LINEAR ALGEBRA

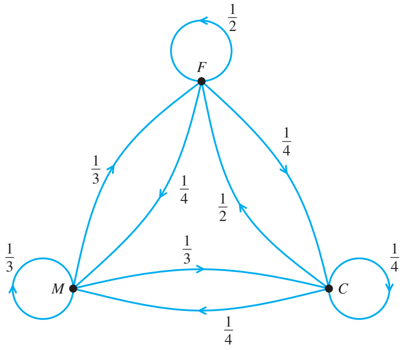
WITH APPLICATIONS

EIGHTH EDITION

STEVEN J. LEON

CHAPTER

1



Matrices and Systems of Equations

Probably the most important problem in mathematics is that of solving a system of linear equations. Well over 75 percent of all mathematical problems encountered in scientific or industrial applications involve solving a linear system at some stage. By using the methods of modern mathematics, it is often possible to take a sophisticated problem and reduce it to a single system of linear equations. Linear systems arise in applications to such areas as business, economics, sociology, ecology, demography, genetics, electronics, engineering, and physics. Therefore, it seems appropriate to begin this book with a section on linear systems.

1.1 Systems of Linear Equations

A *linear equation in n unknowns* is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are real numbers and $x_1, x_2, \ldots, x_n$ are variables. A *linear system* of m equations in n unknowns is then a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

where the a_{ij}'s and the b_i's are all real numbers. We will refer to systems of the form (1) as $m \times n$ linear systems. The following are examples of linear systems:

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

System **(a)** is a 2×2 system, **(b)** is a 2×3 system, and **(c)** is a 3×2 system.

By a solution of an $m \times n$ system, we mean an ordered n-tuple of numbers $(x_1, x_2, \ldots, x_n)$ that satisfies all the equations of the system. For example, the ordered pair $(1, 2)$ is a solution of system **(a)**, since

$$\begin{aligned} 1 \cdot (1) + 2 \cdot (2) &= 5 \\ 2 \cdot (1) + 3 \cdot (2) &= 8 \end{aligned}$$

The ordered triple $(2, 0, 0)$ is a solution of system **(b)**, since

$$\begin{aligned} 1 \cdot (2) - 1 \cdot (0) + 1 \cdot (0) &= 2 \\ 2 \cdot (2) + 1 \cdot (0) - 1 \cdot (0) &= 4 \end{aligned}$$

Actually, system **(b)** has many solutions. If α is any real number, it is easily seen that the ordered triple $(2, \alpha, \alpha)$ is a solution. However, system **(c)** has no solution. It follows from the third equation that the first coordinate of any solution would have to be 4. Using $x_1 = 4$ in the first two equations, we see that the second coordinate must satisfy

$$\begin{aligned} 4 + x_2 &= 2 \\ 4 - x_2 &= 1 \end{aligned}$$

Since there is no real number that satisfies both of these equations, the system has no solution. If a linear system has no solution, we say that the system is *inconsistent*. If the system has at least one solution, we say that it is *consistent*. Thus, system **(c)** is inconsistent, while systems **(a)** and **(b)** are both consistent.

The set of all solutions of a linear system is called the *solution set* of the system. If a system is inconsistent, its solution set is empty. A consistent system will have a nonempty solution set. To solve a consistent system, we must find its solution set.

2×2 Systems

Let us examine geometrically a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Each equation can be represented graphically as a line in the plane. The ordered pair (x_1, x_2) will be a solution of the system if and only if it lies on both lines. For example, consider the three systems

(i) $\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 2 \end{aligned}$ **(ii)** $\begin{aligned} x_1 + x_2 &= 2 \\ x_1 + x_2 &= 1 \end{aligned}$ **(iii)** $\begin{aligned} x_1 + x_2 &= 2 \\ -x_1 - x_2 &= -2 \end{aligned}$

The two lines in system (i) intersect at the point $(2, 0)$. Thus, $\{(2, 0)\}$ is the solution set of (i). In system (ii) the two lines are parallel. Therefore, system (ii) is inconsistent and hence its solution set is empty. The two equations in system (iii) both represent the same line. Any point on this line will be a solution of the system (see Figure 1.1.1).

In general, there are three possibilities: the lines intersect at a point, they are parallel, or both equations represent the same line. The solution set then contains either one, zero, or infinitely many points.

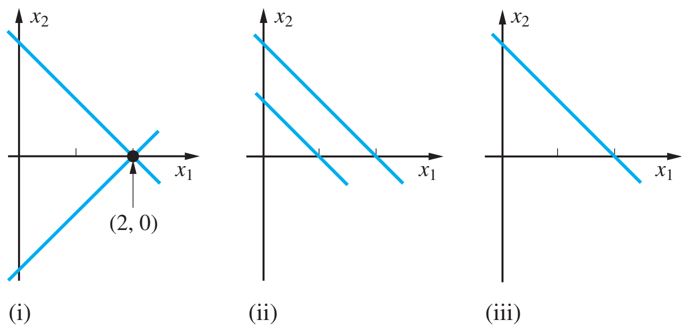


Figure 1.1.1.

The situation is the same for $m \times n$ systems. An $m \times n$ system may or may not be consistent. If it is consistent, it must have either exactly one solution or infinitely many solutions. These are the only possibilities. We will see why this is so in Section 2 when we study the row echelon form. Of more immediate concern is the problem of finding all solutions of a given system. To tackle this problem, we introduce the notion of *equivalent systems*.

Equivalent Systems

Consider the two systems

$$\textbf{(a)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ x_2 &= 3 \\ 2x_3 &= 4 \end{aligned} \qquad\qquad \textbf{(b)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ -3x_1 - x_2 + x_3 &= 5 \\ 3x_1 + 2x_2 + x_3 &= 2 \end{aligned}$$

System **(a)** is easy to solve because it is clear from the last two equations that $x_2 = 3$ and $x_3 = 2$. Using these values in the first equation, we get

$$\begin{aligned} 3x_1 + 2 \cdot 3 - 2 &= -2 \\ x_1 &= -2 \end{aligned}$$

Thus, the solution of the system is $(-2, 3, 2)$. System **(b)** seems to be more difficult to solve. Actually, system **(b)** has the same solution as system **(a)**. To see this, add the first two equations of the system:

$$\begin{array}{r} 3x_1 + 2x_2 - x_3 = -2 \\ -3x_1 - x_2 + x_3 = 5 \\ \hline x_2 = 3 \end{array}$$

If (x_1, x_2, x_3) is any solution of **(b)**, it must satisfy all the equations of the system. Thus, it must satisfy any new equation formed by adding two of its equations. Therefore, x_2 must equal 3. Similarly, (x_1, x_2, x_3) must satisfy the new equation formed by subtracting the first equation from the third:

$$\begin{array}{r} 3x_1 + 2x_2 + x_3 = 2 \\ 3x_1 + 2x_2 - x_3 = -2 \\ \hline 2x_3 = 4 \end{array}$$

Therefore, any solution of system **(b)** must also be a solution of system **(a)**. By a similar argument, it can be shown that any solution of **(a)** is also a solution of **(b)**. This can be done by subtracting the first equation from the second:

$$\begin{array}{rcr} x_2 \qquad\quad & = & 3 \\ 3x_1 + 2x_2 - x_3 & = & -2 \\ \hline -3x_1 - \quad x_2 + x_3 & = & 5 \end{array}$$

Then add the first and third equations:

$$\begin{array}{rcr} 33x_1 + 2x_2 - \quad x_3 & = & -2 \\ 2x_3 & = & 4 \\ \hline 3x_1 + 2x_2 + \quad x_3 & = & 2 \end{array}$$

Thus, (x_1, x_2, x_3) is a solution of system **(b)** if and only if it is a solution of system **(a)**. Therefore, both systems have the same solution set, $\{(-2, 3, 2)\}$.

Definition

Two systems of equations involving the same variables are said to be **equivalent** if they have the same solution set.

Clearly, if we interchange the order in which two equations of a system are written, this will have no effect on the solution set. The reordered system will be equivalent to the original system. For example, the systems

$$\begin{array}{rcl} x_1 + 2x_2 = 4 \\ 3x_1 - \quad x_2 = 2 \\ 4x_1 + \quad x_2 = 6 \end{array} \qquad \text{and} \qquad \begin{array}{rcl} 4x_1 + \quad x_2 = 6 \\ 3x_1 - \quad x_2 = 2 \\ x_1 + 2x_2 = 4 \end{array}$$

both involve the same three equations and, consequently, they must have the same solution set.

If one equation of a system is multiplied through by a nonzero real number, this will have no effect on the solution set, and the new system will be equivalent to the original system. For example, the systems

$$\begin{array}{r} x_1 + x_2 + \quad x_3 = 3 \\ -2x_1 - x_2 + 4x_3 = 1 \end{array} \qquad \text{and} \qquad \begin{array}{r} 2x_1 + 2x_2 + 2x_3 = 6 \\ -2x_1 - \quad x_2 + 4x_3 = 1 \end{array}$$

are equivalent.

If a multiple of one equation is added to another equation, the new system will be equivalent to the original system. This follows since the n-tuple $(x_1, \ldots, x_n)$ will satisfy the two equations

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i \\ a_{j1}x_1 + \cdots + a_{jn}x_n &= b_j \end{aligned}$$

if and only if it satisfies the equations

$$\begin{aligned} a_{i1}x_1 + \cdots + a_{in}x_n &= b_i \\ (a_{j1} + \alpha a_{i1})x_1 + \cdots + (a_{jn} + \alpha a_{in})x_n &= b_j + \alpha b_i \end{aligned}$$

To summarize, there are three operations that can be used on a system to obtain an equivalent system:

I. The order in which any two equations are written may be interchanged.

II. Both sides of an equation may be multiplied by the same nonzero real number.

III. A multiple of one equation may be added to (or subtracted from) another.

Given a system of equations, we may use these operations to obtain an equivalent system that is easier to solve.

$n \times n$ Systems

Let us restrict ourselves to $n \times n$ systems for the remainder of this section. We will show that if an $n \times n$ system has exactly one solution, then operations **I** and **III** can be used to obtain an equivalent "strictly triangular system."

Definition

A system is said to be in **strict triangular form** if, in the kth equation, the coefficients of the first $k - 1$ variables are all zero and the coefficient of x_k is nonzero ($k = 1, \ldots, n$).

EXAMPLE 1 The system

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

is in strict triangular form, since in the second equation the coefficients are 0, 1, -1, respectively, and in the third equation the coefficients are 0, 0, 2, respectively. Because of the strict triangular form, the system is easy to solve. It follows from the third equation that $x_3 = 2$. Using this value in the second equation, we obtain

$$x_2 - 2 = 2 \qquad \text{or} \qquad x_2 = 4$$

Using $x_2 = 4$, $x_3 = 2$ in the first equation, we end up with

$$\begin{aligned} 3x_1 + 2 \cdot 4 + 2 &= 1 \\ x_1 &= -3 \end{aligned}$$

Thus, the solution of the system is $(-3, 4, 2)$. ■

Any $n \times n$ strictly triangular system can be solved in the same manner as the last example. First, the nth equation is solved for the value of x_n. This value is used in the $(n - 1)$st equation to solve for x_{n-1}. The values x_n and x_{n-1} are used in the $(n - 2)$nd equation to solve for x_{n-2}, and so on. We will refer to this method of solving a strictly triangular system as *back substitution*.

EXAMPLE 2 Solve the system

$$\begin{aligned} 2x_1 - x_2 + 3x_3 - 2x_4 &= 1 \\ x_2 - 2x_3 + 3x_4 &= 2 \\ 4x_3 + 3x_4 &= 3 \\ 4x_4 &= 4 \end{aligned}$$

Solution

Using back substitution, we obtain

$$\begin{aligned} 4x_4 &= 4 & x_4 &= 1 \\ 4x_3 + 3 \cdot 1 &= 3 & x_3 &= 0 \\ x_2 - 2 \cdot 0 + 3 \cdot 1 &= 2 & x_2 &= -1 \\ 2x_1 - (-1) + 3 \cdot 0 - 2 \cdot 1 &= 1 & x_1 &= 1 \end{aligned}$$

Thus, the solution is $(1, -1, 0, 1)$. ■

In general, given a system of n linear equations in n unknowns, we will use operations **I** and **III** to try to obtain an equivalent system that is strictly triangular. (We will see in the next section of the book that it is not possible to reduce the system to strictly triangular form in the cases where the system does not have a unique solution.)

EXAMPLE 3 Solve the system

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

Solution

Subtracting 3 times the first row from the second row yields

$$-7x_2 - 6x_3 = -10$$

Subtracting 2 times the first row from the third row yields

$$-x_2 - x_3 = -2$$

If the second and third equations of our system, respectively, are replaced by these new equations, we obtain the equivalent system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -x_2 - x_3 &= -2 \end{aligned}$$

If the third equation of this system is replaced by the sum of the third equation and $-\frac{1}{7}$ times the second equation, we end up with the following strictly triangular system:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -\tfrac{1}{7}x_3 &= -\tfrac{4}{7} \end{aligned}$$

Using back substitution, we get

$$x_3 = 4, \qquad x_2 = -2, \qquad x_1 = 3$$

■

Let us look back at the system of equations in the last example. We can associate with that system a 3×3 array of numbers whose entries are the coefficients of the x_i's:

$$\begin{Bmatrix} 1 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{Bmatrix}$$

We will refer to this array as the *coefficient matrix* of the system. The term *matrix* means simply a rectangular array of numbers. A matrix having m rows and n columns is said to be $m \times n$. A matrix is said to be *square* if it has the same number of rows and columns—that is, if $m = n$.

If we attach to the coefficient matrix an additional column whose entries are the numbers on the right-hand side of the system, we obtain the new matrix

$$\left\{\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right\}$$

We will refer to this new matrix as the *augmented matrix*. In general, when an $m \times r$ matrix B is attached to an $m \times n$ matrix A in this way, the augmented matrix is denoted by $(A|B)$. Thus, if

$$A = \begin{Bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Bmatrix}, \qquad B = \begin{Bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mr} \end{Bmatrix}$$

then

$$(A|B) = \left\{\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1r} \\ \vdots & & & \vdots & & \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mr} \end{array}\right\}$$

With each system of equations, we may associate an augmented matrix of the form

$$\left\{\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right\}$$

The system can be solved by performing operations on the augmented matrix. The x_i's are placeholders that can be omitted until the end of the computation. Corresponding to the three operations used to obtain equivalent systems, the following row operations may be applied to the augmented matrix:

Elementary Row Operations

I. Interchange two rows.
II. Multiply a row by a nonzero real number.
III. Replace a row by its sum with a multiple of another row.

Returning to the example, we find that the first row is used to eliminate the elements in the first column of the remaining rows. We refer to the first row as the *pivotal row*. For emphasis, the entries in the pivotal row are all in bold type and the entire row is color shaded. The first nonzero entry in the pivotal row is called the *pivot*.

$$
\begin{array}{r}
(\text{pivot } a_{11} = 1) \\
\left.\begin{array}{r}\text{entries to be eliminated} \\ a_{21} = 3 \text{ and } a_{31} = 2\end{array}\right\} \rightarrow
\end{array}
\left[\begin{array}{rrr|r}
\mathbf{1} & \mathbf{2} & \mathbf{1} & \mathbf{3} \\
3 & -1 & -3 & -1 \\
2 & 3 & 1 & 4
\end{array}\right]
\begin{array}{l} \leftarrow \text{pivotal row} \\ \\ \\ \end{array}
$$

By using row operation III, 3 times the first row is subtracted from the second row and 2 times the first row is subtracted from the third. When this is done, we end up with the matrix

$$
\left[\begin{array}{rrr|r}
1 & 2 & 1 & 3 \\
\mathbf{0} & \mathbf{-7} & \mathbf{-6} & \mathbf{-10} \\
0 & -1 & -1 & -2
\end{array}\right]
\begin{array}{l} \\ \leftarrow \text{pivotal row} \\ \\ \end{array}
$$

At this step we choose the second row as our new pivotal row and apply row operation III to eliminate the last element in the second column. This time, the pivot is -7 and the quotient $\frac{-1}{-7} = \frac{1}{7}$ is the multiple of the pivotal row that is subtracted from the third row. We end up with the matrix

$$
\left[\begin{array}{rrr|r}
1 & 2 & 1 & 3 \\
0 & -7 & -6 & -10 \\
0 & 0 & -\frac{1}{7} & -\frac{4}{7}
\end{array}\right]
$$

This is the augmented matrix for the strictly triangular system, which is equivalent to the original system. The solution of the system is easily obtained by back substitution.

EXAMPLE 4 Solve the system

$$
\begin{aligned}
4 - x_2 - x_3 + x_4 &= 0 \\
x_1 + x_2 + x_3 + x_4 &= 6 \\
2x_1 + 4x_2 + x_3 - 2x_4 &= -1 \\
3x_1 + x_2 - 2x_3 + 2x_4 &= 3
\end{aligned}
$$

Solution

The augmented matrix for this system is

$$
\left[\begin{array}{rrrr|r}
0 & -1 & -1 & 1 & 0 \\
1 & 1 & 1 & 1 & 6 \\
2 & 4 & 1 & -2 & -1 \\
3 & 1 & -2 & 2 & 3
\end{array}\right]
$$

Since it is not possible to eliminate any entries by using 0 as a pivot element, we will use row operation I to interchange the first two rows of the augmented matrix. The new first row will be the pivotal row and the pivot element will be 1:

$$
(\text{pivot } a_{11} = 1) \quad
\left[\begin{array}{rrrr|r}
\mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{6} \\
0 & -1 & -1 & 1 & 0 \\
2 & 4 & 1 & -2 & -1 \\
3 & 1 & -2 & 2 & 3
\end{array}\right]
\begin{array}{l} \leftarrow \text{pivot row} \\ \\ \\ \\ \end{array}
$$

Row operation III is then used twice to eliminate the two nonzero entries in the first column:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ \mathbf{0} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} & \mathbf{0} \\ 0 & 2 & -1 & -4 & -13 \\ 0 & -2 & -5 & -1 & -15 \end{array}\right]$$

Next, the second row is used as the pivotal row to eliminate the entries in the second column below the pivot element -1:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{-3} & \mathbf{-2} & \mathbf{-13} \\ 0 & 0 & -3 & -3 & -15 \end{array}\right]$$

Finally, the third row is used as the pivotal row to eliminate the last element in the third column:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -3 & -2 & -13 \\ 0 & 0 & 0 & -1 & -2 \end{array}\right]$$

This augmented matrix represents a strictly triangular system. Solving by back substitution, we obtain the solution $(2, -1, 3, 2)$. ■

In general, if an $n \times n$ linear system can be reduced to strictly triangular form, then it will have a unique solution that can be obtained by performing back substitution on the triangular system. We can think of the reduction process as an algorithm involving $n-1$ steps. At the first step, a pivot element is chosen from among the nonzero entries in the first column of the matrix. The row containing the pivot element is called the *pivotal row*. We interchange rows (if necessary) so that the pivotal row is the new first row. Multiples of the pivotal row are then subtracted from each of the remaining $n-1$ rows so as to obtain 0's in the first entries of rows 2 through n. At the second step, a pivot element is chosen from the nonzero entries in column 2, rows 2 through n, of the matrix. The row containing the pivot is then interchanged with the second row of the matrix and is used as the new pivotal row. Multiples of the pivotal row are then subtracted from the remaining $n-2$ rows so as to eliminate all entries below the pivot in the second column. The same procedure is repeated for columns 3 through $n-1$. Note that at the second step row 1 and column 1 remain unchanged, at the third step the first two rows and first two columns remain unchanged, and so on. At each step, the overall dimensions of the system are effectively reduced by 1 (see Figure 1.1.2).

If the elimination process can be carried out as described, we will arrive at an equivalent strictly triangular system after $n-1$ steps. However, the procedure will break down if, at any step, all possible choices for a pivot element are equal to 0. When this happens, the alternative is to reduce the system to certain special echelon, or staircase-shaped, forms. These echelon forms will be studied in the next section. They will also be used for $m \times n$ systems, where $m \neq n$.

Step 1 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right)$

Step 2 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right)$

Step 3 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{array}\right)$

Figure 1.1.2.

SECTION 1.1 EXERCISES

1. Use back substitution to solve each of the following systems of equations:

(a) $\begin{aligned} x_1 - 3x_2 &= 2 \\ 2x_2 &= 6 \end{aligned}$

(b) $\begin{aligned} x_1 + x_2 + x_3 &= 8 \\ 2x_2 + x_3 &= 5 \\ 3x_3 &= 9 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= 5 \\ 3x_2 + x_3 - 2x_4 &= 1 \\ -x_3 + 2x_4 &= -1 \\ 4x_4 &= 4 \end{aligned}$

(d) $\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 5 \\ 2x_2 + x_3 - 2x_4 + x_5 &= 1 \\ 4x_3 + x_4 - 2x_5 &= 1 \\ x_4 - 3x_5 &= 0 \\ 2x_5 &= 2 \end{aligned}$

2. Write out the coefficient matrix for each of the systems in Exercise 1.

3. In each of the following systems, interpret each equation as a line in the plane. For each system, graph the lines and determine geometrically the number of solutions.

(a) $\begin{aligned} x_1 + x_2 &= 4 \\ x_1 - x_2 &= 2 \end{aligned}$

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

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

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

4. Write an augmented matrix for each of the systems in Exercise 3.

5. Write out the system of equations that corresponds to each of the following augmented matrices:

(a) $\left\{\begin{array}{cc|c} 3 & 2 & 8 \\ 1 & 5 & 7 \end{array}\right\}$

(b) $\left\{\begin{array}{ccc|c} 5 & -2 & 1 & 3 \\ 2 & 3 & -4 & 0 \end{array}\right\}$

(c) $\left\{\begin{array}{ccc|c} 2 & 1 & 4 & -1 \\ 4 & -2 & 3 & 4 \\ 5 & 2 & 6 & -1 \end{array}\right\}$

(d) $\left\{\begin{array}{cccc|c} 4 & -3 & 1 & 2 & 4 \\ 3 & 1 & -5 & 6 & 5 \\ 1 & 1 & 2 & 4 & 8 \\ 5 & 1 & 3 & -2 & 7 \end{array}\right\}$

6. Solve each of the following systems:

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

(b) $\begin{aligned} 2x_1 + x_2 &= 8 \\ 4x_1 - 3x_2 &= 6 \end{aligned}$

(c) $\begin{aligned} 4x_1 + 3x_2 &= 4 \\ \tfrac{2}{3}x_1 + 4x_2 &= 3 \end{aligned}$

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

(e) $\begin{aligned} 2x_1 + x_2 + 3x_3 &= 1 \\ 4x_1 + 3x_2 + 5x_3 &= 1 \\ 6x_1 + 5x_2 + 5x_3 &= -3 \end{aligned}$

(f) $\begin{aligned} 3x_1 + 2x_2 + x_3 &= 0 \\ -2x_1 + x_2 - x_3 &= 2 \\ 2x_1 - x_2 + 2x_3 &= -1 \end{aligned}$

(g) $\begin{aligned} \tfrac{1}{3}x_1 + \tfrac{2}{3}x_2 + 2x_3 &= -1 \\ x_1 + 2x_2 + \tfrac{3}{2}x_3 &= \tfrac{3}{2} \\ \tfrac{1}{2}x_1 + 2x_2 + \tfrac{12}{5}x_3 &= \tfrac{1}{10} \end{aligned}$

(h) $\begin{aligned} x_2 + x_3 + x_4 &= 0 \\ 3x_1 + 3x_3 - 4x_4 &= 7 \\ x_1 + x_2 + x_3 + 2x_4 &= 6 \\ 2x_1 + 3x_2 + x_3 + 3x_4 &= 6 \end{aligned}$

7. The two systems

$$\begin{aligned} 2x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &= 5 \end{aligned} \quad \text{and} \quad \begin{aligned} 2x_1 + x_2 &= -1 \\ 4x_1 + 3x_2 &= 1 \end{aligned}$$

have the same coefficient matrix but different right-hand sides. Solve both systems simultaneously by eliminating the first entry in the second row of the augmented matrix

$$\left(\begin{array}{cc|cc} 2 & 1 & 3 & -1 \\ 4 & 3 & 5 & 1 \end{array}\right)$$

and then performing back substitutions for each of the columns corresponding to the right-hand sides.

8. Solve the two systems

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= 9 \end{aligned} \qquad \begin{aligned} x_1 + 2x_2 - 2x_3 &= 9 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= -2 \end{aligned}$$

by doing elimination on a 3×5 augmented matrix and then performing two back substitutions.

9. Given a system of the form

$$\begin{aligned} -m_1x_1 + x_2 &= b_1 \\ -m_2x_1 + x_2 &= b_2 \end{aligned}$$

where m_1, m_2, b_1, and b_2 are constants,

(a) Show that the system will have a unique solution if $m_1 \neq m_2$.

(b) Show that if $m_1 = m_2$, then the system will be consistent only if $b_1 = b_2$.

(c) Give a geometric interpretation of parts (a) and (b).

10. Consider a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

where a_{11}, a_{12}, a_{21}, and a_{22} are constants. Explain why a system of this form must be consistent.

11. Give a geometrical interpretation of a linear equation in three unknowns. Give a geometrical description of the possible solution sets for a 3×3 linear system.

1.2 Row Echelon Form

In Section 1 we learned a method for reducing an $n \times n$ linear system to strict triangular form. However, this method will fail if, at any stage of the reduction process, all the possible choices for a pivot element in a given column are 0.

EXAMPLE 1 Consider the system represented by the augmented matrix

$$\left(\begin{array}{ccccc|c} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 1 & 1 & 2 & 2 & 4 & 1 \end{array}\right) \quad \leftarrow \text{pivotal row}$$

If row operation III is used to eliminate the nonzero entries in the last four rows of the first column, the resulting matrix will be

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{2} & \mathbf{0} \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right) \quad \leftarrow \text{pivotal row}$$

At this stage, the reduction to strict triangular form breaks down. All four possible choices for the pivot element in the second column are 0. How do we proceed from here? Since our goal is to simplify the system as much as possible, it seems natural to move over to the third column and eliminate the last three entries:

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3} \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right)$$

In the fourth column, all the choices for a pivot element are 0; so again we move on to the next column. If we use the third row as the pivotal row, the last two entries in the fifth column are eliminated and we end up with the matrix

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right)$$

The coefficient matrix that we end up with is not in strict triangular form; it is in staircase, or echelon, form. The horizontal and vertical line segments in the array for the coefficient matrix indicate the structure of the staircase form. Note that the vertical drop is 1 for each step, but the horizontal span for a step can be more than 1.

The equations represented by the last two rows are

$$\begin{aligned} 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= -4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= -3 \end{aligned}$$

Since there are no 5-tuples that could satisfy these equations, the system is inconsistent. ■

Suppose now that we change the right-hand side of the system in the last example so as to obtain a consistent system. For example, if we start with

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right)$$

then the reduction process will yield the echelon form augmented matrix

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The last two equations of the reduced system will be satisfied for any 5-tuple. Thus, the solution set will be the set of all 5-tuples satisfying the first three equations,

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ x_3 + x_4 + 2x_5 &= 0 \\ x_5 &= 3 \end{aligned} \tag{1}$$

The variables corresponding to the first nonzero elements in each row of the reduced matrix will be referred to as *lead variables*. Thus, x_1, x_3, and x_5 are the lead variables. The remaining variables corresponding to the columns skipped in the reduction process will be referred to as *free variables*. Hence, x_2 and x_4 are the free variables. If we transfer the free variables over to the right-hand side in (1), we obtain the system

$$\begin{aligned} x_1 + x_3 + x_5 &= 1 - x_2 - x_4 \\ x_3 + 2x_5 &= -x_4 \\ x_5 &= 3 \end{aligned} \tag{2}$$

System (2) is strictly triangular in the unknowns x_1, x_3, and x_5. Thus, for each pair of values assigned to x_2 and x_4, there will be a unique solution. For example, if $x_2 = x_4 = 0$, then $x_5 = 3$, $x_3 = -6$, and $x_1 = 4$, and hence $(4, 0, -6, 0, 3)$ is a solution to the system.

Definition

A matrix is said to be in **row echelon form**

(i) If the first nonzero entry in each nonzero row is 1.

(ii) If row k does not consist entirely of zeros, the number of leading zero entries in row $k + 1$ is greater than the number of leading zero entries in row k.

(iii) If there are rows whose entries are all zero, they are below the rows having nonzero entries.

EXAMPLE 2 The following matrices are in row echelon form:

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

■

EXAMPLE 3 The following matrices are not in row echelon form:

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

The first matrix does not satisfy condition (i). The second matrix fails to satisfy condition (iii), and the third matrix fails to satisfy condition (ii). ■

Definition

The process of using row operations I, II, and III to transform a linear system into one whose augmented matrix is in row echelon form is called **Gaussian elimination**.

Note that row operation II is necessary in order to scale the rows so that the leading coefficients are all 1. If the row echelon form of the augmented matrix contains a row of the form

$$\left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right]$$

the system is inconsistent. Otherwise, the system will be consistent. If the system is consistent and the nonzero rows of the row echelon form of the matrix form a strictly triangular system, the system will have a unique solution.

Overdetermined Systems

A linear system is said to be *overdetermined* if there are more equations than unknowns. Overdetermined systems are *usually* (but not always) inconsistent.

EXAMPLE 4 Solve each of the following overdetermined systems:

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

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

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

Solution

By now the reader should be familiar enough with the elimination process that we can omit the intermediate steps in reducing each of these systems. Thus, we may write

$$\text{System (a):} \quad \left[\begin{array}{rr|r} 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 2 & -2 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right]$$

It follows from the last row of the reduced matrix that the system is inconsistent. The three equations in system (a) represent lines in the plane. The first two lines intersect at the point $(2, -1)$. However, the third line does not pass through this point. Thus, there are no points that lie on all three lines (see Figure 1.2.1).

$$\text{System (b):} \quad \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 2 & -1 & 3 & 5 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Using back substitution, we see that system (b) has exactly one solution: $(0.1, -0.3, 1.5)$. The solution is unique because the nonzero rows of the reduced

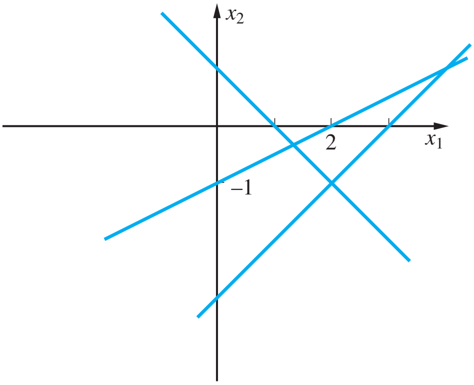


Figure 1.2.1.

matrix form a strictly triangular system.

$$\text{System (c):} \quad \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 3 & 1 & 2 & 3 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Solving for x_2 and x_1 in terms of x_3, we obtain

$$\begin{aligned} x_2 &= -0.2x_3 \\ x_1 &= 1 - 2x_2 - x_3 = 1 - 0.6x_3 \end{aligned}$$

It follows that the solution set is the set of all ordered triples of the form $(1 - 0.6\alpha, -0.2\alpha, \alpha)$, where α is a real number. This system is consistent and has infinitely many solutions because of the free variable x_3. ■

Underdetermined Systems

A system of m linear equations in n unknowns is said to be *underdetermined* if there are fewer equations than unknowns ($m < n$). Although it is possible for underdetermined systems to be inconsistent, they are usually consistent with infinitely many solutions. It is not possible for an underdetermined system to have a unique solution. The reason for this is that any row echelon form of the coefficient matrix will involve $r \leq m$ nonzero rows. Thus, there will be r lead variables and $n - r$ free variables, where $n - r \geq n - m > 0$. If the system is consistent, we can assign the free variables arbitrary values and solve for the lead variables. Therefore, a consistent underdetermined system will have infinitely many solutions.

EXAMPLE 5 Solve the following underdetermined systems:

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

(b)
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 &= 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 &= 2 \end{aligned}$$

Solution

$$\text{System (a):}\quad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 3 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$$

Clearly, system (a) is inconsistent. We can think of the two equations in system (a) and (b) as representing planes in 3-space. Usually, two planes intersect in a line; however, in this case the planes are parallel.

$$\text{System (b):}\quad \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

System (b) is consistent, and since there are two free variables, the system will have infinitely many solutions. In cases such as these it is convenient to continue the elimination process and simplify the form of the reduced matrix even further. We continue eliminating until all the terms above each leading 1 are eliminated. Thus, for system (b), we will continue and eliminate the first two entries in the fifth column and then the first element in the fourth column, as follows:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} \end{array}\right] \rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{2} \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

If we put the free variables over on the right-hand side, it follows that

$$\begin{aligned} x_1 &= 1 - x_2 - x_3 \\ x_4 &= 2 \\ x_5 &= -1 \end{aligned}$$

Thus, for any real numbers α and β, the 5-tuple

$$(1 - \alpha - \beta, \alpha, \beta, 2, -1)$$

is a solution of the system. ■

In the case where the row echelon form of a consistent system has free variables, the standard procedure is to continue the elimination process until all the entries above each leading 1 have been eliminated, as in system (b) of the previous example. The resulting reduced matrix is said to be in *reduced row echelon form*.

Reduced Row Echelon Form

Definition

A matrix is said to be in **reduced row echelon form** if

(i) The matrix is in row echelon form.

(ii) The first nonzero entry in each row is the only nonzero entry in its column.

The following matrices are in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The process of using elementary row operations to transform a matrix into reduced row echelon form is called *Gauss–Jordan reduction*.

EXAMPLE 6 Use Gauss–Jordan reduction to solve the system

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

Solution

$$\left[\begin{array}{rrrr|r} \mathbf{-1} & \mathbf{1} & \mathbf{-1} & \mathbf{3} & \mathbf{0} \\ 3 & 1 & -1 & -1 & 0 \\ 2 & -1 & -2 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} -1 & 1 & -1 & 3 & 0 \\ \mathbf{0} & \mathbf{4} & \mathbf{-4} & \mathbf{8} & \mathbf{0} \\ 0 & 1 & -4 & 5 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrrr|r} -1 & 1 & -1 & 3 & 0 \\ 0 & 4 & -4 & 8 & 0 \\ 0 & 0 & -3 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} & \mathbf{0} \end{array}\right] \quad \text{row echelon form}$$

$$\rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 0 & -2 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ 0 & 0 & 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right] \quad \text{reduced row echelon form}$$

If we set x_4 equal to any real number α, then $x_1 = \alpha$, $x_2 = -\alpha$, and $x_3 = \alpha$. Thus, all ordered 4-tuples of the form $(\alpha, -\alpha, \alpha, \alpha)$ are solutions of the system. ■

APPLICATION 1 Traffic Flow

In the downtown section of a certain city, two sets of one-way streets intersect as shown in Figure 1.2.2. The average hourly volume of traffic entering and leaving this section during rush hour is given in the diagram. Determine the amount of traffic between each of the four intersections.

Solution

At each intersection, the number of automobiles entering must be the same as the number leaving. For example, at intersection A, the number of automobiles entering is $x_1 + 450$ and the number leaving is $x_2 + 610$. Thus,

$$x_1 + 450 = x_2 + 610 \qquad \text{(intersection } A\text{)}$$

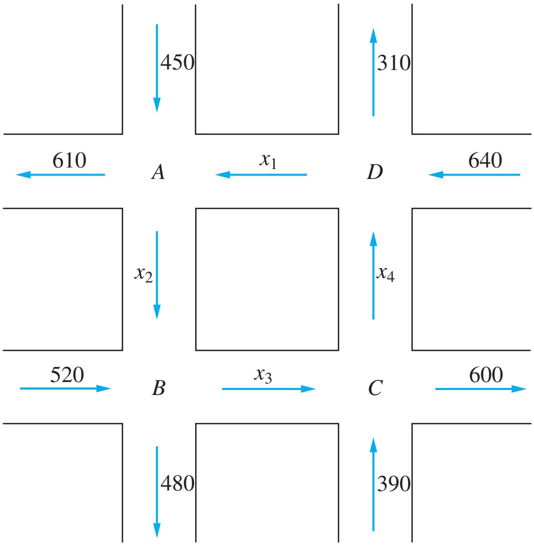


Figure 1.2.2.

Similarly,

$$\begin{aligned} x_2 + 520 &= x_3 + 480 \text{ (intersection } B\text{)} \\ x_3 + 390 &= x_4 + 600 \text{ (intersection } C\text{)} \\ x_4 + 640 &= x_1 + 310 \text{ (intersection } D\text{)} \end{aligned}$$

The augmented matrix for the system is

$$\left(\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 160 \\ 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 1 & -1 & 210 \\ -1 & 0 & 0 & 1 & -330 \end{array}\right)$$

The reduced row echelon form for this matrix is

$$\left(\begin{array}{rrrr|r} 1 & 0 & 0 & -1 & 330 \\ 0 & 1 & 0 & -1 & 170 \\ 0 & 0 & 1 & -1 & 210 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

The system is consistent, and since there is a free variable, there are many possible solutions. The traffic flow diagram does not give enough information to determine x_1, x_2, x_3, and x_4 uniquely. If the amount of traffic were known between any pair of intersections, the traffic on the remaining arteries could easily be calculated. For example, if the amount of traffic between intersections C and D averages 200 automobiles per hour, then $x_4 = 200$. Using this value, we can then solve for x_1, x_2, and x_3:

$$\begin{aligned} x_1 &= x_4 + 330 = 530 \\ x_2 &= x_4 + 170 = 370 \\ x_3 &= x_4 + 210 = 410 \end{aligned}$$

APPLICATION 2 Electrical Networks

In an electrical network, it is possible to determine the amount of current in each branch in terms of the resistances and the voltages. An example of a typical circuit is given in Figure 1.2.3.

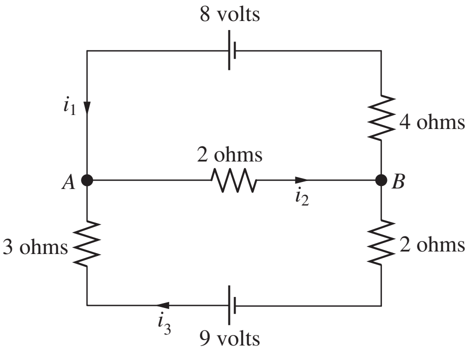
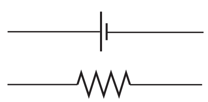


Figure 1.2.3.

The symbols in the figure have the following meanings:



A path along which current may flow

An electrical source

A resistor

The electrical source is usually a battery (with a voltage measured in volts) that drives a charge and produces a current. The current will flow out from the terminal of the battery that is represented by the longer vertical line. The resistances are measured in ohms. The letters represent nodes and the i's represent the currents between the nodes. The currents are measured in amperes. The arrows show the direction of the currents. If, however, one of the currents, say i_2, turns out to be negative, this would mean that the current along that branch is in the direction opposite that of the arrow.

To determine the currents, the following rules are used:

Kirchhoff's Laws

1. At every node, the sum of the incoming currents equals the sum of the outgoing currents.
2. Around every closed loop, the algebraic sum of the voltage gains must equal the algebraic sum of the voltage drops.

The voltage drops E for each resistor are given by *Ohm's law*,

$$E = iR$$

where i represents the current in amperes and R the resistance in ohms.

Let us find the currents in the network pictured in Figure 1.2.3. From the first law, we have

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \qquad \text{(node } A\text{)} \\ -i_1 + i_2 - i_3 &= 0 \qquad \text{(node } B\text{)} \end{aligned}$$

By the second law,

$$\begin{aligned} 4i_1 + 2i_2 &= 8 \qquad \text{(top loop)} \\ 2i_2 + 5i_3 &= 9 \qquad \text{(bottom loop)} \end{aligned}$$

The network can be represented by the augmented matrix

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 4 & 2 & 0 & 8 \\ 0 & 2 & 5 & 9 \end{array}\right]$$

This matrix is easily reduced to row echelon form:

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Solving by back substitution, we see that $i_1 = 1$, $i_2 = 2$, and $i_3 = 1$.

Homogeneous Systems

A system of linear equations is said to be *homogeneous* if the constants on the right-hand side are all zero. Homogeneous systems are always consistent. It is a trivial matter to find a solution; just set all the variables equal to zero. Thus, if an $m \times n$ homogeneous system has a unique solution, it must be the trivial solution $(0, 0, \ldots, 0)$. The homogeneous system in Example 6 consisted of $m = 3$ equations in $n = 4$ unknowns. In the case that $n > m$, there will always be free variables and, consequently, additional nontrivial solutions. This result has essentially been proved in our discussion of underdetermined systems, but, because of its importance, we state it as a theorem.

Theorem 1.2.1 An $m \times n$ homogeneous system of linear equations has a nontrivial solution if $n > m$.

Proof A homogeneous system is always consistent. The row echelon form of the matrix can have at most m nonzero rows. Thus there are at most m lead variables. Since there are n variables altogether and $n > m$, there must be some free variables. The free variables can be assigned arbitrary values. For each assignment of values to the free variables, there is a solution of the system. ∎

APPLICATION 3 Chemical Equations

In the process of photosynthesis, plants use radiant energy from sunlight to convert carbon dioxide (CO_2) and water (H_2O) into glucose ($C_6H_{12}O_6$) and oxygen (O_2). The chemical equation of the reaction is of the form

$$x_1CO_2 + x_2H_2O \rightarrow x_3O_2 + x_4C_6H_{12}O_6$$

To balance the equation, we must choose x_1, x_2, x_3, and x_4 so that the numbers of carbon, hydrogen, and oxygen atoms are the same on each side of the equation. Since carbon dioxide contains one carbon atom and glucose contains six, to balance the carbon atoms we require that

$$x_1 = 6x_4$$

Similarly, to balance the oxygen, we need

$$2x_1 + x_2 = 2x_3 + 6x_4$$

and finally, to balance the hydrogen, we need

$$2x_2 = 12x_4$$

If we move all the unknowns to the left-hand sides of the three equations, we end up with the homogeneous linear system

$$\begin{aligned} x_1 \qquad\qquad\quad - \ \ 6x_4 &= 0 \\ 2x_1 + \ \ x_2 - 2x_3 - \ \ 6x_4 &= 0 \\ 2x_2 \qquad - 12x_4 &= 0 \end{aligned}$$

By Theorem 1.2.1, the system has nontrivial solutions. To balance the equation, we must find solutions (x_1, x_2, x_3, x_4) whose entries are nonnegative integers. If we solve the system in the usual way, we see that x_4 is a free variable and

$$x_1 = x_2 = x_3 = 6x_4$$

In particular, if we take $x_4 = 1$, then $x_1 = x_2 = x_3 = 6$ and the equation takes the form

$$6CO_2 + 6H_2O \rightarrow 6O_2 + C_6H_{12}O_6$$

APPLICATION 4 Economic Models for Exchange of Goods

Suppose that in a primitive society the members of a tribe are engaged in three occupations: farming, manufacturing of tools and utensils, and weaving and sewing of clothing. Assume that initially the tribe has no monetary system and that all goods and services are bartered. Let us denote the three groups by F, M, and C, and suppose that the directed graph in Figure 1.2.4 indicates how the bartering system works in practice.

The figure indicates that the farmers keep half of their produce and give one-fourth of their produce to the manufacturers and one-fourth to the clothing producers. The manufacturers divide the goods evenly among the three groups, one-third going to each group. The group producing clothes gives half of the clothes to the farmers and divides the other half evenly between the manufacturers and themselves. The result is summarized in the following table:

	F	M	C
F	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$
M	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$
C	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$

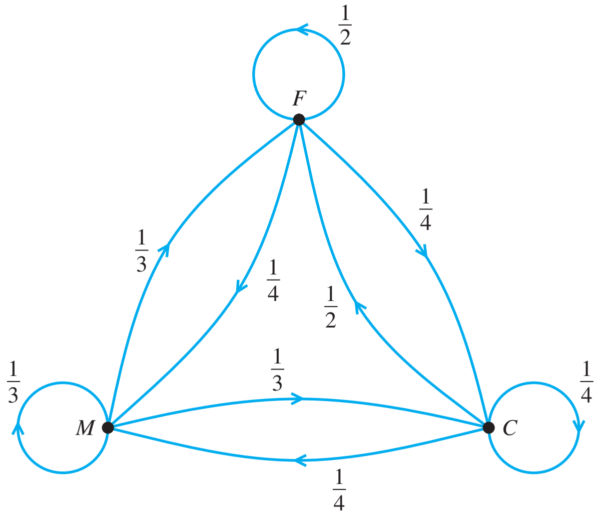


Figure 1.2.4.

The first column of the table indicates the distribution of the goods produced by the farmers, the second column indicates the distribution of the manufactured goods, and the third column indicates the distribution of the clothing.

As the size of the tribe grows, the system of bartering becomes too cumbersome and, consequently, the tribe decides to institute a monetary system of exchange. For this simple economic system, we assume that there will be no accumulation of capital or debt and that the prices for each of the three types of goods will reflect the values of the existing bartering system. The question is how to assign values to the three types of goods that fairly represent the current bartering system.

The problem can be turned into a linear system of equations using an economic model that was originally developed by the Nobel Prize-winning economist Wassily Leontief. For this model, we will let x_1 be the monetary value of the goods produced by the farmers, x_2 be the value of the manufactured goods, and x_3 be the value of all the clothing produced. According to the first row of the table, the value of the goods received by the farmers amounts to half the value of the farm goods produced, plus one-third the value of the manufactured products and half the value of the clothing goods. Thus, the total value of goods received by the farmer is $\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3$. If the system is fair, the total value of goods received by the farmers should equal x_1, the total value of the farm goods produced. Hence, we have the linear equation

$$\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3 = x_1$$

Using the second row of the table and equating the value of the goods produced and received by the manufacturers, we obtain a second equation:

$$\frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = x_2$$

Finally, using the third row of the table, we get

$$\frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = x_3$$

These equations can be rewritten as a homogeneous system:

$$\begin{aligned} -\tfrac{1}{2}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{2}x_3 &= 0 \\ \tfrac{1}{4}x_1 - \tfrac{2}{3}x_2 + \tfrac{1}{4}x_3 &= 0 \\ \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 - \tfrac{3}{4}x_3 &= 0 \end{aligned}$$

The reduced row echelon form of the augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{5}{3} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

There is one free variable: x_3. Setting $x_3 = 3$, we obtain the solution (5, 3, 3), and the general solution consists of all multiples of (5, 3, 3). It follows that the variables x_1, x_2, and x_3 should be assigned values in the ratio

$$x_1 : x_2 : x_3 = 5 : 3 : 3$$

This simple system is an example of the closed Leontief input–output model. Leontief's models are fundamental to our understanding of economic systems. Modern applications would involve thousands of industries and lead to very large linear systems. The Leontief models will be studied in greater detail later, in Section 8 of Chapter 6.

SECTION 1.2 EXERCISES

1. Which of the matrices that follow are in row echelon form? Which are in reduced row echelon form?

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

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

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

(g) $\begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & 3 & 6 \end{bmatrix}$ **(h)** $\begin{bmatrix} 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. The augmented matrices that follow are in row echelon form. For each case, indicate whether the corresponding linear system is consistent. If the system has a unique solution, find it.

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

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

(d) $\left[\begin{array}{ccc|c} 1 & -2 & 2 & -2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$

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

(f) $\left[\begin{array}{ccc|c} 1 & -1 & 3 & 8 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$

3. The augmented matrices that follow are in reduced row echelon form. In each case, find the solution set of the corresponding linear system.

(a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$ (b) $\left[\begin{array}{ccc|c} 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right]$

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

(d) $\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 5 \\ 0 & 0 & 1 & 3 & 4 \end{array}\right]$

(e) $\left[\begin{array}{cccc|c} 1 & 5 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(f) $\left[\begin{array}{ccc|c} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$

4. For each of the systems in Exercise 3, make a list of the lead variables and a second list of the free variables.

5. For each of the systems of equations that follow, use Gaussian elimination to obtain an equivalent system whose coefficient matrix is in row echelon form. Indicate whether the system is consistent. If the system is consistent and involves no free variables, use back substitution to find the unique solution. If the system is consistent and there are free variables, transform it to reduced row echelon form and find all solutions.

(a) $\begin{aligned} x_1 - 2x_2 &= 3 \\ 2x_1 - x_2 &= 9 \end{aligned}$ (b) $\begin{aligned} 2x_1 - 3x_2 &= 5 \\ -4x_1 + 6x_2 &= 8 \end{aligned}$

(c) $\begin{aligned} x_1 + x_2 &= 0 \\ 2x_1 + 3x_2 &= 0 \\ 3x_1 - 2x_2 &= 0 \end{aligned}$

(d) $\begin{aligned} 3x_1 + 2x_2 - x_3 &= 4 \\ x_1 - 2x_2 + 2x_3 &= 1 \\ 11x_1 + 2x_2 + x_3 &= 14 \end{aligned}$

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

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

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

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

(i) $\begin{aligned} -x_1 + 2x_2 - x_3 &= 2 \\ -2x_1 + 2x_2 + x_3 &= 4 \\ 3x_1 + 2x_2 + 2x_3 &= 5 \\ -3x_1 + 8x_2 + 5x_3 &= 17 \end{aligned}$

(j) $\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 1 \\ -x_1 - x_2 + 4x_3 - x_4 &= 6 \\ -2x_1 - 4x_2 + 7x_3 - x_4 &= 1 \end{aligned}$

(k) $\begin{aligned} x_1 + 3x_2 + x_3 + x_4 &= 3 \\ 2x_1 - 2x_2 + x_3 + 2x_4 &= 8 \\ x_1 - 5x_2 + x_4 &= 5 \end{aligned}$

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

6. Use Gauss–Jordan reduction to solve each of the following systems:

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

(b) $\begin{aligned} x_1 + 3x_2 + x_3 + x_4 &= 3 \\ 2x_1 - 2x_2 + x_3 + 2x_4 &= 8 \\ 3x_1 + x_2 + 2x_3 - x_4 &= -1 \end{aligned}$

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

(d) $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ 2x_1 + x_2 - x_3 + 3x_4 &= 0 \\ x_1 - 2x_2 + x_3 + x_4 &= 0 \end{aligned}$

7. Give a geometric explanation of why a homogeneous linear system consisting of two equations in three unknowns must have infinitely many solutions. What are the possible numbers of solutions of a nonhomogeneous 2×3 linear system? Give a geometric explanation of your answer.

8. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ -1 & 4 & 3 & 2 \\ 2 & -2 & a & 3 \end{array}\right)$$

For what values of a will the system have a unique solution?

9. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 5 & 3 & 0 \\ -1 & 1 & \beta & 0 \end{array}\right)$$

(a) Is it possible for the system to be inconsistent? Explain.

(b) For what values of β will the system have infinitely many solutions?

10. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{ccc|c} 1 & 1 & 3 & 2 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & a & b \end{array}\right)$$

(a) For what values of a and b will the system have infinitely many solutions?

(b) For what values of a and b will the system be inconsistent?

11. Given the linear systems

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

solve both systems by incorporating the right-hand sides into a 2×2 matrix B and computing the reduced row echelon form of

$$(A|B) = \left(\begin{array}{cc|cc} 1 & 2 & 2 & 1 \\ 3 & 7 & 8 & 7 \end{array}\right)$$

12. Given the linear systems

(a) $\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ -x_1 - x_2 + 2x_3 &= 3 \\ 2x_1 + 3x_2 &= 0 \end{aligned}$

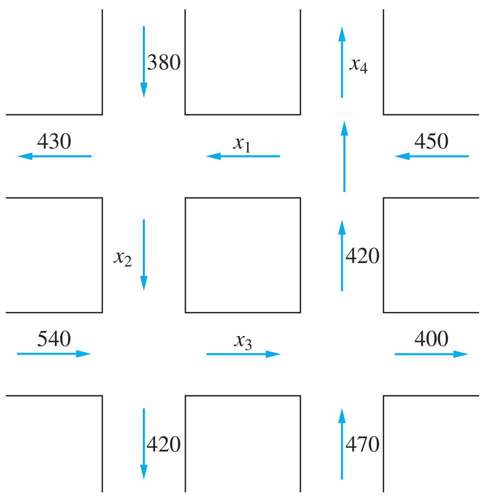
(b) $\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ -x_1 - x_2 + 2x_3 &= 2 \\ 2x_1 + 3x_2 &= -2 \end{aligned}$

solve both systems by computing the row echelon form of an augmented matrix $(A|B)$ and performing back substitution twice.

13. Given a homogeneous system of linear equations, if the system is overdetermined, what are the possibilities as to the number of solutions? Explain.

14. Given a nonhomogeneous system of linear equations, if the system is underdetermined, what are the possibilities as to the number of solutions? Explain.

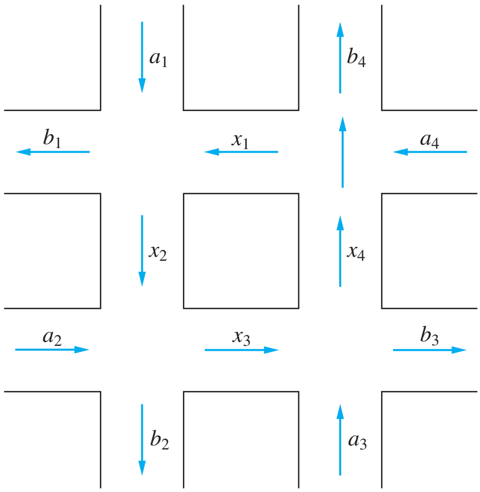
15. Determine the values of x_1, x_2, x_3, and x_4 for the following traffic flow diagram:



16. Consider the traffic flow diagram that follows, where $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4$ are fixed positive integers. Set up a linear system in the unknowns x_1, x_2, x_3, x_4 and show that the system will be consistent if and only if

$$a_1 + a_2 + a_3 + a_4 = b_1 + b_2 + b_3 + b_4$$

What can you conclude about the number of automobiles entering and leaving the traffic network?



17. Let (c_1, c_2) be a solution of the 2×2 system

$$a_{11}x_1 + a_{12}x_2 = 0$$
$$a_{21}x_1 + a_{22}x_2 = 0$$

Show that, for any real number α, the ordered pair $(\alpha c_1, \alpha c_2)$ is also a solution.

18. In Application 3, the solution (6, 6, 6, 1) was obtained by setting the free variable $x_4 = 1$.

(a) Determine the solution corresponding to $x_4 = 0$. What information, if any, does this solution give about the chemical reaction? Is the term "trivial solution" appropriate in this case?

(b) Choose some other values of x_4, such as 2, 4, or 5, and determine the corresponding solutions. How are these nontrivial solutions related?

19. Liquid benzene burns in the atmosphere. If a cold object is placed directly over the benzene, water will condense on the object and a deposit of soot (carbon) will also form on the object. The chemical equation for this reaction is of the form

$$x_1 C_6H_6 + x_2 O_2 \rightarrow x_3 C + x_4 H_2O$$

Determine values of x_1, x_2, x_3, and x_4 to balance the equation.

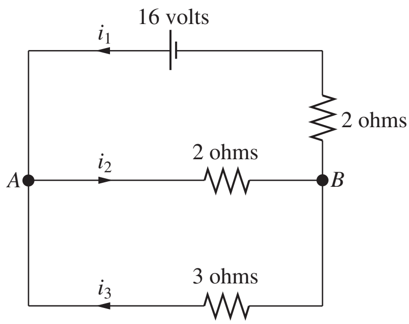
20. Nitric acid is prepared commercially by a series of three chemical reactions. In the first reaction, nitrogen (N_2) is combined with hydrogen (H_2) to form ammonia (NH_3). Next, the ammonia is combined with oxygen (O_2) to form nitrogen dioxide (NO_2) and water. Finally, the NO_2 reacts with some of the water to form nitric acid (HNO_3) and nitric oxide (NO). The amounts of each of the components of these reactions are measured in moles (a standard unit of measurement for chemical reactions). How many moles of nitrogen, hydrogen, and oxygen are necessary to produce 8 moles of nitric acid?

21. In Application 4, determine the relative values of x_1, x_2, and x_3 if the distribution of goods is as described in the following table:

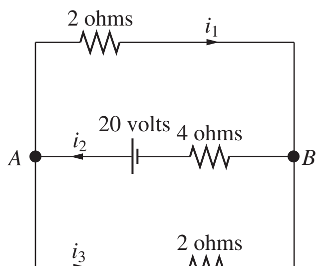
	F	M	C
F	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
M	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$
C	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

22. Determine the amount of each current for the following networks:

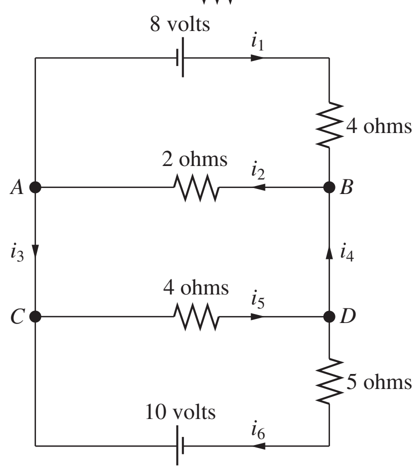
(a)



(b)



(c)



1.3 Matrix Arithmetic

In this section, we introduce the standard notations used for matrices and vectors and define arithmetic operations (addition, subtraction, and multiplication) with matrices. We will also introduce two additional operations: *scalar multiplication* and *transposition*. We will see how to represent linear systems as equations involving matrices and vectors and then derive a theorem characterizing when a linear system is consistent.

The entries of a matrix are called *scalars*. They are usually either real or complex numbers. For the most part, we will be working with matrices whose entries are real numbers. Throughout the first five chapters of the book, the reader may assume that the term *scalar* refers to a real number. However, in Chapter 6 there will be occasions when we will use the set of complex numbers as our scalar field.

Matrix Notation

If we wish to refer to matrices without specifically writing out all their entries, we will use capital letters A, B, C, and so on. In general, a_{ij} will denote the entry of the matrix A that is in the ith row and the jth column. We will refer to this entry as the (i, j) entry of A. Thus, if A is an $m \times n$ matrix, then

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

We will sometimes shorten this to $A = (a_{ij})$. Similarly, a matrix B may be referred to as (b_{ij}), a matrix C as (c_{ij}), and so on.

Vectors

Matrices that have only one row or one column are of special interest, since they are used to represent solutions of linear systems. A solution of a system of m linear equations in n unknowns is an n-tuple of real numbers. We will refer to an n-tuple of real numbers as a *vector*. If an n-tuple is represented in terms of a $1 \times n$ matrix, then we will refer to it as a *row vector*. Alternatively, if the n-tuple is represented by an $n \times 1$ matrix, then we will refer to it as a *column vector*. For example, the solution of the linear system

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

can be represented by the row vector (2, 1) or the column vector $\begin{Bmatrix} 2 \\ 1 \end{Bmatrix}$.

In working with matrix equations, it is generally more convenient to represent the solutions in terms of column vectors ($n \times 1$ matrices). The set of all $n \times 1$ matrices of real numbers is called *Euclidean n-space* and is usually denoted by $\mathbb{R}^n$. Since we will be working almost exclusively with column vectors in the future, we will generally omit the word "column" and refer to the elements of $\mathbb{R}^n$ as simply *vectors*, rather than as column vectors. The standard notation for a column vector is a boldface lowercase letter, as in

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

For row vectors, there is no universal standard notation. In this book, we will represent both row and column vectors with boldface lowercase letters, and to distinguish a row vector from a column vector we will place a horizontal arrow above the letter. Thus, the horizontal arrow indicates a horizontal array (row vector) rather than a vertical array (column vector).

For example,

$$\vec{\mathbf{x}} = (x_1, x_2, x_3, x_4) \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

are row and column vectors with four entries each.

Given an $m \times n$ matrix A, it is often necessary to refer to a particular row or column. The standard notation for the jth column vector of A is $\mathbf{a}_j$. There is no universally accepted standard notation for the ith row vector of a matrix A. In this book, since we use horizontal arrows to indicate row vectors, we denote the ith row vector of A by $\vec{\mathbf{a}}_i$.

If A is an $m \times n$ matrix, then the row vectors of A are given by

$$\vec{\mathbf{a}}_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, \ldots, m$$

and the column vectors are given by

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad j = 1, \ldots, n$$

The matrix A can be represented in terms of either its column vectors or its row vectors:

$$A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \quad \text{or} \quad A = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$$

Similarly, if B is an $n \times r$ matrix, then

$$B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r) = \begin{bmatrix} \vec{\mathbf{b}}_1 \\ \vec{\mathbf{b}}_2 \\ \vdots \\ \vec{\mathbf{b}}_n \end{bmatrix}$$

EXAMPLE 1 If

$$A = \begin{pmatrix} 3 & 2 & 5 \\ -1 & 8 & 4 \end{pmatrix}$$

then

$$\mathbf{a}_1 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 2 \\ 8 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$$

and

$$\vec{\mathbf{a}}_1 = (3, 2, 5), \quad \vec{\mathbf{a}}_2 = (-1, 8, 4)$$

■

Equality

For two matrices to be equal, they must have the same dimensions and their corresponding entries must agree.

Definition

Two $m \times n$ matrices A and B are said to be **equal** if $a_{ij} = b_{ij}$ for each i and j.

Scalar Multiplication

If A is a matrix and α is a scalar, then αA is the matrix formed by multiplying each of the entries of A by α.

Definition

If A is an $m \times n$ matrix and α is a scalar, then αA is the $m \times n$ matrix whose (i, j) entry is αa_{ij}.

For example, if

$$A = \begin{pmatrix} 4 & 8 & 2 \\ 6 & 8 & 10 \end{pmatrix}$$

then

$$\frac{1}{2}A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 4 & 5 \end{pmatrix} \quad \text{and} \quad 3A = \begin{pmatrix} 12 & 24 & 6 \\ 18 & 24 & 30 \end{pmatrix}$$

Matrix Addition

Two matrices with the same dimensions can be added by adding their corresponding entries.

Definition

If $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose (i, j) entry is $a_{ij} + b_{ij}$ for each ordered pair (i, j).

For example,

$$\begin{bmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 4 & 3 \\ 5 & 7 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 2 \\ 1 \\ 8 \end{bmatrix} + \begin{bmatrix} -8 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ 4 \\ 10 \end{bmatrix}$$

If we define $A - B$ to be $A + (-1)B$, then it turns out that $A - B$ is formed by subtracting the corresponding entry of B from each entry of A. Thus,

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

If O represents the matrix, with the same dimensions as A, whose entries are all 0, then

$$A + O = O + A = A$$

We will refer to O as the *zero matrix*. It acts as an additive identity on the set of all $m \times n$ matrices. Furthermore, each $m \times n$ matrix A has an additive inverse. Indeed,

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

It is customary to denote the additive inverse by $-A$. Thus,

$$-A = (-1)A$$

Matrix Multiplication and Linear Systems

We have yet to define the most important operation: the multiplication of two matrices. Much of the motivation behind the definition comes from the applications to linear systems of equations. If we have a system of one linear equation in one unknown, it can be written in the form

$$ax = b \tag{2}$$

We generally think of a, x, and b as being scalars; however, they could also be treated as 1×1 matrices. Our goal now is to generalize equation (2) so that we can represent an $m \times n$ linear system by a single matrix equation of the form

$$A\mathbf{x} = \mathbf{b}$$

where A is an $m \times n$ matrix, $\mathbf{x}$ is an unknown vector in $\mathbb{R}^n$, and $\mathbf{b}$ is in $\mathbb{R}^m$. We consider first the case of one equation in several unknowns.

Case 1. One Equation in Several Unknowns
Let us begin by examining the case of one equation in several variables. Consider, for example, the equation

$$3x_1 + 2x_2 + 5x_3 = 4$$

If we set

$$A = \begin{pmatrix} 3 & 2 & 5 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = \begin{pmatrix} 3 & 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 3x_1 + 2x_2 + 5x_3$$

then the equation $3x_1 + 2x_2 + 5x_3 = 4$ can be written as the matrix equation

$$A\mathbf{x} = 4$$

For a linear equation with n unknowns of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

if we let

$$A = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

then the system can be written in the form $A\mathbf{x} = \mathbf{b}$.

For example, if

$$A = \begin{pmatrix} 2 & 1 & -3 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ -2 \end{pmatrix}$$

then

$$A\mathbf{x} = 2 \cdot 3 + 1 \cdot 2 + (-3) \cdot 1 + 4 \cdot (-2) = -3$$

Note that the result of multiplying a row vector on the left by a column vector on the right is a scalar. Consequently, this type of multiplication is often referred to as a *scalar product*.

Case 2. *M* Equations in *N* Unknowns
Consider now an $m \times n$ linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3}$$

It is desirable to write the system (3) in a form similar to (2)—that is, as a matrix equation

$$A\mathbf{x} = \mathbf{b} \tag{4}$$

where $A = (a_{ij})$ is known, $\mathbf{x}$ is an $n \times 1$ matrix of unknowns, and $\mathbf{b}$ is an $m \times 1$ matrix representing the right-hand side of the system. Thus, if we set

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \tag{5}$$

then the linear system of equations (3) is equivalent to the matrix equation (4).

Given an $m \times n$ matrix A and a vector $\mathbf{x}$ in $\mathbb{R}^n$, it is possible to compute a product $A\mathbf{x}$ by (5). The product $A\mathbf{x}$ will be an $m \times 1$ matrix—that is, a vector in $\mathbb{R}^m$. The rule for determining the ith entry of $A\mathbf{x}$ is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

which is equal to $\vec{\mathbf{a}}_i\mathbf{x}$, the scalar product of the ith row vector of A and the column vector $\mathbf{x}$. Thus,

$$A\mathbf{x} = \begin{bmatrix} \vec{\mathbf{a}}_1\mathbf{x} \\ \vec{\mathbf{a}}_2\mathbf{x} \\ \vdots \\ \vec{\mathbf{a}}_n\mathbf{x} \end{bmatrix}$$

EXAMPLE 2

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 5 & 3 & 7 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} 4x_1 + 2x_2 + x_3 \\ 5x_1 + 3x_2 + 7x_3 \end{bmatrix}$$

■

EXAMPLE 3

$$A = \begin{bmatrix} -3 & 1 \\ 2 & 5 \\ 4 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} -3 \cdot 2 + 1 \cdot 4 \\ 2 \cdot 2 + 5 \cdot 4 \\ 4 \cdot 2 + 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 24 \\ 16 \end{bmatrix}$$ ■

EXAMPLE 4 Write the following system of equations as a matrix equation of the form $A\mathbf{x} = \mathbf{b}$:

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

Solution

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

An alternative way to represent the linear system (3) as a matrix equation is to express the product $A\mathbf{x}$ as a sum of column vectors:

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Thus, we have

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \tag{6}$$

Using this formula, we can represent the system of equations (3) as a matrix equation of the form

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{7}$$

EXAMPLE 5 The linear system

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

can be written as a matrix equation

$$x_1 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$ ■

Definition If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are vectors in $\mathbb{R}^m$ and $c_1, c_2, \ldots, c_n$ are scalars, then a sum of the form

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n$$

is said to be a **linear combination** of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

It follows from equation (6) that the product $A\mathbf{x}$ is a linear combination of the column vectors of A. Some books even use this linear combination representation as the definition of matrix vector multiplication.

If A is an $m \times n$ matrix and $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

EXAMPLE 6 If we choose $x_1 = 2$, $x_2 = 3$, and $x_3 = 4$ in Example 5, then

$$\begin{Bmatrix} 5 \\ 6 \end{Bmatrix} = 2\begin{Bmatrix} 2 \\ 5 \end{Bmatrix} + 3\begin{Bmatrix} 3 \\ -4 \end{Bmatrix} + 4\begin{Bmatrix} -2 \\ 2 \end{Bmatrix}$$

Thus, the vector $\begin{Bmatrix} 5 \\ 6 \end{Bmatrix}$ is a linear combination of the three column vectors of the coefficient matrix. It follows that the linear system in Example 5 is consistent and

$$\mathbf{x} = \begin{Bmatrix} 2 \\ 3 \\ 4 \end{Bmatrix}$$

is a solution of the system. ■

The matrix equation (7) provides a nice way of characterizing whether a linear system of equations is consistent. Indeed, the following theorem is a direct consequence of (7).

Theorem 1.3.1 **Consistency Theorem for Linear Systems**

A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be written as a linear combination of the column vectors of A.

EXAMPLE 7 The linear system

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ 2x_1 + 4x_2 &= 1 \end{aligned}$$

is inconsistent, since the vector $\begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$ cannot be written as a linear combination of the column vectors $\begin{Bmatrix} 1 \\ 2 \end{Bmatrix}$ and $\begin{Bmatrix} 2 \\ 4 \end{Bmatrix}$. Note that any linear combination of these vectors would be of the form

$$x_1\begin{Bmatrix} 1 \\ 2 \end{Bmatrix} + x_2\begin{Bmatrix} 2 \\ 4 \end{Bmatrix} = \begin{Bmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{Bmatrix}$$

and hence the second entry of the vector must be double the first entry. ■

Matrix Multiplication

More generally, it is possible to multiply a matrix A times a matrix B if the number of columns of A equals the number of rows of B. The first column of the product is determined by the first column of B; that is, the first column of AB is $A\mathbf{b}_1$, the second column of AB is $A\mathbf{b}_2$, and so on. Thus the product AB is the matrix whose columns are $A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_n$:

$$AB = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_n)$$

The (i, j) entry of AB is the ith entry of the column vector $A\mathbf{b}_j$. It is determined by multiplying the ith row vector of A times the jth column vector of B.

Definition

If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times r$ matrix, then the product $AB = C = (c_{ij})$ is the $m \times r$ matrix whose entries are defined by

$$c_{ij} = \vec{\mathbf{a}}_i \mathbf{b}_j = \sum_{k=1}^{n} a_{ik} b_{kj}$$

EXAMPLE 8 If

$$A = \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix}$$

then

$$\begin{aligned} AB &= \begin{bmatrix} 3 & -2 \\ \mathbf{2} & \mathbf{4} \\ 1 & -3 \end{bmatrix} \begin{bmatrix} -2 & 1 & \mathbf{3} \\ 4 & 1 & \mathbf{6} \end{bmatrix} \\ &= \begin{bmatrix} 3\cdot(-2) - 2\cdot 4 & 3\cdot 1 - 2\cdot 1 & 3\cdot 3 - 2\cdot 6 \\ 2\cdot(-2) + 4\cdot 4 & 2\cdot 1 + 4\cdot 1 & \mathbf{2\cdot 3 + 4\cdot 6} \\ 1\cdot(-2) - 3\cdot 4 & 1\cdot 1 - 3\cdot 1 & 1\cdot 3 - 3\cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} -14 & 1 & -3 \\ 12 & 6 & \mathbf{30} \\ -14 & -2 & -15 \end{bmatrix} \end{aligned}$$

The shading indicates how the (2, 3) entry of the product AB is computed as a scalar product of the second row vector of A and the third column vector of B. It is also possible to multiply B times A, however, the resulting matrix BA is not equal to AB. In fact, AB and BA do not even have the same dimensions, as the following multiplication shows:

$$\begin{aligned} BA &= \begin{bmatrix} -2\cdot 3 + 1\cdot 2 + 3\cdot 1 & -2\cdot(-2) + 1\cdot 4 + 3\cdot(-3) \\ 4\cdot 3 + 1\cdot 2 + 6\cdot 1 & 4\cdot(-2) + 1\cdot 4 + 6\cdot(-3) \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 \\ 20 & -22 \end{bmatrix} \end{aligned}$$

■

EXAMPLE 9 If

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix}$$

then it is impossible to multiply A times B, since the number of columns of A does not equal the number of rows of B. However, it is possible to multiply B times A.

$$BA = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 17 & 26 \\ 15 & 24 \end{bmatrix}$$ ■

If A and B are both $n \times n$ matrices, then AB and BA will also be $n \times n$ matrices, but, in general, they will not be equal. *Multiplication of matrices is not commutative*.

EXAMPLE 10 If

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

then

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

and

$$BA = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

Hence, $AB \neq BA$. ■

APPLICATION 1 Production Costs

A company manufactures three products. Its production expenses are divided into three categories. In each category, an estimate is given for the cost of producing a single item of each product. An estimate is also made of the amount of each product to be produced per quarter. These estimates are given in Tables 1 and 2. At its stockholders' meeting, the company would like to present a single table showing the total costs for each quarter in each of the three categories: raw materials, labor, and overhead.

Table 1 Production Costs per Item (dollars)

	Product		
Expenses	**A**	**B**	**C**
Raw materials	0.10	0.30	0.15
Labor	0.30	0.40	0.25
Overhead and miscellaneous	0.10	0.20	0.15

Table 2 Amount Produced per Quarter

	Season			
Product	**Summer**	**Fall**	**Winter**	**Spring**
A	4000	4500	4500	4000
B	2000	2600	2400	2200
C	5800	6200	6000	6000

Solution

Let us consider the problem in terms of matrices. Each of the two tables can be represented by a matrix, namely,

$$M = \begin{bmatrix} 0.10 & 0.30 & 0.15 \\ 0.30 & 0.40 & 0.25 \\ 0.10 & 0.20 & 0.15 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 4000 & 4500 & 4500 & 4000 \\ 2000 & 2600 & 2400 & 2200 \\ 5800 & 6200 & 6000 & 6000 \end{bmatrix}$$

If we form the product MP, the first column of MP will represent the costs for the summer quarter:

$$\begin{array}{ll} \text{Raw materials:} & (0.10)(4000) + (0.30)(2000) + (0.15)(5800) = 1870 \\ \text{Labor:} & (0.30)(4000) + (0.40)(2000) + (0.25)(5800) = 3450 \\ \text{Overhead and} & \\ \text{miscellaneous:} & (0.10)(4000) + (0.20)(2000) + (0.15)(5800) = 1670 \end{array}$$

The costs for the fall quarter are given in the second column of MP:

$$\begin{array}{ll} \text{Raw materials:} & (0.10)(4500) + (0.30)(2600) + (0.15)(6200) = 2160 \\ \text{Labor:} & (0.30)(4500) + (0.40)(2600) + (0.25)(6200) = 3940 \\ \text{Overhead and} & \\ \text{miscellaneous:} & (0.10)(4500) + (0.20)(2600) + (0.15)(6200) = 1900 \end{array}$$

Columns 3 and 4 of MP represent the costs for the winter and spring quarters, respectively. Thus, we have

$$MP = \begin{bmatrix} 1870 & 2160 & 2070 & 1960 \\ 3450 & 3940 & 3810 & 3580 \\ 1670 & 1900 & 1830 & 1740 \end{bmatrix}$$

The entries in row 1 of MP represent the total cost of raw materials for each of the four quarters. The entries in rows 2 and 3 represent the total cost for labor and overhead, respectively, for each of the four quarters. The yearly expenses in each category may be obtained by adding the entries in each row. The numbers in each of the columns may be added to obtain the total production costs for each quarter. Table 3 summarizes the total production costs. ■

Table 3

	Season				
	Summer	**Fall**	**Winter**	**Spring**	**Year**
Raw materials	1,870	2,160	2,070	1,960	8,060
Labor	3,450	3,940	3,810	3,580	14,780
Overhead and miscellaneous	1,670	1,900	1,830	1,740	7,140
Total production costs	6,990	8,000	7,710	7,280	29,980

Notational Rules

Just as in ordinary algebra, if an expression involves both multiplication and addition and there are no parentheses to indicate the order of the operations, multiplications are carried out before additions. This is true for both scalar and matrix multiplications. For example, if

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

then

$$A + BC = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 7 & 7 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 11 \\ 0 & 6 \end{bmatrix}$$

and

$$3A + B = \begin{bmatrix} 9 & 12 \\ 3 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 15 \\ 5 & 7 \end{bmatrix}$$

The Transpose of a Matrix

Given an $m \times n$ matrix A, it is often useful to form a new $n \times m$ matrix whose columns are the rows of A.

Definition

The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix B defined by

$$b_{ji} = a_{ij} \tag{8}$$

for $j = 1, \ldots, n$ and $i = 1, \ldots, m$. The transpose of A is denoted by A^T.

It follows from (8) that the jth row of A^T has the same entries, respectively, as the jth column of A, and the ith column of A^T has the same entries, respectively, as the ith row of A.

EXAMPLE 11

(a) If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

(b) If $B = \begin{bmatrix} -3 & 2 & 1 \\ 4 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$, then $B^T = \begin{bmatrix} -3 & 4 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$.

(c) If $C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, then $C^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. ■

The matrix C in Example 11 is its own transpose. This frequently happens with matrices that arise in applications.

Definition

An $n \times n$ matrix A is said to be **symmetric** if $A^T = A$.

The following are some examples of symmetric matrices:

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

APPLICATION 2 Information Retrieval

The growth of digital libraries on the Internet has led to dramatic improvements in the storage and retrieval of information. Modern retrieval methods are based on matrix theory and linear algebra.

In a typical situation, a database consists of a collection of documents and we wish to search the collection and find the documents that best match some particular search conditions. Depending on the type of database, we could search for such items as research articles in journals, Web pages on the Internet, books in a library, or movies in a film collection.

To see how the searches are done, let us assume that our database consists of m documents and that there are n dictionary words that can be used as keywords for searches. Not all words are allowable, since it would not be practical to search for common words such as articles or prepositions. If the key dictionary words are ordered alphabetically, then we can represent the database by an $m \times n$ matrix A. Each document is represented by a column of the matrix. The first entry in the jth column of A would be a number representing the relative frequency of the first key dictionary word in the jth document. The entry a_{2j} represents the relative frequency of the second word in the jth document, and so on. The list of keywords to be used in the search is represented by a vector $\mathbf{x}$ in $\mathbb{R}^m$. The ith entry of $\mathbf{x}$ is taken to be 1 if the ith word in the list of keywords is on our search list; otherwise, we set $x_i = 0$. To carry out the search, we simply multiply A^T times $\mathbf{x}$.

Simple Matching Searches

The simplest type of search determines how many of the key search words are in each document; it does not take into account the relative frequencies of the words. Suppose, for example, that our database consists of these book titles:

B1. *Applied Linear Algebra*

B2. *Elementary Linear Algebra*

B3. *Elementary Linear Algebra with Applications*

B4. *Linear Algebra and Its Applications*

B5. *Linear Algebra with Applications*

B6. *Matrix Algebra with Applications*
B7. *Matrix Theory*

The collection of keywords is given by the following alphabetical list:

algebra, *application*, *elementary*, *linear*, *matrix*, *theory*

For a simple matching search, we just use 0's and 1's, rather than relative frequencies for the entries of the database matrix. Thus, the (i, j) entry of the matrix will be 1 if the ith word appears in the title of the jth book and 0 if it does not. We will assume that our search engine is sophisticated enough to equate various forms of a word. So, for example, in our list of titles the words *applied* and *applications* are both counted as forms of the word *application*. The database matrix for our list of books is the array defined by Table 4.

Table 4 Array Representation for Database of Linear Algebra Books

	Books						
Key Words	**B1**	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**
algebra	1	1	1	1	1	1	0
application	1	0	1	1	1	1	0
elementary	0	1	1	0	0	0	0
linear	1	1	1	1	1	0	0
matrix	0	0	0	0	0	1	1
theory	0	0	0	0	0	0	1

If the words we are searching for are *applied, linear*, and *algebra*, then the database matrix and search vector are respectively given by

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

If we set $\mathbf{y} = A^T\mathbf{x}$, then

$$\mathbf{y} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 3 \\ 3 \\ 3 \\ 2 \\ 0 \end{bmatrix}$$

The value of y_1 is the number of search word matches in the title of the first book, the value of y_2 is the number of matches in the second book title, and so on. Since

$y_1 = y_3 = y_4 = y_5 = 3$, the titles of books B1, B3, B4, and B5 must contain all three search words. If the search is set up to find titles matching all search words, then the search engine will report the titles of the first, third, fourth, and fifth books.

Relative-Frequency Searches

Searches of noncommercial databases generally find all documents containing the key search words and then order the documents based on the relative frequencies of the keywords. In this case, the entries of the database matrix should represent the relative frequencies of the keywords in the documents. For example, suppose that in the dictionary of all key words of the database the 6th word is *algebra* and the 8th word is *applied*, where all words are listed alphabetically. If, say, document 9 in the database contains a total of 200 occurrences of keywords from the dictionary, and if the word *algebra* occurred 10 times in the document and the word *applied* occurred 6 times, then the relative frequencies for these words would be $\frac{10}{200}$ and $\frac{6}{200}$, and the corresponding entries in the database matrix would be

$$a_{69} = 0.05 \quad \text{and} \quad a_{89} = 0.03$$

To search for these two words, we take our search vector $\mathbf{x}$ to be the vector whose entries x_6 and x_8 are both equal to 1 and whose remaining entries are all 0. We then compute

$$\mathbf{y} = A^T\mathbf{x}$$

The entry of $\mathbf{y}$ corresponding to document 9 is

$$y_9 = a_{69} \cdot 1 + a_{89} \cdot 1 = 0.08$$

Note that 16 of the 200 words (8% of the words) in document 9 match the key search words. If y_j is the largest entry of $\mathbf{y}$, this would indicate that the jth document in the database is the one that contains the keywords with the greatest relative frequencies.

Advanced Search Methods

A search for the keywords *linear* and *algebra* could easily turn up hundreds of documents, some of which may not even be about linear algebra. If we were to increase the number of search words and require that all search words be matched, then we would run the risk of excluding some crucial linear algebra documents. Rather than match all words of the expanded search list, our database search should give priority to those documents which match most of the keywords with high relative frequencies. To accomplish this, we need to find the columns of the database matrix A that are "closest" to the search vector $\mathbf{x}$. One way to measure how close two vectors are is to define *the angle between the vectors*. We will do this in Section 1 of Chapter 5.

We will also revisit the information retrieval application after we have learned about the *singular value decomposition* (Chapter 6, Section 5). This decomposition can be used to find a simpler approximation to the database matrix, which will speed up the searches dramatically. Often it has the added advantage of filtering out *noise*; that is, using the approximate version of the database matrix may automatically have the effect of eliminating documents that use keywords in unwanted contexts. For example, a dental student and a mathematics student could both use *calculus* as one of their search words. Since the list of mathematics search words does not contain any other dental terms, a mathematics search using an approximate database matrix is likely to eliminate all documents relating to dentistry. Similarly, the mathematics documents would be filtered out in the dental student's search.

Web Searches and Page Ranking

Modern Web searches could easily involve billions of documents with hundreds of thousands of keywords. Indeed, as of July 2008, there were more than 1 trillion Web pages on the Internet, and it is not uncommon for search engines to acquire or update as many as 10 million Web pages in a single day. Although the database matrix for pages on the Internet is extremely large, searches can be simplified dramatically, since the matrices and search vectors are *sparse*; that is, most of the entries in any column are 0's.

For Internet searches, the better search engines will do simple matching searches to find all pages matching the keywords, but they will not order them on the basis of the relative frequencies of the keywords. Because of the commercial nature of the Internet, people who want to sell products may deliberately make repeated use of keywords to ensure that their Web site is highly ranked in any relative-frequency search. In fact, it is easy to surreptitiously list a keyword hundreds of times. If the font color of the word matches the background color of the page, then the viewer will not be aware that the word is listed repeatedly.

For Web searches, a more sophisticated algorithm is necessary for ranking the pages that contain all of the key search words. In Chapter 6, we will study a special type of matrix model for assigning probabilities in certain random processes. This type of model is referred to as a *Markov process* or a *Markov chain*. In Section 3 of Chapter 6, we will see how to use Markov chains to model Web surfing and obtain rankings of Web pages.

References

1. Berry, Michael W., and Murray Browne, *Understanding Search Engines*: *Mathematical Modeling and Text Retrieval*, SIAM, Philadelphia, 1999.

SECTION 1.3 EXERCISES

1. If

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

compute

(a) $2A$ **(b)** $A + B$

(c) $2A - 3B$ **(d)** $(2A)^T - (3B)^T$

(e) AB **(f)** BA

(g) $A^T B^T$ **(h)** $(BA)^T$

2. For each of the pairs of matrices that follow, determine whether it is possible to multiply the first matrix times the second. If it is possible, perform the multiplication.

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

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

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

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

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

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

3. For which of the pairs in Exercise 2 is it possible to multiply the second matrix times the first, and what would the dimension of the product matrix be?

4. Write each of the following systems of equations as a matrix equation.

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

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

(c)
$$\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 2 \\ 3x_1 - 2x_2 - x_3 &= 0 \end{aligned}$$

5. If

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

verify that

(a) $5A = 3A + 2A$ **(b)** $6A = 3(2A)$

(c) $(A^T)^T = A$

6. If

$$A = \begin{bmatrix} 4 & 1 & 6 \\ 2 & 3 & 5 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 3 & 0 \\ -2 & 2 & -4 \end{bmatrix}$$

verify that

(a) $A + B = B + A$

(b) $3(A + B) = 3A + 3B$

(c) $(A + B)^T = A^T + B^T$

7. If

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}$$

verify that

(a) $3(AB) = (3A)B = A(3B)$

(b) $(AB)^T = B^T A^T$

8. If

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

verify that

(a) $(A + B) + C = A + (B + C)$

(b) $(AB)C = A(BC)$

(c) $A(B + C) = AB + AC$

(d) $(A + B)C = AC + BC$

9. Let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$$

(a) Write $\mathbf{b}$ as a linear combination of the column vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

(b) Use the result from part (a) to determine a solution of the linear system $A\mathbf{x} = \mathbf{b}$. Does the system have any other solutions? Explain.

(c) Write $\mathbf{c}$ as a linear combination of the column vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

10. For each of the choices of A and $\mathbf{b}$ that follow, determine whether the system $A\mathbf{x} = \mathbf{b}$ is consistent by examining how $\mathbf{b}$ relates to the column vectors of A. Explain your answers in each case.

(a) $A = \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$

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

11. Let A be a 5×3 matrix. If

$$\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_2 + \mathbf{a}_3$$

then what can you conclude about the number of solutions of the linear system $A\mathbf{x} = \mathbf{b}$? Explain.

12. Let A be a 3×4 matrix. If

$$\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4$$

then what can you conclude about the number of solutions of the linear system $A\mathbf{x} = \mathbf{b}$? Explain.

13. Let $A\mathbf{x} = \mathbf{b}$ be a linear system whose augmented matrix $(A\,|\,\mathbf{b})$ has reduced row echelon form

$$\left[\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 1 & -2 \\ 0 & 0 & 1 & 2 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

(a) Find all solutions to the system.

(b) If

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{a}_3 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$

determine $\mathbf{b}$.

14. Let A be an $m \times n$ matrix. Explain why the matrix multiplications $A^T A$ and AA^T are possible.

15. A matrix A is said to be *skew symmetric* if $A^T = -A$. Show that if a matrix is skew symmetric, then its diagonal entries must all be 0.

16. In Application 2, suppose that we are searching the database of seven linear algebra books for the search words *elementary, matrix, algebra*. Form a search vector $\mathbf{x}$, and then compute a vector $\mathbf{y}$ that represents the results of the search. Explain the significance of the entries of the vector $\mathbf{y}$.

17. Let A be a 2×2 matrix with $a_{11} \neq 0$ and let $\alpha = a_{21}/a_{11}$. Show that A can be factored into a product of the form

$$\begin{bmatrix} 1 & 0 \\ \alpha & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & b \end{bmatrix}$$

What is the value of b?

1.4 Matrix Algebra

The algebraic rules used for real numbers may or may not work when matrices are used. For example, if a and b are real numbers then

$$a + b = b + a \quad \text{and} \quad ab = ba$$

For real numbers, the operations of addition and multiplication are both commutative. The first of these algebraic rules works when we replace a and b by square matrices A and B; that is,

$$A + B = B + A$$

However, we have already seen that matrix multiplication is not commutative. This fact deserves special emphasis.

Warning: In general, $AB \neq BA$. Matrix multiplication is *not* commutative.

In this section we examine which algebraic rules work for matrices and which do not.

Algebraic Rules

The following theorem provides some useful rules for doing matrix algebra:

Theorem 1.4.1 Each of the following statements is valid for any scalars α and β and for any matrices A, B, and C for which the indicated operations are defined.

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $(AB)C = A(BC)$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AC + BC$
6. $(\alpha\beta)A = \alpha(\beta A)$

(b) If

$$\mathbf{a}_1 = \begin{Bmatrix} 1 \\ 1 \\ 3 \\ 4 \end{Bmatrix} \quad \text{and} \quad \mathbf{a}_3 = \begin{Bmatrix} 2 \\ -1 \\ 1 \\ 3 \end{Bmatrix}$$

determine **b**.

14. Let A be an $m \times n$ matrix. Explain why the matrix multiplications $A^T A$ and AA^T are possible.

15. A matrix A is said to be *skew symmetric* if $A^T = -A$. Show that if a matrix is skew symmetric, then its diagonal entries must all be 0.

16. In Application 2, suppose that we are searching the database of seven linear algebra books for the search words *elementary, matrix, algebra*. Form a search vector **x**, and then compute a vector **y** that represents the results of the search. Explain the significance of the entries of the vector **y**.

17. Let A be a 2×2 matrix with $a_{11} \neq 0$ and let $\alpha = a_{21}/a_{11}$. Show that A can be factored into a product of the form

$$\begin{Bmatrix} 1 & 0 \\ \alpha & 1 \end{Bmatrix} \begin{Bmatrix} a_{11} & a_{12} \\ 0 & b \end{Bmatrix}$$

What is the value of b?

1.4 Matrix Algebra

The algebraic rules used for real numbers may or may not work when matrices are used. For example, if a and b are real numbers then

$$a + b = b + a \quad \text{and} \quad ab = ba$$

For real numbers, the operations of addition and multiplication are both commutative. The first of these algebraic rules works when we replace a and b by square matrices A and B; that is,

$$A + B = B + A$$

However, we have already seen that matrix multiplication is not commutative. This fact deserves special emphasis.

Warning: In general, $AB \neq BA$. Matrix multiplication is *not* commutative.

In this section we examine which algebraic rules work for matrices and which do not.

Algebraic Rules

The following theorem provides some useful rules for doing matrix algebra:

Theorem 1.4.1 Each of the following statements is valid for any scalars α and β and for any matrices A, B, and C for which the indicated operations are defined.

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $(AB)C = A(BC)$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AC + BC$
6. $(\alpha\beta)A = \alpha(\beta A)$

7. $\alpha(AB) = (\alpha A)B = A(\alpha B)$
8. $(\alpha + \beta)A = \alpha A + \beta A$
9. $\alpha(A + B) = \alpha A + \alpha B$

We will prove two of the rules and leave the rest for the reader to verify.

Proof of Rule 4 Assume that $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ and $C = (c_{ij})$ are both $n \times r$ matrices. Let $D = A(B + C)$ and $E = AB + AC$. It follows that

$$d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

and

$$e_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

But

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

so that $d_{ij} = e_{ij}$ and hence $A(B + C) = AB + AC$. ■

Proof of Rule 3 Let A be an $m \times n$ matrix, B an $n \times r$ matrix, and C an $r \times s$ matrix. Let $D = AB$ and $E = BC$. We must show that $DC = AE$. By the definition of matrix multiplication,

$$d_{il} = \sum_{k=1}^{n} a_{ik}b_{kl} \quad \text{and} \quad e_{kj} = \sum_{l=1}^{r} b_{kl}c_{lj}$$

The (i, j) entry of DC is

$$\sum_{l=1}^{r} d_{il}c_{lj} = \sum_{l=1}^{r} \left(\sum_{k=1}^{n} a_{ik}b_{kl} \right) c_{lj}$$

and the (i, j) entry of AE is

$$\sum_{k=1}^{n} a_{ik}e_{kj} = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{r} b_{kl}c_{lj} \right)$$

Since

$$\sum_{l=1}^{r} \left(\sum_{k=1}^{n} a_{ik}b_{kl} \right) c_{lj} = \sum_{l=1}^{r} \left(\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj} \right) = \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{r} b_{kl}c_{lj} \right)$$

it follows that

$$(AB)C = DC = AE = A(BC)$$ ■

The algebraic rules given in Theorem 1.4.1 seem quite natural, since they are similar to the rules that we use with real numbers. However, there are important differences between the rules for matrix algebra and the algebraic rules for real numbers. Some of these differences are illustrated in Exercises 1 through 5 at the end of this section.

EXAMPLE 1 If

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

verify that $A(BC) = (AB)C$ and $A(B + C) = AB + AC$.

Solution

$$A(BC) = \begin{Bmatrix} 1 & 2 \\ 3 & 4 \end{Bmatrix} \begin{Bmatrix} 4 & 1 \\ 1 & 2 \end{Bmatrix} = \begin{Bmatrix} 6 & 5 \\ 16 & 11 \end{Bmatrix}$$

$$(AB)C = \begin{Bmatrix} -4 & 5 \\ -6 & 11 \end{Bmatrix} \begin{Bmatrix} 1 & 0 \\ 2 & 1 \end{Bmatrix} = \begin{Bmatrix} 6 & 5 \\ 16 & 11 \end{Bmatrix}$$

Thus,

$$A(BC) = \begin{Bmatrix} 6 & 5 \\ 16 & 11 \end{Bmatrix} = (AB)C$$

$$A(B + C) = \begin{Bmatrix} 1 & 2 \\ 3 & 4 \end{Bmatrix} \begin{Bmatrix} 3 & 1 \\ -1 & 3 \end{Bmatrix} = \begin{Bmatrix} 1 & 7 \\ 5 & 15 \end{Bmatrix}$$

$$AB + AC = \begin{Bmatrix} -4 & 5 \\ -6 & 11 \end{Bmatrix} + \begin{Bmatrix} 5 & 2 \\ 11 & 4 \end{Bmatrix} = \begin{Bmatrix} 1 & 7 \\ 5 & 15 \end{Bmatrix}$$

Therefore,

$$A(B + C) = AB + AC$$

■

Notation

Since $(AB)C = A(BC)$, we may simply omit the parentheses and write ABC. The same is true for a product of four or more matrices. In the case where an $n \times n$ matrix is multiplied by itself a number of times, it is convenient to use exponential notation. Thus, if k is a positive integer, then

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

EXAMPLE 2 If

$$A = \begin{Bmatrix} 1 & 1 \\ 1 & 1 \end{Bmatrix}$$

then

$$A^2 = \begin{Bmatrix} 1 & 1 \\ 1 & 1 \end{Bmatrix} \begin{Bmatrix} 1 & 1 \\ 1 & 1 \end{Bmatrix} = \begin{Bmatrix} 2 & 2 \\ 2 & 2 \end{Bmatrix}$$

$$A^3 = AAA = AA^2 = \begin{Bmatrix} 1 & 1 \\ 1 & 1 \end{Bmatrix} \begin{Bmatrix} 2 & 2 \\ 2 & 2 \end{Bmatrix} = \begin{Bmatrix} 4 & 4 \\ 4 & 4 \end{Bmatrix}$$

and, in general,

$$A^n = \begin{Bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{Bmatrix}$$

■

APPLICATION 1 A Simple Model for Marital Status Computations

In a certain town, 30 percent of the married women get divorced each year and 20 percent of the single women get married each year. There are 8000 married women and 2000 single women. Assuming that the total population of women remains constant, how many married women and how many single women will there be after 1 year? After 2 years?

Solution

Form a matrix A as follows: The entries in the first row of A will be the percentages of married and single women, respectively, who are married after 1 year. The entries in the second row will be the percentages of women who are single after 1 year. Thus,

$$A = \begin{Bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{Bmatrix}$$

If we let $\mathbf{x} = \begin{Bmatrix} 8000 \\ 2000 \end{Bmatrix}$, the number of married and single women after 1 year can be computed by multiplying A times $\mathbf{x}$.

$$A\mathbf{x} = \begin{Bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{Bmatrix} \begin{Bmatrix} 8000 \\ 2000 \end{Bmatrix} = \begin{Bmatrix} 6000 \\ 4000 \end{Bmatrix}$$

After 1 year, there will be 6000 married women and 4000 single women. To find the number of married and single women after 2 years, compute

$$A^2\mathbf{x} = A(A\mathbf{x}) = \begin{Bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{Bmatrix} \begin{Bmatrix} 6000 \\ 4000 \end{Bmatrix} = \begin{Bmatrix} 5000 \\ 5000 \end{Bmatrix}$$

After 2 years, half of the women will be married and half will be single. In general, the number of married and single women after n years can be determined by computing $A^n\mathbf{x}$. ■

APPLICATION 2 Ecology: Demographics of the Loggerhead Sea Turtle

The management and preservation of many wildlife species depends on our ability to model population dynamics. A standard modeling technique is to divide the life cycle of a species into a number of stages. The models assume that the population sizes for each stage depend only on the female population and that the probability of survival of an individual female from one year to the next depends only on the stage of the life cycle and not on the actual age of an individual. For example, let us consider a four-stage model for analyzing the population dynamics of the loggerhead sea turtle (see Figure 1.4.1).

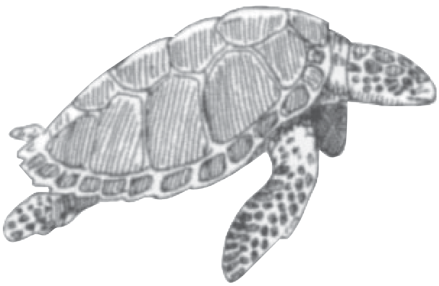


Figure 1.4.1. Loggerhead Sea Turtle

At each stage, we estimate the probability of survival over a 1-year period. We also estimate the ability to reproduce in terms of the expected number of eggs laid in a given year. The results are summarized in Table 1. The approximate ages for each stage are listed in parentheses next to the stage description.

Table 1 Four-Stage Model for Loggerhead Sea Turtle Demographics

Stage Number	Description (age in years)	Annual survivorship	Eggs laid per year
1	Eggs, hatchlings (<1)	0.67	0
2	Juveniles and subadults (1–21)	0.74	0
3	Novice breeders (22)	0.81	127
4	Mature breeders (23–54)	0.81	79

If d_i represents the duration of the ith stage and s_i is the annual survivorship rate for that stage, then it can be shown that the proportion remaining in stage i the following year will be

$$p_i = \left(\frac{1 - s_i^{d_i - 1}}{1 - s_i^{d_i}} \right) s_i \tag{1}$$

and the proportion of the population that will survive and move into stage $i + 1$ the following year will be

$$q_i = \frac{s_i^{d_i}(1 - s_i)}{1 - s_i^{d_i}} \tag{2}$$

If we let e_i denote the average number of eggs laid by a member of stage i ($i = 2, 3, 4$) in 1 year and form the matrix

$$L = \begin{Bmatrix} p_1 & e_2 & e_3 & e_4 \\ q_1 & p_2 & 0 & 0 \\ 0 & q_2 & p_3 & 0 \\ 0 & 0 & q_3 & p_4 \end{Bmatrix} \tag{3}$$

then L can be used to predict the turtle populations at each stage in future years. A matrix of the form (3) is called a *Leslie matrix*, and the corresponding population model is sometimes referred to as a *Leslie population model*. Using the figures from Table 1, the Leslie matrix for our model is

$$L = \begin{Bmatrix} 0 & 0 & 127 & 79 \\ 0.67 & 0.7394 & 0 & 0 \\ 0 & 0.0006 & 0 & 0 \\ 0 & 0 & 0.81 & 0.8077 \end{Bmatrix}$$

Suppose that the initial populations at each stage were 200,000, 300,000, 500, and 1500, respectively. If we represent these initial populations by a vector $\mathbf{x}_0$, the populations at each stage after 1 year are determined with the matrix equation

$$\mathbf{x}_1 = L\mathbf{x}_0 = \begin{Bmatrix} 0 & 0 & 127 & 79 \\ 0.67 & 0.7394 & 0 & 0 \\ 0 & 0.0006 & 0 & 0 \\ 0 & 0 & 0.81 & 0.8077 \end{Bmatrix} \begin{Bmatrix} 200{,}000 \\ 300{,}000 \\ 500 \\ 1{,}500 \end{Bmatrix} = \begin{Bmatrix} 182{,}000 \\ 355{,}820 \\ 180 \\ 1{,}617 \end{Bmatrix}$$

(The computations have been rounded to the nearest integer.) To determine the population vector after 2 years, we multiply again by the matrix L:

$$\mathbf{x}_2 = L\mathbf{x}_1 = L^2\mathbf{x}_0$$

In general, the population after k years is determined by computing $\mathbf{x}_k = L^k\mathbf{x}_0$. To see longer range trends, we compute $\mathbf{x}_{10}$, $\mathbf{x}_{25}$, and $\mathbf{x}_{50}$. The results are summarized in Table 2. The model predicts that the total number of breeding-age turtles will decrease by 80 percent over a 50-year period.

Table 2 Loggerhead Sea Turtle Population Projections

Stage Number	Initial population	10 years	25 years	50 years
1	200,000	114,264	74,039	35,966
2	300,000	329,212	213,669	103,795
3	500	214	139	68
4	1,500	1,061	687	334

A seven-stage model describing the population dynamics is presented in reference [1] to follow. We will use the seven-stage model in the computer exercises at the end of this chapter. Reference [2] is the original paper by Leslie.

References

1. Crouse, Deborah T., Larry B. Crowder, and Hal Caswell, "A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation," *Ecology*, 68(5), 1987.
2. Leslie, P. H., "On the Use of Matrices in Certain Population Mathematics," *Biometrika*, 33, 1945.

The Identity Matrix

Just as the number 1 acts as an identity for the multiplication of real numbers, there is a special matrix I that acts as an identity for matrix multiplication; that is,

$$IA = AI = A \tag{4}$$

for any $n \times n$ matrix A. It is easy to verify that, if we define I to be an $n \times n$ matrix with 1's on the main diagonal and 0's elsewhere, then I satisfies equation (4) for any $n \times n$ matrix A. More formally, we have the following definition:

Definition

The $n \times n$ **identity matrix** is the matrix $I = (\delta_{ij})$, where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

As an example, let us verify equation (4) in the case $n = 3$. We have

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

and

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

In general, if B is any $m \times n$ matrix and C is any $n \times r$ matrix, then

$$BI = B \quad \text{and} \quad IC = C$$

The column vectors of the $n \times n$ identity matrix I are the standard vectors used to define a coordinate system in Euclidean n-space. The standard notation for the jth column vector of I is $\mathbf{e}_j$, rather than the usual $\mathbf{i}_j$. Thus, the $n \times n$ identity matrix can be written

$$I = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$$

Matrix Inversion

A real number a is said to have a multiplicative inverse if there exists a number b such that $ab = 1$. Any nonzero number a has a multiplicative inverse $b = \frac{1}{a}$. We generalize the concept of multiplicative inverses to matrices with the following definition:

Definition

> An $n \times n$ matrix A is said to be **nonsingular** or **invertible** if there exists a matrix B such that $AB = BA = I$. The matrix B is said to be a **multiplicative inverse** of A.

If B and C are both multiplicative inverses of A, then

$$B = BI = B(AC) = (BA)C = IC = C$$

Thus, a matrix can have at most one multiplicative inverse. We will refer to the multiplicative inverse of a nonsingular matrix A as simply the *inverse* of A and denote it by A^{-1}.

EXAMPLE 3 The matrices

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix}$$

are inverses of each other, since

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

■

EXAMPLE 4 The 3×3 matrices

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

are inverses, since

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

and

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

■

EXAMPLE 5 The matrix

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

has no inverse. Indeed, if B is any 2×2 matrix, then

$$BA = \begin{Bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{Bmatrix} \begin{Bmatrix} 1 & 0 \\ 0 & 0 \end{Bmatrix} = \begin{Bmatrix} b_{11} & 0 \\ b_{21} & 0 \end{Bmatrix}$$

Thus, BA cannot equal I. ■

Definition An $n \times n$ matrix is said to be **singular** if it does not have a multiplicative inverse.

Note

Only square matrices have multiplicative inverses. One should not use the terms *singular* and *nonsingular* when referring to nonsquare matrices.

Often we will be working with products of nonsingular matrices. It turns out that any product of nonsingular matrices is nonsingular. The following theorem characterizes how the inverse of the product of a pair of nonsingular matrices A and B is related to the inverses of A and B:

Theorem 1.4.2 If A and B are nonsingular $n \times n$ matrices, then AB is also nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}B = I$$

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I$$

■

It follows by induction that, if $A_1, \ldots, A_k$ are all nonsingular $n \times n$ matrices, then the product $A_1 A_2 \cdots A_k$ is nonsingular and

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1} A_1^{-1}$$

In the next section, we will learn how to determine whether a matrix has a multiplicative inverse. We will also learn a method for computing the inverse of a nonsingular matrix.

Algebraic Rules for Transposes

There are four basic algebraic rules involving transposes:

Algebraic Rules for Transposes

1. $(A^T)^T = A$
2. $(\alpha A)^T = \alpha A^T$
3. $(A + B)^T = A^T + B^T$
4. $(AB)^T = B^T A^T$

The first three rules are straightforward. We leave it to the reader to verify that they are valid. To prove the fourth rule, we need only show that the (i, j) entries of $(AB)^T$ and $B^T A^T$ are equal. If A is an $m \times n$ matrix, then, for the multiplications to be possible, B must have n rows. The (i, j) entry of $(AB)^T$ is the (j, i) entry of AB. It is computed by multiplying the jth row vector of A times the ith column vector of B:

$$\vec{\mathbf{a}}_j \mathbf{b}_i = (a_{j1}, a_{j2}, \ldots, a_{jn}) \begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{bmatrix} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni} \tag{5}$$

The (i, j) entry of $B^T A^T$ is computed by multiplying the ith row of B^T times the jth column of A^T. Since the ith row of B^T is the transpose of the ith column of B and the jth column of A^T is the transpose of the jth row of A, it follows that the (i, j) entry of $B^T A^T$ is given by

$$\mathbf{b}_i^T \vec{\mathbf{a}}_j^T = (b_{1i}, b_{2i}, \ldots, b_{ni}) \begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{bmatrix} = b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ni}a_{jn} \tag{6}$$

It follows from (5) and (6) that the (i, j) entries of $(AB)^T$ and $B^T A^T$ are equal.

The next example illustrates the idea behind the last proof.

EXAMPLE 6 Let

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

Note that, on the one hand, the (3, 2) entry of AB is computed taking the scalar product of the third row of A and the second column of B:

$$AB = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 5 \\ \mathbf{2} & \mathbf{4} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} & 2 \\ 2 & \mathbf{1} & 1 \\ 5 & \mathbf{4} & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 5 \\ 34 & 23 & 14 \\ 15 & \mathbf{8} & 9 \end{bmatrix}$$

When the product is transposed, the (3, 2) entry of AB becomes the (2, 3) entry of $(AB)^T$:

$$(AB)^T = \begin{bmatrix} 10 & 34 & 15 \\ 6 & 23 & \mathbf{8} \\ 5 & 14 & 9 \end{bmatrix}$$

On the other hand, the (2, 3) entry of $B^T A^T$ is computed taking the scalar product of the second row of B^T and the third column of A^T:

$$B^T A^T = \begin{bmatrix} 1 & 2 & 5 \\ \mathbf{0} & \mathbf{1} & \mathbf{4} \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & \mathbf{2} \\ 2 & 3 & \mathbf{4} \\ 1 & 5 & \mathbf{1} \end{bmatrix} = \begin{bmatrix} 10 & 34 & 15 \\ 6 & 23 & \mathbf{8} \\ 5 & 14 & 9 \end{bmatrix}$$

In both cases, the arithmetic for computing the (3, 2) entry is the same. ■

Symmetric Matrices and Networks

Recall that a matrix A is symmetric if $A^T = A$. One type of application that leads to symmetric matrices is problems involving networks. These problems are often solved with the techniques of an area of mathematics called *graph theory*.

APPLICATION 3 Networks and Graphs

Graph theory is an important areas of applied mathematics. It is used to model problems in virtually all the applied sciences. Graph theory is particularly useful in applications involving communication networks.

A *graph* is defined to be a set of points called *vertices*, together with a set of unordered pairs of vertices, which are referred to as *edges*. Figure 1.4.2 gives a geometrical representation of a graph. We can think of the vertices V_1, V_2, V_3, V_4, and V_5 as corresponding to the nodes in a communication network.

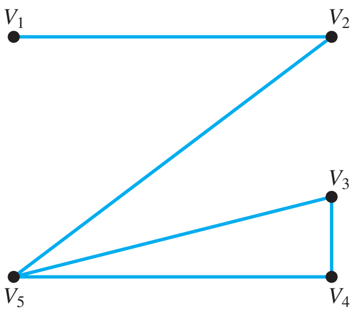


Figure 1.4.2.

The line segments joining the vertices correspond to the edges:

$$\{V_1, V_2\},\ \{V_2, V_5\},\ \{V_3, V_4\},\ \{V_3, V_5\},\ \{V_4, V_5\}$$

Each edge represents a direct communication link between two nodes of the network.

An actual communication network could involve a large number of vertices and edges. Indeed, if there are millions of vertices, a graphical picture of the network would be quite confusing. An alternative is to use a matrix representation for the network. If the graph contains a total of n vertices, we can define an $n \times n$ matrix A by

$$a_{ij} = \begin{cases} 1 & \text{if } \{V_i, V_j\} \text{ is an edge of the graph} \\ 0 & \text{if there is no edge joining } V_i \text{ and } V_j \end{cases}$$

The matrix A is called the *adjacency matrix* of the graph. The adjacency matrix for the graph in Figure 1.4.2 is given by

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

Note that the matrix A is symmetric. Indeed, any adjacency matrix must be symmetric, for if $\{V_i, V_j\}$ is an edge of the graph, then $a_{ij} = a_{ji} = 1$ and $a_{ij} = a_{ji} = 0$ if there is no edge joining V_i and V_j. In either case, $a_{ij} = a_{ji}$.

We can think of a *walk* in a graph as a sequence of edges linking one vertex to another. For example, in Figure 1.4.2 the edges $\{V_1, V_2\}$, $\{V_2, V_5\}$ represent a walk from vertex V_1 to vertex V_5. The *length of the walk* is said to be 2, since it consists of two edges. A simple way to describe the walk is to indicate the movement between vertices by arrows. Thus, $V_1 \to V_2 \to V_5$ denotes a walk of length 2 from V_1 to V_5. Similarly, $V_4 \to V_5 \to V_2 \to V_1$ represents a walk of length 3 from V_4 to V_1. It is possible to traverse the same edges more than once in a walk. For example, $V_5 \to V_3 \to V_5 \to V_3$ is a walk of length 3 from V_5 to V_3. In general, by taking powers of the adjacency matrix, we can determine the number of walks of any specified length between two vertices.

Theorem 1.4.3 If A is an $n \times n$ adjacency matrix of a graph and $a_{ij}^{(k)}$ represents the (i, j) entry of A^k, then $a_{ij}^{(k)}$ is equal to the number of walks of length k from V_i to V_j.

Proof The proof is by mathematical induction. In the case $k = 1$, it follows from the definition of the adjacency matrix that a_{ij} represents the number of walks of length 1 from V_i to V_j. Assume for some m that each entry of A^m is equal to the number of walks of length m between the corresponding vertices. Thus, $a_{il}^{(m)}$ is the number of walks of length m from V_i to V_l. Now, on the one hand, if there is an edge $\{V_l, V_j\}$, then $a_{il}^{(m)} a_{lj} = a_{il}^{(m)}$ is the number of walks of length $m + 1$ from V_i to V_j of the form

$$V_i \to \cdots \to V_l \to V_j$$

On the other hand, if $\{V_l, V_j\}$ is not an edge, then there are no walks of length $m + 1$ of this form from V_i to V_j and

$$a_{il}^{(m)} a_{lj} = a_{il}^{(m)} \cdot 0 = 0$$

It follows that the total number of walks of length $m + 1$ from V_i to V_j is given by

$$a_{i1}^{(m)} a_{1j} + a_{i2}^{(m)} a_{2j} + \cdots + a_{in}^{(m)} a_{nj}$$

But this is just the (i, j) entry of A^{m+1}. ■

EXAMPLE 7 To determine the number of walks of length 3 between any two vertices of the graph in Figure 1.4.2, we need only compute

$$A^3 = \begin{Bmatrix} 0 & 2 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 2 & 4 \\ 0 & 4 & 4 & 4 & 2 \end{Bmatrix}$$

Thus, the number of walks of length 3 from V_3 to V_5 is $a_{35}^{(3)} = 4$. Note that the matrix A^3 is symmetric. This reflects the fact that there are the same number of walks of length 3 from V_i to V_j as there are from V_j to V_i. ■

SECTION 1.4 EXERCISES

1. Explain why each of the following algebraic rules will not work in general when the real numbers a and b are replaced by $n \times n$ matrices A and B.
   (a) $(a+b)^2 = a^2 + 2ab + b^2$
   (b) $(a+b)(a-b) = a^2 - b^2$

2. Will the rules in Exercise 1 work if a is replaced by an $n \times n$ matrix A and b is replaced by the $n \times n$ identity matrix I?

3. Find nonzero 2×2 matrices A and B such that $AB = O$.

4. Find nonzero matrices A, B, and C such that
$$AC = BC \quad \text{and} \quad A \neq B$$

5. The matrix
$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$
has the property that $A^2 = O$. Is it possible for a nonzero symmetric 2×2 matrix to have this property? Prove your answer.

6. Prove the associative law of multiplication for 2×2 matrices; that is, let
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$
$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$$
and show that
$$(AB)C = A(BC)$$

7. Let
$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$
Compute A^2 and A^3. What will A^n turn out to be?

8. Let
$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$
Compute A^2 and A^3. What will A^{2n} and A^{2n+1} turn out to be?

9. Let
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
Show that $A^n = O$ for $n \geq 4$.

10. Let A and B be symmetric $n \times n$ matrices. For each of the following, determine whether the given matrix must be symmetric or could be nonsymmetric:
   (a) $C = A + B$ **(b)** $D = A^2$
   (c) $E = AB$ **(d)** $F = ABA$
   (e) $G = AB + BA$ **(f)** $H = AB - BA$

11. Let C be a nonsymmetric $n \times n$ matrix. For each of the following, determine whether the given matrix must be symmetric or could be nonsymmetric:
   (a) $A = C + C^T$ **(b)** $B = C - C^T$
   (c) $D = C^T C$ **(d)** $E = C^T C - CC^T$
   (e) $F = (I + C)(I + C^T)$
   (f) $G = (I + C)(I - C^T)$

12. Let
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$
Show that if $d = a_{11}a_{22} - a_{21}a_{12} \neq 0$, then
$$A^{-1} = \frac{1}{d}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

13. Use the result from Exercise 12 to find the inverse of each of the following matrices:
   (a) $\begin{pmatrix} 7 & 2 \\ 3 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$ **(c)** $\begin{pmatrix} 4 & 3 \\ 2 & 2 \end{pmatrix}$

14. Let A and B be $n \times n$ matrices. Show that if
$$AB = A \quad \text{and} \quad B \neq I$$
then A must be singular.

15. Let A be a nonsingular matrix. Show that A^{-1} is also nonsingular and $(A^{-1})^{-1} = A$.

16. Prove that if A is nonsingular, then A^T is nonsingular and
$$(A^T)^{-1} = (A^{-1})^T$$
[*Hint*: $(AB)^T = B^T A^T$.]

17. Let A be an $n \times n$ matrix and let $\mathbf{x}$ and $\mathbf{y}$ be vectors in $\mathbb{R}^n$. Show that if $A\mathbf{x} = A\mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$, then the matrix A must be singular.

18. Let A be a nonsingular $n \times n$ matrix. Use mathematical induction to prove that A^m is nonsingular and

$$(A^m)^{-1} = (A^{-1})^m$$

for $m = 1, 2, 3, \ldots$.

19. Let A be an $n \times n$ matrix. Show that if $A^2 = O$, then $I - A$ is nonsingular and $(I - A)^{-1} = I + A$.

20. Let A be an $n \times n$ matrix. Show that if $A^{k+1} = O$, then $I - A$ is nonsingular and

$$(I - A)^{-1} = I + A + A^2 + \cdots + A^k$$

21. Given

$$R = \begin{Bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{Bmatrix}$$

show that R is nonsingular and $R^{-1} = R^T$.

22. An $n \times n$ matrix A is said to be an *involution* if $A^2 = I$. Show that if G is any matrix of the form

$$G = \begin{Bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{Bmatrix}$$

then G is an involution.

23. Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$ (i.e., $\mathbf{u}^T\mathbf{u} = 1$) and let $H = I - 2\mathbf{u}\mathbf{u}^T$. Show that H is an involution.

24. A matrix A is said to be *idempotent* if $A^2 = A$. Show that each of the following matrices are idempotent:

(a) $\begin{Bmatrix} 1 & 0 \\ 1 & 0 \end{Bmatrix}$ **(b)** $\begin{Bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{Bmatrix}$

(c) $\begin{Bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{Bmatrix}$

25. Let A be an idempotent matrix.

(a) Show that $I - A$ is also idempotent.

(b) Show that $I + A$ is nonsingular and $(I + A)^{-1} = I - \frac{1}{2}A$

26. Let D be an $n \times n$ diagonal matrix whose diagonal entries are either 0 or 1.

(a) Show that D is idempotent.

(b) Show that if X is a nonsingular matrix and $A = XDX^{-1}$, then A is idempotent.

27. Let A be an involution matrix, and let

$$B = \frac{1}{2}(I + A) \quad \text{and} \quad C = \frac{1}{2}(I - A)$$

Show that B and C are both idempotent and $BC = O$.

28. Let A be an $m \times n$ matrix. Show that A^TA and AA^T are both symmetric.

29. Let A and B be symmetric $n \times n$ matrices. Prove that $AB = BA$ if and only if AB is also symmetric.

30. Let A be an $n \times n$ matrix and let

$$B = A + A^T \quad \text{and} \quad C = A - A^T$$

(a) Show that B is symmetric and C is skew symmetric.

(b) Show that every $n \times n$ matrix can be represented as a sum of a symmetric matrix and a skew-symmetric matrix.

31. In Application 1, how many married women and how many single women will there be after 3 years?

32. Consider the matrix

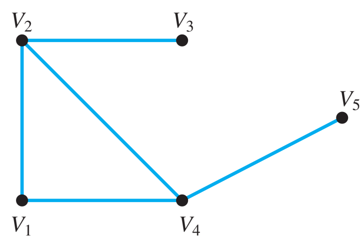
$$A = \begin{Bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{Bmatrix}$$

(a) Draw a graph that has A as its adjacency matrix. Be sure to label the vertices of the graph.

(b) By inspecting the graph, determine the number of walks of length 2 from V_2 to V_3 and from V_2 to V_5.

(c) Compute the second row of A^3, and use it to determine the number of walks of length 3 from V_2 to V_3 and from V_2 to V_5.

33. Consider the graph



(a) Determine the adjacency matrix A of the graph.

(b) Compute A^2. What do the entries in the first row of A^2 tell you about walks of length 2 that start from V_1?

(c) Compute A^3. How many walks of length 3 are there from V_2 to V_4? How many walks of length less than or equal to 3 are there from V_2 to V_4?

For each of the conditional statements that follow, answer true *if the statement is always true and answer* false *otherwise. In the case of a true statement, explain or prove your answer. In the case of a false statement, give an example to show that the statement is not always true.*

34. If $A\mathbf{x} = B\mathbf{x}$ for some nonzero vector $\mathbf{x}$, then the matrices A and B must be equal.

35. If A and B are singular $n \times n$ matrices, then $A + B$ is also singular.

36. If A and B are nonsingular matrices, then $(AB)^T$ is nonsingular and

$$((AB)^T)^{-1} = (A^{-1})^T(B^{-1})^T$$

1.5 Elementary Matrices

In this section, we view the process of solving a linear system in terms of matrix multiplications rather than row operations. Given a linear system $A\mathbf{x} = \mathbf{b}$, we can multiply both sides by a sequence of special matrices to obtain an equivalent system in row echelon form. The special matrices we will use are called *elementary matrices*. We will use them to see how to compute the inverse of a nonsingular matrix and also to obtain an important matrix factorization. We begin by considering the effects of multiplying both sides of a linear system by a nonsingular matrix.

Equivalent Systems

Given an $m \times n$ linear system $A\mathbf{x} = \mathbf{b}$, we can obtain an equivalent system by multiplying both sides of the equation by a nonsingular $m \times m$ matrix M:

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

$$MA\mathbf{x} = M\mathbf{b} \tag{2}$$

Clearly, any solution of (1) will also be a solution of (2). On the other hand, if $\hat{\mathbf{x}}$ is a solution of (2), then

$$M^{-1}(MA\hat{\mathbf{x}}) = M^{-1}(M\mathbf{b})$$

$$A\hat{\mathbf{x}} = \mathbf{b}$$

and it follows that the two systems are equivalent.

To transform the system $A\mathbf{x} = \mathbf{b}$ to a simpler form that is easier to solve, we can apply a sequence of nonsingular matrices $E_1, \ldots, E_k$ to both sides of the equation. The new system will then be of the form

$$U\mathbf{x} = \mathbf{c}$$

where $U = E_k \cdots E_1 A$ and $\mathbf{c} = E_k \cdots E_2 E_1 \mathbf{b}$. The transformed system will be equivalent to the original, provided that $M = E_k \cdots E_1$ is nonsingular. However, M is nonsingular, since it is a product of nonsingular matrices.

We will show next that any of the three elementary row operations can be accomplished by multiplying A on the left by a nonsingular matrix.

Elementary Matrices

If we start with the identity matrix I and then perform exactly one elementary row operation, the resulting matrix is called an *elementary* matrix.

There are three types of elementary matrices corresponding to the three types of elementary row operations.

Type I An elementary matrix of type I is a matrix obtained by interchanging two rows of I.

EXAMPLE 1 The matrix

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix of type I since it was obtained by interchanging the first two rows of I. If A is a 3×3 matrix, then

$$E_1A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$AE_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}$$

Multiplying A on the left by E_1 interchanges the first and second rows of A. Right multiplication of A by E_1 is equivalent to the elementary column operation of interchanging the first and second columns. ■

Type II An elementary matrix of type II is a matrix obtained by multiplying a row of I by a nonzero constant.

EXAMPLE 2

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is an elementary matrix of type II. If A is a 3×3 matrix, then

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 3a_{31} & 3a_{32} & 3a_{33} \end{bmatrix}$$

$$AE_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 3a_{13} \\ a_{21} & a_{22} & 3a_{23} \\ a_{31} & a_{32} & 3a_{33} \end{bmatrix}$$

Multiplication on the left by E_2 performs the elementary row operation of multiplying the third row by 3, while multiplication on the right by E_2 performs the elementary column operation of multiplying the third column by 3. ■

Type III An elementary matrix of type III is a matrix obtained from I by adding a multiple of one row to another row.

EXAMPLE 3

$$E_3 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix of type III. If A is a 3×3 matrix, then

$$E_3 A = \begin{bmatrix} a_{11} + 3a_{31} & a_{12} + 3a_{32} & a_{13} + 3a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$A E_3 = \begin{bmatrix} a_{11} & a_{12} & 3a_{11} + a_{13} \\ a_{21} & a_{22} & 3a_{21} + a_{23} \\ a_{31} & a_{32} & 3a_{31} + a_{33} \end{bmatrix}$$

Multiplication on the left by E_3 adds 3 times the third row to the first row. Multiplication on the right adds 3 times the first column to the third column. ■

In general, suppose that E is an $n \times n$ elementary matrix. We can think of E as being obtained from I by either a row operation or a column operation. If A is an $n \times r$ matrix, *premultiplying A by E has the effect of performing that same row operation on A*. If B is an $m \times n$ matrix, *postmultiplying B by E is equivalent to performing that same column operation on B*.

Theorem 1.5.1 If E is an elementary matrix, then E is nonsingular and E^{-1} is an elementary matrix of the same type.

Proof If E is the elementary matrix of type I formed from I by interchanging the ith and jth rows, then E can be transformed back into I by interchanging these same rows again. Therefore, $EE = I$ and hence E is its own inverse. If E is the elementary matrix of type II formed by multiplying the ith row of I by a nonzero scalar α, then E can be transformed into the identity matrix by multiplying either its ith row or its ith column by $1/\alpha$. Thus,

$$E^{-1} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & 1/\alpha & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \quad i\text{th row}$$

Finally, if E is the elementary matrix of type III formed from I by adding m times the

ith row to the jth row, that is,

$$E = \begin{bmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{matrix} \\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\ \end{matrix}$$

then E can be transformed back into I either by subtracting m times the ith row from the jth row or by subtracting m times the jth column from the ith column. Thus,

$$E^{-1} = \begin{bmatrix} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & -m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix}$$

■

Definition

A matrix B is **row equivalent** to a matrix A if there exists a finite sequence $E_1, E_2, \ldots, E_k$ of elementary matrices such that

$$B = E_k E_{k-1} \cdots E_1 A$$

In other words, B is row equivalent to A if B can be obtained from A by a finite number of row operations. In particular, if two augmented matrices $(A \mid \mathbf{b})$ and $(B \mid \mathbf{c})$ are row equivalent, then $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ are equivalent systems.

The following properties of row equivalent matrices are easily established:

I. If A is row equivalent to B, then B is row equivalent to A.

II. If A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.

Property (I) can be proved using Theorem 1.5.1. The details of the proofs of (I) and (II) are left as an exercise for the reader.

Theorem 1.5.2 Equivalent Conditions for Nonsingularity

Let A be an $n \times n$ matrix. The following are equivalent:

(a) A is nonsingular.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0}$.

(c) A is row equivalent to I.

Proof We prove first that statement (a) implies statement (b). If A is nonsingular and $\hat{\mathbf{x}}$ is a solution of $A\mathbf{x} = \mathbf{0}$, then

$$\hat{\mathbf{x}} = I\hat{\mathbf{x}} = (A^{-1}A)\hat{\mathbf{x}} = A^{-1}(A\hat{\mathbf{x}}) = A^{-1}\mathbf{0} = \mathbf{0}$$

Thus, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Next, we show that statement (b) implies statement (c). If we use elementary row operations, the system can be transformed into the form $U\mathbf{x} = \mathbf{0}$, where U is in row echelon form. If one of the diagonal elements of U were 0, the last row of U would consist entirely of 0's. But then $A\mathbf{x} = \mathbf{0}$ would be equivalent to a system with more unknowns than equations and, hence, by Theorem 1.2.1, would have a nontrivial solution. Thus, U must be a strictly triangular matrix with diagonal elements all equal to 1. It then follows that I is the reduced row echelon form of A and hence A is row equivalent to I.

Finally, we will show that statement (c) implies statement (a). If A is row equivalent to I, there exist elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$A = E_k E_{k-1} \cdots E_1 I = E_k E_{k-1} \cdots E_1$$

But since E_i is invertible, $i = 1, \ldots, k$, the product $E_k E_{k-1} \cdots E_1$ is also invertible. Hence, A is nonsingular and

$$A^{-1} = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$$ ■

Corollary 1.5.3 The system $A\mathbf{x} = \mathbf{b}$ of n linear equations in n unknowns has a unique solution if and only if A is nonsingular.

Proof If A is nonsingular and $\hat{\mathbf{x}}$ is any solution of $A\mathbf{x} = \mathbf{b}$, then

$$A\hat{\mathbf{x}} = \mathbf{b}$$

Multiplying both sides of this equation by A^{-1}, we see that $\hat{\mathbf{x}}$ must be equal to $A^{-1}\mathbf{b}$.

Conversely, if $A\mathbf{x} = \mathbf{b}$ has a unique solution $\hat{\mathbf{x}}$, then we claim that A cannot be singular. Indeed, if A were singular, then the equation $A\mathbf{x} = \mathbf{0}$ would have a solution $\mathbf{z} \neq \mathbf{0}$. But this would imply that $\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$ is a second solution of $A\mathbf{x} = \mathbf{b}$, since

$$A\mathbf{y} = A(\hat{\mathbf{x}} + \mathbf{z}) = A\hat{\mathbf{x}} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Therefore, if $A\mathbf{x} = \mathbf{b}$ has a unique solution, then A must be nonsingular. ■

If A is nonsingular, then A is row equivalent to I and hence there exist elementary matrices $E_1, \ldots, E_k$ such that

$$E_k E_{k-1} \cdots E_1 A = I$$

Multiplying both sides of this equation on the right by A^{-1}, we obtain

$$E_k E_{k-1} \cdots E_1 I = A^{-1}$$

Thus, the same series of elementary row operations that transforms a nonsingular matrix A into I will transform I into A^{-1}. This gives us a method for computing A^{-1}. If we augment A by I and perform the elementary row operations that transform A into I on the augmented matrix, then I will be transformed into A^{-1}. That is, the reduced row echelon form of the augmented matrix $(A|I)$ will be $\left(I|A^{-1}\right)$.

EXAMPLE 4 Compute A^{-1} if

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

Solution

$$\left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right]$$

Thus,

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix}$$ ■

EXAMPLE 5 Solve the system

$$\begin{aligned} x_1 + 4x_2 + 3x_3 &= 12 \\ -x_1 - 2x_2 \qquad &= -12 \\ 2x_1 + 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

The coefficient matrix of this system is the matrix A of the last example. The solution of the system is then

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 12 \\ -12 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -\frac{8}{3} \end{bmatrix}$$ ■

Diagonal and Triangular Matrices

An $n \times n$ matrix A is said to be *upper triangular* if $a_{ij} = 0$ for $i > j$ and *lower triangular* if $a_{ij} = 0$ for $i < j$. Also, A is said to be *triangular* if it is either upper triangular or lower triangular. For example, the 3×3 matrices

$$\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ 6 & 0 & 0 \\ 1 & 4 & 3 \end{bmatrix}$$

are both triangular. The first is upper triangular and the second is lower triangular.

A triangular matrix may have 0's on the diagonal. However, for a linear system $A\mathbf{x} = \mathbf{b}$ to be in strict triangular form, the coefficient matrix A must be upper triangular with nonzero diagonal entries.

An $n \times n$ matrix A is *diagonal* if $a_{ij} = 0$ whenever $i \neq j$. The matrices

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

are all diagonal. A diagonal matrix is both upper triangular and lower triangular.

Triangular Factorization

If an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then it is possible to represent the reduction process in terms of a matrix factorization. We illustrate how this is done in the next example.

EXAMPLE 6 Let

$$A = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix}$$

and let us use only row operation III to carry out the reduction process. At the first step, we subtract $\frac{1}{2}$ times the first row from the second and then we subtract twice the first row from the third.

$$\begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & -9 & 5 \end{bmatrix}$$

To keep track of the multiples of the first row that were subtracted, we set $l_{21} = \frac{1}{2}$ and $l_{31} = 2$. We complete the elimination process by eliminating the -9 in the $(3, 2)$ position:

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

Let $l_{32} = -3$, the multiple of the second row subtracted from the third row. If we call the resulting matrix U and set

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}$$

then it is easily verified that

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{pmatrix} = A \qquad \blacksquare$$

The matrix L in the previous example is lower triangular with 1's on the diagonal. We say that L is *unit lower triangular*. The factorization of the matrix A into a product of a unit lower triangular matrix L times a strictly upper triangular matrix U is often referred to as an *LU factorization*.

To see why the factorization in Example 6 works, let us view the reduction process in terms of elementary matrices. The three row operations that were applied to the matrix A can be represented in terms of multiplications by elementary matrices

$$E_3E_2E_1A = U \tag{3}$$

where

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{pmatrix}$$

correspond to the row operations in the reduction process. Since each of the elementary matrices is nonsingular, we can multiply equation (3) by their inverses.

$$A = E_1^{-1}E_2^{-1}E_3^{-1}U$$

[We multiply in reverse order because $(E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}$.] However, when the inverses are multiplied in this order, the multipliers l_{21}, l_{31}, l_{32} fill in below the diagonal in the product:

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{pmatrix} = L$$

In general, if an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then A has an LU factorization. The matrix L is unit lower triangular, and if $i > j$, then l_{ij} is the multiple of the jth row subtracted from the ith row during the reduction process.

The LU factorization is a very useful way of viewing the elimination process. We will find it particularly useful in Chapter 7 when we study computer methods for solving linear systems. Many of the major topics in linear algebra can be viewed in terms of matrix factorizations. We will study other interesting and important factorizations in Chapters 5 through 7.

SECTION 1.5 EXERCISES

1. Which of the matrices that follow are elementary matrices? Classify each elementary matrix by type.

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

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. Find the inverse of each matrix in Exercise 1. For each elementary matrix, verify that its inverse is an elementary matrix of the same type.

3. For each of the following pairs of matrices, find an elementary matrix E such that $EA = B$:

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

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

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

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

4. For each of the following pairs of matrices, find an elementary matrix E such that $AE = B$:

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

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

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

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

5. Let

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

$$C = \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -3 \\ 2 & 2 & 6 \end{bmatrix}$$

(a) Find an elementary matrix E such that $EA = B$.

(b) Find an elementary matrix F such that $FB = C$.

(c) Is C row equivalent to A? Explain.

6. Let

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

(a) Find elementary matrices E_1, E_2, E_3 such that

$$E_3E_2E_1A = U$$

where U is an upper triangular matrix.

(b) Determine the inverses of E_1, E_2, E_3 and set $L = E_1^{-1}E_2^{-1}E_3^{-1}$. What type of matrix is L? Verify that $A = LU$.

7. Let

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 4 \end{bmatrix}$$

(a) Express A as a product of elementary matrices.

(b) Express A^{-1} as a product of elementary matrices.

8. Compute the LU factorization of each of the following matrices:

(a) $\begin{bmatrix} 3 & 1 \\ 9 & 5 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 4 \\ -2 & 1 \end{bmatrix}$

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

9. Let

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

(a) Verify that

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

(b) Use A^{-1} to solve $A\mathbf{x} = \mathbf{b}$ for the following choices of $\mathbf{b}$:

(i) $\mathbf{b} = (1, 1, 1)^T$ **(ii)** $\mathbf{b} = (1, 2, 3)^T$
(iii) $\mathbf{b} = (-2, 1, 0)^T$

10. Find the inverse of each of the following matrices:

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

(c) $\begin{bmatrix} 2 & 6 \\ 3 & 8 \end{bmatrix}$ **(d)** $\begin{bmatrix} 3 & 0 \\ 9 & 3 \end{bmatrix}$

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

(g) $\begin{pmatrix} -1 & -3 & -3 \\ 2 & 6 & 1 \\ 3 & 8 & 3 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -2 & -3 \end{pmatrix}$

11. Given

$$A = \begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

compute A^{-1} and use it to

(a) find a 2×2 matrix X such that $AX = B$.

(b) find a 2×2 matrix Y such that $YA = B$.

12. Let

$$A = \begin{pmatrix} 5 & 3 \\ 3 & 2 \end{pmatrix}, B = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}, C = \begin{pmatrix} 4 & -2 \\ -6 & 3 \end{pmatrix}$$

Solve each of the following matrix equations:

(a) $AX + B = C$ **(b)** $XA + B = C$

(c) $AX + B = X$ **(d)** $XA + C = X$

13. Is the transpose of an elementary matrix an elementary matrix of the same type? Is the product of two elementary matrices an elementary matrix?

14. Let U and R be $n \times n$ upper triangular matrices and set $T = UR$. Show that T is also upper triangular and that $t_{jj} = u_{jj}r_{jj}$ for $j = 1, \ldots, n$.

15. Let A be a 3×3 matrix and suppose that

$$2\mathbf{a}_1 + \mathbf{a}_2 - 4\mathbf{a}_3 = \mathbf{0}$$

How many solutions will the system $A\mathbf{x} = \mathbf{0}$ have? Explain. Is A nonsingular? Explain.

16. Let A be a 3×3 matrix and suppose that

$$\mathbf{a}_1 = 3\mathbf{a}_2 - 2\mathbf{a}_3$$

Will the system $A\mathbf{x} = \mathbf{0}$ have a nontrivial solution? Is A nonsingular? Explain your answers.

17. Let A and B be $n \times n$ matrices and let $C = A - B$. Show that if $A\mathbf{x}_0 = B\mathbf{x}_0$ and $\mathbf{x}_0 \neq \mathbf{0}$, then C must be singular.

18. Let A and B be $n \times n$ matrices and let $C = AB$. Prove that if B is singular, then C must be singular. [*Hint*: Use Theorem 1.5.2.]

19. Let U be an $n \times n$ upper triangular matrix with nonzero diagonal entries.

(a) Explain why U must be nonsingular.

(b) Explain why U^{-1} must be upper triangular.

20. Let A be a nonsingular $n \times n$ matrix and let B be an $n \times r$ matrix. Show that the reduced row echelon form of $(A|B)$ is $(I|C)$, where $C = A^{-1}B$.

21. In general, matrix multiplication is not commutative (i.e., $AB \neq BA$). However, in certain special cases the commutative property does hold. Show that

(a) if D_1 and D_2 are $n \times n$ diagonal matrices, then $D_1D_2 = D_2D_1$.

(b) if A is an $n \times n$ matrix and

$$B = a_0I + a_1A + a_2A^2 + \cdots + a_kA^k$$

where $a_0, a_1, \ldots, a_k$ are scalars, then $AB = BA$.

22. Show that if A is a symmetric nonsingular matrix, then A^{-1} is also symmetric.

23. Prove that if A is row equivalent to B, then B is row equivalent to A.

24. **(a)** Prove that if A is row equivalent to B and B is row equivalent to C, then A is row equivalent to C.

(b) Prove that any two nonsingular $n \times n$ matrices are row equivalent.

25. Let A and B be $m \times n$ matrices. Prove that if B is row equivalent to A and U is any row echelon form A, then B is row equivalent to U.

26. Prove that B is row equivalent to A if and only if there exists a nonsingular matrix M such that $B = MA$.

27. Is it possible for a singular matrix B to be row equivalent to a nonsingular matrix A? Explain.

28. Given a vector $\mathbf{x} \in \mathbb{R}^{n+1}$, the $(n+1) \times (n+1)$ matrix V defined by

$$v_{ij} = \begin{cases} 1 & \text{if } j = 1 \\ x_i^{j-1} & \text{for } j = 2, \ldots, n+1 \end{cases}$$

is called the Vandermonde matrix.

(a) Show that if

$$V\mathbf{c} = \mathbf{y}$$

and

$$p(x) = c_1 + c_2x + \cdots + c_{n+1}x^n$$

then

$$p(x_i) = y_i, \qquad i = 1, 2, \ldots, n+1$$

(b) Suppose that $x_1, x_2, \ldots, x_{n+1}$ are all distinct. Show that if $\mathbf{c}$ is a solution to $V\mathbf{x} = \mathbf{0}$, then the coefficients $c_1, c_2, \ldots, c_n$ must all be zero and hence V must be nonsingular.

For each of the following, answer true *if the statement is always true and answer* false *otherwise. In the case of a true statement, explain or prove your answer. In the case of a false statement, give an example to show that the statement is not always true.*

29. If A is row equivalent to I and $AB = AC$, then B must equal C.

30. If E and F are elementary matrices and $G = EF$, then G is nonsingular.

31. If A is a 4×4 matrix and $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_3 + 2\mathbf{a}_4$, then A must be singular.

32. If A is row equivalent to both B and C, then A is row equivalent to $B + C$.

1.6 Partitioned Matrices

Often it is useful to think of a matrix as being composed of a number of submatrices. A matrix C can be partitioned into smaller matrices by drawing horizontal lines between the rows and vertical lines between the columns. The smaller matrices are often referred to as *blocks*. For example, let

$$C = \begin{pmatrix} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{pmatrix}$$

If lines are drawn between the second and third rows and between the third and fourth columns, then C will be divided into four submatrices, C_{11}, C_{12}, C_{21}, and C_{22}:

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ \hline 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{array}\right)$$

One useful way of partitioning a matrix is into columns. For example, if

$$B = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{pmatrix}$$

then we can partition B into three column submatrices:

$$B = (\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \left(\begin{array}{c|c|c} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{array}\right)$$

Suppose that we are given a matrix A with three columns; then the product AB can be viewed as a block multiplication. Each block of B is multiplied by A, and the result is a matrix with three blocks: $A\mathbf{b}_1$, $A\mathbf{b}_2$, and $A\mathbf{b}_3$; that is,

$$AB = A(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = (A\mathbf{b}_1, A\mathbf{b}_2, A\mathbf{b}_3)$$

For example, if

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