following transformation T is linear.

$$T(x, y) = (x - y, 3x)$$

SOLUTION

T maps $\mathbf{R}^2 \to \mathbf{R}^2$. For example, $T(5, 1) = (4, 15)$. The image of $(5, 1)$ is $(4, 15)$.

We first show that T preserves addition. Let (x_1, y_1) and (x_2, y_2) be elements of $\mathbf{R}^2$. Then

$$\begin{aligned} T((x_1, y_1) + (x_2, y_2)) &= T(x_1 + x_2, y_1 + y_2) && \text{by vector addition} \\ &= (x_1 + x_2 - y_1 - y_2, 3x_1 + 3x_2) && \text{by definition of } T \\ &= (x_1 - y_1, 3x_1) + (x_2 - y_2, 3x_2) && \text{by vector addition} \\ &= T(x_1, y_1) + T(x_2, y_2) && \text{by definition of } T \end{aligned}$$

Thus T preserves vector addition.

We now show that T preserves scalar multiplication. Let c be a scalar.

$$\begin{aligned} T(c(x_1, y_1)) &= T(cx_1, cy_1) && \text{by scalar multiplication of a vector} \\ &= (cx_1 - cy_1, 3cx_1) && \text{by definition of } T \\ &= c(x_1 - y_1, 3x_1) && \text{by scalar multiplication of a vector} \\ &= cT(x_1, y_1) && \text{by definition of } T \end{aligned}$$

Thus T preserves scalar multiplication. T is linear.

Note that T in Example 1 maps $\mathbf{R}^2$ into $\mathbf{R}^2$. The domain and codomain are the same. A transformation for which the domain and codomain are the same is often referred to as an *operator*.

The following example illustrates a transformation that is not linear.

EXAMPLE 2 Show that the following transformation $T: \mathbf{R}^3 \to \mathbf{R}^2$ is not linear.

$$T(x, y, z) = (xy, z)$$

SOLUTION

Let us first test addition. Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be elements of $\mathbf{R}^3$. Then

$$\begin{aligned} T((x_1, y_1, z_1) + (x_2, y_2, z_2)) &= T(x_1 + x_2, y_1 + y_2, z_1 + z_2) \\ &= ((x_1 + x_2)(y_1 + y_2), z_1 + z_2) \\ &= (x_1y_1 + x_2y_2 + x_1y_2 + x_2y_1, z_1 + z_2) \end{aligned}$$

and

$$\begin{aligned} T(x_1, y_1, z_1) + T(x_2, y_2, z_2) &= (x_1y_1, z_1) + (x_2y_2, z_2) \\ &= (x_1y_1 + x_2y_2, z_1 + z_2) \end{aligned}$$

Thus, in general,

$$T((x_1, y_1, z_1) + (x_2, y_2, z_2)) \neq T(x_1, y_1, z_1) + T(x_2, y_2, z_2)$$

Since vector addition is not preserved, T is not linear.

(It is not necessary to check the second linearity condition. The fact that one condition is not satisfied is sufficient to prove that T is not linear. It can be shown, in fact, that this particular transformation does not preserve scalar multiplication either.)

In the previous section we used ad hoc ways of arriving at matrices that described certain transformations such as rotations, dilations, and reflections. We now introduce a method for constructing a matrix representation for any linear transformation on $\mathbf{R}^n$. We pave the way with the following example.

EXAMPLE 3 Determine a matrix A that describes the linear transformation

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ 3y \end{bmatrix}$$

SOLUTION

It can be shown that T is linear. The domain of T is $\mathbf{R}^2$. We find the effect of T on the standard basis of $\mathbf{R}^2$.

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

These vectors will be the columns of the matrix A that describe the transformation. We get

$$A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$$

T can be written as a matrix transformation,

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

(We can check that this matrix does work: $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + y \\ 3y \end{bmatrix}$.)

We now arrive at the general result: We see *why* the above method works.

Matrix Representation

Let T be a linear transformation on $\mathbf{R}^n$. Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be the standard basis of $\mathbf{R}^n$ and $\mathbf{u}$ be an arbitrary vector in $\mathbf{R}^n$, written in column form.

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \text{ and } \mathbf{u} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

We can express $\mathbf{u}$ in terms of $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

$$\mathbf{u} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$$

Since T is a linear transformation

$$\begin{aligned} T(\mathbf{u}) &= T(c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n) \\ &= c_1T(\mathbf{e}_1) + \cdots + c_nT(\mathbf{e}_n) \quad \text{(see Exercise 36)} \\ &= [T(\mathbf{e}_1) \ldots T(\mathbf{e}_n)]\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \end{aligned}$$

where $[T(\mathbf{e}_1) \ldots\ T(\mathbf{e}_n)]$ is a matrix with columns $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$.

Thus the linear transformation T is defined by the matrix

$$A = [T(\mathbf{e}_1) \ldots\ T(\mathbf{e}_n)]$$

A is called the *standard matrix* of T.

In this discussion we see the importance of a basis for working with vector spaces. We mentioned earlier how a basis, in a sense, represents the whole space. Here we see that the effect of a linear transformation on a basis leads to a matrix representation, a representation of the transformation on the whole space.

We derived the matrix for a rotation in the last section using an ad hoc method. You are asked to confirm this matrix in this standard manner, using bases, in the exercises that follow. We now use this method to derive the matrix of a reflection in the line $y = -x$.

EXAMPLE 4 The transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ -x \end{bmatrix}$ defines a reflection in the line $y = -x$, Figure 2.21(a). It can be shown that T is linear. Determine the standard matrix of this transformation. Find the image of $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

SOLUTION

We find the effect of T on the standard basis.

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

The standard matrix is thus

$$A = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

The transformation can be written

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Applying the transformation to the point $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$, we get $T\left(\begin{bmatrix} 4 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -4 \end{bmatrix}$.

See Figure 2.21(b).

We now discuss the use of these transformations in computer graphics.

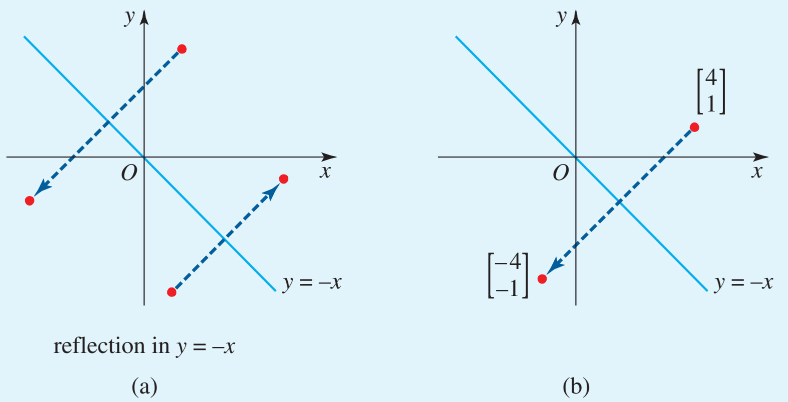


Figure 2.21

Transformations in Computer Graphics

Computer graphics is the field that studies the creation and manipulation of pictures with the aid of computers. The impact of computer graphics is felt in many homes through video games; its uses in research, industry, and business are vast and are ever expanding. Architects use computer graphics to explore designs, molecular biologists display and manipulate pictures of molecules to gain insight into their structure, pilots are trained using graphics flight simulators, and transportation engineers use computer-generated transforms in their planning work—to mention a few applications.

The manipulation of pictures in computer graphics is carried out using sequences of transformations. Rotations, reflections, dilations, and contractions are defined by matrices. A sequence of such transformations can be performed by a single transformation defined by the product of the matrices. Unfortunately, translation, as it now stands, uses matrix addition, and any sequence of transformations involving translations cannot be combined in this manner into a single matrix. However, if coordinates called **homogeneous coordinates** are used to describe points in a plane, then translations can also be accomplished through matrix multiplication, and any sequence of these transformations can be defined in terms of a single matrix. In homogeneous coordinates, a third component of 1 is added to each coordinate, and rotation, reflection, dilation/contraction, and translation R, F, D, and T are defined by the following matrices.

$$\underset{\text{point}}{\overset{\mathrm{X}}{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}} \quad A = \underset{\text{rotation}}{\overset{R}{\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}}} \; B = \underset{\text{reflection}}{\overset{F}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}}$$

$$C = \underset{\substack{\text{dilation/contraction} \\ (r > 0)}}{\overset{D}{\begin{bmatrix} r & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{bmatrix}}} \quad E = \underset{\text{translation}}{\overset{T}{\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}}}$$

Thus, for example, a dilation D followed by a translation T and then a rotation R would be defined by $R \circ T \circ D(\mathrm{X}) = AEC(\mathrm{X})$. The composite transformation $R \circ T \circ D$ would be described by the single matrix AEC.

Some programming languages provide subroutines for rotation, translation, and dilation/contraction (and also scale and shear, see Exercises 19 and 22 following) that can be used to move pictures on the screen. To accomplish this movement, the subroutines convert screen coordinates into homogeneous coordinates and use the matrices that define these transformations in homogeneous coordinates.

We now illustrate how the transformations are used to rotate a geometrical figure about a point other than the origin.

EXAMPLE 5 Determine the matrix that defines a rotation of points in a plane through an angle θ about a point $P(h, k)$. Use this general result to find the matrix that defines a rotation of the points through an angle of $\pi/2$ about the point $(5, 4)$. Find the image of the triangle having the following vertices $A(1, 2)$, $B(2, 8)$, and $C(3, 2)$ under this rotation. See Figure 2.22.

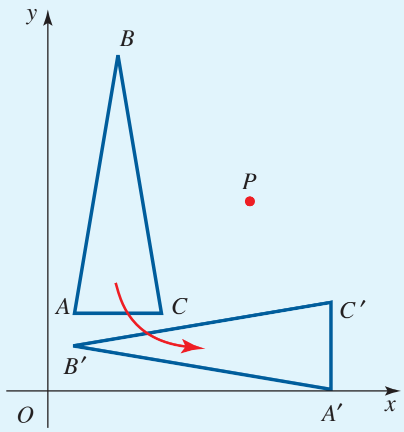


Figure 2.22

SOLUTION

The rotation about P can be accomplished by a sequence of three of the above transformations:

(a) A translation T_1 that takes P to the origin O.
(b) A rotation R about the origin through an angle θ.
(c) A translation T_2 that takes O back to P.

The matrices that describe these transformations are as follows.

$$\underset{T_1}{\begin{bmatrix} 1 & 0 & -h \\ 0 & 1 & -k \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{R}{\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{T_2}{\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}}$$

The rotation R_p about P can be accomplished as follows.

$$R_p\left(\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}\right) = T_2 \circ R \circ T_1\left(\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} 1 & 0 & -h \\ 0 & 1 & -k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta & -h\cos\theta + k\sin\theta + h \\ \sin\theta & \cos\theta & -h\sin\theta - k\cos\theta + k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

To get the specific matrix that defines the rotation of the plane through an angle $\pi/2$ about the point $P(5, 4)$, for example, let $h = 5$, $k = 4$, and $\theta = \pi/2$. The rotation matrix is

$$M = \begin{bmatrix} 0 & -1 & 9 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

To find the images of the vertices of the triangle ABC, write these vertices in column form as homogeneous coordinates and multiply by M. On performing the matrix multiplications, we get

$$\overset{A}{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}} \mapsto \overset{A'}{\begin{bmatrix} 7 \\ 0 \\ 1 \end{bmatrix}} \qquad \overset{B}{\begin{bmatrix} 2 \\ 8 \\ 1 \end{bmatrix}} \mapsto \overset{B'}{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}} \qquad \overset{C}{\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}} \mapsto \overset{C'}{\begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}}$$

The triangle with vertices $A(1, 2)$, $B(2, 8)$, $C(3, 2)$ is transformed into the triangle with vertices $A'(7, 0)$, $B'(1, 1)$, $C'(7, 2)$. See Figure 2.22.

Fractal Pictures of Nature

Computer graphics systems based on traditional Euclidean geometry are suitable for creating pictures of manmade objects such as machinery, buildings, and airplanes. Images of such objects can be created using lines, circles, and so on. However, these techniques are not appropriate when it comes to constructing images of natural objects such as animals, trees, and landscapes. In the words of mathematician Benoit B. Mandelbrot, "Clouds are not spheres, mountains are not cones, coastlines are not circles, and bark is not smooth, nor does lightning travel in straight lines." However, nature does wear its irregularities in an unexpectedly orderly fashion; it is full of shapes that repeat themselves on different scales within the same object. In 1975 Mandelbrot introduced a new geometry, which he called **fractal geometry**, that can be used to describe natural phenomena. A **fractal** is a convenient label for irregular and fragmented self-similar shapes. Fractal objects contain structures nested within one another. Each smaller structure being a miniature, though not necessarily identical version, of the larger form. The story behind the word fractal is interesting.

Mandelbrot came across the Latin adjective *fractus,* from the verb *frangere,* to break, in his son's Latin book. The resonance of the main English cognates fracture and fraction seemed appropriate and he coined the word fractal!

We now discuss methods that have been developed by a research team at the Georgia Institute of Technology for forming images of natural objects using fractals. These fractal images of nature are generated using affine transformations. Figure 2.23 shows a fractal image of a fern being gradually generated. Let us see how this is done.

Figure 2.23

Consider the following four affine transformations $T_1, \ldots, T_4$. Associate probabilities $p_1, \ldots, p_4$ with these transformations.

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0.86 & 0.03 \\ -0.03 & 0.86 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 1.5 \end{bmatrix},\ p_1 = 0.83$$

$$T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0.2 & -0.25 \\ 0.21 & 0.23 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 1.5 \end{bmatrix},\ p_2 = 0.08$$

$$T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -0.15 & 0.27 \\ 0.25 & 0.26 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 0.45 \end{bmatrix},\ p_3 = 0.08$$

$$T_4\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0.17 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ p_4 = 0.01$$

The following algorithm is used on a computer to produce the image of the fern.

1. Let $x = 0$, $y = 0$.
2. Use a random generator to select one of the affine transformations T_i according to the given probabilities.
3. Let $(x', y') = T_i(x, y)$.
4. Plot (x', y').
5. Let $(x, y) = (x', y')$.
6. Repeat Steps 2, 3, 4, and 5 twenty thousand times.

As Step 4 is executed, each of twenty thousand times, the image of the fern gradually appears.

Each affine transformation T_i involves six parameters *a, b, c, d, e, f*, and a probability p_i, as follows

$$T_i\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix},\ p_i$$

The affine transformations and corresponding probabilities that generate a fractal are written as rows of a matrix, called an *iterated function system* (IFS). The IFS for the fern is as follows.

IFS for a fern

$$\begin{array}{cccccccc} T & a & b & c & d & e & f & p \end{array}$$
$$\begin{bmatrix} 1 & 0.86 & 0.03 & -0.03 & 0.86 & 0 & 1\ .5 & 0.83 \\ 2 & 0.2 & -0.25 & 0.21 & 0.23 & 0 & 1\ .5 & 0.08 \\ 3 & -0.15 & 0.27 & 0.25 & 0.26 & 0 & 0.45 & 0.08 \\ 4 & 0 & 0 & 0 & 0.17 & 0 & 0 & 0.01 \end{bmatrix}$$

The appropriate affine transformations that produce a given fractal object are found by determining transformations that map the object (called the **attractor**) into various disjoint images, the union of which is the whole fractal. A theorem called the **Collage Theorem** then guarantees that the transformations can be grouped into an IFS that produces the fractal.

Different probabilities do not in general lead to different images, but they do affect the rate at which the image is produced. Appropriate probabilities are

$$p_i = \frac{\text{area of the image under transformation } T_i}{\text{area of image of object}}$$

These techniques are very valuable because they can be used to produce an image to any desired degree of accuracy using a highly compressed data set. A fractal image containing infinitely many points, whose organization is too complicated to describe directly, can be reproduced using mathematical formulas.

EXERCISE SET 2.6

Check for Linearity

1. Prove that the transformation T: $\mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(x, y) = (2x, x - y)$ is linear. Find the images of the elements $(1, 2)$ and $(-1, 4)$ under this transformation.

2. Prove that $T(x, y) = (3x + y, 2y, x - y)$ defines a linear transformation T: $\mathbf{R}^2 \to \mathbf{R}^3$. Find the images of $(1, 2)$ and $(2, -5)$.

3. Prove that T: $\mathbf{R}^3 \to \mathbf{R}^3$ defined by $T(x, y, z) = (0, y, 0)$ is linear. This transformation is also called a projection. Why is this term appropriate?

4. Prove that the following transformations T: $\mathbf{R}^3 \to \mathbf{R}^2$ are not linear.
   (a) $T(x, y, z) = (3x, y^2)$
   (b) $T(x, y, z) = (x + 2, 4y)$

5. Prove that T: $\mathbf{R}^2 \to \mathbf{R}$ defined by $T(x, y) = x + a$, where a is a nonzero scalar, is not linear.

In Exercises 6–11, determine whether the given transformations are linear.

6. $T(x, y, z) = (2x, y)$ of $\mathbf{R}^3 \to \mathbf{R}^2$.

7. $T(x, y) = x - y$ of $\mathbf{R}^2 \to \mathbf{R}$.

8. $T(x, y) = (x, y, z)$ of $\mathbf{R}^2 \to \mathbf{R}^3$, when **(a)** $z = 0$, **(b)** $z = 1$.

9. $T(x) = (x, 2x, 3x)$ of $\mathbf{R} \to \mathbf{R}^3$.

10. $T(x, y) = (x^2, y)$ of $\mathbf{R}^2 \to \mathbf{R}^2$.

11. $T(x, y, z) = (x + 2y, x + y + z, 3z)$ of $\mathbf{R}^3 \to \mathbf{R}^3$.

Standard Matrix of a Linear Transformation

12. Find the standard matrix of each of the following linear transformations on $\mathbf{R}^2$.

(a) $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ x - y \end{bmatrix}$ **(b)** $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - y \\ x + y \end{bmatrix}$

(c) $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - 5y \\ 3y \end{bmatrix}$ **(d)** $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2y \\ -3x \end{bmatrix}$

13. Find the standard matrix A of each of the following linear transformations. Verify your answers by computing $A\begin{bmatrix} x \\ y \end{bmatrix}$ for (a) and (b), $A\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ for (c) and (d).

(a) $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ x + y \end{bmatrix}$ **(b)** $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 2x \\ 3x \end{bmatrix}$

(c) $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x \\ x + y \end{bmatrix}$ **(d)** $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + 2z \\ y + 3z \end{bmatrix}$

14. Derive the rotation matrix by finding the effect of a rotation on the standard basis of $\mathbf{R}^2$.

15. We have seen examples of rotations, dilations, contractions, and reflections. How would you describe the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}$? Show that T is linear. Determine the standard matrix of T. Find the image of the point $\begin{bmatrix} -3 \\ 5 \end{bmatrix}$.

Projection

16. Find the standard matrix of the operator $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 0 \end{bmatrix}$.

Observe that this transformation projects all points onto the x-axis. It is called a *projection operator*.

17. Determine the matrix that defines projection onto the y-axis.

18. Determine the matrix that defines projection onto the line $y = x$. Find the image of $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ under this projection.

Scaling

19. Find the standard matrix of the operator $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} ax \\ by \end{bmatrix}$,

where a and b are positive scalars. T is called a *scaling of factor a in the x-direction and factor b in the y-direction*. A scaling distorts a figure, since x and y do not change in the same manner. Sketch the image of the unit square under this transformation when $a = 3$ and $b = 2$.

20. Find the equation of the image of the line $y = 2x$ under a scaling of factor 2 in the x-direction and factor 3 in the y-direction.

21. Find the equation of the image of the unit circle, $x^2 + y^2 = 1$, under a scaling of factor 4 in the x-direction and factor 3 in the y-direction.

Shear

22. Find the standard matrix of the operator $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + cy \\ y \end{bmatrix}$, where c is a scalar. T is called a *shear of factor c in the x-direction.* Sketch the image of the unit square under a shear of factor 2 in the x-direction. Observe how the x-value of each point is increased by a factor of $2y$, causing a shearing of the figure.

23. Sketch the image of the unit square under a shear of factor 0.5 in the y-direction.

24. Find the equation of the image of the line $y = 3x$ under a shear of factor 5 in the x-direction.

General Matrix Transformations

25. Construct single 2×2 matrices that define the following transformations on $\mathbf{R}^2$. Find the image of the point $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ under each transformation.

(a) A dilation of factor 3, then a shear of factor 2 in the x-direction.

(b) A scaling of factor 3 in the x-direction, of factor 2 in the y-direction, then a reflection in the line $y = x$.

(c) A dilation of factor 2, then a shear of factor 3 in the x-direction, then a rotation through $\pi/2$ counterclockwise.

26. Find the matrix that maps $\mathbf{R}^2 \to \mathbf{R}^2$ such that $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \mapsto \begin{bmatrix} 7 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

27. Transformations T_1 and T_2 are said to be *commutative* if $T_2 \circ T_1(\mathbf{u}) = T_1 \circ T_2(\mathbf{u})$ for all vectors $\mathbf{u}$. Let R be a rotation, D dilation, F reflection, S scaling, H shear, and A affine transformation. Which pairs of transformations are commutative?

28. Find the matrix that defines a rotation of three-space through an angle of $\pi/2$ about the z-axis. (You may consider either direction.)

29. Find the matrix that defines an expansion of three-dimensional space outward from the origin, so that each point moves to three times as far away.

30. Determine the matrix that can be used to define a rotation through $\pi/2$ about the point $(5, 1)$. Find the image of the unit square under this rotation.

31. Consider the general translation defined by the following matrix T. Does this transformation have an inverse? If so, find it.

$$T = \begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}$$

32. Consider the general scaling defined by the following matrix $S(c \neq 0, d \neq 0)$. Does this transformation have an inverse? If so find it.

$$S = \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

33. Find the image of the triangle having the following vertices A, B, and C (in homogeneous coordinates), under the sequence of transformations T followed by R, followed by S. Sketch the original and final triangle.

$$A\begin{bmatrix} 1 \\ 6 \\ 1 \end{bmatrix}, B\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, C\begin{bmatrix} 4 \\ 6 \\ 1 \end{bmatrix};$$

$$T = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}, R = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$S = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Miscellaneous Results

34. Let T be a linear transformation. Use the fact that T preserves addition and scalar multiplication to show that

(a) $T(-\mathbf{v}) = -T(\mathbf{v})$

(b) $T(\mathbf{v} - \mathbf{w}) = T(\mathbf{v}) - T(\mathbf{w})$

35. Let T be a transformation between vector spaces, $\mathbf{u}$ and $\mathbf{v}$ vectors in the domain, and a and b scalars. Prove that T is linear if and only if

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

This can be used as an alternative definition of linear transformation.

36. Let T be a linear transformation with domain U. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors in U, and $c_1, \ldots, c_m$ be scalars. Prove that

$$T(c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m) = c_1T(\mathbf{v}_1) + \cdots + c_mT(\mathbf{v}_m)$$

(Can you prove this result by induction?)

37. Prove that the composition of two linear transformations is a linear transformation.

38. State (with a brief explanation) whether the following statements are true or false.

(a) $T(x, y) = (x, y)$ is a linear transformation.

(b) Let $T: U \to V$ be a linear transformation. Distinct vectors in U always have distinct images in V.

(c) If the dimension of V is greater than that of U, there are no linear transformations from U to V.

39. Prove that translations and affine transformations are not linear.

*2.7 The Leontief Input-Output Model in Economics

In this section we introduce the Leontief model that is used to analyze the interdependence of economies. The importance of this model to current economic planning was mentioned in the introduction to this chapter.

Consider an economic situation that involves n interdependent industries. The output of any one industry is needed as input by other industries, and even possibly by the industry itself. We shall see how a mathematical model involving a system of linear equations can be constructed to analyze such a situation. Let us assume, for the sake of simplicity, that each industry produces one commodity. Let a_{ij} denote the amount of input of a certain commodity i to produce unit output of commodity j. In our model let the amounts of input and output be measured in dollars. Thus, for example, $a_{34} = 0.45$ means that 45 cents' worth of commodity 3 is required to produce one dollar's worth of commodity 4.

$$a_{ij} = \text{amount of commodity } i \text{ in one dollar of commodity } j$$

The elements a_{ij}, called **input coefficients**, define a matrix A called the **input-output matrix**, which describes the interdependence of the industries.

* Sections and chapters marked with an asterisk are optional. The instructor can use these sections to build around the core material to give the course the desired flavor.

EXAMPLE 1 National input-output matrices are used to describe interindustry relations that constitute the economic fabric of countries. We now display part of the matrix that describes the interdependency of the U.S. economy for 1972. The economic structure is actually described in terms of the flow among 79 producing sectors; the matrix is thus a 79×79 matrix. We cannot, of course, display the whole matrix. We list 10 sectors to give the reader a feel for the catagories involved.

1. Livestock and livestock products
2. Agricultural crops
3. Forestry and fishery products
4. Agricultural, forestry, and fishery services
5. Iron and ferroalloy ores mining
6. Nonferrous metal ores mining
7. Coal mining
8. Crude petroleum and natural gas
9. Stone and clay mining and quarrying
10. Chemical and fertilizer mineral mining

The matrix A based on these sectors is

	1	**2**	**3**	**4 . . .**
1	0.26110	0.02481	0	0.05278 . . .
2	0.23277	0.03218	0	0.01444
3	0	0	0.00467	0.00294
4	0.02821	0.03673	0.02502	0.02959
5	0	0	0	0
6	0	0	0	0
7	0	0.00002	0	0
8	0	0	0	0
9	0.00001	0.00251	0	0.00034
10	0	0.00130	0	0
⋮	⋮			

Thus, for example, $a_{72} = 0.00002$ implies that \$0.00002 from the coal mining sector (sector 7) goes into producing each \$1 from the the agricultural crops sector (sector 2).

We now extend the model to include an open sector. The products of industries may go not only into other producing industries, but also into other nonproducing sectors of the economy such as consumers and governments. All such nonproducing sectors are grouped into what is called the **open sector**. The open sector in the above model of the 1972 U.S. economy included, for example, federal, state, and local government purchases. Let

$d_i =$ demand of the open sector from industry i.

$x_i =$ total output of industry i necessary to meet demands of all n industries and the open sector.

a_{ij} is the amount required from industry i to produce unit output in industry j. Thus $a_{ij}x_i$ will be the amount required to produce x_i units of output in industry j. We get

$$\underbrace{x_i}_{\substack{\text{total output}\\ \text{of industry } i}} = \underbrace{a_{i1}x_1}_{\substack{\text{demand of}\\ \text{industry 1}}} + \underbrace{a_{i2}x_2}_{\substack{\text{demand of}\\ \text{industry 2}}} + \cdots + \underbrace{a_{in}x_n}_{\substack{\text{demand of}\\ \text{industry } n}} + \underbrace{d_i}_{\substack{\text{demand of}\\ \text{open sector}}}$$

The output levels required of the entire set of n industries in order to meet these demands are given by the system of n linear equations

$$\begin{aligned} x_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + d_1 \\ x_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + d_2 \\ &\vdots \\ x_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + d_n \end{aligned}$$

This system of equations can be written in matrix form

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}$$

Let us introduce the following notation.

$$\underbrace{X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\text{the output matrix}} \quad \text{and} \quad \underbrace{D = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}}_{\text{the demand matrix}}$$

The system of equations can now be written as a single matrix equation with the terms having the following significance:

$$\underbrace{X}_{\substack{\text{total}\\ \text{output}}} = \underbrace{AX}_{\substack{\text{interindustry}\\ \text{portion of output}}} + \underbrace{D}_{\substack{\text{open sector}\\ \text{portion of output}}}$$

When the model is applied to the economy of a country, X represents the total output of each of the producing sectors of the economy and AX describes the contributions made by the various sectors to fulfilling the intersectional input requirements of the economy. D is equal to $(X - AX)$, the difference between total output X and industry transaction AX.

D is thus the GNP of the economy

In practice, the equation $X = AX + D$ is applied in a variety of ways, depending on which variables are considered known and which are not known. For example, an analyst seeking to determine the implications of a change in government purchases or consumer demands on the economy described by A might assign values to D and solve the equation for X. The equation could be used in this manner to predict the amount of outputs from each sector needed to attain various GNPs. (Example 2 illustrates this application of the model.) On the other hand, an economist knowing the limited production capacity of an economic system described by A would consider X as known and solve the equation for D, to predict the maximum GNP the system can achieve. (Exercises 7, 8, and 9 following illustrate this application of the model.)

EXAMPLE 2 Consider an economy consisting of three industries having the following input-output matrix A. Determine the output levels required of the industries to meet the demands of the other industries and of the open sector in each case.

$$A = \begin{bmatrix} 0.2 & 0.2 & 0.4 \\ 0.6 & 0.6 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}, \quad D = \begin{bmatrix} 9 \\ 12 \\ 16 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 9 \\ 8 \end{bmatrix}, \text{ and } \begin{bmatrix} 12 \\ 18 \\ 32 \end{bmatrix} \text{ in turn}$$

The units of D are millions of dollars.

SOLUTION

We wish to compute the output levels X that correspond to the various open sector demands D. X is given by the equation $X = AX + D$. Rewrite as follows.

$$X - AX = D$$
$$(I - A)X = D$$

To solve this equation for X, we can use either Gauss-Jordan elimination or the matrix inverse method. In practice, the matrix inverse method is used; a discussion of the merits of this approach is given below. We get

$$X = (I - A)^{-1}D$$

This is the equation that is used to determine X when A and D are known. For our matrix A we get

$$I - A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.2 & 0.2 & 0.4 \\ 0.6 & 0.6 & 0 \\ 0 & 0 & 0.2 \end{bmatrix} = \begin{bmatrix} 0.8 & -0.2 & -0.4 \\ -0.6 & 0.4 & 0 \\ 0 & 0 & 0.8 \end{bmatrix}$$

$(I - A)^{-1}$ is computed using Gauss-Jordan elimination.

$$(I - A)^{-1} = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 4 & 1.5 \\ 0 & 0 & 1.25 \end{bmatrix}$$

We can efficiently compute $X = (I - A)^{-1} D$ for each of the three values of D by forming a matrix having the various values of D as columns:

$$X = \underbrace{\begin{bmatrix} 2 & 1 & 1 \\ 3 & 4 & 1.5 \\ 0 & 0 & 1.25 \end{bmatrix}}_{(I - A)^{-1}} \underbrace{\begin{bmatrix} 9 & 6 & 12 \\ 12 & 9 & 18 \\ 16 & 8 & 32 \end{bmatrix}}_{\text{various values of } D} = \underbrace{\begin{bmatrix} 46 & 29 & 74 \\ 99 & 66 & 156 \\ 20 & 10 & 40 \end{bmatrix}}_{\text{corresponding outputs}}$$

The output levels necessary to meet the demands

$$\begin{bmatrix} 9 \\ 12 \\ 16 \end{bmatrix}, \quad \begin{bmatrix} 6 \\ 9 \\ 8 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 12 \\ 18 \\ 32 \end{bmatrix} \quad \text{are} \quad \begin{bmatrix} 46 \\ 99 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} 29 \\ 66 \\ 10 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 74 \\ 156 \\ 40 \end{bmatrix}$$

respectively. The units are millions of dollars.

Numerical Considerations In practice, analyses of this type usually involve many sectors (as we saw in the example of the U.S. economy), implying large input-output matrices. There is usually a great deal of computation involved in implementing the model and an efficient algorithm is needed. The elements of an input-output matrix A are usually zero or very small. This characteristic of A has led to an appropriate numerical method for computing $(I - A)^{-1}$ that makes the matrix inverse method more efficient for solving the system of equations $(I - A)X = D$ than an elimination method. We now describe this method for computing $(I - A)^{-1}$. Consider the following matrix multiplication for any positive integer m.

$$\begin{aligned} &(I - A)(I + A + A^2 + \cdots + A^m) \\ &\quad = I(I + A + A^2 + \cdots + A^m) - A(I + A + A^2 + \cdots + A^m) \\ &\quad = (I + A + A^2 + \cdots + A^m) - (A + A^2 + A^3 + \cdots + A^{m+1}) \\ &\quad = I - A^{m+1} \end{aligned}$$

The elements of successive powers of A become small rapidly and A^{m+1} approaches the zero matrix. Thus, for an appropriately large m,

$$(I - A)(I + A + A^2 + \cdots + A^m) = I$$

This implies that

$$(I - A)^{-1} = I + A + A^2 + \cdots + A^m$$

This expression is used on a computer to compute $(I - A)^{-1}$ in this model.

Readers who are interested in finding out more about applications of this model should read "The World Economy of the Year 2000" by Wassily W. Leontief, page 166, *Scientific American,* September 1980. The article describes the application of this model to a world economy. The model was commissioned by the United Nations with special financial support from the Netherlands. In the model the world is divided into 15 distinct geographic regions, each one described by an individual input-output matrix. The regions are then linked by a larger matrix that is used in an input-output model. Overall more than 200 economic sectors are included in the model. By feeding in various values, economists use the model to create scenarios of future world economic conditions.

EXERCISE SET 2.7

1. Consider the following input-output matrix that defines the interdependency of five industries.

$$\begin{array}{l} \\ \text{1. Auto} \\ \text{2. Steel} \\ \text{3. Electricity} \\ \text{4. Coal} \\ \text{5. Chemical} \end{array} \begin{array}{c} \begin{array}{ccccc} \quad 1 & \quad 2 & \quad 3 & \quad 4 & \quad 5 \end{array} \\ \begin{bmatrix} 0.15 & 0.10 & 0.05 & 0.05 & 0.10 \\ 0.40 & 0.20 & 0.10 & 0.10 & 0.10 \\ 0.10 & 0.25 & 0.20 & 0.10 & 0.20 \\ 0.10 & 0.20 & 0.30 & 0.15 & 0.10 \\ 0.05 & 0.10 & 0.05 & 0.02 & 0.05 \end{bmatrix} \end{array}$$

Determine

(a) the amount of electricity consumed in producing \$1 worth of steel.

(b) the amount of steel consumed in producing \$1 worth in the auto industry.

(c) the largest consumer of coal.

(d) the largest consumer of electricity.

(e) on which industry the auto industry is most dependent.

In Exercises 2–6 consider the economies consisting of either two or three industries. Determine the output levels required of each industry in each situation to meet the demands of the other industries and of the open sector.

2. $A = \begin{bmatrix} 0.20 & 0.60 \\ 0.40 & 0.10 \end{bmatrix}$,

$D = \begin{bmatrix} 24 \\ 12 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 12 \end{bmatrix}$ in turn

3. $A = \begin{bmatrix} 0.10 & 0.40 \\ 0.30 & 0.20 \end{bmatrix}$,

$D = \begin{bmatrix} 6 \\ 12 \end{bmatrix}, \begin{bmatrix} 18 \\ 6 \end{bmatrix}$, and $\begin{bmatrix} 24 \\ 12 \end{bmatrix}$ in turn

4. $A = \begin{bmatrix} 0.30 & 0.60 \\ 0.35 & 0.10 \end{bmatrix}$,

$D = \begin{bmatrix} 42 \\ 84 \end{bmatrix}, \begin{bmatrix} 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 14 \\ 7 \end{bmatrix}$, and $\begin{bmatrix} 42 \\ 42 \end{bmatrix}$ in turn

5. $A = \begin{bmatrix} 0.20 & 0.20 & 0.10 \\ 0 & 0.40 & 0.20 \\ 0 & 0.20 & 0.60 \end{bmatrix}$,

$D = \begin{bmatrix} 4 \\ 8 \\ 8 \end{bmatrix}, \begin{bmatrix} 0 \\ 8 \\ 16 \end{bmatrix}$, and $\begin{bmatrix} 8 \\ 24 \\ 8 \end{bmatrix}$ in turn

6. $A = \begin{bmatrix} 0.20 & 0.20 & 0 \\ 0.40 & 0.40 & 0.60 \\ 0.40 & 0.10 & 0.40 \end{bmatrix}$,

$D = \begin{bmatrix} 36 \\ 72 \\ 36 \end{bmatrix}, \begin{bmatrix} 36 \\ 0 \\ 18 \end{bmatrix}, \begin{bmatrix} 36 \\ 0 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 18 \\ 18 \end{bmatrix}$ in turn

In Exercises 7–9 consider the economies consisting of either two or three industries. The output levels of the industries are given. Determine the amounts available for the open sector from each industry.

7. $A = \begin{bmatrix} 0.20 & 0.40 \\ 0.50 & 0.10 \end{bmatrix}, \quad X = \begin{bmatrix} 8 \\ 10 \end{bmatrix}$

8. $A = \begin{bmatrix} 0.10 & 0.20 & 0.30 \\ 0 & 0.10 & 0.40 \\ 0.50 & 0.40 & 0.20 \end{bmatrix}, \quad X = \begin{bmatrix} 10 \\ 10 \\ 20 \end{bmatrix}$

9. $A = \begin{bmatrix} 0.10 & 0.10 & 0.20 \\ 0.20 & 0.10 & 0.30 \\ 0.40 & 0.30 & 0.15 \end{bmatrix}, \quad X = \begin{bmatrix} 6 \\ 4 \\ 5 \end{bmatrix}$

10. Let a_{ij} be an arbitrary element of an input-output matrix. Why would you expect a_{ij} to satisfy the condition $0 \le a_{ij} \le 1$?

11. In an economically feasible situation the sum of the elements of each column of the input-output matrix is less than or equal to unity. Explain why this should be so.

12. Consider a two-industry economy described by an input-output matrix A whose columns add up to one. We can express such a matrix in the form

$$A = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix}.$$

(a) Show that the matrix $I - A$ has no inverse.

(b) Illustrate this result for the matrix $A = \begin{bmatrix} 0.2 & 0.7 \\ 0.8 & 0.3 \end{bmatrix}$.

(c) What is the implication for an economy described by such a matrix A? (*Hint*: Consider the equation $X = (I - A)^{-1}D$.)

*2.8 Markov Chains, Population Movements, and Genetics

Certain matrices, called **stochastic matrices**, are important in the study of random phenomena where the exact outcome is not known but probabilities can be determined. In this section, we introduce stochastic matrices, derive some of their properties, and give examples of their application. One example is an analysis of population movement between cities and suburbs in the United States. The second example illustrates the use of stochastic matrices in genetics.

At this time we remind the reader of some basic ideas of probability. If the outcome of an event is *sure to occur*, we say that the probability of that outcome is 1. On the other hand, if it *will not occur*, we say that the probability is 0. Other probabilities are represented by fractions between 0 and 1; *the larger the fraction, the greater the probability p of that outcome occurring*. Thus we have the restriction $0 \le p \le 1$ on a probability p.

If any one of n completely independent outcomes is equally likely to happen, and if m of these outcomes are of interest to us, then the probability p that one of these outcomes will occur is defined to be the fraction m/n.

As an example, consider the event of drawing a single card from a deck of 52 playing cards. What is the probability that the outcome will be an ace or a king? First of all we see that there are 52 possible outcomes. There are 4 aces and 4 kings in the deck; there are 8 outcomes of interest. Thus the probability of drawing an ace or a king is $\frac{8}{52}$, or $\frac{2}{13}$.

We now introduce matrices whose elements are probabilities.

DEFINITION A **stochastic matrix** is a square matrix whose elements are probabilities and whose columns add up to 1.

The following matrices are stochastic matrices.

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{2}{3} \end{bmatrix} \qquad \begin{bmatrix} 0 & \frac{3}{4} \\ 1 & \frac{1}{4} \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & \frac{3}{4} \\ 0 & \frac{1}{2} & \frac{1}{8} \\ 0 & \frac{1}{2} & \frac{1}{8} \end{bmatrix}$$

The following matrices are not stochastic.

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

the sum of the elements in the first column is not 1

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

the 2 in the 1st row is not a probability since it is greater than 1

A general 2×2 stochastic matrix can be written

$$\begin{bmatrix} x & y \\ 1 - x & 1 - y \end{bmatrix}$$

where $0 \le x \le 1$ and $0 \le y \le 1$.

Stochastic matrices have the following useful property. (The reader is asked to prove this result for 2×2 stochastic matrices in the exercises that follow.)

THEOREM 2.10

If A and B are stochastic matrices of the same size, then AB is a stochastic matrix.

Thus if A is stochastic, then $A^2, A^3, A^4, \ldots$ are all stochastic.

EXAMPLE 1 Stochastic matrices are used by city planners to analyze trends in land use. Such a matrix has been used by the city of Toronto, for example. The researchers collect data and write them in the form of a stochastic matrix P. The rows and columns of P represent land uses. We illustrate typical categories for a five-year period in the matrix that follows. The element p_{ij} is the probability that land that was in use j in 2005 was in use i in 2010.

$$\begin{array}{c} \text{Use in 2005} \\ \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} .4 & .15 & .1 & .05 & .05 \\ .1 & .35 & .3 & .15 & .35 \\ .15 & .15 & .5 & .35 & .20 \\ .1 & .30 & .1 & .4 & .25 \\ .25 & .05 & 0 & .05 & .15 \end{bmatrix} \end{array} \begin{array}{l} \text{Use in 2010} \\ \text{1. Residential} \\ \text{2. Office} \\ \text{3. Commercial} \\ \text{4. Parking} \\ \text{5. Vacant} \end{array}$$

Let us interpret some of the information contained in this matrix. For example, $p_{42} = 0.30$. This tells us that land that was office space in 2005 had a probability of 0.30 of becoming a parking area by 2010. The fourth row of P gives the probabilities that various areas of the city have become parking areas by 2010. These relatively large figures reveal the increasingly dominant role of parking in land use.

The diagonal elements give the probabilities that land use remained in the same category. For example, $p_{22} = 0.35$ is the probability that office land remained office land. The relatively high figures of these diagonal elements reflect the tendency for land to remain in the same broad category of usage.

Perhaps the most interesting statistic in the preceding matrix is that office land within the city in 2005 has such a high probability of becoming residential in 2010; $p_{12} = 0.15$.

EXAMPLE 2 In this example we develop a model of population movement between cities and surrounding suburbs in the United States. The numbers given are based on statistics in *Statistical Abstract of the United States*.

It is estimated that the number of people living in cities in the United States during 2015 was 83 million. The number of people living in the surrounding suburbs was 178 million. Let us represent this information by the matrix $X_0 = \begin{bmatrix} 83 \\ 178 \end{bmatrix}$.

Consider the population flow from cities to suburbs. During 2015 the probability of a person staying in the city was 0.97. Thus the probability of moving to the suburbs was 0.03 (assuming that all those who moved went to the suburbs). Consider now the reverse population flow, from suburbia to city. The probability of a person moving to the city was 0.01; the probability of remaining in suburbia was 0.99. These probabilities can be written as the elements of a stochastic matrix P:

$$\begin{array}{c} \text{(from)} \qquad \text{(to)} \\ \text{city} \quad \text{suburb} \quad\quad \end{array}$$
$$P = \begin{bmatrix} 0.97 & 0.01 \\ 0.03 & 0.99 \end{bmatrix} \begin{array}{l} \textit{city} \\ \textit{suburb} \end{array}$$

The probability of moving from location A to location B is given by the element in column A and row B. In this context, the stochastic matrix is called a *matrix of transition probabilities*.

Now consider the population distribution in 2016, one year later:

$$\begin{aligned}\text{city population in 2016} &= \begin{matrix}\text{people who remained}\\ \text{from 2015}\end{matrix} + \begin{matrix}\text{people who moved}\\ \text{in from the suburbs}\end{matrix}\\ &= (0.97 \times 83) + (0.01 \times 178)\\ &= 82.29 \text{ million.}\end{aligned}$$

$$\begin{aligned}\text{suburban population in 2016} &= \begin{matrix}\text{people who moved in}\\ \text{from the city}\end{matrix} + \begin{matrix}\text{people who stayed}\\ \text{from 2015}\end{matrix}\\ &= (0.03 \times 83) + (0.99 \times 178)\\ &= 178.71 \text{ million.}\end{aligned}$$

Note that we can arrive at these numbers using matrix multiplication:

$$\begin{bmatrix}0.97 & 0.01\\ 0.03 & 0.99\end{bmatrix}\begin{bmatrix}83\\ 178\end{bmatrix} = \begin{bmatrix}82.29\\ 178.71\end{bmatrix}$$

Using 2015 as the base year, let X_1 be the population in 2016, one year later. We can write

$$X_1 = PX_0$$

Assume that the population flow represented by the matrix P is unchanged over the years. The population distribution X_2 after 2 years is given by

$$X_2 = PX_1$$

After 3 years, the population distribution is given by

$$X_3 = PX_2$$

After n years, we get

$$X_n = PX_{n-1}$$

The predictions of this model (displaying elements to four decimal places) are

$$\begin{aligned}&X_0 = \begin{bmatrix}83\\ 178\end{bmatrix}\begin{matrix}\textit{city}\\ \textit{suburb}\end{matrix}, \quad X_1 = \begin{bmatrix}82.2900\\ 178.7100\end{bmatrix},\\ &X_2 = \begin{bmatrix}81.6084\\ 179.3916\end{bmatrix}, \quad X_3 = \begin{bmatrix}80.9541\\ 180.0459\end{bmatrix},\\ &X_4 = \begin{bmatrix}80.3259\\ 180.6741\end{bmatrix},\end{aligned}$$

and so on.

Observe how the city population is decreasing annually, while that of the suburbs is increasing. We return to this model in Section 5 of the "Determinants and Eigenvectors" chapter. There we find that the sequence $X_0, X_1, X_2, \ldots$ approaches $\begin{bmatrix}65.2500\\ 195.7500\end{bmatrix}$. If conditions do not change, city population will gradually approach 65.2500 million, while the population of suburbia will approach 195.7500 million.

Further, note that the sequence $X_1, X_2, X_3, \ldots X_n$ can be directly computed from X_0, as follows:

$$X_1 = PX_0, \qquad X_2 = P^2X_0, \qquad X_3 = P^3X_0, \ldots, X_n = P^nX_0$$

The matrix P^n is a stochastic matrix that takes X_0 into X_n, in n steps. This result can be generalized. That is, P^n can be used in this manner to predict the distribution n stages later, from any given distribution.

$$X_{i+n} = P^nX_i$$

P^n is called the *n-step transition matrix*. The (i, j)th element of P^n gives the probability of going from state j to state i in n steps. For example, it can be shown that (writing to 2 decimal places)

$$\begin{array}{rcl} & \begin{array}{cc} \text{(from)} & \qquad \text{(to)} \end{array} & \\ & \begin{array}{cc} \text{city} & \text{suburb} \end{array} & \\ P^4 = & \begin{bmatrix} .89 & .04 \\ .11 & .96 \end{bmatrix} & \begin{array}{l} \textit{city} \\ \textit{suburb} \end{array} \end{array}$$

Thus, for instance, the probability of living in the city in 2015 and being in the suburbs 4 years later is 0.11.

The probabilities in this model depend only on the current state of a person—whether the person is living in the city or in suburbia. This type of model, where the probability of going from one state to another depends only on the current state rather than on a more complete historical description, is called a *Markov Chain*.*

A modification that allows for possible annual population growth or decrease would give improved estimates of future population distributions. The reader is asked to build such a factor into the model in the exercises that follow.

These concepts can be extended to Markov processes involving more than two states. The following example illustrates a Markov chain involving three states.

EXAMPLE 3 Markov chains are useful tools for scientists in many fields. We now discuss the role of Markov chains in *genetics*.

Genetics is the branch of biology that deals with heredity. It is the study of units called **genes**, which determine the characteristics living things inherit from their parents. The inheritance of such traits as sex, height, eye color, and hair color of human beings, and such traits as petal color and leaf shape of plants, are governed by genes. Because many diseases are inherited, genetics is important in medicine. In agriculture, breeding methods based on genetic principles led to important advances in both plant and animal breeding. High-yield hybrid corn ranks as one of the most important contributions of genetics to increasing food production. We shall discuss a mathematical model developed for analyzing the behavior of traits involving a pair of genes. We illustrate the concepts involved in terms of crossing a pair of guinea pigs.

The traits that we shall study in guinea pigs are the traits of long hair and short hair. The length of hair is governed by a pair of genes that we shall denote A and a. A guinea pig may have any one of the combinations AA, Aa, or aa. (aA is genetically the same as Aa.) Each of these classes is called a **genotype**. The AA type of guinea pig is indistinguishable in appearance from the Aa type—both have long hair—while the aa type has short hair. The A gene is said to **dominate** the a gene. An animal is called **dominant** if it has AA genes, **hybrid** with Aa genes, and **recessive** with aa genes.

*Andrei Andreyevich Markov (1856–1922) was educated and taught at the University of St. Petersburg, Russia. He made contributions to the mathematical fields of number theory, probability, and function theory. It was said that "he gave distinguished lectures with irreproachable strictness of argument, and developed in his students that mathematical cast of mind that takes nothing for granted." Markov was personally interested in his students, being faculty advisor to a math circle. He developed chains to analyze literary texts, where the states were vowels and consonants. Markov was a man of strong opinions who was involved in politics. When the establishment celebrated the 300th anniversary of the House of Romanov, Markov organized his own celebration of the 200th anniversary of the law of large numbers!

When two guinea pigs are crossed, the offspring inherits one gene from each parent in a random manner. Given the genotypes of the parents, we can determine the probabilities of the genotype of the offspring. Consider a given population of guinea pigs. Let us perform a series of experiments in which we *keep crossing offspring with dominant animals only.* Thus we keep crossing AA, Aa, and aa with AA. What are the probabilities of the offspring being AA, Aa, or aa in each of these cases?

Consider the crossing of AA with AA. The offspring will have one gene from each parent, so it will be of type AA. Thus the probabilities of AA, Aa, and aa resulting are 1, 0, and 0, respectively. All offspring have long hair.

Next consider the crossing of Aa with AA. Taking one gene from each parent, we have the possibilities of AA, AA (taking A from the first parent and each A in turn from the second parent), aA, and aA (taking the a from the first parent and each A in turn from the second parent). Thus the probabilities of AA, Aa, and aa, respectively, are $\frac{1}{2}$, $\frac{1}{2}$, and 0, respectively. All offspring again have long hair.

Finally, on crossing aa with AA there is only one possibility, namely aA. Thus the probabilities of AA, Aa, and aa are 0, 1, and 0, respectively.

All offspring resulting from these experiments have long hair. This series of experiments is a Markov chain having transition matrix

$$P = \begin{array}{c} \begin{array}{ccc} AA & Aa & aa \end{array} \\ \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array} \begin{array}{l} \\ AA \\ Aa \\ aa \end{array}$$

Consider an initial population of guinea pigs made up of an equal number of each genotype. Let the initial distribution be $X_0 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$, representing the fraction of guinea pigs of each type initially. The components of $X_1, X_2, X_3, \ldots$ will give the fractions of following generations that are of types AA, Aa, and aa, respectively. We get

$$X_0 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \begin{array}{l} AA \\ Aa, \\ aa \end{array} \quad X_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{4} \\ 0 \end{bmatrix}, \quad X_3 = \begin{bmatrix} \frac{7}{8} \\ \frac{1}{8} \\ 0 \end{bmatrix}, \quad X_4 = \begin{bmatrix} \frac{15}{16} \\ \frac{1}{16} \\ 0 \end{bmatrix},$$

and so on.

Observe that the aa type disappears after the initial generation and that the Aa type becomes a smaller and smaller fraction of each successive generation. The sequence in fact approaches the matrix

$$X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{array}{l} AA \\ Aa \\ aa \end{array}$$

The genotype AA in this model is called an **absorbing state**.

Here we have considered the case of crossing offspring with a dominant animal. The reader is asked to construct a similar model that describes the crossing of offspring with a hybrid in the exercises that follow. Some of the offspring will have long hair and some short hair in that series of experiments.

EXERCISE SET 2.8

Stochastic Matrices

1. State which of the following matrices are stochastic and which are not. Explain why a matrix is not stochastic.

(a) $\begin{bmatrix} \frac{1}{4} & 0 \\ \frac{3}{4} & 1 \end{bmatrix}$ **(b)** $\begin{bmatrix} \frac{1}{2} & -1 \\ \frac{1}{2} & 2 \end{bmatrix}$

(c) $\begin{bmatrix} \frac{1}{3} & \frac{1}{7} \\ \frac{2}{3} & \frac{5}{7} \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & \frac{3}{8} & 0 \\ \frac{1}{2} & \frac{1}{8} & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$ **(f)** $\begin{bmatrix} 0 & \frac{1}{5} & \frac{3}{4} \\ \frac{5}{6} & \frac{2}{5} & \frac{3}{4} \\ \frac{1}{6} & \frac{2}{5} & -\frac{1}{2} \end{bmatrix}$

2. Prove that the product of two 2×2 stochastic matrices is a stochastic matrix.

3. A stochastic matrix, the sum of whose rows is 1, is called a **doubly stochastic matrix**. Give examples of 2×2 and 3×3 doubly stochastic matrices. Is the product of two doubly stochastic matrices doubly stochastic?

Population Movement Models

4. Use the stochastic matrix of Example 1 of this section to answer the following questions:

(a) What is the probability that land used for residential in 2005 was used for offices in 2010?

(b) What is the probability that land used for parking in 2005 was in a residential area in 2010?

(c) Vacant land in 2005 had the highest probability of becoming what kind of land in 2010?

(d) Which was the most stable usage of land over the period 2005–2010?

5. In the model of Example 2, determine

(a) the probability of moving from the city to the suburbs in two years.

(b) the probability of moving from the suburbs to the city in three years.

6. Construct a model of population flow between metropolitan and nonmetropolitan areas of the United States, given that their respective populations in 2015 were 261 million and 48 million. The probabilities are given by the matrix

$$\begin{array}{cc} \text{(from)} & \text{(to)} \\ \begin{array}{cc} \text{metro} & \text{nonmetro} \end{array} & \\ \begin{bmatrix} 0.99 & 0.02 \\ 0.01 & 0.98 \end{bmatrix} & \begin{array}{l} \text{metro} \\ \text{nonmetro} \end{array} \end{array}$$

Predict the population distributions of metropolitan and nonmetropolitan areas for the years 2016 through 2020 (in millions, to four decimal places). If a person was living in a metropolitan area in 2015, what is the probability that the person will still be living in a metropolitan area in 2020?

7. Construct a model of population flows between cities, suburbs, and nonmetropolitan areas of the United States. Their respective populations in 2015 were 83 million, 178 million, and 48 million. The stochastic matrix giving the probabilities of the moves is

$$\begin{array}{cc} \text{(from)} & \text{(to)} \\ \begin{array}{ccc} \text{city} & \text{suburb} & \text{nonmetro} \end{array} & \\ \begin{bmatrix} 0.96 & 0.01 & 0.015 \\ 0.03 & 0.98 & 0.005 \\ 0.01 & 0.01 & 0.98 \end{bmatrix} & \begin{array}{l} \text{city} \\ \text{suburb} \\ \text{nonmetro} \end{array} \end{array}$$

This model is a refinement on the model of the previous exercise in that the metropolitan population is broken down into city and suburb. It is also a more complete model than that of Example 2 of this section, which did not allow for any population outside cities and suburbs.

Predict the populations of city, suburban, and nonmetropolitan areas for 2016, 2017, and 2018. If a person was living in the city in 2015, what is the probability that the person will be living in a nonmetropolitan area in 2017?

8. In the period prior to 2015, the total population of the United States increased by 1% per annum. Assume that the population increases annually by 1% during the years immediately following. Build this factor into the model of Example 2 and predict the populations of city and suburbia in 2020.

9. Assume that births, deaths, and immigration increased the population in U.S. cities by 1.2% during the period prior to 2015, and that the populations of the suburbs increased by 0.8% due to these factors. Allow for these increases in the model of Example 2 and predict the populations for the year 2020.

10. Consider the population movement model for flow between cities and suburbs of Example 2 of this section. Determine the population distributions for 2012 to 2014—prior to 2015. Is the chain going from 2015 into the past a Markov Chain? What are the characteristics of the matrix that takes one from distribution to distribution into the past.

Miscellaneous Stochastic Models

11. The following stochastic matrix gives occupational transition probabilities.

$$\begin{array}{ccc} \text{(initial generation)} & & \\ \begin{array}{cc} \text{white-collar} & \text{manual} \end{array} & & \\ \begin{bmatrix} 1 & 0.2 \\ 0 & 0.8 \end{bmatrix} & \begin{array}{l} \text{white-collar} \\ \text{manual} \end{array} & \text{(next generation)} \end{array}$$

(a) If the father is a manual worker, what is the probability that the son will be a white-collar worker?

(b) If there are 10,000 in the white-collar category and 20,000 in the manual category, what will the distribution be one generation later?

12. The following matrix gives occupational transition probabilities.

$$\begin{array}{cc} \multicolumn{2}{c}{\text{(initial generation)}} \\ \text{nonfarming} & \text{farming} \end{array}$$
$$\begin{bmatrix} 1 & 0.4 \\ 0 & 0.6 \end{bmatrix} \begin{array}{l} \text{nonfarming} \\ \text{farming} \end{array} \text{(next generation)}$$

(a) If the father is a farmer, what is the probability that the son will be a farmer?

(b) If there are 10,000 in the nonfarming category and 1,000 in the farming category at a certain time, what will the distribution be one generation later? Four generations later?

(c) If the father is a farmer, what is the probability that the grandson will be a farmer?

13. A market analysis of car purchasing trends in a certain region has concluded that a family purchases a new car once every 3 years on an average. The buying patterns are described by the matrix

$$P = \begin{array}{c} \begin{array}{cc} \text{small} & \text{large} \end{array} \\ \begin{bmatrix} 80\% & 40\% \\ 20\% & 60\% \end{bmatrix} \end{array} \begin{array}{l} \text{small} \\ \text{large} \end{array}$$

The elements of P are to be interpreted as follows: The first column indicates that of the current small cars, 80% will be replaced with small cars, 20% with a large car. The second column implies that 40% of the current large cars will be replaced with small cars while 60% will be replaced with large cars. Write the elements of P as follows to get a stochastic matrix that defines a Markov chain.

$$P = \begin{bmatrix} \frac{80}{100} & \frac{40}{100} \\ \frac{20}{100} & \frac{60}{100} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.4 \\ 0.2 & 0.6 \end{bmatrix}$$

If there are currently 40,000 small cars and 50,000 large cars in the region, what is your prediction of the distribution in 12 years' time?

14. The conclusion of an analysis of voting trends in a certain state is that the voting patterns of successive generations are described by the following matrix P.

$$P = \begin{array}{c} \begin{array}{ccc} \text{Dem.} & \text{Rep.} & \text{Ind.} \end{array} \\ \begin{bmatrix} 80\% & 20\% & 60\% \\ 15\% & 70\% & 30\% \\ 5\% & 10\% & 10\% \end{bmatrix} \end{array} \begin{array}{l} \text{Democrat} \\ \text{Republican} \\ \text{Independent} \end{array}$$

Among the Democrats of one generation, 80% of the next generation are Democrats, 15% are Republican, and 5% are Independents. Express P as a stochastic matrix that defines a Markov chain model of the voting patterns. If there are 2.5 million registered Democrats, 1.5 million registered Republicans, and 0.25 million registered Independents at a certain time, what is the distribution likely to be in the next generation?

15. Determine the transition matrix for a Markov chain that describes the crossing of offspring of guinea pigs with hybrids only. There is no absorbing state in this model. Let the initial matrix $X_0 = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$ be the fraction of guinea pigs of each type initially. Determine the distributions for the next three generations.

16. Let P be a 2×2 symmetric stochastic matrix. Prove that P must be of the form $\begin{bmatrix} x & 1-x \\ 1-x & x \end{bmatrix}$, where $0 \le x \le 1$. If P describes population movement between two states, what can you say about the movement?

*2.9 A Communication Model and Group Relationships in Sociology

Many branches of the physical sciences, social sciences, and business use models from *graph theory* to analyze relationships. We introduce the reader to this important area of mathematics, which uses linear algebra, with an example from the field of communication.

Consider a communication network involving five stations, labeled $P_1, \ldots, P_5$. The communication links could be roads, phone lines, or Internet links, for example. Certain stations are linked by two-way communication, others by one-way links. Still others may have only indirect communication by way of intermediate stations. Suppose the network of interest is described in Figure 2.24. Lines joining stations represent direct communication links; the arrows give the directions of those links. For example, stations P_1 and

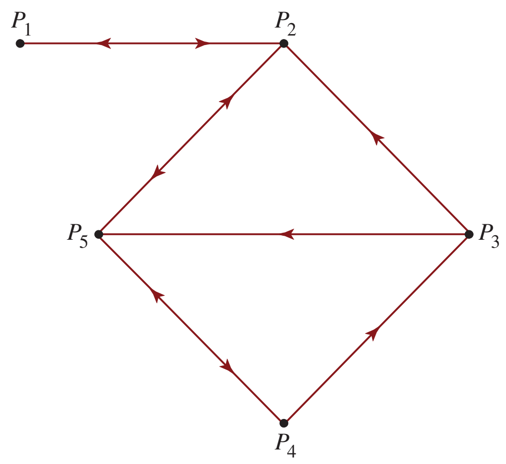


Figure 2.24

P_2 have two-way direct communication. Station P_4 can send a message to P_1 by way of stations P_3 and P_2 or by way of P_5 and P_2. This communication network is an example of a digraph.

DEFINITION A **digraph** is a finite collection of **vertices** $P_1, P_2, \ldots, P_n$, together with **directed arcs** joining certain pairs of vertices. A **path** between vertices is a sequence of arcs that allows one to proceed in a continuous manner from one vertex to another. The **length** of a path is its number of arcs. A path of length n is called an ***n*-path**.

In the above communication network, there are five vertices, namely $P_1, \ldots, P_5$. Suppose that we are interested in sending a message from P_3 to P_1. From the figure we see that there are various paths that can be taken. The path $P_3 \rightarrow P_2 \rightarrow P_1$ is of length 2 (a 2-path), while the path $P_3 \rightarrow P_5 \rightarrow P_2 \rightarrow P_1$ is of length 3 (a 3-path). The path $P_3 \rightarrow P_5 \rightarrow P_4 \rightarrow P_3 \rightarrow P_2 \rightarrow P_1$, a 5-path, takes in the vertex P_3 twice. Such paths could be of interest if, for example, P_3 wanted to consult with P_5 and P_4 before sending a message to P_1. In this section, however, we shall be interested in finding the shortest route to send a message from one station to another.

Communication networks can be vast, involving many stations. It is impractical to get information about large networks from diagrams, as we have done here. The mathematical theory that we now present, from graph theory, can be used to give information about large networks. The mathematics can be implemented on the computer.

A digraph can be described by a matrix A, called its adjacency matrix. This matrix consists of zeros and ones and is defined as follows.

DEFINITION Consider a digraph with vertices $P_1, \ldots, P_n$. The **adjacency matrix** A of the digraph is such that

$$a_{ij} = \begin{cases} 1 & \text{if there is an arc from vertex } P_i \text{ to } P_j \\ 0 & \text{if there is no arc from vertex } P_i \text{ to } P_j \\ 0 & \text{if } i = j \end{cases}$$

The adjacency matrix of the communication network is

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

For example, $a_{12} = 1$ since there is an arc from P_1 to P_2; $a_{13} = 0$ since there is no arc from P_1 to P_3.

The network is completely described by the adjacency matrix. This matrix is a mathematical "picture" of the network that can be given to a computer. We could look at the sketch of the network given in Figure 2.24 and decide how best to send a message from P_3 to P_1. How can we extract such information from the adjacency matrix? The following theorem from graph theory gives information about paths within digraphs.

THEOREM 2.11

If A is the adjacency matrix of a digraph, let $a_{ij}^{(m)}$ be the element in row i and column j of A^m.

$$\text{The number of } m\text{-paths from } P_i \text{ to } P_j = a_{ij}^{(m)}$$

Proof Consider a digraph with n vertices. a_{i1} is the number of arcs from P_i to P_1, and a_{1j} is the number of arcs from P_1 to P_j. Thus $a_{i1}\,a_{1j}$ is the number of 2-paths from P_i to P_j, passing through P_1. Summing up over all such possible intermediate stations, we see that the total number of 2-paths from P_i to P_j is

$$a_{i1}a_{1j} + a_{i2}a_{2j} + \cdots + a_{in}a_{nj}$$

This is the element in row i, column j of A^2. Thus $a_{ij}^{(2)}$ is the number of 2-paths from P_i *to* P_j.

Let us now look at 3-paths. Interpret a 3-path as a 2-path followed by an arc. The number of 2-paths from P_i to P_1 followed by an arc from P_1 to P_j is $a_{i1}^{(2)}\,a_{1j}$. Summing up over all such possible intermediate stations we see that the total number of 3-paths from P_i *to* P_j is

$$a_{i1}^{(2)}a_{1j} + a_{i2}^{(2)}a_{2j} + \cdots + a_{in}^{(2)}a_{nj}$$

This is the element in row i, column j, of the matrix product A^2A; that is of A^3. Thus $a_{ij}^{(3)}$ is the number of 3-paths from P_i to P_j.

Continuing thus, we can interpret a 4-path as a 3-path followed by an arc, and so on, arriving at the result that $a_{ij}^{(m)}$ is the number of m-paths from P_i to P_j.

We now illustrate the application of this theorem to the communication network of Figure 2.24. Successive powers of the adjacency matrix are determined.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix} \qquad A^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 2 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 2 & 0 & 1 & 0 \\ 2 & 0 & 1 & 0 & 3 \\ 1 & 2 & 1 & 1 & 2 \\ 2 & 1 & 1 & 1 & 3 \\ 0 & 4 & 0 & 2 & 1 \end{bmatrix} \qquad A^4 = \begin{bmatrix} 2 & 0 & 1 & 0 & 3 \\ 0 & 6 & 0 & 3 & 1 \\ 2 & 4 & 1 & 2 & 4 \\ 1 & 6 & 1 & 3 & 3 \\ 4 & 1 & 2 & 1 & 6 \end{bmatrix} \qquad A^5 = \cdots$$

Let us use these matrices to discuss paths from P_4 to P_1.

A gives $a_{41} = 0$. There is no direct communication.
A^2 gives $a_{41}^{(2)} = 0$. There are no 2-paths from P_4 to P_1.
A^3 gives $a_{41}^{(3)} = 2$. There are two distinct 3-paths from P_4 to P_1.

These are the shortest paths from P_4 to P_1. If we check with Figure 2.24, we see that this is the case. The two 3-paths are

$$P_4 \rightarrow P_3 \rightarrow P_2 \rightarrow P_1 \quad \text{and} \quad P_4 \rightarrow P_5 \rightarrow P_2 \rightarrow P_1$$

As a second example, let us determine the length of the shortest path from P_1 to P_3.

A gives $a_{13} = 0$. There is no direct communication.
A^2 gives $a_{13}^{(2)} = 0$. There are no 2-paths from P_1 to P_3.
A^3 gives $a_{13}^{(3)} = 0$. There are no 3-paths from P_1 to P_3.
A^4 gives $a_{13}^{(4)} = 1$. There is a single 4-path from P_1 to P_3.

This result is confirmed when we examine the digraph. The quickest way to send a message from P_1 to P_3 is the 4-path

$$P_1 \rightarrow P_2 \rightarrow P_5 \rightarrow P_4 \rightarrow P_3$$

This model that we have discussed gives the lengths of the shortest paths of a digraph; it does not give the intermediate stations that make up that path. Mathematicians have not, as yet, been able to derive this information from the adjacency matrix. An algorithm for finding the shortest paths for a specific digraph, using a search procedure, has been developed by a Dutch computer scientist named Edsger Dijkstra. See, for example, *Algorithmics: Theory and Practice* by Gilles Brassard and Paul Bratley, Prentice Hall, 1988, for a discussion of this algorithm. The following discussion leads to an application where the lengths of the shortest paths, not the actual paths, are important.

Distance in a Digraph

The **distance** from one vertex of a digraph to another is the length of the shortest path from that vertex to the other. If there is no path from the one vertex to the other, we say that the distance is **undefined**. The distances between the various vertices of a digraph form the elements of a matrix. The **distance matrix** D is defined as follows:

$$d_{ij} = \begin{cases} \text{number of arcs in shortest path from vertex } P_i \text{ to vertex } P_j \\ 0 \quad \text{if } i = j \\ x \quad \text{if there is no path from } P_i \text{ to } P_j \end{cases}$$

If the digraph is small, the distance matrix can be constructed by observation. Powers of the adjacency matrix are used to construct the distance matrix of a large digraph. The distance matrix of the previous communication network in Figure 2.24 is

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

Note that the distance from P_i to P_j is not necessarily equal to the distance from P_j to P_i in a digraph, implying that a distance matrix in graph theory is not necessarily symmetric.

We now illustrate how a digraph and a distance matrix can be used to analyze group relationships in sociology.

Group Relationships in Sociology

Consider a group of five people. A sociologist is interested in finding out which one of the five has the most influence over, or dominates, the other members. The group is asked to fill out the following questionnaire:

- Your name ____________________
- Person whose opinion you value most ____________________

These answers are then tabulated. Let us for convenience label the group members $M_1, M_2, \ldots, M_5$. Suppose the results are as given in Table 2.1.

Table 2.1

Group member	Person whose opinion valued
M_1	M_4
M_2	M_1
M_3	M_2
M_4	M_2
M_5	M_4

The sociologist makes the assumption that the person whose opinion a member values most is the person who influences that member most. Thus, influence goes from the right column to the left column in the above table. We can represent these results by a digraph. The group members are represented by vertices, direct influence by an arc, the direction of influence being the direction of the arc. See Figure 2.25. Construct the distance matrix of this digraph, and add up all the elements in each row.

$$D = \begin{bmatrix} 0 & 1 & 2 & 2 & 3 \\ 2 & 0 & 1 & 1 & 2 \\ x & x & 0 & x & x \\ 1 & 2 & 3 & 0 & 1 \\ x & x & x & x & 0 \end{bmatrix} \quad \begin{array}{c} \text{row sums} \\ 8 \\ 6 \\ 4x \\ 7 \\ 4x \end{array}$$

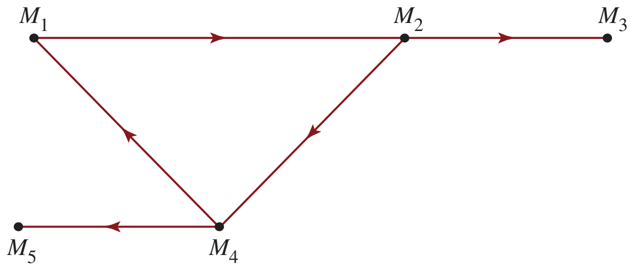


Figure 2.25

In this graph, arcs correspond to direct influence, and 2-paths, 3-paths, etc., correspond to indirect influence. Thus, presumably, the smaller the distance from M_i to M_j, the greater the influence M_i has on M_j. The sum of the elements in row i gives the total distance of M_i to the other vertices. This leads to the following interpretation of row sums.

The smaller row sum i, the greater the influence of person M_i on the group.

We see that the smallest row sum is 6, for row 2. Thus M_2 is the most influential person in the group, followed by M_4 and then M_1. M_3 and M_5 have no influence on other members.

Readers who are interested in learning more about graph theory are referred to the following two books: *Introduction to Graph Theory* by Robin J. Wilson, John Wiley and Sons, 1987, and *Discrete Mathematical Structures* by Fred S. Roberts, Prentice Hall, 1976. The former book is a beautiful introduction to the mathematics of graph theory; the latter has a splendid collection of applications. "Predicting Chemistry by Topology" by Dennis H. Rouvray, *Scientific American*, 40, September 1986, contains a fascinating account of how graph theoretical methods are being used to predict chemical properties of molecules that have not yet been synthesized.

We complete this section with a discussion of research that is currently being done in the fields of transportation and sociology.

The Worldwide Air Transportation Network

The global structure of the worldwide air transportation network has been analyzed by a team of researchers led by Luis Amaral of Northwestern University and published in *The Proceedings of the National Academy of Sciences*, Volume 102, pages 7794–7799, 2005. The network is interpreted as a digraph with airports as nodes and nonstop passenger flights between airports as arcs. The study involved 3,883 airports with 531,574 flights. (When there is more than one major airport for a region, they are all grouped together. For example, Newark, JFK, and LaGuardia airports are all assigned to New York.)

The lengths of the shortest paths (distances) for pairs of airports were computed. The average minimum number of flights that one needs to take to get from any airport to any other airport in the world was found to be 4.4. The farthest cities in the network are Mount Pleasant in the Falkland Islands and Wasu, Papua, New Guinea. A journey in either direction involves 15 different flights. (The *diameter* of the digraph is 15.)

The report discusses two other indexes for this network. The *degree* of an airport is the number of nonstop flights leaving that airport (the number of arcs leaving the node). This is taken to be a measure of the connectedness of the airport. The *betweenness* index of an airport is the number of shortest paths connecting two other airports that involve a transfer at the given airport. This is a measure of the centrality of the airport. It is found that the most connected airports are not necessarily the most central. Although most connected airports are located in western Europe and North America, the most central airports are distributed uniformly across all of the continents. Significantly, each continent has at least one central airport—e.g., Johannesburg in Africa, and Buenos Aires and Sao Paulo in South America. Interestingly Anchorage (Alaska) and Port Moresby (Papua, New Guinea) have small degrees, as might be expected, but are among the most central airports in the world; Anchorage has few flights out, but many shortest paths between other airports would be disrupted if Anchorage closed down. A list of the most central airports and their betweenness index and degree follows.

Rank	Airport	Betweenness	Degree
1	Paris	58.8	250
2	Anchorage	55.2	39
3	London	54.7	242
4	Singapore	47.5	92
5	New York	47.2	179
6	Los Angeles	44.8	133
7	Port Moresby	43.4	38
8	Frankfurt	41.5	237
9	Tokyo	39.1	111
10	Moscow	34.5	186

Chicago is ranked 13th and Miami 25th, with betweenness indexes of 28.8 and 20.1, respectively, and degrees of 184 and 110. Surprisingly, Atlanta is not listed in the top 25 airports. Readers who are interested in reading more about connectivity of networks should read Section 5 of the "Numerical Methods" chapter.

Distance in Social Networks*

A team of scientists at Columbia University—Duncan J. Watts, Peter Dodds, and Ruby Muhamad—are studying communication on the Web. The *Small World* project is an online experiment to test the idea that any two people in the world can be connected via "six degrees of separation." Their initial results indicate that most anyone can reach a distant stranger in an average of six relays. Participants were given basic facts (name, location, profession, some educational background) about 18 target individuals in 13 countries and asked to send a message to those targets. The participants sent the message to someone they thought was "closer" (in a message sense) to the target. These recipients were urged to do the same. Calculations based on the results of the experiment, which allowed for attrition, predicted that on average it takes 5 message relays to reach a target in the same country, while it takes 7 to find a target in another country and 6 worldwide. Thinking in terms of a graph, where the vertices are people and directed arcs are transmitted messages, the average distance between two people in the same country is 5, the average distance between two people in different countries is 7, while the worldwide average is 6. The experiment is now being refined and will be repeated. Readers can find the report of this experiment in the journal *Science*, Volume 301, pages 827–829, 2003.

* I am grateful to Duncan J. Watts for helpful comments on this project.

EXERCISE SET 2.9

Digraphs and Adjacency Matrices

1. Determine the adjacency matrix and the distance matrix of each of the digraphs in Figure 2.26.

2. The **diameter** of a digraph is the largest of the distances between the vertices. Determine the diameters of the digraphs in Figure 2.26.

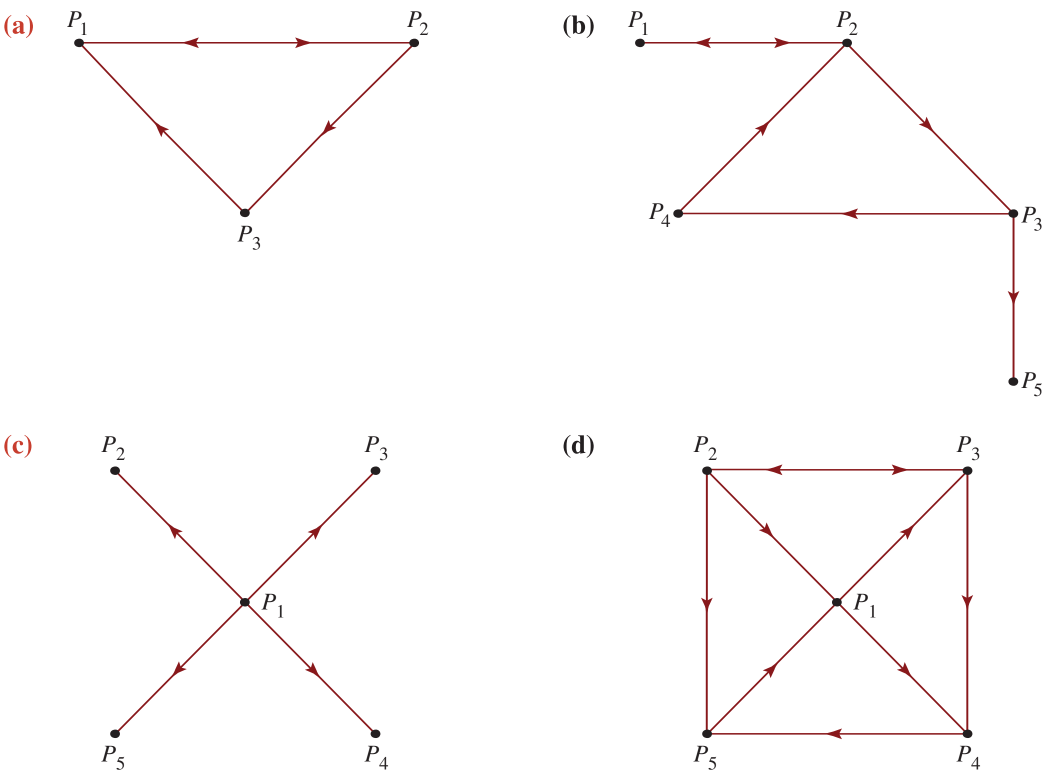


Figure 2.26

3. Sketch the digraphs that have the following adjacency matrices.

(a) $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$ **(d)** $\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$

Applications of Digraphs

4. The network given in Figure 2.27 describes a system of streets in a city downtown area. Many of the streets are one-way. Interpret the network as a digraph. Give its adjacency matrix and distance matrix.

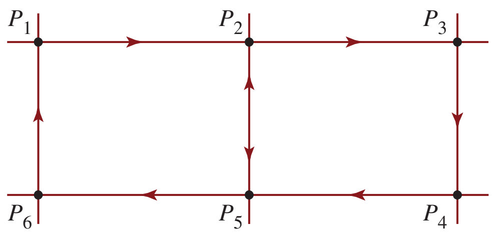


Figure 2.27

5. Graph theory is being used in mathematical models to better understand the delicate balance of nature. Figure 2.28 gives the digraph that describes the food web of an ecological community in the Ocala National Forest, in central Florida. Determine the adjacency matrix.

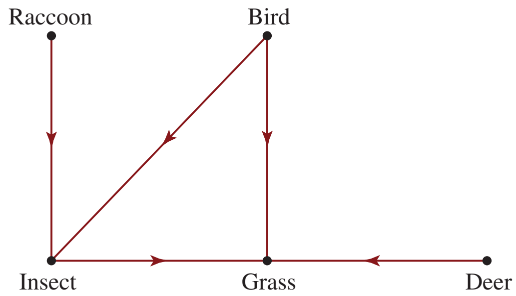


Figure 2.28

6. When all arcs in a network are two-way, it is customary not to include arrows, since they are not necessary. The term **graph** is then used. Scientists are using graphs to predict the chemical properties of molecules. The graphs in Figure 2.29 describe the molecular structures of butane and isobutane. Both have the same chemical formula C_4H_{10}. Determine the adjacency matrices of these graphs. Note that the matrices are symmetric. Would you expect the adjacency matrix of any graph to be symmetric?

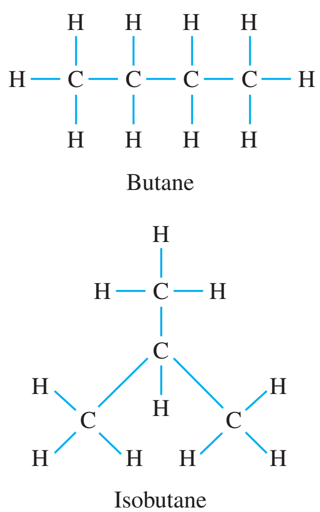


Figure 2.29

7. The following matrix defines a communication network. Sketch the network. Determine the shortest path for sending a message from

(a) P_2 to P_5 **(b)** P_3 to P_2

Find the distance matrix of the digraph.

$$\begin{bmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

8. In each of the following exercises, the matrix A is the adjacency matrix for a communication network. Sketch the networks. Powers of the adjacency matrices are given. Interpret the elements that have been circled.

(a) $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

$A^2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

$A^3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

$A^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$

$A^3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

$A^2 = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

$A^3 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

$A^2 = \begin{bmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

$A^3 = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

(e) $A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$ $A^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$

$A^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 1 & 1 & 1 \\ 2 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$

Information about Digraphs

9. Let A be the adjacency matrix of a digraph. What do you know about the digraph in each of the following cases?

(a) The third row of A is all zeros.

(b) The fourth column of A is all zeros.

(c) The sum of the elements in the fifth row of A is 3.

(d) The sum of the elements in the second column of A is 2.

(e) The second row of A^3 is all zeros.

(f) The third column of A^4 is all zeros.

10. Consider digraphs with adjacency matrices having the following characteristics. What can you tell about each digraph?

(a) The second row is all zeros.

(b) The third column is all zeros.

(c) The fourth row has all ones except for zero in the diagonal location.

(d) The fifth column has all ones except for zero in the diagonal location.

(e) The sum of the elements in the third row is 5.

(f) The sum of the elements in the second column is 4.

(g) The number of ones in the matrix is 7.

(h) The sum of the elements in row 2 of the fourth power is 3.

(i) The sum of the elements in column 3 of the fifth power is 4.

(j) The fourth row of the square of the matrix is all zeros.

(k) The third column of the fourth power is all zeros.

11. Let A be the adjacency matrix of a digraph. Sketch the digraph if A^2 is as follows. Use the digraph to find A^3.

$$A^2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

[*Hint*: You are given all the 2-paths in the digraph.]

12. Let A be the adjacency matrix of a digraph. Sketch all the possible digraphs described by A if

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Group Relationships

13. The following tables represent information obtained from questionnaires given to groups of people. In each case construct the digraph that describes the leadership structure within the group. Rank the members according to their influence on the group.

(a)

Group member	Person whose opinion valued
M_1	M_4
M_2	M_1
M_3	M_2
M_4	M_2

(b)

Group member	Person whose opinion valued
M_1	M_5
M_2	M_1
M_3	M_2
M_4	M_3
M_5	M_4

(c)

Group member	Person whose opinion valued
M_1	M_4
M_2	M_1 and M_5
M_3	M_2
M_4	M_3
M_5	M_1

(d)

Group member	Person whose opinion valued
M_1	M_5
M_2	M_1
M_3	M_1 and M_4
M_4	M_5
M_5	M_3

14. The following matrices describe the relationship "friendship" between groups of people. $a_{ij} = 1$ if M_i is a friend of M_j; $a_{ij} = 0$ otherwise. Draw the digraphs that describe these relationships. Note that all the matrices are symmetric. What is the significance of this symmetry? Can such a relationship be described by a matrix that is not symmetric?

(a) $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$

15. A structure in a digraph that is of interest to social scientists is a clique. A **clique** is the largest subset of a digraph consisting of three or more vertices, each pair of which is mutually related. The application of this concept to the relationship "friendship" is immediate: three or more people form a clique if they are all friends and if they do not all have any mutual friendships with any single person outside that set. Give an example of a digraph that contains a clique.

Miscellaneous Results

16. Prove that the adjacency matrix of a digraph is necessarily square.

17. Let A be the adjacency matrix of a digraph. The matrix A^t is a square matrix consisting of zeros and ones. It is also the adjacency matrix of a digraph. How are the digraphs of A and A^t related?

18. If the adjacency matrix of a digraph is symmetric, what does this tell you about the digraph?

19. Prove that the shortest path from one vertex of a digraph to another vertex cannot contain any repeated vertices.

20. In a graph with n vertices, what is the greatest possible distance between two vertices.

21. Let A be the adjacency matrix of a communication digraph. Let $C = AA^t$. Show that c_{ij} = number of stations that can receive a message directly from both stations i and j.

22. The **reachability matrix** R of a digraph is defined as follows:

$$r_{ij} = \begin{cases} 1 & \text{if there is a path from vertex } P_i \text{ to } P_j \\ 1 & \text{if } i = j \\ 0 & \text{if there is no path from } P_i \text{ to } P_j \end{cases}$$

Determine the reachability matrices of the digraphs of Exercise 8.

23. The reachability matrix of a digraph can be constructed using information from the adjacency matrix and its various powers. How many powers of the adjacency matrix of a digraph having n vertices would have to be calculated to get all the information needed?

24. **(a)** If the adjacency matrix of a digraph is symmetric, does this mean that the reachability matrix is symmetric?

(b) If the reachability matrix of a digraph is symmetric, does this mean that the adjacency matrix is symmetric?

25. Let R be the reachability matrix of a communication digraph. Let $r(i)$ be the sum of the elements of row i of R and $c(j)$ be the sum of the elements of column j of R. What information about the digraph do $r(i)$ and $c(j)$ give?

26. Let R be the reachability matrix of a digraph. What information about the digraph does the element in row i, column j of R^2 give?

27. Let R be the reachability matrix of a digraph. What information about the digraph does R^t give?

28. The adjacency matrix A and reachability matrix R of a digraph are both made up of elements that are either zero or one. Can

(a) A and R have the same number of ones?

(b) A have more ones than R?

(c) A have fewer ones than R?

CHAPTER 2 REVIEW EXERCISES

1. Let $A = \begin{bmatrix} 2 & 0 \\ 7 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 7 & 0 \\ -1 & 3 \end{bmatrix}$,

$C = \begin{bmatrix} 6 & -1 & 3 \\ 5 & 0 & -2 \end{bmatrix}$, and $D = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

Compute the following (if they exist).

(a) $2AB$ **(b)** $AB + C$ **(c)** $BA + AB$

(d) $AD - 3D$ **(e)** $AC + BC$ **(f)** $2DA + B$

2. Let A be a 2×2 matrix, B a 2×2 matrix, C a 2×3 matrix, D a 3×2 matrix, and E a 3×1 matrix. Determine which of the following matrix expressions exist and give the sizes of the resulting matrices when they do exist.

(a) AB **(b)** $(A^2)C$

(c) $B^3 + 3(CD)$ **(d)** $DC + BA$

(e) $DA - 2(DB)$ **(f)** $C - 3D$

(g) $3(BA)(CD) + (4A)(BC)D$

3. If $A = \begin{bmatrix} 1 & -3 \\ 0 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & -3 \\ 5 & 0 & -1 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & -4 & 5 \\ 7 & 1 & 0 \end{bmatrix}$, determine the following elements of $D = 2AB - 3C$, without computing the complete matrix.

(a) d_{12} **(b)** d_{23}

4. **(a)** Let $A = \begin{bmatrix} 3 & 1 \\ 7 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 & 6 \\ 2 & 0 & -1 \end{bmatrix}$. Compute the product AB using the columns of B.

(b) Let $P = \begin{bmatrix} 1 & 2 & 3 \\ 5 & -1 & 4 \end{bmatrix}$, $Q = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$.

Express the product PQ as a linear combination of the columns of P.

(c) Let $A = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 4 & 1 \\ 2 & 5 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 & 2 \\ -1 & -3 & 7 \\ 0 & 1 & -2 \end{bmatrix}$.

Find all the partitions of B that can be used to calculate AB for the partition of A given here:

$$A = \left[\begin{array}{cc|c} 3 & -1 & 2 \\ 0 & 4 & 1 \\ \hline 2 & 5 & 0 \end{array}\right]$$

5. If $A = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -2 & 1 \\ 3 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}$, compute each of the following.

(a) $(A^t)^2$ **(b)** $A^t - B^2$

(c) $AB^3 - 2C^2$ **(d)** $A^2 - 3A + 4I_2$

6. Consider the following system of equations. You are given two solutions, X_1 and X_2. Generate four other solutions using the operations of addition and scalar multiplication. Find a solution for which $x = 1$, $y = 9$.

$$\begin{aligned} x + 2y + 5z - 3w &= 0 \\ x - 2y - 3z - w &= 0 \\ 5x - 4y - 3z - 8w &= 0 \\ -4x + 8y + 12z + 4w &= 0 \end{aligned}$$

$$X_1 = \begin{bmatrix} 5 \\ 3 \\ -1 \\ 2 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 2 \end{bmatrix}$$

7. Find the subspace of solutions for the following homogeneous system of linear equations. Give a basis and the dimension of the subspace.

$$\begin{aligned} x_1 + x_2 + 2x_3 + 2x_4 &= 0 \\ x_1 + 2x_2 + 6x_3 + x_4 &= 0 \\ 3x_1 + 2x_2 + 2x_3 + 7x_4 &= 0 \end{aligned}$$

8. Consider the following nonhomogeneous system of linear equations. For convenience, its general solution is given. **(a)** Write down the corresponding homogeneous system and give its general solution. **(b)** Give a basis for this subspace of solutions to the homogeneous system and a written description of the subspace. **(c)** Give a written description of the subspace of solutions to the nonhomogeneous system.

$$\begin{aligned} x_1 + 2x_2 - 8x_3 &= 8 \\ x_2 - 3x_3 &= 2 \\ x_1 + x_2 - 5x_3 &= 6 \end{aligned}$$

General solution $(2r + 4, 3r + 2, r)$.

9. Determine the inverse of each of the following matrices, if it exists, using the method of Gauss-Jordan elimination.

(a) $\begin{bmatrix} 1 & 4 \\ 2 & -1 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & 3 & 3 \\ 1 & 2 & 3 \\ 1 & 4 & 6 \end{bmatrix}$ **(c)** $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$

10. Use the matrix inverse method to solve the following system of equations.

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 1 \\ 2x_1 + 5x_2 - 3x_3 &= 5 \\ -3x_1 + 2x_2 - 4x_3 &= 7 \end{aligned}$$

11. Find A such that $3A^{-1} = \begin{bmatrix} 5 & -6 \\ -2 & 3 \end{bmatrix}$.

12. Verify the associative property of multiplication $A(BC) = (AB)C$.

13. **(a)** Let T_1 and T_2 be the following row operations. T_1: interchange rows 1 and 3 of I_3. T_2: add -4 times row 2 of I_3 to row 1. Give the elementary matrices corresponding to T_1 and T_2.

(b) Determine the row operations defined by the elementary matrices

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

14. If n is a nonnegative integer and c is a scalar, prove that $(cA)^n = c^n A^n$.

15. Let A be a matrix such that $AA^t = O$. Show that $A = O$.

16. A matrix is said to be *normal* if $AA^t = A^tA$. Prove that all symmetric matrices are normal.

17. A matrix A is *idempotent* if $A^2 = A$. Prove that if A is idempotent then A^t is also idempotent.

18. A matrix A is *nilpotent* if $A^p = 0$ for some positive integer p. The least such integer p is called the *degree of nilpotency*. Prove that if A is nilpotent, then A^t is also nilpotent with the same degree of nilpotency.

19. Prove that if A is symmetric and invertible, then A^{-1} is also symmetric.

20. Prove that a matrix with a row of zeros or a column of zeros has no inverse.

21. Consider the transformation T defined by the following matrix A. Determine the images of the vectors $\mathbf{x}$ and $\mathbf{y}$.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 4 & 1 \end{bmatrix};\ \mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix},\ \mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}$$

22. Determine whether the following transformations are linear.

(a) $T(x, y) = (2x, y, y - x)$ of $\mathbf{R}^2 \to \mathbf{R}^3$
(b) $T(x, y) = (x + y, 2y + 3)$ of $\mathbf{R}^2 \to \mathbf{R}^2$

23. Find the matrix that maps $\mathbf{R}^2 \to \mathbf{R}^2$ such that $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \mapsto \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -2 \end{bmatrix} \mapsto \begin{bmatrix} -1 \\ 11 \end{bmatrix}$.

24. Find a single matrix that defines a rotation of the plane through an angle of $\frac{\pi}{6}$ about the origin, while at the same time moves points to three times their original distance from the origin.

25. Determine the matrix that defines a reflection about the line $y = -x$.

26. Determine the matrix that defines a projection onto the line $y = -x$.

27. Find the equation of the image of the line $y = -5x + 1$ under a scaling of factor 5 in the x-direction and factor 2 in the y-direction.

28. Find the equation of the image of the line $y = 2x + 3$ under a shear of factor 3 in the y-direction.

29. Construct a 2×2 matrix that defines a shear of factor 3 in the y-direction, followed by a scaling of factor 2 in the x-direction, followed by a reflection about the y-axis.

30. Compute $A + B$ and AB for the following matrices, and show that A is hermitian.

$$A = \begin{bmatrix} 2 & 4 - 3i \\ 4 + 3i & -1 \end{bmatrix},\ B = \begin{bmatrix} 3 + i & 1 - 2i \\ 2 + 7i & -2 + i \end{bmatrix}$$

31. Prove that every real symmetric matrix is hermitian.

32. The following matrix A describes the pottery contents of various graves. Determine possible chronological orderings of the graves and then the pottery types.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

33. The following stochastic matrix P gives the probabilities for a certain region of college- and noncollege-educated households having at least one college-educated child. By a college-educated household we understand that at least one parent is college educated, while by noncollege-educated we mean that neither parent is college educated.

$$\begin{array}{cccl} & \multicolumn{2}{c}{\text{household}} & \\ & \text{coll ed} & \text{noncoll ed} & \\ P = & \left[\begin{array}{c} 0.9 \\ 0.1 \end{array}\right. & \left.\begin{array}{c} 0.25 \\ 0.75 \end{array}\right] & \begin{array}{l} \text{college-educated} \\ \text{noncollege-educated} \end{array} \text{child} \end{array}$$

If there are currently 300,000 college-educated households and 750,000 noncollege-educated households, what is the predicted distribution for two generations hence? What is the probability that a couple that has no college education will have at least one grandchild with a college education?

34. Let A be the adjacency matrix of a digraph. What do you know about the digraph in each of the following cases?

(a) All the elements in the fourth column of A are zero.
(b) The sum of the elements in the third row of A is 2.
(c) The sum of the elements in the second row of A^3 is 4.
(d) The third column of A^2 is all zeros.
(e) The element in the $(4, 4)$ location of A^3 is 2.
(f) The number of nonzero elements in A^4 is 3.