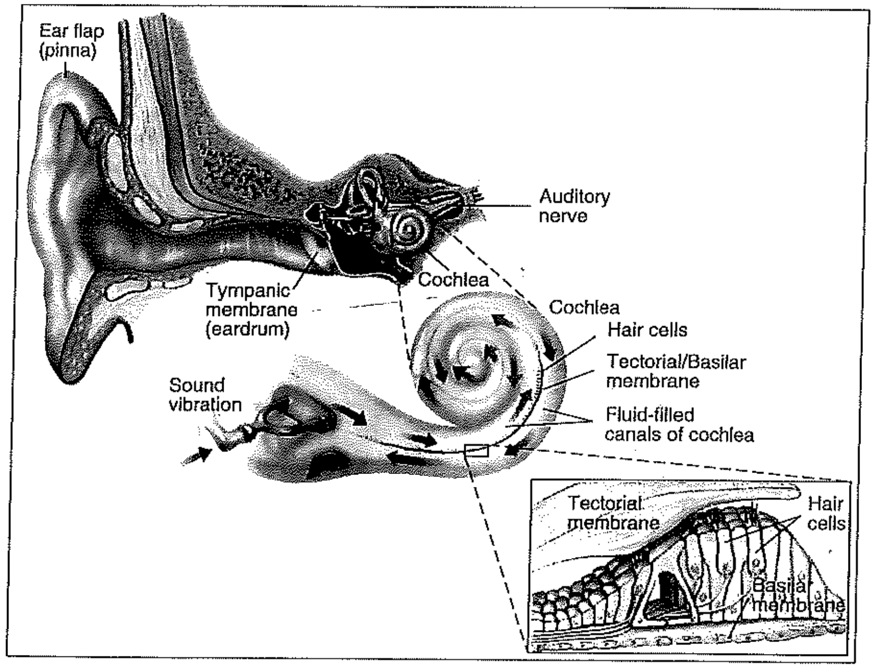
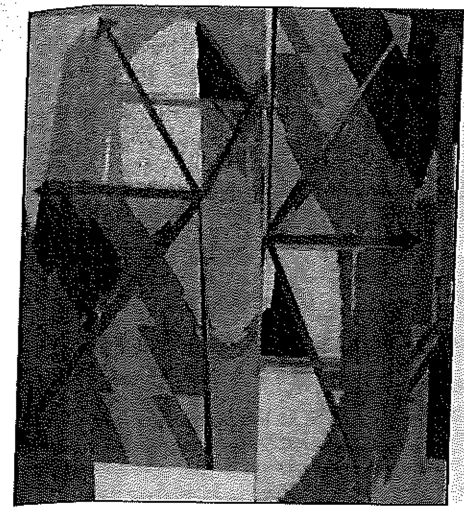
11.19 Harvesting of Animal Populations (p. 764)

Sustainable harvesting of animal populations requires knowledge of the demographics of the population. To maximize the yield of a periodic harvest, different sustainable harvesting strategies can be compared through matrix techniques that describe the population's growth dynamics.



11.20 A Least-Squares Model for Human Hearing (p. 773)

The inner ear contains a structure with thousands of hairlike sensory receptors. These receptors, driven by the vibrations of the eardrum, respond to different frequencies according to their locations and produce electrical impulses that travel to the brain through the auditory nerve. In this way the inner ear acts as a signal processor that decomposes a complicated sound wave into a spectrum of different frequencies.



CHAPTER 1

SYSTEMS OF LINEAR EQUATIONS AND MATRICES

1.1 INTRODUCTION TO SYSTEMS OF LINEAR EQUATIONS

The study of systems of linear equations and their solutions is one of the major topics in linear algebra. In this section we shall introduce some basic terminology and discuss a method for solving such systems.

LINEAR EQUATIONS

A line in the xy-plane can be represented algebraically by an equation of the form

$$a_1x + a_2y = b$$

An equation of this kind is called a linear equation in the variables x and y. More generally, we define a ***linear equation*** in the n variables $x_1, x_2, \ldots, x_n$ to be one that can be expressed in 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 constants. The variables in a linear equation are sometimes called the ***unknowns***.

Example 1 The following are linear equations:

$$x + 3y = 7 \qquad x_1 - 2x_2 - 3x_3 + x_4 = 7$$
$$y = \tfrac{1}{2}x + 3z + 1 \qquad x_1 + x_2 + \cdots + x_n = 1$$

Observe that a linear equation does not involve any products or roots of variables. All variables occur only to the first power and do not appear as arguments for trigonometric, logarithmic, or exponential functions. The following are *not* linear equations:

$$x + 3y^2 = 7 \qquad 3x + 2y - z + xz = 4$$
$$y - \sin x = 0 \qquad \sqrt{x_1} + 2x_2 + x_3 = 1$$

A ***solution*** of a linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ is a sequence of n numbers $s_1, s_2, \ldots, s_n$ such that the equation is satisfied when we substitute $x_1 = s_1$, $x_2 = s_2, \ldots, x_n = s_n$. The set of all solutions of the equation is called its ***solution set*** or sometimes the ***general solution*** of the equation.

Example 2 Find the solution set of

(a) $4x - 2y = 1$ (b) $x_1 - 4x_2 + 7x_3 = 5$

Solution (a). To find solutions of (a), we can assign an arbitrary value to x and solve for y, or choose an arbitrary value for y and solve for x. If we follow the first approach and assign x an arbitrary value t, we obtain

$$x = t, \qquad y = 2t - \tfrac{1}{2}$$

These formulas describe the solution set in terms of the arbitrary parameter t. Particular numerical solutions can be obtained by substituting specific values for t. For example, $t = 3$ yields the solution $x = 3$, $y = \frac{11}{2}$; and $t = -\frac{1}{2}$ yields the solution $x = -\frac{1}{2}$, $y = -\frac{3}{2}$.

If we follow the second approach and assign y the arbitrary value t, we obtain

$$x = \tfrac{1}{2}t + \tfrac{1}{4}, \qquad y = t$$

Although these formulas are different from those obtained above, they yield the same solution set as t varies over all possible real numbers. For example, the previous formulas gave the solution $x = 3, y = \frac{11}{2}$ when $t = 3$, while the formulas immediately above yield that solution when $t = \frac{11}{2}$.

Solution (b). To find the solution set of (b) we can assign arbitrary values to any two variables and solve for the third variable. In particular, if we assign arbitrary values s and t to x_2 and x_3, respectively, and solve for x_1, we obtain

$$x_1 = 5 + 4s - 7t, \qquad x_2 = s, \qquad x_3 = t$$

LINEAR SYSTEMS

A finite set of linear equations in the variables $x_1, x_2, \ldots, x_n$ is called a ***system of linear equations*** or a ***linear system.*** A sequence of numbers $s_1, s_2, \ldots, s_n$ is called a ***solution*** of the system if $x_1 = s_1$, $x_2 = s_2, \ldots, x_n = s_n$ is a solution of every equation in the system. For example, the system

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

has the solution $x_1 = 1$, $x_2 = 2$, $x_3 = -1$ since these values satisfy both equations. However, $x_1 = 1$, $x_2 = 8$, $x_3 = 1$ is not a solution since these values satisfy only the first of the two equations in the system.

Not all systems of linear equations have solutions. For example, if we multiply the second equation of the system

$$\begin{aligned} x + y &= 4 \\ 2x + 2y &= 6 \end{aligned}$$

by $\frac{1}{2}$, it becomes evident that there are no solutions since the resulting equivalent system

$$\begin{aligned} x + y &= 4 \\ x + y &= 3 \end{aligned}$$

has contradictory equations.

A system of equations that has no solutions is said to be ***inconsistent;*** if there is at least one solution of the system, it is called ***consistent.*** To illustrate the possibilities that can occur in solving systems of linear equations, consider a general system of two linear equations in the unknowns x and y:

$$\begin{aligned} a_1x + b_1y &= c_1 \qquad (a_1, b_1 \text{ not both zero}) \\ a_2x + b_2y &= c_2 \qquad (a_2, b_2 \text{ not both zero}) \end{aligned}$$

The graphs of these equations are lines; call them l_1 and l_2. Since a point (x, y) lies on a line if and only if the numbers x and y satisfy the equation of the line, the solutions of the system of equations correspond to points of intersection of l_1 and l_2. There are three possibilities (Figure 1):

- The lines l_1 and l_2 may be parallel, in which case there is no intersection and consequently no solution to the system.
- The lines l_1 and l_2 may intersect at only one point, in which case the system has exactly one solution.
- The lines l_1 and l_2 may coincide, in which case there are infinitely many points of intersection and consequently infinitely many solutions to the system.

Although we have considered only two equations with two unknowns here, we will show later that the same three possibilities hold for arbitrary linear systems:

Every system of linear equations has either no solutions, exactly one solution, or infinitely many solutions.

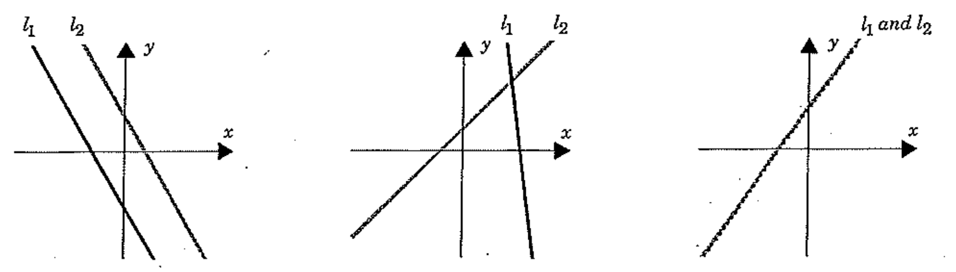


Figure 1 (a) No solution (b) One solution (c) Infinitely many solutions

An arbitrary system of m linear equations in n unknowns can be written as

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

where $x_1, x_2, \ldots, x_n$ are the unknowns and the subscripted a's and b's denote constants. For example, a general system of three linear equations in four unknowns can be written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + a_{14}x_4 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + a_{24}x_4 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + a_{34}x_4 &= b_3 \end{aligned}$$

The double subscripting on the coefficients of the unknowns is a useful device that is used to specify the location of the coefficient in the system. The first subscript on the coefficient a_{ij} indicates the equation in which the coefficient occurs, and the second subscript indicates which unknown it multiplies. Thus, a_{12} is in the first equation and multiplies unknown x_2.

AUGMENTED MATRICES

If we mentally keep track of the location of the +'s, the x's, and the ='s, a system of m linear equations in n unknowns can be abbreviated by writing only the rectangular array of numbers:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

This is called the ***augmented matrix*** for the system. (The term *matrix* is used in mathematics to denote a rectangular array of numbers. Matrices arise in many contexts, which we will consider in more detail in later sections.) For example, the augmented matrix for the system of equations

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 9 \\ 2x_1 + 4x_2 - 3x_3 &= 1 \\ 3x_1 + 6x_2 - 5x_3 &= 0 \end{aligned}$$

is

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

REMARK. When constructing an augmented matrix, the unknowns must be written in the same order in each equation and the constants must be on the right.

The basic method for solving a system of linear equations is to replace the given system by a new system that has the same solution set but which is easier to solve. This new system is generally obtained in a series of steps by applying the following three types of operations to eliminate unknowns systematically.

1. Multiply an equation through by a nonzero constant.
2. Interchange two equations.
3. Add a multiple of one equation to another.

Since the rows (horizontal lines) of an augmented matrix correspond to the equations in the associated system, these three operations correspond to the following operations on the rows of the augmented matrix.

1. Multiply a row through by a nonzero constant.
2. Interchange two rows.
3. Add a multiple of one row to another row.

ELEMENTARY ROW OPERATIONS

These are called ***elementary row operations.*** The following example illustrates how these operations can be used to solve systems of linear equations. Since a systematic procedure for finding solutions will be derived in the next section, it is not necessary to worry about how the steps in this example were selected. The main effort at this time should be devoted to understanding the computations and the discussion.

Example 3 In the left column below we solve a system of linear equations by operating on the equations in the system, and in the right column we solve the same system by operating on the rows of the augmented matrix.

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + 6y - 5z &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

Add -2 times the first equation to the second to obtain | Add -2 times the first row to the second to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3x + 6y - 5z &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

Add -3 times the first equation to the third to obtain | Add -3 times the first row to the third to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3y - 11z &= -27 \end{aligned} \qquad \begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

Multiply the second equation by $\frac{1}{2}$ to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ 3y - 11z &= -27 \end{aligned}$$

Multiply the second row by $\frac{1}{2}$ to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

Add -3 times the second equation to the third to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ -\tfrac{1}{2}z &= -\tfrac{3}{2} \end{aligned}$$

Add -3 times the second row to the third to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

Multiply the third equation by -2 to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ z &= 3 \end{aligned}$$

Multiply the third row by -2 to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Add -1 times the second equation to the first to obtain

$$\begin{aligned} x \qquad + \tfrac{11}{2}z &= \tfrac{35}{2} \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ z &= 3 \end{aligned}$$

Add -1 times the second row to the first to obtain

$$\begin{bmatrix} 1 & 0 & \frac{11}{2} & \frac{35}{2} \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Add $-\frac{11}{2}$ times the third equation to the first and $\frac{7}{2}$ times the third equation to the second to obtain

$$\begin{aligned} x \qquad &= 1 \\ y \quad &= 2 \\ z &= 3 \end{aligned}$$

Add $-\frac{11}{2}$ times the third row to the first and $\frac{7}{2}$ times the third row to the second to obtain

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

The solution

$$x = 1, \quad y = 2, \quad z = 3$$

is now evident.

EXERCISE SET 1.1

1. Which of the following are linear equations in x_1, x_2, and x_3?
   (a) $x_1 + 5x_2 - \sqrt{2}x_3 = 1$ (b) $x_1 + 3x_2 + x_1x_3 = 2$ (c) $x_1 = -7x_2 + 3x_3$
   (d) $x_1^{-2} + x_2 + 8x_3 = 5$ (e) $x_1^{3/5} - 2x_2 + x_3 = 4$ (f) $\pi x_1 - \sqrt{2}x_2 + \frac{1}{3}x_3 = 7^{1/3}$

2. Given that k is a constant, which of the following are linear equations?
   (a) $x_1 - x_2 + x_3 = \sin k$ (b) $kx_1 - \dfrac{1}{k}x_2 = 9$ (c) $2^k x_1 + 7x_2 - x_3 = 0$

3. Find the solution set of each of the following linear equations.
   (a) $7x - 5y = 3$ (b) $3x_1 - 5x_2 + 4x_3 = 7$
   (c) $-8x_1 + 2x_2 - 5x_3 + 6x_4 = 1$ (d) $3v - 8w + 2x - y + 4z = 0$

4. Find the augmented matrix for each of the following systems of linear equations.

   (a) $\begin{aligned} 3x_1 - 2x_2 &= -1 \\ 4x_1 + 5x_2 &= 3 \\ 7x_1 + 3x_2 &= 2 \end{aligned}$ (b) $\begin{aligned} 2x_1 \qquad + 2x_3 &= 1 \\ 3x_1 - x_2 + 4x_3 &= 7 \\ 6x_1 + x_2 - x_3 &= 0 \end{aligned}$ (c) $\begin{aligned} x_1 + 2x_2 \qquad - x_4 + x_5 &= 1 \\ 3x_2 + x_3 \qquad - x_5 &= 2 \\ x_3 + 7x_4 \qquad &= 1 \end{aligned}$ (d) $\begin{aligned} x_1 \qquad &= 1 \\ x_2 \quad &= 2 \\ x_3 &= 3 \end{aligned}$

5. Find a system of linear equations corresponding to the augmented matrix.

   (a) $\begin{bmatrix} 2 & 0 & 0 \\ 3 & -4 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 0 & -2 & 5 \\ 7 & 1 & 4 & -3 \\ 0 & -2 & 1 & 7 \end{bmatrix}$

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

6. (a) Find a linear equation in the variables x and y that has the general solution $x = 5 + 2t$, $y = t$.
   (b) Show that $x = t$, $y = \frac{1}{2}t - \frac{5}{2}$ is also the general solution of the equation in part (a).

7. The curve $y = ax^2 + bx + c$ shown in Figure 2 passes through the points (x_1, y_1), (x_2, y_2), and (x_3, y_3). Show that the coefficients a, b, and c are a solution of the system of linear equations whose augmented matrix is

$$\begin{bmatrix} x_1^2 & x_1 & 1 & y_1 \\ x_2^2 & x_2 & 1 & y_2 \\ x_3^2 & x_3 & 1 & y_3 \end{bmatrix}$$

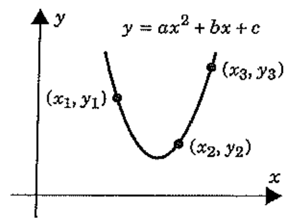


Figure 2

8. For which value(s) of the constant k does the following system of linear equations have no solutions? Exactly one solution? Infinitely many solutions?

$$\begin{aligned} x - y &= 3 \\ 2x - 2y &= k \end{aligned}$$

9. Consider the system of equations

$$\begin{aligned} ax + by &= k \\ cx + dy &= l \\ ex + fy &= m \end{aligned}$$

   Discuss the relative positions of the lines $ax + by = k$, $cx + dy = l$, and $ex + fy = m$ when
   (a) the system has no solutions
   (b) the system has exactly one solution
   (c) the system has infinitely many solutions

10. Show that if the system of equations in Exercise 9 is consistent, then at least one equation can be discarded from the system without altering the solution set.

11. Let $k = l = m = 0$ in Exercise 9; show that the system must be consistent. What can be said about the point of intersection of the three lines if the system has exactly one solution?

12. Consider the system of equations

$$\begin{aligned} x + y + 2z &= a \\ x \quad + z &= b \\ 2x + y + 3z &= c \end{aligned}$$

Show that in order for this system to be consistent, a, b, and c must satisfy $c = a + b$.

13. Prove: If the linear equations $x_1 + kx_2 = c$ and $x_1 + lx_2 = d$ have the same solution set, then the equations are identical.

1.2 GAUSSIAN ELIMINATION

In this section we shall give a systematic procedure for solving systems of linear equations; it is based on the idea of reducing the augmented matrix to a form that is simple enough that the system of equations can be solved by inspection.

REDUCED ROW-ECHELON FORM

In Example 3 of the preceding section, we solved the given linear system by reducing the augmented matrix to

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

from which the solution of the system was evident. This is an example of a matrix that is in ***reduced row-echelon form.*** To be of this form, a matrix must have the following properties.

1. *If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. (We call this a **leading** 1.)*
2. *If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.*
3. *In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.*
4. *Each column that contains a leading 1 has zeros everywhere else.*

A matrix having properties **1**, **2**, and **3** (but not necessarily 4) is said to be in ***row-echelon form.***

Example 1 The following matrices are in reduced row-echelon form.

$$\begin{bmatrix} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & -2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

The following matrices are in row-echelon form but not in reduced row-echelon form.

$$\begin{bmatrix} 1 & 4 & 3 & 7 \\ 0 & 1 & 6 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 6 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The reader should check to see that each of the matrices above satisfies all the necessary requirements.

REMARK. As the preceding example illustrates, a matrix in row-echelon form has zeros below each leading 1, whereas a matrix in reduced row-echelon form has zeros both above and below each leading 1.

If, by a sequence of elementary row operations, the augmented matrix for a system of linear equations is put in reduced row-echelon form, then the solution set of the system will be evident by inspection or after a few simple steps. The next example illustrates this.

Example 2 Suppose that the augmented matrix for a system of linear equations has been reduced by row operations to the given reduced row-echelon form. Solve the system.

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

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

Solution (a). The corresponding system of equations is

$$\begin{aligned} x_1 \qquad\quad &= 5 \\ x_2 \quad &= -2 \\ x_3 &= 4 \end{aligned}$$

By inspection, $x_1 = 5$, $x_2 = -2$, $x_3 = 4$.

Solution (b). The corresponding system of equations is

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

Since x_1, x_2, and x_3 correspond to leading 1's in the augmented matrix, we call them ***leading variables***. The nonleading variables (in this case x_4) are called ***free variables***. Solving for the leading variables in terms of the free variable gives

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

From this form of the equations we see that the free variable x_4 can be assigned an arbitrary value, say t, which then determines the values of the leading variables x_1, x_2, and x_3. Thus there are infinitely many solutions, and the general solution is given by the formulas

$$x_1 = -1 - 4t, \quad x_2 = 6 - 2t, \quad x_3 = 2 - 3t, \quad x_4 = t$$

Solution (c). The corresponding system of equations is

$$\begin{aligned} x_1 + 6x_2 \qquad\qquad + 4x_5 &= -2 \\ x_3 \qquad + 3x_5 &= 1 \\ x_4 + 5x_5 &= 2 \end{aligned}$$

Here the leading variables are x_1, x_3, and x_4, and the free variables are x_2 and x_5. Solving for the leading variables in terms of the free variables gives

$$\begin{aligned} x_1 &= -2 - 6x_2 - 4x_5 \\ x_3 &= 1 - 3x_5 \\ x_4 &= 2 - 5x_5 \end{aligned}$$

Since x_5 can be assigned an arbitrary value, t, and x_2 can be assigned an arbitrary value, s, there are infinitely many solutions. The general solution is given by the formulas

$$x_1 = -2 - 6s - 4t, \quad x_2 = s, \quad x_3 = 1 - 3t, \quad x_4 = 2 - 5t, \quad x_5 = t$$

Solution (d). The last equation in the corresponding system of equations is

$$0x_1 + 0x_2 + 0x_3 = 1$$

Since this equation cannot be satisfied, there is no solution to the system.

GAUSSIAN ELIMINATION

We have just seen how easy it is to solve a system of linear equations once its augmented matrix is in reduced row-echelon form. Now we shall give a step-by-step procedure that can be used to reduce any matrix to reduced row-echelon form. As we state each step in the procedure, we shall illustrate the idea by reducing the following matrix to reduced row-echelon form.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step 1. Locate the leftmost column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

↑ **Leftmost nonzero column**

Step 2. Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The first and second rows in the preceding matrix were interchanged.

Step 3. If the entry that is now at the top of the column found in Step 1 is a, multiply the first row by $1/a$ in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The first row of the preceding matrix was multiplied by $\frac{1}{2}$.

Step 4. Add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zeros.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

-2 times the first row of the preceding matrix was added to the third row.

Step 5. Now cover the top row in the matrix and begin again with Step 1 applied to the submatrix that remains. Continue in this way until the *entire* matrix is in row-echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

↑ **Leftmost nonzero column in the submatrix**

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

The first row in the submatrix was multiplied by $-\frac{1}{2}$ to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

-5 times the first row of the submatrix was added to the second row of the submatrix to introduce a zero below the leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

The top row in the submatrix was covered, and we returned again to Step 1.

↑ **Leftmost nonzero column in the new submatrix**

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

The first (and only) row in the new submatrix was multiplied by 2 to introduce a leading 1.

The *entire* matrix is now in row-echelon form. To find the reduced row-echelon form we need the following additional step.

Step 6. Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

$\frac{7}{2}$ times the third row of the preceding matrix was added to the second row.

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

-6 times the third row was added to the first row.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

5 times the second row was added to the first row.

The last matrix is in reduced row-echelon form.

The above procedure for reducing a matrix to reduced row-echelon form is called ***Gauss–Jordan elimination**** (see page 13). If we use only the first five steps, the procedure produces a row-echelon form and is called ***Gaussian elimination.***

REMARK. It can be shown that *every matrix has a unique reduced row-echelon form*; that is, one will arrive at the same reduced row-echelon form for a given matrix no matter how the row operations are varied. (A proof of this result can be found in the article "The Reduced Row Echelon Form of a Matrix is Unique: A Simple Proof," by Thomas Yuster, *Mathematics Magazine*, Vol. 57, No. 2, 1984, pp. 93–94.) In contrast, *a row-echelon form of a given matrix is not unique*: different sequences of row operations can produce different row-echelon forms.

Example 3 Solve by Gauss–Jordan elimination.

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \qquad\qquad + 2x_5 \qquad\qquad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1 \\ 5x_3 + 10x_4 \qquad\qquad + 15x_6 &= 5 \\ 2x_1 + 6x_2 \qquad\qquad + 8x_4 + 4x_5 + 18x_6 &= 6 \end{aligned}$$

The augmented matrix for the system is

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$$

---

* *Karl Friedrich Gauss* (1777–1855) was a German mathematician and scientist. Sometimes called the "prince of mathematicians," Gauss ranks with Isaac Newton and Archimedes as one of the three greatest mathematicians who ever lived. In the entire history of mathematics there may never have been a child so precocious as Gauss—by his own account he worked out the rudiments of arithmetic before he could talk. One day, before he was even three years old, his genius became apparent to his parents in a very dramatic way. His father was preparing the weekly payroll for the laborers under his charge while the boy watched quietly from a corner. At the end of the long and tedious calculation, Gauss informed his father that there was an error in the result and stated the answer, which he had worked out in his head. To the astonishment of his parents, a check of the computations showed Gauss to be correct!

In his doctoral dissertation Gauss gave the first complete proof of the fundamental theorem of algebra, which states that every polynomial equation has as many solutions as its degree. At age 19 he solved a problem that baffled Euclid, inscribing a regular polygon of seventeen sides in a circle using straightedge and compass; and in 1801, at age 24, he published his first masterpiece, *Disquisitiones Arithmeticae*, considered by many to be one of the most brilliant achievements in mathematics. In that paper Gauss systematized the study of number theory (properties of the integers) and formulated the basic concepts that form the foundation of that subject.

Among his myriad achievements, Gauss discovered the Gaussian or "bell-shaped" curve that is fundamental in probability, gave the first geometric interpretation of complex numbers and established their fundamental role in mathematics, developed methods of characterizing surfaces intrinsically by means of the curves that they contain, developed the theory of conformal (angle-preserving) maps, and discovered non-Euclidean geometry 30 years before the ideas were published by others. In physics he made major contributions to the theory of lenses and capillary action, and with Wilhelm Weber he did fundamental work in electromagnetism. Gauss invented the heliotrope, bifilar magnetometer, and an electrotelegraph.

Gauss was deeply religious and aristocratic in demeanor. He mastered foreign languages with ease, read extensively, and enjoyed mineralogy and botany as hobbies. He disliked teaching and was usually cool and discouraging to other mathematicians, possibly because he had already anticipated their work. It has been said that if Gauss had published all of his discoveries, the current state of mathematics would be advanced by 50 years. He was without a doubt the greatest mathematician of the modern era.

Wilhelm Jordan (1842–1899) was a German engineer who specialized in geodesy. His contribution to solving linear systems appeared in his popular book, *Handbuch der Vermessungskunde* (*Handbook of Geodesy*), in 1888.

Adding -2 times the first row to the second and fourth rows gives

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & -2 & 0 & -3 & -1 \\ 0 & 0 & 5 & 10 & 0 & 15 & 5 \\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix}$$

Multiplying the second row by -1 and then adding -5 times the new second row to the third row and -4 times the new second row to the fourth row gives

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

Interchanging the third and fourth rows and then multiplying the third row of the resulting matrix by $\frac{1}{6}$ gives the row-echelon form

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

Adding -3 times the third row to the second row and then adding 2 times the second row of the resulting matrix to the first row yields the reduced row-echelon form

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

The corresponding system of equations is

$$\begin{aligned} x_1 + 3x_2 \qquad + 4x_4 + 2x_5 \qquad &= 0 \\ x_3 + 2x_4 \qquad\qquad &= 0 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

(We have discarded the last equation, $0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 = 0$, since it will be satisfied automatically by the solutions of the remaining equations.) Solving for the leading variables, we obtain

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

If we assign the free variables x_2, x_4, and x_5 arbitrary values r, s, and t, respectively, the general solution is given by the formulas

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = \tfrac{1}{3}$$

BACK-SUBSTITUTION

Example 4 It is sometimes preferable to solve a system of linear equations by using Gaussian elimination to bring the augmented matrix into row-echelon form without continuing all the way to the reduced row-echelon form. When this is done, the corresponding system of equations can be solved by a technique called ***back-substitution.*** We shall illustrate this method using the system of equations in Example 3.

From the computations in Example 3, a row-echelon form of the augmented matrix is

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

To solve the corresponding system of equations

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \qquad + 2x_5 \qquad &= 0 \\ x_3 + 2x_4 \qquad + 3x_6 &= 1 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

we proceed as follows:

Step 1. Solve the equations for the leading variables.

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 \\ x_3 &= 1 - 2x_4 - 3x_6 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Step 2. Beginning with the bottom equation and working upward, successively substitute each equation into all the equations above it.

Substituting $x_6 = \frac{1}{3}$ into the second equation yields

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Substituting $x_3 = -2x_4$ into the first equation yields

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Step 3. Assign arbitrary values to the free variables, if any.

If we assign x_2, x_4, and x_5 the arbitrary values r, s, and t, respectively, the general solution is given by the formulas

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = \tfrac{1}{3}$$

This agrees with the solution obtained in Example 3.

REMARK. The arbitrary values that are assigned to the free variables are often called ***parameters***. Although we shall generally use the letters $r, s, t, \ldots$ for the parameters, any letters that do not conflict with the variable names may be used.

Example 5 Solve

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + 6y - 5z &= 0 \end{aligned}$$

by Gaussian elimination and back-substitution.

Solution. This is the system in Example 3 of Section 1.1. In that example we converted the augmented matrix

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

to the row-echelon form

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

The system corresponding to this matrix is

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ z &= 3 \end{aligned}$$

Solving for the leading variables yields

$$\begin{aligned} x &= 9 - y - 2z \\ y &= -\tfrac{17}{2} + \tfrac{7}{2}z \\ z &= 3 \end{aligned}$$

Substituting the bottom equation into those above yields

$$\begin{aligned} x &= 3 - y \\ y &= 2 \\ z &= 3 \end{aligned}$$

and substituting the second equation into the top yields

$$\begin{aligned} x &= 1 \\ y &= 2 \\ z &= 3 \end{aligned}$$

This agrees with the result found by Gauss–Jordan elimination in Example 3 of Section 1.1.

HOMOGENEOUS LINEAR SYSTEMS

A system of linear equations is said to be ***homogeneous*** if the constant terms are all zero; that is, the system has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad & \;\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Every homogeneous system of linear equations is consistent, since all such systems have $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ as a solution. This solution is called the ***trivial solution***; if there are other solutions, they are called ***nontrivial solutions***.

Because a homogeneous linear system always has the trivial solution, there are only two possibilities for its solutions:

- The system has only the trivial solution.
- The system has infinitely many solutions in addition to the trivial solution.

In the special case of a homogeneous linear system of two equations in two unknowns, say

$$\begin{aligned} a_1x + b_1y &= 0 \qquad (a_1, b_1 \text{ not both zero}) \\ a_2x + b_2y &= 0 \qquad (a_2, b_2 \text{ not both zero}) \end{aligned}$$

the graphs of the equations are lines through the origin, and the trivial solution corresponds to the point of intersection at the origin (Figure 1).

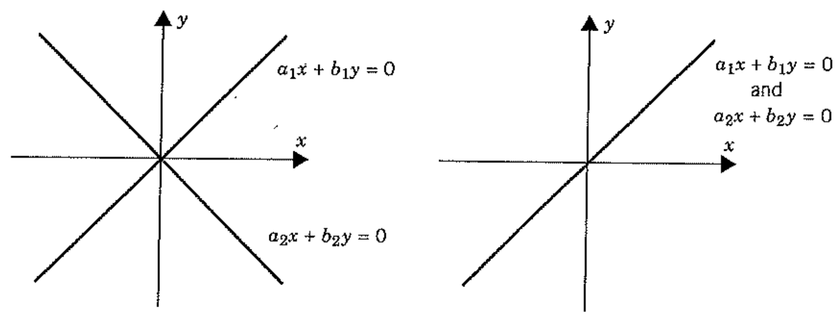


Figure 1 Only the trivial solution | Infinitely many solutions

There is one case in which a homogeneous system is assured of having nontrivial solutions, namely, whenever the system involves more unknowns than equations. To see why, consider the following example of four equations in five unknowns.

Example 6 Solve the following homogeneous system of linear equations by Gauss–Jordan elimination.

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

Solution. The augmented matrix for the system is

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

Reducing this matrix to reduced row-echelon form, we obtain

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

The corresponding system of equations is

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

Solving for the leading variables yields

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

Thus, the general solution is

$$x_1 = -s - t, \quad x_2 = s, \quad x_3 = -t, \quad x_4 = 0, \quad x_5 = t$$

Note that the trivial solution is obtained when $s = t = 0$.

Example 6 illustrates two important points about solving homogeneous systems of linear equations. First, none of the three elementary row operations alters the final column of zeros in the augmented matrix, so that the system of equations corresponding to the reduced row-echelon form of the augmented matrix must also be a homogeneous system [see system (2)]. Second, depending on whether the reduced row-echelon form of the augmented matrix has any zero rows, the number of equations in the reduced system is the same as or less than the number of equations in the original system [compare systems (1) and (2)]. Thus, if the given homogeneous system has m equations in n unknowns with $m < n$, and if there are r nonzero rows in the reduced row-echelon form of the augmented matrix, we will have $r < n$. It follows that the system of equations corresponding to the reduced row-echelon form of the augmented matrix will have the form

$$\begin{aligned} \cdots x_{k_1} \qquad\qquad & + \Sigma(\ ) = 0 \\ \cdots x_{k_2} \qquad & + \Sigma(\ ) = 0 \\ \ddots \qquad & \quad \vdots \\ x_{k_r} & + \Sigma(\ ) = 0 \end{aligned} \tag{3}$$

where $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ are the leading variables and $\Sigma(\ )$ denotes sums (possibly all different) that involve the $n - r$ free variables [compare system (3) with system (2) above]. Solving for the leading variables gives

$$\begin{aligned} x_{k_1} &= -\Sigma(\ ) \\ x_{k_2} &= -\Sigma(\ ) \\ &\ \vdots \\ x_{k_r} &= -\Sigma(\ ) \end{aligned}$$

As in Example 6, we can assign arbitrary values to the free variables on the right-hand side and thus obtain infinitely many solutions to the system.

In summary, we have the following important theorem.

Theorem 1.2.1. *A homogeneous system of linear equations with more unknowns than equations has infinitely many solutions.*

REMARK. Note that Theorem 1.2.1 applies only to homogeneous systems. A nonhomogeneous system with more unknowns than equations need not be consistent (Exercise 34); however, if the system is consistent, it will have infinitely many solutions. This will be proved later.

COMPUTER SOLUTION OF LINEAR SYSTEMS

In applications it is not uncommon to encounter large linear systems that must be solved by computer. Most computer algorithms for solving such systems are based on Gaussian elimination or Gauss–Jordan elimination, but the basic procedures are often modified to deal with such issues as

- Reducing roundoff errors
- Minimizing the use of computer memory space
- Solving the system with maximum speed

Some of these matters will be considered in Chapter 9. For hand computations fractions are an annoyance that often cannot be avoided. However, in some cases it is possible to avoid them by varying the elementary row operations in the right way. Thus, once the methods of Gaussian elimination and Gauss–Jordan elimination have been mastered, the reader may wish to vary the steps in specific problems to avoid fractions (see Exercise 18).

EXERCISE SET 1.2

1. Which of the following 3×3 matrices are in reduced row-echelon form?

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

(f) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (g) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (h) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$ (i) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ (j) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

2. Which of the following 3×3 matrices are in row-echelon form?

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

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

3. In each part determine whether the matrix is in row-echelon form, reduced row-echelon form, both, or neither.

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

(d) $\begin{bmatrix} 1 & -7 & 5 & 5 \\ 0 & 1 & 3 & 2 \end{bmatrix}$ (e) $\begin{bmatrix} 1 & 3 & 0 & 2 & 0 \\ 1 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

4. In each part suppose that the augmented matrix for a system of linear equations has been reduced by row operations to the given reduced row-echelon form. Solve the system.

(a) $\begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 7 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 & -7 & 8 \\ 0 & 1 & 0 & 3 & 2 \\ 0 & 0 & 1 & 1 & -5 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -6 & 0 & 0 & 3 & -2 \\ 0 & 0 & 1 & 0 & 4 & 7 \\ 0 & 0 & 0 & 1 & 5 & 8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

5. In each part suppose that the augmented matrix for a system of linear equations has been reduced by row operations to the given row-echelon form. Solve the system.

(a) $\begin{bmatrix} 1 & -3 & 4 & 7 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & 1 & 5 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 8 & -5 & 6 \\ 0 & 1 & 4 & -9 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 7 & -2 & 0 & -8 & -3 \\ 0 & 0 & 1 & 1 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & -3 & 7 & 1 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6. Solve each of the following systems by Gauss–Jordan elimination.

(a) $\begin{aligned} x_1 + x_2 + 2x_3 &= 8 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 3x_1 - 7x_2 + 4x_3 &= 10 \end{aligned}$ (b) $\begin{aligned} 2x_1 + 2x_2 + 2x_3 &= 0 \\ -2x_1 + 5x_2 + 2x_3 &= 1 \\ 8x_1 + x_2 + 4x_3 &= -1 \end{aligned}$

(c) $\begin{aligned} x - y + 2z - w &= -1 \\ 2x + y - 2z - 2w &= -2 \\ -x + 2y - 4z + w &= 1 \\ 3x \qquad\qquad - 3w &= -3 \end{aligned}$ (d) $\begin{aligned} -2b + 3c &= 1 \\ 3a + 6b - 3c &= -2 \\ 6a + 6b + 3c &= 5 \end{aligned}$

7. Solve each of the systems in Exercise 6 by Gaussian elimination.

8. Solve each of the following systems by Gauss–Jordan elimination.

(a) $\begin{aligned} 2x_1 - 3x_2 &= -2 \\ 2x_1 + x_2 &= 1 \\ 3x_1 + 2x_2 &= 1 \end{aligned}$ (b) $\begin{aligned} 3x_1 + 2x_2 - x_3 &= -15 \\ 5x_1 + 3x_2 + 2x_3 &= 0 \\ 3x_1 + x_2 + 3x_3 &= 11 \\ -6x_1 - 4x_2 + 2x_3 &= 30 \end{aligned}$

(c) $\begin{aligned} 4x_1 - 8x_2 &= 12 \\ 3x_1 - 6x_2 &= 9 \\ -2x_1 + 4x_2 &= -6 \end{aligned}$ (d) $\begin{aligned} 10y - 4z + w &= 1 \\ x + 4y - z + w &= 2 \\ 3x + 2y + z + 2w &= 5 \\ -2x - 8y + 2z - 2w &= -4 \\ x - 6y + 3z \qquad &= 1 \end{aligned}$

9. Solve each of the systems in Exercise 8 by Gaussian elimination.

10. Solve each of the following systems by Gauss–Jordan elimination.

(a) $\begin{aligned} 5x_1 - 2x_2 + 6x_3 &= 0 \\ -2x_1 + x_2 + 3x_3 &= 1 \end{aligned}$ (b) $\begin{aligned} x_1 - 2x_2 + x_3 - 4x_4 &= 1 \\ x_1 + 3x_2 + 7x_3 + 2x_4 &= 2 \\ x_1 - 12x_2 - 11x_3 - 16x_4 &= 5 \end{aligned}$ (c) $\begin{aligned} w + 2x - y &= 4 \\ x - y &= 3 \\ w + 3x - 2y &= 7 \\ 2u + 4v + w + 7x \qquad &= 7 \end{aligned}$

11. Solve each of the systems in Exercise 10 by Gaussian elimination.

12. Without using pencil and paper, determine which of the following homogeneous systems have nontrivial solutions.

(a) $\begin{aligned} 2x_1 - 3x_2 + 4x_3 - x_4 &= 0 \\ 7x_1 + x_2 - 8x_3 + 9x_4 &= 0 \\ 2x_1 + 8x_2 + x_3 - x_4 &= 0 \end{aligned}$ (b) $\begin{aligned} x_1 + 3x_2 - x_3 &= 0 \\ x_2 - 8x_3 &= 0 \\ 4x_3 &= 0 \end{aligned}$

(c) $\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 0 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0 \end{aligned}$ (d) $\begin{aligned} 3x_1 - 2x_2 &= 0 \\ 6x_1 - 4x_2 &= 0 \end{aligned}$

13. Solve the following homogeneous systems of linear equations by any method.

(a) $\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 2x_2 \quad\quad &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$ (b) $\begin{aligned} 3x_1 + x_2 + x_3 + x_4 &= 0 \\ 5x_1 - x_2 + x_3 - x_4 &= 0 \end{aligned}$ (c) $\begin{aligned} 2x + 2y + 4z &= 0 \\ w \quad - y - 3z &= 0 \\ 2w + 3x + y + z &= 0 \\ -2w + x + 3y - 2z &= 0 \end{aligned}$

14. Solve the following homogeneous systems of linear equations by any method.

(a) $\begin{aligned} 2x - y - 3z &= 0 \\ -x + 2y - 3z &= 0 \\ x + y + 4z &= 0 \end{aligned}$ (b) $\begin{aligned} v + 3w - 2x &= 0 \\ 2u + v - 4w + 3x &= 0 \\ 2u + 3v + 2w - x &= 0 \\ -4u - 3v + 5w - 4x &= 0 \end{aligned}$ (c) $\begin{aligned} x_1 + 3x_2 \quad\quad + x_4 &= 0 \\ x_1 + 4x_2 + 2x_3 \quad &= 0 \\ -2x_2 - 2x_3 - x_4 &= 0 \\ 2x_1 - 4x_2 + x_3 + x_4 &= 0 \\ x_1 - 2x_2 - x_3 + x_4 &= 0 \end{aligned}$

15. Solve the following systems by any method.

(a) $\begin{aligned} 2I_1 - I_2 + 3I_3 + 4I_4 &= 9 \\ I_1 \quad - 2I_3 + 7I_4 &= 11 \\ 3I_1 - 3I_2 + I_3 + 5I_4 &= 8 \\ 2I_1 + I_2 + 4I_3 + 4I_4 &= 10 \end{aligned}$ (b) $\begin{aligned} Z_3 + Z_4 + Z_5 &= 0 \\ -Z_1 - Z_2 + 2Z_3 - 3Z_4 + Z_5 &= 0 \\ Z_1 + Z_2 - 2Z_3 \quad - Z_5 &= 0 \\ 2Z_1 + 2Z_2 - Z_3 \quad + Z_5 &= 0 \end{aligned}$

16. Solve the following systems, where a, b, and c are constants.

(a) $\begin{aligned} 2x + y &= a \\ 3x + 6y &= b \end{aligned}$ (b) $\begin{aligned} x_1 + x_2 + x_3 &= a \\ 2x_1 \quad + 2x_3 &= b \\ 3x_2 + 3x_3 &= c \end{aligned}$

17. For which values of a will the following system have no solutions? Exactly one solution? Infinitely many solutions?

$$\begin{aligned} x + 2y - \quad 3z &= 4 \\ 3x - y + \quad 5z &= 2 \\ 4x + y + (a^2 - 14)z &= a + 2 \end{aligned}$$

18. Reduce

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

to reduced row-echelon form without introducing any fractions.

19. Find two different row-echelon forms of

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

20. Solve the following system of nonlinear equations for the unknown angles α, β, and γ, where $0 \le \alpha \le 2\pi$, $0 \le \beta \le 2\pi$, and $0 \le \gamma < \pi$.

$$\begin{aligned} 2\sin\alpha - \cos\beta + 3\tan\gamma &= 3 \\ 4\sin\alpha + 2\cos\beta - 2\tan\gamma &= 2 \\ 6\sin\alpha - 3\cos\beta + \tan\gamma &= 9 \end{aligned}$$

21. Solve the following system of nonlinear equations for x, y, and z.

$$\begin{aligned} x^2 + y^2 + z^2 &= 6 \\ x^2 - y^2 + 2z^2 &= 2 \\ 2x^2 + y^2 - z^2 &= 3 \end{aligned}$$

22. Show that the following nonlinear system has eighteen solutions if $0 \le \alpha \le 2\pi$, $0 \le \beta \le 2\pi$, and $0 \le \gamma \le 2\pi$.

$$\begin{aligned} \sin\alpha + 2\cos\beta + 3\tan\gamma &= 0 \\ 2\sin\alpha + 5\cos\beta + 3\tan\gamma &= 0 \\ -\sin\alpha - 5\cos\beta + 5\tan\gamma &= 0 \end{aligned}$$

23. For which value(s) of γ does the following system of equations have nontrivial solutions?

$$\begin{aligned} (\lambda - 3)x + \quad\quad y &= 0 \\ x + (\lambda - 3)y &= 0 \end{aligned}$$

24. Consider the system of equations

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \\ ex + fy &= 0 \end{aligned}$$

Discuss the relative positions of the lines $ax + by = 0$, $cx + dy = 0$, and $ex + fy = 0$ when

(a) the system has only the trivial solution (b) the system has nontrivial solutions

25. Figure 2 shows the graph of a cubic equation $y = ax^3 + bx^2 + cx + d$. Find the coefficients a, b, c, and d.

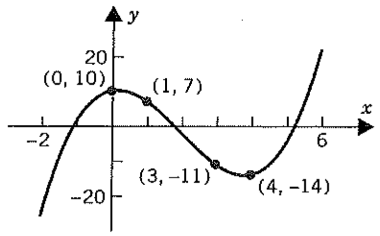


Figure 2

26. Recall from plane geometry that three points, not all lying on a straight line, determine a unique circle. It is shown in analytic geometry that a circle in the xy-plane has an equation of the form

$$ax^2 + ay^2 + bx + cy + d = 0$$

Find an equation of the circle shown in Figure 3.

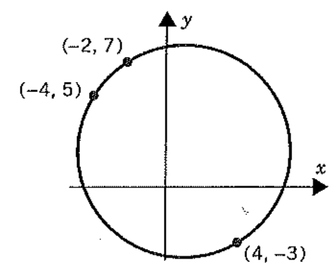


Figure 3

27. Describe the possible reduced row-echelon forms of

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

28. Show that if $ad - bc \neq 0$, then the reduced row-echelon form of

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{is} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

29. Use Exercise 28 to show that if $ad - bc \neq 0$, then the system

$$\begin{aligned} ax + by &= k \\ cx + dy &= l \end{aligned}$$

has exactly one solution.

30. Solve the system

$$\begin{aligned} 2x_1 - x_2 &= \lambda x_1 \\ 2x_1 + x_2 + x_3 &= \lambda x_2 \\ -2x_1 + 2x_2 + x_3 &= \lambda x_3 \end{aligned}$$

for x_1, x_2, and x_3 if
(a) $\lambda = 1$ (b) $\lambda = 2$

31. Consider the system of equations

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}$$

(a) Show that if $x = x_0$, $y = y_0$ is any solution of the system and k is any constant, then $x = kx_0$, $y = ky_0$ is also a solution.
(b) Show that if $x = x_0$, $y = y_0$ and $x = x_1$, $y = y_1$ are any two solutions, then $x = x_0 + x_1$, $y = y_0 + y_1$ is also a solution.

32. Consider the systems of equations

$$\text{(I)}\ \begin{aligned} ax + by &= k \\ cx + dy &= l \end{aligned} \qquad \text{(II)}\ \begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}$$

(a) Show that if $x = x_1$, $y = y_1$ and $x = x_2$, $y = y_2$ are both solutions of I, then $x = x_1 - x_2$, $y = y_1 - y_2$ is a solution of II.
(b) Show that if $x = x_1$, $y = y_1$ is a solution of I and $x = x_0$, $y = y_0$ is a solution of II, then $x = x_1 + x_0$, $y = y_1 + y_0$ is a solution of I.

33. (a) In the system of equations numbered (3), explain why it would be incorrect to denote the leading variables by $x_1, x_2, \ldots, x_r$ rather than $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ as we have done.
(b) The system of equations numbered (2) is a specific case of (3). What value does r have in this case? What are $x_{k_1}, x_{k_2}, \ldots, x_{k_r}$ in this case? Write out the sums denoted by $\Sigma(\ )$ in (3).

34. Find an inconsistent linear system that has more unknowns than equations.

1.3 MATRICES AND MATRIX OPERATIONS

Rectangular arrays of real numbers arise in many contexts other than as augmented matrices for systems of linear equations. In this section we shall consider such arrays as objects in their own right and develop some of their properties for use in our later work.

MATRIX NOTATION AND TERMINOLOGY

Definition. A ***matrix*** is a rectangular array of numbers. The numbers in the array are called the ***entries*** in the matrix.

Example 1 Some examples of matrices are

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{bmatrix}, \quad [2 \quad 1 \quad 0 \quad -3], \quad \begin{bmatrix} -\sqrt{2} & \pi & e \\ 3 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad [4]$$

The ***size*** of a matrix is described in terms of the number of rows (horizontal lines) and columns (vertical lines) it contains. For example, the first matrix in Example 1 has three rows and two columns, so its size is 3 by 2 (written 3×2). In a size description, the first number always denotes the number of rows and the second denotes the number of columns. The remaining matrices in Example 1 have sizes 1×4, 3×3, 2×1, and 1×1, respectively. A matrix with only one column is called a ***column matrix*** (or a ***column vector***), and a matrix with only one row is called a ***row matrix*** (or a ***row vector***). Thus, in Example 1 the 2×1 matrix is a column matrix, the 1×4 matrix is a row matrix, and the 1×1 matrix is both a row matrix and a column matrix. (The term *vector* has another meaning that we will discuss in subsequent chapters.)

REMARK. It is common practice to omit the brackets on a 1×1 matrix. Thus, we might write 4 rather than [4]. Although this makes it impossible to tell whether 4 denotes the number "four" or the 1×1 matrix whose entry is "four," this rarely causes problems, since it is usually possible to tell which is meant from the context in which the symbol appears.

We shall use capital letters to denote matrices and lowercase letters to denote numerical quantities; thus, we might write

$$A = \begin{bmatrix} 2 & 1 & 7 \\ 3 & 4 & 2 \end{bmatrix} \quad \text{or} \quad C = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

When discussing matrices, it is common to refer to numerical quantities as ***scalars***. Unless stated otherwise, *scalars will be real numbers*; complex scalars will be considered in Chapter 10.

The entry that occurs in row i and column j of a matrix A will be denoted by a_{ij}.

Thus, a general 3×4 matrix might be written as

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

and a general $m \times n$ matrix as

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

When compactness of notation is desired, the preceding matrix can be written as

$$[a_{ij}]_{m \times n} \quad \text{or} \quad [a_{ij}]$$

the first notation being used when it is important in the discussion to know the size and the second when the size need not be emphasized. Usually, we shall match the letter denoting a matrix with the letter denoting its entries; thus, for a matrix B we would generally use b_{ij} for the entry in row i and column j and for a matrix C we would use c_{ij}.

The entry in row i and column j of a matrix A is also commonly denoted by the symbol $(A)_{ij}$. Thus, for matrix (1) above, we have

$$(A)_{ij} = a_{ij}$$

and for the matrix

$$A = \begin{bmatrix} 2 & -3 \\ 7 & 0 \end{bmatrix}$$

we have $(A)_{11} = 2$, $(A)_{12} = -3$, $(A)_{21} = 7$, and $(A)_{22} = 0$.

Row and column matrices are of special importance, and it is common practice to denote them by boldface lowercase letters rather than capital letters. For such matrices double subscripting of the entries is unnecessary. Thus, a general $1 \times n$ row matrix **a** and a general $m \times 1$ column matrix **b** would be written as

$$\mathbf{a} = [a_1 \quad a_2 \quad \cdots \quad a_n] \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

A matrix A with n rows and n columns is called a ***square matrix of order n***, and the entries $a_{11}, a_{22}, \ldots, a_{nn}$ are said to be on the ***main diagonal*** of A (see the shaded entries in Figure 1).

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

Figure 1

OPERATIONS ON MATRICES

So far, we have used matrices to abbreviate the work in solving systems of linear equations. For other applications, however, it is desirable to develop an "arithmetic of matrices" in which matrices can be added, subtracted, and multiplied in a useful way. The remainder of this section will be devoted to developing this arithmetic.

Definition. Two matrices are defined to be ***equal*** if they have the same size and their corresponding entries are equal.

In matrix notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then $A = B$ if and only if $(A)_{ij} = (B)_{ij}$, or equivalently, $a_{ij} = b_{ij}$ for all i and j.

Example 2 Consider the matrices

$$A = \begin{bmatrix} 2 & 1 \\ 3 & x \end{bmatrix} \quad B = \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 0 \end{bmatrix}$$

If $x = 5$, then $A = B$, but for all other values of x the matrices A and B are not equal, since not all of their corresponding entries are equal. There is no value of x for which $A = C$ since A and C have different sizes.

Definition. If A and B are matrices of the same size, then the ***sum*** $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A, and the ***difference*** $A - B$ is the matrix obtained by subtracting the entries of B from the corresponding entries of A. Matrices of different sizes cannot be added or subtracted.

In matrix notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{and} \quad (A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}$$

Example 3 Consider the matrices

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

Then

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

The expressions $A + C$, $B + C$, $A - C$, and $B - C$ are undefined.

Definition. If A is any matrix and c is any scalar, then the ***product*** cA is the matrix obtained by multiplying each entry of A by c.

In matrix notation, if $A = [a_{ij}]$, then

$$(cA)_{ij} = c(A)_{ij} = ca_{ij}$$

Example 4 For the matrices

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 & 7 \\ -1 & 3 & -5 \end{bmatrix} \quad C = \begin{bmatrix} 9 & -6 & 3 \\ 3 & 0 & 12 \end{bmatrix}$$

we have

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

It is common practice to denote $(-1)B$ by $-B$.

If $A_1, A_2, \ldots, A_n$ are matrices of the same size and $c_1, c_2, \ldots, c_n$ are scalars, then an expression of the form

$$c_1A_1 + c_2A_2 + \cdots + c_nA_n$$

is called a ***linear combination*** of $A_1, A_2, \ldots, A_n$ with ***coefficients*** $c_1, c_2, \ldots, c_n$. For example, if A, B, and C are the matrices in Example 4, then

$$\begin{aligned} 2A - B + \tfrac{1}{3}C &= 2A + (-1)B + \tfrac{1}{3}C \\ &= \begin{bmatrix} 4 & 6 & 8 \\ 2 & 6 & 2 \end{bmatrix} + \begin{bmatrix} 0 & -2 & -7 \\ 1 & -3 & 5 \end{bmatrix} + \begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 7 & 2 & 2 \\ 4 & 3 & 11 \end{bmatrix} \end{aligned}$$

is the linear combination of A, B, and C with scalar coefficients 2, -1, and $\frac{1}{3}$.

Thus far we have defined multiplication of a matrix by a scalar but not the multiplication of two matrices. Since matrices are added by adding corresponding entries and subtracted by subtracting corresponding entries, it would seem natural to define multiplication of matrices by multiplying corresponding entries. However, it turns out that such a definition would not be very useful for most problems. Experience has led mathematicians to the following less natural but more useful definition of matrix multiplication.

Definition. If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the ***product*** AB is the $m \times n$ matrix whose entries are determined as follows. To find the entry in row i and column j of AB, single out row i from the matrix A and column j from the matrix B. Multiply the corresponding entries from the row and column together and then add up the resulting products.

Example 5 Consider the matrices

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

Since A is a 2×3 matrix and B is a 3×4 matrix, the product AB is a 2×4 matrix. To determine, for example, the entry in row 2 and column 3 of AB, we single out row 2 from A and column 3 from B. Then, as illustrated below, we multiply corresponding entries together and add up these products.

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \boxed{26} & \square \end{bmatrix}$$

$$(2 \cdot 4) + (6 \cdot 3) + (0 \cdot 5) = 26$$

The entry in row 1 and column 4 of AB is computed as follows.

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \boxed{13} \\ \square & \square & \square & \square \end{bmatrix}$$

$$(1 \cdot 3) + (2 \cdot 1) + (4 \cdot 2) = 13$$

The computations for the remaining products are

$$\begin{aligned} (1 \cdot 4) + (2 \cdot 0) + (4 \cdot 2) &= 12 \\ (1 \cdot 1) - (2 \cdot 1) + (4 \cdot 7) &= 27 \\ (1 \cdot 4) + (2 \cdot 3) + (4 \cdot 5) &= 30 \\ (2 \cdot 4) + (6 \cdot 0) + (0 \cdot 2) &= 8 \\ (2 \cdot 1) - (6 \cdot 1) + (0 \cdot 7) &= -4 \\ (2 \cdot 3) + (6 \cdot 1) + (0 \cdot 2) &= 12 \end{aligned} \qquad AB = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

The definition of matrix multiplication requires that the number of columns of the first factor A be the same as the number of rows of the second factor B in order to form the product AB. If this condition is not satisfied, the product is undefined. A convenient way to determine whether a product of two matrices is defined is to write down the size of the first factor and, to the right of it, write down the size of the second factor. If, as in Figure 2, the inside numbers are the same, then the product is defined. The outside numbers then give the size of the product.

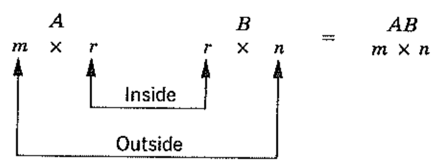


Figure 2

Example 6 Suppose that A, B, and C are matrices with the following sizes:

A	B	C
3×4	4×7	7×3

Then AB is defined and is a 3×7 matrix; CA is defined and is a 7×4 matrix; and BC is defined and is a 4×3 matrix. The products AC, CB, and BA are all undefined.

If $A = [a_{ij}]$ is a general $m \times r$ matrix and $B = [b_{ij}]$ is a general $r \times n$ matrix, then as illustrated by the shading in Figure 3, the entry $(AB)_{ij}$ in row i and column j of AB is given by

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj} \tag{2}$$

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ir} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn} \end{bmatrix}$$

Figure 3

PARTITIONED MATRICES

A matrix can be subdivided or ***partitioned*** into smaller matrices by inserting horizontal and vertical rules between selected rows and columns. For example, below are three possible partitions of a general 3×4 matrix A—the first is a partition of A into four ***submatrices*** A_{11}, A_{12}, A_{21}, and A_{22}; the second is a partition of A into its row matrices $\mathbf{r}_1, \mathbf{r}_2$, and $\mathbf{r}_3$; and the third is a partition of A into its column matrices $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$, and $\mathbf{c}_4$:

$$A = \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$A = \left[\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix}$$

$$A = \left[\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3 \quad \mathbf{c}_4]$$

MATRIX MULTIPLICATION BY COLUMNS AND BY ROWS

Sometimes it may be desirable to find a particular row or column of a matrix product AB without computing the entire product. The following results, whose proofs are left as exercises, are useful for that purpose:

$$j\text{th column matrix of } AB = A[j\text{th column matrix of } B] \tag{3}$$

$$i\text{th row matrix of } AB = [i\text{th row matrix of } A]B \tag{4}$$

Example 7 If A and B are the matrices in Example 5, then from (3) the second column matrix of AB can be obtained by the computation

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 27 \\ -4 \end{bmatrix}$$

(Second column of B) (Second column of AB)

and from (4) the first row matrix of AB can be obtained by the computation

$$[1 \quad 2 \quad 4] \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = [12 \quad 27 \quad 30 \quad 13]$$

(First row of A) (First row of AB)

If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ denote the row matrices of A and $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ denote the column matrices of B, then it follows from Formulas (3) and (4) that

$$AB = A[\mathbf{b}_1 \mid \mathbf{b}_2 \mid \cdots \mid \mathbf{b}_n] = [A\mathbf{b}_1 \mid A\mathbf{b}_2 \mid \cdots \mid A\mathbf{b}_n] \tag{5}$$

(AB computed column by column)

$$AB = \begin{bmatrix} \mathbf{a}_1 \\ \hline \mathbf{a}_2 \\ \hline \vdots \\ \hline \mathbf{a}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{a}_1 B \\ \hline \mathbf{a}_2 B \\ \hline \vdots \\ \hline \mathbf{a}_m B \end{bmatrix} \tag{6}$$

(AB computed row by row)

REMARK. Formulas (5) and (6) are special cases of a more general procedure for multiplying partitioned matrices (see Exercises 15–17).

MATRIX PRODUCTS AS LINEAR COMBINATIONS

Row and column matrices provide an alternative way of thinking about matrix multiplication. For example, suppose that

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

Then

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \tag{7}$$

In words, (7) tells us that *the product $A\mathbf{x}$ of a matrix A with a column matrix $\mathbf{x}$ is a linear combination of the column matrices of A with the coefficients coming from the matrix $\mathbf{x}$*. In the exercises we ask the reader to show that *the product $\mathbf{y}A$ of a $1 \times m$ matrix $\mathbf{y}$ with an $m \times n$ matrix A is a linear combination of the row matrices of A with scalar coefficients coming from $\mathbf{y}$.*

Example 8 The matrix product

$$\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

can be written as the linear combination

$$2\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

and the matrix product

$$[1 \quad -9 \quad -3]\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix} = [-16 \quad -18 \quad 35]$$

can be written as the linear combination

$$1[-1 \quad 3 \quad 2] - 9[1 \quad 2 \quad -3] - 3[2 \quad 1 \quad -2] = [-16 \quad -18 \quad 35]$$

It follows from (5) and (7) that *the jth column matrix of a product AB is a linear combination of the column matrices of A with the coefficients coming from the jth column of B.*

Example 9 We showed in Example 5 that

$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix}\begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

The column matrices of AB can be expressed as linear combinations of the column matrices of A as follows:

$$\begin{bmatrix} 12 \\ 8 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 27 \\ -4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 7\begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 30 \\ 26 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 5\begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 13 \\ 12 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \end{bmatrix} + 2\begin{bmatrix} 4 \\ 0 \end{bmatrix}$$

MATRIX FORM OF A LINEAR SYSTEM

Matrix multiplication has an important application to systems of linear equations. Consider any system of m linear equations in n unknowns.

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

Since two matrices are equal if and only if their corresponding entries are equal, we can replace the m equations in this system by the single matrix equation

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

The $m \times 1$ matrix on the left side of this equation can be written as a product to give

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

If we designate these matrices by A, $\mathbf{x}$, and $\mathbf{b}$, respectively, the original system of m equations in n unknowns has been replaced by the single matrix equation

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

The matrix A in this equation is called the ***coefficient matrix*** of the system. The augmented matrix for the system is obtained by adjoining $\mathbf{b}$ to A as the last column; thus the augmented matrix is

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

TRANSPOSE OF A MATRIX

We conclude this section by defining two matrix operations that have no analogs in the real numbers.

Definition. If A is any $m \times n$ matrix, then the ***transpose of A***, denoted by A^T, is defined to be the $n \times m$ matrix that results from interchanging the rows and columns of A; that is, the first column of A^T is the first row of A, the second column of A^T is the second row of A, and so forth.

Example 10 The following are some examples of matrices and their transposes.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 5 & 6 \end{bmatrix} \quad C = [1 \;\; 3 \;\; 5] \quad D = [4]$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{bmatrix} \quad B^T = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix} \quad C^T = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \quad D^T = [4]$$

Observe that not only are the columns of A^T the rows of A, but the rows of A^T are columns of A. Thus, the entry in row i and column j of A^T is the entry in row j and column i of A; that is,

$$(A^T)_{ij} = (A)_{ji} \tag{8}$$

Note the reversal of the subscripts.

In the special case where A is a square matrix, the transpose of A can be obtained by interchanging entries that are symmetrically positioned about the main diagonal (Figure 4). Stated another way, A^T can be obtained by "reflecting" A about its main diagonal.

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 7 & 0 \\ -5 & 8 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 4 \\ 3 & 7 & 0 \\ -5 & 8 & 6 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 3 & -5 \\ -2 & 7 & 8 \\ 4 & 0 & 6 \end{bmatrix}$$

Interchange entries that are symmetrically positioned about the main diagonal.

Figure 4

TRACE OF A SQUARE MATRIX

Definition. If A is a square matrix, then the ***trace of A***, denoted by $\text{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A. The trace of A is undefined if A is not a square matrix.

Example 11 The following are examples of matrices and their traces.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 2 & 7 & 0 \\ 3 & 5 & -8 & 4 \\ 1 & 2 & 7 & -3 \\ 4 & -2 & 1 & 0 \end{bmatrix}$$

$$\text{tr}(A) = a_{11} + a_{22} + a_{33} \qquad \text{tr}(B) = -1 + 5 + 7 + 0 = 11$$

EXERCISE SET 1.3

1. Suppose that A, B, C, D, and E are matrices with the following sizes:

A	B	C	D	E
(4×5)	(4×5)	(5×2)	(4×2)	(5×4)

Determine which of the following matrix expressions are defined. For those which are defined, give the size of the resulting matrix.
(a) BA (b) $AC + D$ (c) $AE + B$ (d) $AB + B$
(e) $E(A + B)$ (f) $E(AC)$ (g) E^TA (h) $(A^T + E)D$

2. Solve the following matrix equation for a, b, c, and d.

$$\begin{bmatrix} a-b & b+c \\ 3d+c & 2a-4d \end{bmatrix} = \begin{bmatrix} 8 & 1 \\ 7 & 6 \end{bmatrix}$$

3. Consider the matrices

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

Compute the following (where possible).
(a) $D + E$ (b) $D - E$ (c) $5A$ (d) $-7C$
(e) $2B - C$ (f) $4E - 2D$ (g) $-3(D + 2E)$ (h) $A - A$
(i) $\text{tr}(D)$ (j) $\text{tr}(D - 3E)$ (k) $4\,\text{tr}(7B)$ (l) $\text{tr}(A)$

4. Using the matrices in Exercise 3, compute the following (where possible).
(a) $2A^T + C$ (b) $D^T - E^T$ (c) $(D - E)^T$ (d) $B^T + 5C^T$
(e) $\frac{1}{2}C^T - \frac{1}{4}A$ (f) $B - B^T$ (g) $2E^T - 3D^T$ (h) $(2E^T - 3D^T)^T$

5. Using the matrices in Exercise 3, compute the following (where possible).
(a) AB (b) BA (c) $(3E)D$ (d) $(AB)C$
(e) $A(BC)$ (f) CC^T (g) $(DA)^T$ (h) $(C^TB)A^T$
(i) $\text{tr}(DD^T)$ (j) $\text{tr}(4E^T - D)$ (k) $\text{tr}(C^TA^T + 2E^T)$

6. Using the matrices in Exercise 3, compute the following (where possible).
(a) $(2D^T - E)A$ (b) $(4B)C + 2B$ (c) $(-AC)^T + 5D^T$
(d) $(BA^T - 2C)^T$ (e) $B^T(CC^T - A^TA)$ (f) $D^TE^T - (ED)^T$

7. Let

$$A = \begin{bmatrix} 3 & -2 & 7 \\ 6 & 5 & 4 \\ 0 & 4 & 9 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 6 & -2 & 4 \\ 0 & 1 & 3 \\ 7 & 7 & 5 \end{bmatrix}$$

Use the method of Example 7 to find
(a) the first row of AB (b) the third row of AB (c) the second column of AB
(d) the first column of BA (e) the third row of AA (f) the third column of AA

8. Let A and B be the matrices in Exercise 7.
(a) Express each column matrix of AB as a linear combination of the column matrices of A.
(b) Express each column matrix of BA as a linear combination of the column matrices of B.

9. Let

$$\mathbf{y} = [y_1 \;\; y_2 \;\; \cdots \;\; y_m] \quad \text{and} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Show that the product $\mathbf{y}A$ can be expressed as a combination of the row matrices of A with the scalar coefficients from $\mathbf{y}$.

10. Let A and B be the matrices in Exercise 7.
(a) Use the result in Exercise 9 to express each row matrix of AB as a linear combination of the row matrices of B.
(b) Use the result in Exercise 9 to express each row matrix of BA as a linear combination of the row matrices of A.

11. Let C, D, and E be the matrices in Exercise 3. Using as few computations as possible, determine the entry in row 2 and column 3 of $C(DE)$.

12. (a) Show that if AB and BA are both defined, then AB and BA are square matrices.
(b) Show that if A is an $m \times n$ matrix and $A(BA)$ is defined, then B is an $n \times m$ matrix.

13. In each part find matrices A, $\mathbf{x}$, and $\mathbf{b}$ that express the given system of linear equations as a single matrix equation $A\mathbf{x} = \mathbf{b}$.

(a)
$$\begin{aligned} 2x_1 - 3x_2 + 5x_3 &= 7 \\ 9x_1 - x_2 + x_3 &= -1 \\ x_1 + 5x_2 + 4x_3 &= 0 \end{aligned}$$

(b)
$$\begin{aligned} 4x_1 \qquad\quad - 3x_3 + x_4 &= 1 \\ 5x_1 + x_2 \qquad\quad - 8x_4 &= 3 \\ 2x_1 - 5x_2 + 9x_3 - x_4 &= 0 \\ 3x_2 - x_3 + 7x_4 &= 2 \end{aligned}$$

14. In each part, express the matrix equation as a system of linear equations.

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

(b) $$\begin{bmatrix} 3 & -2 & 0 & 1 \\ 5 & 0 & 2 & -2 \\ 3 & 1 & 4 & 7 \\ -2 & 5 & 1 & 6 \end{bmatrix}\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

15. If A and B are partitioned into submatrices, for example,

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right]$$

then AB can be expressed as

$$AB = \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right]$$

provided the sizes of the submatrices of A and B are such that the indicated operations can be performed. This method of multiplying partitioned matrices is called ***block multiplication.*** In each part compute the product by block multiplication. Check your results by multiplying directly.

(a) $$A = \left[\begin{array}{cc|cc} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array}\right], \qquad B = \left[\begin{array}{cc|c} 2 & 1 & 4 \\ -3 & 5 & 2 \\ \hline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array}\right]$$

(b) $$A = \left[\begin{array}{ccc|c} -1 & 2 & 1 & 5 \\ \hline 0 & -3 & 4 & 2 \\ 1 & 5 & 6 & 1 \end{array}\right], \qquad B = \left[\begin{array}{cc|c} 2 & 1 & 4 \\ -3 & 5 & 2 \\ 7 & -1 & 5 \\ \hline 0 & 3 & -3 \end{array}\right]$$

16. Adapt the method of Exercise 15 to compute the following products by block multiplication.

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

(b) $$\left[\begin{array}{cc} 2 & -5 \\ 1 & 3 \\ 0 & 5 \\ \hline 1 & 4 \end{array}\right]\begin{bmatrix} 2 & -1 & 3 & -4 \\ 0 & 1 & 5 & 7 \end{bmatrix}$$

(c) $$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 & 2 \end{array}\right]\left[\begin{array}{cc} 3 & 3 \\ -1 & 4 \\ 1 & 5 \\ \hline 2 & -2 \\ 1 & 6 \end{array}\right]$$

17. In each part determine whether block multiplication can be used to compute AB from the given partitions If so, compute the product by block multiplication.

(a) $$A = \left[\begin{array}{ccc|c} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array}\right], \qquad B = \left[\begin{array}{cc|c} 2 & 1 & 4 \\ -3 & 5 & 2 \\ \hline 7 & -1 & 5 \\ 0 & 3 & -3 \end{array}\right]$$

(b) $$A = \left[\begin{array}{cccc} -1 & 2 & 1 & 5 \\ 0 & -3 & 4 & 2 \\ \hline 1 & 5 & 6 & 1 \end{array}\right], \qquad B = \left[\begin{array}{c|c|c} 2 & 1 & 4 \\ -3 & 5 & 2 \\ 7 & -1 & 5 \\ 0 & 3 & -3 \end{array}\right]$$

18. (a) Show that if A has a row of zeros and B is any matrix for which AB is defined, then AB also has a row of zeros.
(b) Find a similar result involving a column of zeros.

19. Let A be any $m \times n$ matrix and let 0 be the $m \times n$ matrix each of whose entries is zero. Show that if $kA = 0$, then $k = 0$ or $A = 0$.

20. Let I be the $n \times n$ matrix whose entry in row i and column j is

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

Show that $AI = IA = A$ for every $n \times n$ matrix A.

21. In each part find a 6×6 matrix $[a_{ij}]$ that satisfies the stated condition. Make your answers as general as possible by using letters rather than specific numbers for the nonzero entries.
(a) $a_{ij} = 0$ if $i \neq j$ (b) $a_{ij} = 0$ if $i > j$ (c) $a_{ij} = 0$ if $i < j$ (d) $a_{ij} = 0$ if $|i - j| > 1$

22. Find the 4×4 matrix $A = [a_{ij}]$ whose entries satisfy the stated condition.
(a) $a_{ij} = i + j$ (b) $a_{ij} = i^{j-1}$ (c) $a_{ij} = \begin{cases} 1 & \text{if} \quad |i - j| > 1 \\ -1 & \text{if} \quad |i - j| \leq 1 \end{cases}$

23. Prove: If A is an $m \times n$ matrix, then

$$\operatorname{tr}(AA^T) = \operatorname{tr}(A^TA) = s$$

where s is the sum of the squares of the entries of A.

24. Use the result in Exercise 23 to prove the following.
(a) If A is an $m \times n$ matrix such that $AA^T = 0$ or $A^TA = 0$, then $A = 0$.
(b) If A is an $n \times n$ matrix such that $A = A^T$ and $A^2 = 0$, then $A = 0$.

1.4 INVERSES; RULES OF MATRIX ARITHMETIC

In this section we shall discuss some properties of the arithmetic operations on matrices. We shall see that many of the basic rules of arithmetic for real numbers also hold for matrices but a few do not.

PROPERTIES OF MATRIX OPERATIONS

For real numbers a and b, we always have $ab = ba$, which is called the *commutative law for multiplication*. For matrices, however, AB and BA need not be equal. Equality can fail to hold for three reasons. It can happen, for example, that AB is defined but BA is undefined. This is the case if A is a 2×3 matrix and B is a 3×4 matrix. Also, it can happen that AB and BA are both defined but have different sizes. This is the situation if A is a 2×3 matrix and B is a 3×2 matrix. Finally, as Example 1 shows, it is possible to have $AB \neq BA$ even if both AB and BA are defined and have the same size.

Example 1 Consider the matrices

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

Multiplying gives

$$AB = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix} \qquad BA = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$$

Thus, $AB \neq BA$.

Although the commutative law for multiplication is not valid in matrix arithmetic, many familiar laws of arithmetic are valid for matrices. Some of the most important ones and their names are summarized in the following theorem.

Theorem 1.4.1. *Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.*

(a) $A + B = B + A$ (***Commutative law for addition***)
(b) $A + (B + C) = (A + B) + C$ (***Associative law for addition***)
(c) $A(BC) = (AB)C$ (***Associative law for multiplication***)
(d) $A(B + C) = AB + AC$ (***Left distributive law***)
(e) $(B + C)A = BA + CA$ (***Right distributive law***)
(f) $A(B - C) = AB - AC$
(g) $(B - C)A = BA - CA$
(h) $a(B + C) = aB + aC$
(i) $a(B - C) = aB - aC$
(j) $(a + b)C = aC + bC$
(k) $(a - b)C = aC - bC$
(l) $a(bC) = (ab)C$
(m) $a(BC) = (aB)C = B(aC)$

To prove the equalities in this theorem we must show that the matrix on the left side has the same size as the matrix on the right side and that corresponding entries on the two sides are equal. With the exception of the associative law in part (*c*), the proofs all follow the same general pattern. We shall prove part (*d*) as an illustration. The proof of the associative law, which is more complicated, is outlined in the exercises.

Proof (*d*). We must show that $A(B + C)$ and $AB + AC$ have the same size and that corresponding entries are equal. To form $A(B + C)$, the matrices B and C must have the same size, say $m \times n$, and the matrix A must then have m columns, so its size must be of the form $r \times m$. This makes $A(B + C)$ an $r \times n$ matrix. It follows that $AB + AC$ is also an $r \times n$ matrix and, consequently, $A(B + C)$ and $AB + AC$ have the same size.

Suppose that $A = [a_{ij}]$, $B = [b_{ij}]$, and $C = [c_{ij}]$. We want to show that corresponding entries of $A(B + C)$ and $AB + AC$ are equal; that is,

$$[A(B + C)]_{ij} = [AB + AC]_{ij}$$

for all values of i and j. But from the definitions of matrix addition and matrix multiplication we have

$$\begin{aligned} [A(B + C)]_{ij} &= a_{i1}(b_{1j} + c_{1j}) + a_{i2}(b_{2j} + c_{2j}) + \cdots + a_{im}(b_{mj} + c_{mj}) \\ &= (a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}) + (a_{i1}c_{1j} + a_{i2}c_{2j} + \cdots + a_{im}c_{mj}) \\ &= [AB]_{ij} + [AC_{ij}] = [AB + AC]_{ij} \end{aligned}$$

REMARK. Although the operations of matrix addition and matrix multiplication were defined for pairs of matrices, associative laws (*b*) and (*c*) enable us to denote sums and products of three matrices as $A + B + C$ and ABC without inserting any parentheses. This is justified by the fact that no matter how parentheses are inserted, the associative laws guarantee that the same end result will be obtained. In general, *given any sum or any product of matrices, pairs of parentheses can be inserted or deleted anywhere within the expression without affecting the end result.*

Example 2 As an illustration of the associative law for matrix multiplication, consider

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

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad BC = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}$$

Thus,

$$(AB)C = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

and

$$A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

so $(AB)C = A(BC)$, as guaranteed by Theorem 1.4.1c.

ZERO MATRICES A matrix, all of whose entries are zero, such as

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad [0]$$

is called a ***zero matrix***. A zero matrix will be denoted by 0; if it is important to emphasize the size, we shall write $0_{m\times n}$ for the $m \times n$ zero matrix.

If A is any matrix and 0 is the zero matrix with the same size, it is obvious that $A + 0 = 0 + A = A$. The matrix 0 plays much the same role in these matrix equations as the number 0 plays in the numerical equations $a + 0 = 0 + a = a$.

Since we already know that some of the rules of arithmetic for real numbers do not carry over to matrix arithmetic, it would be foolhardy to assume that all the properties of the real number zero carry over to zero matrices. For example, consider the following two standard results in the arithmetic of real numbers.

- If $ab = ac$ and $a \neq 0$, then $b = c$. (This is called the *cancellation law.*)
- If $ad = 0$, then at least one of the factors on the left is 0.

As the next example shows, the corresponding results are not generally true in matrix arithmetic.

Example 3 Consider the matrices

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

Here

$$AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Although $A \neq 0$, it is *incorrect* to cancel the A from both sides of the equation $AB = AC$ and write $B = C$. Thus, the cancellation law fails to hold for matrices. Also, $AD = 0$, yet $A \neq 0$ and $D \neq 0$.

In spite of the above example, there are a number of familiar properties of the real number 0 that *do* carry over to zero matrices. Some of the more important ones are summarized in the next theorem. The proofs are left as exercises.

Theorem 1.4.2. *Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.*

(a) $A + 0 = 0 + A = A$
(b) $A - A = 0$
(c) $0 - A = -A$
(d) $A0 = 0;\quad 0A = 0$

IDENTITY MATRICES Of special interest are square matrices with 1's on the main diagonal and 0's off the main diagonal, such as

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{and so on.}$$

A matrix of this form is called an ***identity matrix*** and is denoted by I. If it is important to emphasize the size, we shall write I_n for the $n \times n$ identity matrix.

If A is an $m \times n$ matrix, then, as illustrated in the next example,

$$AI_n = A \quad \text{and} \quad I_m A = A$$

Thus, an identity matrix plays much the same role in matrix arithmetic as the number 1 plays in the numerical relationships $a \cdot 1 = 1 \cdot a = a$.

Example 4 Consider the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

Then

$$I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

and

$$AI_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

As the next theorem shows, identity matrices arise naturally in studying reduced row-echelon forms of *square* matrices.

Theorem 1.4.3. *If R is the reduced row-echelon form of an $n \times n$ matrix A, then either R has a row of zeros or R is the identity matrix I_n.*

Proof. Suppose that the reduced row-echelon form of A is

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

Either the last row in this matrix consists entirely of zeros or it does not. If not, the matrix contains no zero rows, and consequently each of the n rows has a leading entry of 1. Since these leading 1's occur progressively further to the right as we move down the matrix, each of these 1's must occur on the main diagonal. Since the other entries in the same column as one of these 1's are zero, R must be I_n. Thus, either R has a row of zeros or $R = I_n$.

INVERSE OF A MATRIX

Definition. If A is a square matrix, and if a matrix B of the same size can be found such that $AB = BA = I$, then A is said to be ***invertible*** and B is called an ***inverse*** of A.

Example 5 The matrix

$$B = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix} \quad \text{is an inverse of} \quad A = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$$

since

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

and

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

Example 6 The matrix

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

is not invertible. To see why, let

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

be any 3×3 matrix. The third column of BA is

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

Thus,

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

PROPERTIES OF INVERSES

It is reasonable to ask whether an invertible matrix can have more than one inverse. The next theorem shows that the answer is no—*an invertible matrix has exactly one inverse.*

Theorem 1.4.4. *If B and C are both inverses of the matrix A, then $B = C$.*

Proof. Since B is an inverse of A, we have $BA = I$. Multiplying both sides on the right by C gives $(BA)C = IC = C$. But $(BA)C = B(AC) = BI = B$, so that $C = B$.

As a consequence of this important result, we can now speak of "the" inverse of an invertible matrix. If A is invertible, then its inverse will be denoted by the symbol A^{-1}. Thus,

$$AA^{-1} = I \quad \text{and} \quad A^{-1}A = I$$

The inverse of A plays much the same role in matrix arithmetic that reciprocal a^{-1} plays in the numerical relationships $aa^{-1} = 1$ and $a^{-1}a = 1$.

In the next section we shall develop a method for finding inverses of invertible matrices of any size; however, the following theorem gives conditions under which a 2×2 matrix is invertible and provides a simple formula for the inverse.

Theorem 1.4.5. *The matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if $ad - bc \neq 0$, in which case the inverse is given by the formula

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

Proof. We leave it for the reader to verify that $AA^{-1} = I_2$ and $A^{-1}A = I_2$.

Theorem 1.4.6. *If A and B are invertible matrices of the same size, then:*

(a) AB is invertible.
(b) $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. If we can show that $(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I$, then we will have simultaneously shown that the matrix AB is invertible and that $(AB)^{-1} = B^{-1}A^{-1}$. But $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. A similar argument shows that $(B^{-1}A^{-1})(AB) = I$.

Although we will not prove it, this result can be extended to include three or more factors; that is,

A product of any number of invertible matrices is invertible, and the inverse of the product is the product of the inverses in the reverse order.

Example 7 Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix} \qquad AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix}$$

Applying the formula in Theorem 1.4.5, we obtain

$$A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} \qquad B^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix} \qquad (AB)^{-1} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Also,

$$B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Therefore, $(AB)^{-1} = B^{-1}A^{-1}$ as guaranteed by Theorem 1.4.6.

POWERS OF A MATRIX

Next, we shall define powers of a square matrix and discuss their properties.

Definition. If A is a square matrix, then we define the nonnegative integer powers of A to be

$$A^0 = I \qquad A^n = \underbrace{AA \cdots A}_{n \text{ factors}} \qquad (n > 0)$$

Moreover, if A is invertible, then we define the negative integer powers to be

$$A^{-n} = (A^{-1})^n = \underbrace{A^{-1}A^{-1} \cdots A^{-1}}_{n \text{ factors}}$$

Because this definition parallels that for real numbers, the usual laws of exponents hold. (We omit the details.)

Theorem 1.4.7. *If A is a square matrix and r and s are integers, then*

$$A^rA^s = A^{r+s} \qquad (A^r)^s = A^{rs}$$

The next theorem provides some useful properties of negative exponents.

Theorem 1.4.8. *If A is an invertible matrix, then:*
(a) A^{-1} is invertible and $(A^{-1})^{-1} = A$.
(b) A^n is invertible and $(A^n)^{-1} = (A^{-1})^n$ for $n = 0, 1, 2, \ldots$
(c) For any nonzero scalar k, the matrix kA is invertible and $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Proof

(*a*) Since $AA^{-1} = A^{-1}A = I$, the matrix A^{-1} is invertible and $(A^{-1})^{-1} = A$.
(*b*) This part is left as an exercise.
(*c*) If k is any nonzero scalar, results (*l*) and (*m*) of Theorem 1.4.1 enable us to write

$$(kA)\left(\frac{1}{k}A^{-1}\right) = \frac{1}{k}(kA)A^{-1} = \left(\frac{1}{k}k\right)AA^{-1} = (1)I = I$$

Similarly, $\left(\frac{1}{k}A^{-1}\right)(kA) = I$ so that kA is invertible and $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Example 8 Let A and A^{-1} be as in Example 7, that is,

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

Then

$$A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$$

$$A^{-3} = (A^{-1})^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$$

POLYNOMIAL EXPRESSIONS INVOLVING MATRICES

If A is a square matrix, say $m \times m$, and if

$$p(x) = a_0 + a_1x + \cdots + a_nx^n \qquad (1)$$

is any polynomial, then we define

$$p(A) = a_0I + a_1A + \cdots + a_nA^n$$

where I is the $m \times m$ identity matrix. In words, $p(A)$ is the $m \times m$ matrix that results when A is substituted for x in (1) and a_0 is replaced by a_0I.

Example 9 If

$$p(x) = 2x^2 - 3x + 4 \quad \text{and} \quad A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}$$

then

$$p(A) = 2A^2 - 3A + 4I = 2\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}^2 - 3\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} + 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 8 \\ 0 & 18 \end{bmatrix} - \begin{bmatrix} -3 & 6 \\ 0 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 9 & 2 \\ 0 & 13 \end{bmatrix}$$

PROPERTIES OF THE TRANSPOSE

The next theorem lists the main properties of the transpose operation.

Theorem 1.4.9. *If the sizes of the matrices are such that the stated operations can be performed, then*
(a) $((A)^T)^T = A$
(b) $(A + B)^T = A^T + B^T$ *and* $(A - B)^T = A^T - B^T$
(c) $(kA)^T = kA^T$, *where k is any scalar*
(d) $(AB)^T = B^TA^T$

Keeping in mind that transposing a matrix interchanges its rows and columns, parts (*a*), (*b*), and (*c*) should be self-evident. For example, part (*a*) states that interchanging rows and columns twice leaves a matrix unchanged; part (*b*) asserts that adding and then interchanging rows and columns yields the same result as first interchanging rows and columns, then adding; and part (*c*) asserts that multiplying by a scalar and then interchanging rows and columns yields the same result as first interchanging rows and columns, then multiplying by the scalar. Part (*d*) is not so obvious, so we give its proof.

Proof (*d*). Let

$$A = [a_{ij}]_{m\times r} \quad \text{and} \quad B = [b_{ij}]_{r\times n}$$

so that the products AB and B^TA^T can both be formed. We leave it for the reader to check that $(AB)^T$ and B^TA^T have the same size, namely $n \times m$. Thus, it only remains to show that corresponding entries of $(AB)^T$ and B^TA^T are the same; that is,

$$\left((AB)^T\right)_{ij} = (B^TA^T)_{ij} \tag{2}$$

Applying Formula (8) of Section 1.3 to the left side of this equation and using the definition of matrix multiplication, we obtain

$$\left((AB)^T\right)_{ij} = (AB)_{ji} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jr}b_{ri} \tag{3}$$

To evaluate the right side of (2) it will be convenient to let a'_{ij} and b'_{ij} denote the ijth

entries of A^T and B^T, respectively, so

$$a'_{ij} = a_{ji} \quad \text{and} \quad b'_{ij} = b_{ji}$$

From these relationships and the definition of matrix multiplication we obtain

$$\begin{aligned}(B^TA^T)_{ij} &= b'_{i1}a'_{1j} + b'_{i2}a'_{2j} + \cdots + b'_{ir}a'_{rj} \\ &= b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ri}a_{jr} \\ &= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jr}b_{ri}\end{aligned}$$

This, together with (3), proves (2).

Although we shall not prove it, part (*d*) of this theorem can be extended to include three or more factors; that is,

The transpose of a product of any number of matrices is equal to the product of their transposes in the reverse order.

REMARK. Note the similarity between this result and the result following Theorem 1.4.6 about the inverse of a product of matrices.

INVERTIBILITY OF A TRANSPOSE

The following theorem establishes a relationship between the inverse of an invertible matrix and the inverse of its transpose.

Theorem 1.4.10. *If A is an invertible matrix, then* A^T *is also invertible and*

$$(A^T)^{-1} = (A^{-1})^T \tag{4}$$

Proof. We can prove the invertibility of A^T and obtain (4) by showing that

$$A^T(A^{-1})^T = (A^{-1})^TA^T = I$$

But from part (*d*) of Theorem 1.4.9 and the fact that $I^T = I$ we have

$$A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$$
$$(A^{-1})^TA^T = (AA^{-1})^T = I^T = I$$

which completes the proof.

Example 10 Consider the matrices

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

Applying Theorem 1.4.5 yields

$$A^{-1} = \begin{bmatrix} 1 & 3 \\ -2 & -5 \end{bmatrix} \quad (A^T)^{-1} = \begin{bmatrix} 1 & -2 \\ 3 & -5 \end{bmatrix}$$

As guaranteed by Theorem 1.4.10, these matrices satisfy (4).

EXERCISE SET 1.4

1. Let

$$A=\begin{bmatrix}2&-1&3\\0&4&5\\-2&1&4\end{bmatrix},\quad B=\begin{bmatrix}8&-3&-5\\0&1&2\\4&-7&6\end{bmatrix},\quad C=\begin{bmatrix}0&-2&3\\1&7&4\\3&5&9\end{bmatrix},\quad a=4,\quad b=-7$$

Show that
(a) $A+(B+C)=(A+B)+C$ (b) $(AB)C=A(BC)$ (c) $(a+b)C=aC+bC$
(d) $a(B-C)=aB-aC$

2. Using the matrices and scalars in Exercise 1, show that
(a) $a(BC)=(aB)C=B(aC)$ (b) $A(B-C)=AB-AC$ (c) $(B+C)A=BA+CA$
(d) $a(bC)=(ab)C$

3. Using the matrices and scalars in Exercise 1, show that
(a) $(A^T)^T=A$ (b) $(A+B)^T=A^T+B^T$ (c) $(aC)^T=aC^T$ (d) $(AB)^T=B^TA^T$

4. Use Theorem 1.4.5 to compute the inverses of the following matrices.

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

5. Verify that the three matrices A, B, and C in Exercise 4 satisfy the relationships $(AB)^{-1}=B^{-1}A^{-1}$ and $(ABC)^{-1}=C^{-1}B^{-1}A^{-1}$.

6. Let A and B be square matrices of the same size. Is $(AB)^2=A^2B^2$ a valid matrix identity? Justify your answer.

7. In each part use the given information to find A.

(a) $A^{-1}=\begin{bmatrix}2&-1\\3&5\end{bmatrix}$ (b) $(7A)^{-1}=\begin{bmatrix}-3&7\\1&-2\end{bmatrix}$

(c) $(5A^T)^{-1}=\begin{bmatrix}-3&-1\\5&2\end{bmatrix}$ (d) $(I+2A)^{-1}=\begin{bmatrix}-1&2\\4&5\end{bmatrix}$

8. Let A be the matrix

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

Compute A^3, A^{-3}, and A^2-2A+I.

9. Let A be the matrix

$$\begin{bmatrix}3&1\\2&1\end{bmatrix}$$

In each part find $p(A)$.
(a) $p(x)=x-2$ (b) $p(x)=2x^2-x+1$ (c) $p(x)=x^3-2x+4$

10. Let $p_1(x)=x^2-9$, $p_2(x)=x+3$, and $p_3=x-3$.
(a) Show that $p_1(A)=p_2(A)p_3(A)$ for the matrix A in Exercise 9.
(b) Show that $p_1(A)=p_2(A)p_3(A)$ for any square matrix A.

11. Find the inverse of

$$\begin{bmatrix}\cos\theta&\sin\theta\\-\sin\theta&\cos\theta\end{bmatrix}$$

12. (a) Find 2×2 matrices A and B such that $(A+B)^2\neq A^2+2AB+B^2$.
(b) Show that if A and B are square matrices such that $AB=BA$, then

$$(A+B)^2=A^2+2AB+B^2$$

(c) Find an expansion of $(A+B)^2$ that is valid for all square matrices A and B having the same size.

13. Consider the matrix

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

where $a_{11}a_{22}\cdots a_{nn}\neq 0$. Show that A is invertible and find its inverse.

14. Show that if a square matrix A satisfies $A^2-3A+I=0$, then $A^{-1}=3I-A$.

15. (a) Show that a matrix with a row of zeros cannot have an inverse.
(b) Show that a matrix with a column of zeros cannot have an inverse.

16. Is the sum of two invertible matrices necessarily invertible?

17. Let A and B be square matrices such that $AB=0$. Show that if A is invertible, then $B=0$.

18. In Theorem 1.4.2 why didn't we write part (*d*) as $A0=0=0A$?

19. The real equation $a^2=1$ has exactly two solutions. Find at least eight different 3×3 matrices that satisfy the matrix equation $A^2=I_3$. [***Hint.*** Look for solutions in which all the entries off the main diagonal are zero.]

20. (a) Find a nonzero 3×3 matrix A such that $A^T=A$.
(b) Find a nonzero 3×3 matrix A such that $A^T=-A$.

21. A square matrix A is called ***symmetric*** if $A^T=A$ and ***skew-symmetric*** if $A^T=-A$. Show that if B is a square matrix, then
(a) BB^T and $B+B^T$ are symmetric (b) $B-B^T$ is skew-symmetric

22. If A is a square matrix and n is a positive integer, is it true that $(A^n)^T=(A^T)^n$? Justify your answer.

23. Let A be the matrix

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

Determine whether A is invertible, and if so, find its inverse. [***Hint.*** Solve $AX=I$ by equating corresponding entries on the two sides.]

24. Prove:
(a) part (*b*) of Theorem 1.4.1 (b) part (*i*) of Theorem 1.4.1 (c) part (*m*) of Theorem 1.4.1

25. Apply parts (*d*) and (*m*) of Theorem 1.4.1 to the matrices A, B, and $(-1)C$ to derive the result in part (*f*).

26. Prove Theorem 1.4.2.

27. Consider the laws of exponents $A^rA^s = A^{r+s}$ and $(A^r)^s = A^{rs}$.
(a) Show that if A is any square matrix, these laws are valid for all nonnegative integer values of r and s.
(b) Show that if A is invertible, these laws hold for all negative integer values of r and s.

28. Show that if A is invertible and k is any nonzero scalar, then $(kA)^n = k^nA^n$ for all integer values of n.

29. (a) Show that if A is invertible and $AB = AC$, then $B = C$.
(b) Explain why part (a) and Example 3 do not contradict one another.

30. Prove part (*c*) of Theorem 1.4.1. [***Hint.*** Assume that A is $m \times n$, B is $n \times p$, and C is $p \times q$. The ijth entry on the left side is $l_{ij} = a_{i1}[BC]_{1j} + a_{i2}[BC]_{2j} + \cdots + a_{in}[BC]_{nj}$ and the ijth entry on the right side is $r_{ij} = [AB]_{i1}c_{1j} + [AB]_{i2}c_{2j} + \cdots + [AB]_{ip}c_{pj}$. Verify that $l_{ij} = r_{ij}$.]

1.5 ELEMENTARY MATRICES AND A METHOD FOR FINDING A^{-1}

In this section we shall develop an algorithm for finding the inverse of an invertible matrix, and we shall discuss some basic properties of invertible matrices.

ELEMENTARY MATRICES

Definition. An $n \times n$ matrix is called an ***elementary matrix*** if it can be obtained from the $n \times n$ identity matrix I_n by performing a single elementary row operation.

Example 1 Listed below are four elementary matrices and the operations that produce them.

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

Multiply the second row of I_2 by -3.	Interchange the second and fourth rows of I_4.	Add 3 times the third row of I_3 to the first row.	Multiply the first row of I_3 by 1.

When a matrix A is multiplied on the *left* by an elementary matrix E, the effect is to perform an elementary row operation on A. This is the content of the following theorem, the proof of which is left for the exercises.

Theorem 1.5.1. *If the elementary matrix E results from performing a certain row operation on I_m and if A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.*

Example 2 Consider the matrix

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

and consider the elementary matrix

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

which results from adding 3 times the first row of I_3 to the third row. The product EA is

$$EA = \begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & -1 & 3 & 6 \\ 4 & 4 & 10 & 9 \end{bmatrix}$$

which is precisely the same matrix that results when we add 3 times the first row of A to the third row.

REMARK. Theorem 1.5.1 is primarily of theoretical interest and will be used for developing some results about matrices and systems of linear equations. Computationally, it is preferable to perform row operations directly rather than multiplying on the left by an elementary matrix.

If an elementary row operation is applied to an identity matrix I to produce an elementary matrix E, then there is a second row operation that, when applied to E, produces I back again. For example, if E is obtained by multiplying the ith row of I by a nonzero constant c, then I can be recovered if the ith row of E is multiplied by $1/c$. The various possibilities are listed in Table 1.

TABLE 1

Row Operation on I That Produces E	Row Operation on E That Reproduces I
Multiply row i by $c \neq 0$	Multiply row i by $1/c$
Interchange rows i and j	Interchange rows i and j
Add c times row i to row j	Add $-c$ times row i to row j

The operations on the right side of this table are called the ***inverse operations*** of the corresponding operations on the left.

Example 3 In each of the following, an elementary row operation is applied to the 2×2 identity matrix to obtain an elementary matrix E, then E is restored to the identity matrix by applying the inverse row operation.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Multiply the second row by 7. — Multiply the second row by $\frac{1}{7}$.

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

Interchange the first and second rows. — Interchange the first and second rows.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Add 5 times the second row to the first. — Add -5 times the second row to the first.

The next theorem gives an important property of elementary matrices.

Theorem 1.5.2. *Every elementary matrix is invertible, and the inverse is also an elementary matrix.*

Proof. If E is an elementary matrix, then E results from performing some row operation on I. Let E_0 be the matrix that results when the inverse of this operation is performed on I. Applying Theorem 1.5.1 and using the fact that inverse row operations cancel the effect of each other, it follows that

$$E_0E = I \qquad \text{and} \qquad EE_0 = I$$

Thus, the elementary matrix E_0 is the inverse of E.

The next theorem establishes some fundamental relationships between invertibility, homogeneous linear systems, reduced row-echelon forms, and elementary matrices. These results are extremely important and will be used many times in later sections.

Theorem 1.5.3. *If A is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false.*
(a) A is invertible.
(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
(c) The reduced row-echelon form of A is I_n.
(d) A is expressible as a product of elementary matrices.

Proof. We shall prove the equivalence by establishing the chain of implications: $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$.

$(a) \Rightarrow (b)$: Assume A is invertible and let $\mathbf{x}_0$ be any solution of $A\mathbf{x} = \mathbf{0}$; thus, $A\mathbf{x}_0 = \mathbf{0}$. Multiplying both sides of this equation by the matrix A^{-1} gives $A^{-1}(A\mathbf{x}_0) = A^{-1}\mathbf{0}$, or $(A^{-1}A)\mathbf{x}_0 = \mathbf{0}$, or $I\mathbf{x}_0 = \mathbf{0}$, or $\mathbf{x}_0 = \mathbf{0}$. Thus, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

$(b) \Rightarrow (c)$: Let $A\mathbf{x} = \mathbf{0}$ be the matrix form of the system

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & 0 \end{array} \tag{1}$$

and assume that the system has only the trivial solution. If we solve by Gauss–Jordan elimination, then the system of equations corresponding to the reduced row-echelon form of the augmented matrix will be

$$\begin{array}{llll} x_1 & & & = 0 \\ & x_2 & & = 0 \\ & & \ddots & \\ & & & x_n = 0 \end{array} \tag{2}$$

Thus, the augmented matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{bmatrix}$$

for (1) can be reduced to the augmented matrix

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

for (2) by a sequence of elementary row operations. If we disregard the last column (of zeros) in each of these matrices, we can conclude that the reduced row-echelon form of A is I_n.

$(c)\Rightarrow(d)$: Assume that the reduced row-echelon form of A is I_n, so that A can be reduced to I_n by a finite sequence of elementary row operations. By Theorem 1.5.1 each of these operations can be accomplished by multiplying on the left by an appropriate elementary matrix. Thus, we can find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2 E_1 A = I_n \tag{3}$$

By Theorem 1.5.2, $E_1, E_2, \ldots, E_k$ are invertible. Multiplying both sides of Equation (3) on the left successively by $E_k^{-1}, \ldots, E_2^{-1}, E_1^{-1}$ we obtain

$$A = E_1^{-1}E_2^{-1}\cdots E_k^{-1}I_n = E_1^{-1}E_2^{-1}\cdots E_k^{-1} \tag{4}$$

By Theorem 1.5.2, this equation expresses A as a product of elementary matrices.

$(d)\Rightarrow(a)$: If A is a product of elementary matrices, then from Theorems 1.4.6 and 1.5.2 the matrix A is a product of invertible matrices, and hence is invertible.

ROW EQUIVALENCE

If a matrix B can be obtained from a matrix A by performing a finite sequence of elementary row operations, then obviously we can get from B back to A by performing the inverses of these elementary row operations in reverse order. Matrices that can be obtained from one another by a finite sequence of elementary row operations are said to be ***row equivalent.*** With this terminology it follows from parts (*a*) and (*c*) of Theorem 1.5.3 that an $n \times n$ matrix A is invertible if and only if it is row equivalent to the $n \times n$ identity matrix.

A METHOD FOR INVERTING MATRICES

As our first application of Theorem 1.5.3, we shall establish a method for determining the inverse of an invertible matrix. Inverting the left and the right sides of (4) yields $A^{-1} = E_k \cdots E_2 E_1$, or equivalently,

$$A^{-1} = E_k \cdots E_2 E_1 I_n \tag{5}$$

which tells us that A^{-1} can be obtained by multiplying I_n successively on the left by the elementary matrices $E_1, E_2, \ldots, E_k$. Since each multiplication on the left by one of these elementary matrices performs a row operation, it follows, by comparing Equations (3) and (5), that *the sequence of row operations that reduces A to I_n will reduce I_n to A^{-1}.* Thus, we have the following result:

To find the inverse of an invertible matrix A, we must find a sequence of elementary row operations that reduces A to the identity and then perform this same sequence of operations on I_n to obtain A^{-1}.

A simple method for carrying out this procedure is given in the following example.

Example 4 Find the inverse of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$

Solution. We want to reduce A to the identity matrix by row operations and simultaneously apply these operations to I to produce A^{-1}. To accomplish this we shall adjoin the identity matrix to the right side of A, thereby producing a matrix of the form

$$[A \mid I]$$

then we shall apply row operations to this matrix until the left side is reduced to I; these operations will convert the right side to A^{-1}, so that the final matrix will have the form

$$[I \mid A^{-1}]$$

The computations are as follows:

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

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

We added -2 times the first row to the second and -1 times the first row to the third.

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

We added 2 times the second row to the third.

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

We multiplied the third row by -1.

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

We added 3 times the third row to the second and -3 times the third row to the first.

$$\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

We added -2 times the second row to the first.

Thus,

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

Often it will not be known in advance whether a given matrix is invertible. If an $n \times n$ matrix A is not invertible, then it cannot be reduced to I_n by elementary row operations [part (c) of Theorem 1.5.3]. Stated another way, the reduced row-echelon form of A has at least one row of zeros. Thus, if the procedure in the last example is attempted on a matrix that is not invertible, then at some point in the computations a row of zeros will occur on the *left side*. It can then be concluded that the given matrix is not invertible, and the computations can be stopped.

Example 5 Consider the matrix

$$A = \begin{bmatrix} 1 & 6 & 4 \\ 2 & 4 & -1 \\ -1 & 2 & 5 \end{bmatrix}$$

Applying the procedure of Example 4 yields

$$\left[\begin{array}{rrr|rrr} 1 & 6 & 4 & 1 & 0 & 0 \\ 2 & 4 & -1 & 0 & 1 & 0 \\ -1 & 2 & 5 & 0 & 0 & 1 \end{array}\right]$$

$$\left[\begin{array}{rrr|rrr} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 8 & 9 & 1 & 0 & 1 \end{array}\right]$$

We added -2 times the first row to the second and added the first row to the third.

$$\left[\begin{array}{rrr|rrr} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right]$$

We added the second row to the third.

Since we have obtained a row of zeros on the left side, A is not invertible.

Example 6 In Example 4 we showed that

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$$

is an invertible matrix. From Theorem 1.5.3 it follows that the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ 2x_1 + 5x_2 + 3x_3 &= 0 \\ x_1 \qquad\quad + 8x_3 &= 0 \end{aligned}$$

has only the trivial solution.

EXERCISE SET 1.5

1. Which of the following are elementary matrices?

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

(e) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (f) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 9 \\ 0 & 0 & 1 \end{bmatrix}$ (g) $\begin{bmatrix} 2 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

2. Find a row operation that will restore the given elementary matrix to an identity matrix.

(a) $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & 0 & -\frac{1}{7} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

3. Consider the matrices

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

Find elementary matrices E_1, E_2, E_3, and E_4 such that
(a) $E_1A = B$ (b) $E_2B = A$ (c) $E_3A = C$ (d) $E_4C = A$

4. In Exercise 3 is it possible to find an elementary matrix E such that $EB = C$? Justify your answer.

In Exercises 5–7 use the method shown in Examples 4 and 5 to find the inverse of the given matrix if the matrix is invertible and check your answer by multiplication.

5. (a) $\begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix}$ (b) $\begin{bmatrix} -3 & 6 \\ 4 & 5 \end{bmatrix}$ (c) $\begin{bmatrix} 6 & -4 \\ -3 & 2 \end{bmatrix}$

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

7. (a) $\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{10} \\ \frac{1}{5} & -\frac{4}{5} & \frac{1}{10} \end{bmatrix}$ (b) $\begin{bmatrix} \sqrt{2} & 3\sqrt{2} & 0 \\ -4\sqrt{2} & \sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 3 & 5 & 0 \\ 1 & 3 & 5 & 7 \end{bmatrix}$

(d) $\begin{bmatrix} -8 & 17 & 2 & \frac{1}{3} \\ 4 & 0 & \frac{2}{5} & -9 \\ 0 & 0 & 0 & 0 \\ -1 & 13 & 4 & 2 \end{bmatrix}$ (e) $\begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 3 & 0 \\ 2 & 1 & 5 & -3 \end{bmatrix}$

8. Find the inverse of each of the following 4×4 matrices, where k_1, k_2, k_3, k_4, and k are all nonzero.

(a) $\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 0 & 0 & k_1 \\ 0 & 0 & k_2 & 0 \\ 0 & k_3 & 0 & 0 \\ k_4 & 0 & 0 & 0 \end{bmatrix}$ (c) $\begin{bmatrix} k & 0 & 0 & 0 \\ 1 & k & 0 & 0 \\ 0 & 1 & k & 0 \\ 0 & 0 & 1 & k \end{bmatrix}$

9. Consider the matrix

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

(a) Find elementary matrices E_1 and E_2 such that $E_2E_1A = I$.
(b) Write A^{-1} as a product of two elementary matrices.
(c) Write A as a product of two elementary matrices.

10. In each part perform the stated row operation on

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

by multiplying A on the left by a suitable elementary matrix. Check your answer in each case by performing the row operation directly on A.
(a) Interchange the first and third rows.
(b) Multiply the second row by $\frac{1}{3}$.
(c) Add twice the second row to the first row.

11. Express the matrix

$$A = \begin{bmatrix} 0 & 1 & 7 & 8 \\ 1 & 3 & 3 & 8 \\ -2 & -5 & 1 & -8 \end{bmatrix}$$

in the form $A = EFGR$, where E, F, and G are elementary matrices, and R is in row-echelon form.

12. Show that if

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

is an elementary matrix, then at least one entry in the third row must be a zero.

13. Show that

$$A = \begin{bmatrix} 0 & a & 0 & 0 & 0 \\ b & 0 & c & 0 & 0 \\ 0 & d & 0 & e & 0 \\ 0 & 0 & f & 0 & g \\ 0 & 0 & 0 & h & 0 \end{bmatrix}$$

is not invertible for any values of the entries.

14. Prove that if A is an $m \times n$ matrix, there is an invertible matrix C such that CA is in reduced row-echelon form.

15. Prove that if A is an invertible matrix and B is row equivalent to A, then B is also invertible.

16. (a) Prove: If A and B are $m \times n$ matrices, then A and B are row equivalent if and only if A and B have the same reduced row-echelon form.
(b) Show that A and B are row equivalent, and find a sequence of elementary row operations that produces B from A.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 1 \\ 2 & 1 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 2 & -2 \\ 1 & 1 & 4 \end{bmatrix}$$

17. Prove Theorem 1.5.1.

1.6 FURTHER RESULTS ON SYSTEMS OF EQUATIONS AND INVERTIBILITY

In this section we shall establish more results about systems of linear equations and invertibility of matrices. Our work will lead to a totally new method for solving n equations in n unknowns.

A BASIC THEOREM

We begin by proving a fundamental result about linear systems that was anticipated in the first section of this text.

> **Theorem 1.6.1.** *Every system of linear equations has either no solutions, exactly one solution, or infinitely many solutions.*

Proof. If $A\mathbf{x} = \mathbf{b}$ is a system of linear equations, exactly one of the following is true: (a) the system has no solutions, (b) the system has exactly one solution, or (c) the system has more than one solution. The proof will be complete if we can show that the system has infinitely many solutions in case (c).

Assume that $A\mathbf{x} = \mathbf{b}$ has more than one solution, and let $\mathbf{x}_0 = \mathbf{x}_1 - \mathbf{x}_2$, where $\mathbf{x}_1$ and $\mathbf{x}_2$ are any two distinct solutions. Because $\mathbf{x}_1$ and $\mathbf{x}_2$ are distinct, the matrix $\mathbf{x}_0$ is nonzero; moreover,

$$A\mathbf{x}_0 = A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

If we now let k be any scalar, then

$$\begin{aligned} A(\mathbf{x}_1 + k\mathbf{x}_0) &= A\mathbf{x}_1 + A(k\mathbf{x}_0) = A\mathbf{x}_1 + k(A\mathbf{x}_0) \\ &= \mathbf{b} + k\mathbf{0} = \mathbf{b} + \mathbf{0} = \mathbf{b} \end{aligned}$$

But this says that $\mathbf{x}_1 + k\mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{b}$. Since $\mathbf{x}_0$ is nonzero and there are infinitely many choices for k, the system $A\mathbf{x} = \mathbf{b}$ has infinitely many solutions.

SOLVING LINEAR SYSTEMS BY MATRIX INVERSION

Thus far, we have studied two methods for solving linear systems: Gaussian elimination and Gauss–Jordan elimination. The following theorem provides a new method for solving certain linear systems.

> **Theorem 1.6.2.** *If A is an invertible $n \times n$ matrix, then for each $n \times 1$ matrix $\mathbf{b}$, the system of equations $A\mathbf{x} = \mathbf{b}$ has exactly one solution, namely, $\mathbf{x} = A^{-1}\mathbf{b}$.*

Theorem 1.6.4. *If A is an $n \times n$ matrix, then the following are equivalent.*
(a) *A is invertible.*
(b) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*
(c) *The reduced row-echelon form of A is* I_n.
(d) *A is expressible as a product of elementary matrices.*
(e) $A\mathbf{x} = \mathbf{b}$ *is consistent for every* $n \times 1$ *matrix* $\mathbf{b}$.
(f) $A\mathbf{x} = \mathbf{b}$ *has exactly one solution for every* $n \times 1$ *matrix* $\mathbf{b}$.

Proof. Since we proved in Theorem 1.5.3 that (a), (b), (c), and (d) are equivalent, it will be sufficient to prove that (a)⇒(f)⇒(e)⇒(a).

(a)⇒(f): This was already proved in Theorem 1.6.2.

(f)⇒(e): This is self-evident: If $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$, then $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(e)⇒(a): If the system $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$, then in particular, the systems

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad A\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

are consistent. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ be solutions of the respective systems, and let us form an $n \times n$ matrix C having these solutions as columns. Thus, C has the form

$$C = [\mathbf{x}_1 \mid \mathbf{x}_2 \mid \cdots \mid \mathbf{x}_n]$$

As discussed in Section 1.3, the successive columns of the product AC will be

$$A\mathbf{x}_1, A\mathbf{x}_2, \ldots, A\mathbf{x}_n$$

Thus,

$$AC = [A\mathbf{x}_1 \mid A\mathbf{x}_2 \mid \cdots \mid A\mathbf{x}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

By part (b) of Theorem 1.6.3 it follows that $C = A^{-1}$. Thus, A is invertible.

We know from earlier work that invertible matrix factors produce an invertible product. The following theorem looks at the converse: It shows that if the product of square matrices is invertible, then the factors themselves must be invertible.

Theorem 1.6.5. *Let A and B be square matrices of the same size. If AB is invertible, then A and B must also be invertible.*

In our later work the following fundamental problem will occur frequently in various contexts.

A Fundamental Problem. Let A be a fixed $m \times n$ matrix. Find all $m \times 1$ matrices $\mathbf{b}$ such that the system of equations $A\mathbf{x} = \mathbf{b}$ is consistent.

If A is an invertible matrix, Theorem 1.6.2 completely solves this problem by asserting that for *every* $m \times 1$ matrix $\mathbf{b}$, the linear system $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$. If A is not square, or if A is square but not invertible, then Theorem 1.6.2 does not apply. In these cases the matrix $\mathbf{b}$ must satisfy certain conditions in order for $A\mathbf{x} = \mathbf{b}$ to be consistent. The following example illustrates how Gaussian elimination can be used to determine such conditions.

Example 3 What conditions must b_1, b_2, and b_3 satisfy in order for the system of equations

$$\begin{aligned} x_1 + x_2 + 2x_3 &= b_1 \\ x_1 \phantom{{}+x_2} + x_3 &= b_2 \\ 2x_1 + x_2 + 3x_3 &= b_3 \end{aligned}$$

to be consistent?

Solution. The augmented matrix is

$$\begin{bmatrix} 1 & 1 & 2 & b_1 \\ 1 & 0 & 1 & b_2 \\ 2 & 1 & 3 & b_3 \end{bmatrix}$$

which can be reduced to row-echelon form as follows.

$$\begin{bmatrix} 1 & 1 & 2 & b_1 \\ 0 & -1 & -1 & b_2 - b_1 \\ 0 & -1 & -1 & b_3 - 2b_1 \end{bmatrix}$$

−1 times the first row was added to the second and −2 times the first row was added to the third.

$$\begin{bmatrix} 1 & 1 & 2 & b_1 \\ 0 & 1 & 1 & b_1 - b_2 \\ 0 & -1 & -1 & b_3 - 2b_1 \end{bmatrix}$$

The second row was multiplied by −1.

$$\begin{bmatrix} 1 & 1 & 1 & b_2 \\ 0 & 1 & 1 & b_1 - b_2 \\ 0 & 0 & 0 & b_3 - b_2 - b_1 \end{bmatrix}$$

The second row was added to the third.

It is now evident from the third row in the matrix that the system has a solution if and only if b_1, b_2, and b_3 satisfy the condition

$$b_3 - b_2 - b_1 = 0 \quad \text{or} \quad b_3 = b_1 + b_2$$

To express this condition another way, $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is a matrix of the form

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_1 + b_2 \end{bmatrix}$$

where b_1 and b_2 are arbitrary.

Example 4 What conditions must b_1, b_2, and b_3 satisfy in order for the system of equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= b_1 \\ 2x_1 + 5x_2 + 3x_3 &= b_2 \\ x_1 \qquad + 8x_3 &= b_3 \end{aligned}$$

to be consistent?

Solution. The augmented matrix is

$$\begin{bmatrix} 1 & 2 & 3 & b_1 \\ 2 & 5 & 3 & b_2 \\ 1 & 0 & 8 & b_3 \end{bmatrix}$$

Reducing this to reduced row-echelon form yields (verify)

$$\begin{bmatrix} 1 & 0 & 0 & -40b_1 + 16b_2 + 9b_3 \\ 0 & 1 & 0 & 13b_1 - 5b_2 - 3b_3 \\ 0 & 0 & 1 & 5b_1 - 2b_2 - b_3 \end{bmatrix} \tag{2}$$

In this case there are no restrictions on b_1, b_2, and b_3; that is, the given system $A\mathbf{x} = \mathbf{b}$ has the unique solution

$$x_1 = -40b_1 + 16b_2 + 9b_3, \quad x_2 = 13b_1 - 5b_2 - 3b_3, \quad x_3 = 5b_1 - 2b_2 - b_3 \tag{3}$$

for all $\mathbf{b}$.

REMARK. Because the system $A\mathbf{x} = \mathbf{b}$ in the preceding example is consistent for all $\mathbf{b}$, it follows from Theorem 1.6.4 that A is invertible. We leave it for the reader to verify that the formulas in (3) can also be obtained by calculating $\mathbf{x} = A^{-1}\mathbf{b}$.

EXERCISE SET 1.6

In Exercises 1–8 solve the system by inverting the coefficient matrix and using Theorem 1.6.2.

1. $\begin{aligned} x_1 + x_2 &= 2 \\ 5x_1 + 6x_2 &= 9 \end{aligned}$

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

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

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

5. $\begin{aligned} x + y + z &= 5 \\ x + y - 4z &= 10 \\ -4x + y + z &= 0 \end{aligned}$

6. $\begin{aligned} - x - 2y - 3z &= 0 \\ w + x + 4y + 4z &= 7 \\ w + 3x + 7y + 9z &= 4 \\ -w - 2x - 4y - 6z &= 6 \end{aligned}$

7. $\begin{aligned} 3x_1 + 5x_2 &= b_1 \\ x_1 + 2x_2 &= b_2 \end{aligned}$

8. $\begin{aligned} x_1 + 2x_2 + 3x_3 &= b_1 \\ 2x_1 + 5x_2 + 5x_3 &= b_2 \\ 3x_1 + 5x_2 + 8x_3 &= b_3 \end{aligned}$

9. Solve the following general system by inverting the coefficient matrix and using Theorem 1.6.2.

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

Use the resulting formulas to find the solution if
(a) $b_1 = -1$, $b_2 = 3$, $b_3 = 4$ (b) $b_1 = 5$, $b_2 = 0$, $b_3 = 0$ (c) $b_1 = -1$, $b_2 = -1$, $b_3 = 3$

10. Solve the three systems in Exercise 9 using the method of Example 2.

In Exercises 11–14 use the method of Example 2 to solve the systems in all the parts simultaneously.

11. $\begin{aligned} x_1 - 5x_2 &= b_1 \\ 3x_1 + 2x_2 &= b_2 \end{aligned}$

(a) $b_1 = 1$, $b_2 = 4$
(b) $b_1 = -2$, $b_2 = 5$

12. $\begin{aligned} -x_1 + 4x_2 + x_3 &= b_1 \\ x_1 + 9x_2 - 2x_3 &= b_2 \\ 6x_1 + 4x_2 - 8x_3 &= b_3 \end{aligned}$

(a) $b_1 = 0$, $b_2 = 1$, $b_3 = 0$
(b) $b_1 = -3$, $b_2 = 4$, $b_3 = -5$

13. $\begin{aligned} 4x_1 - 7x_2 &= b_1 \\ x_1 + 2x_2 &= b_2 \end{aligned}$

(a) $b_1 = 0$, $b_2 = 1$
(b) $b_1 = -4$, $b_2 = 6$
(c) $b_1 = -1$, $b_2 = 3$
(d) $b_1 = -5$, $b_2 = 1$

14. $\begin{aligned} x_1 + 3x_2 + 5x_3 &= b_1 \\ -x_1 - 2x_2 \qquad &= b_2 \\ 2x_1 + 5x_2 + 4x_3 &= b_3 \end{aligned}$

(a) $b_1 = 1$, $b_2 = 0$, $b_3 = -1$
(b) $b_1 = 0$, $b_2 = 1$, $b_3 = 1$
(c) $b_1 = -1$, $b_2 = -1$, $b_3 = 0$

15. The method of Example 2 can be used for linear systems with infinitely many solutions. Use that method to solve the systems in both parts at the same time.

(a) $\begin{aligned} x_1 - 2x_2 + x_3 &= -2 \\ 2x_1 - 5x_2 + x_3 &= 1 \\ 3x_1 - 7x_2 + 2x_3 &= -1 \end{aligned}$ (b) $\begin{aligned} x_1 - 2x_2 + x_3 &= 1 \\ 2x_1 - 5x_2 + x_3 &= -1 \\ 3x_1 - 7x_2 + 2x_3 &= 0 \end{aligned}$

In Exercises 16–19 find conditions that b's must satisfy for the system to be consistent.

16. $\begin{aligned} 6x_1 - 4x_2 &= b_1 \\ 3x_1 - 2x_2 &= b_2 \end{aligned}$

17. $\begin{aligned} x_1 - 2x_2 + 5x_3 &= b_1 \\ 4x_1 - 5x_2 + 8x_3 &= b_2 \\ -3x_1 + 3x_2 - 3x_3 &= b_3 \end{aligned}$

18. $\begin{aligned} x_1 - 2x_2 - x_3 &= b_1 \\ -4x_1 + 5x_2 + 2x_3 &= b_2 \\ -4x_1 + 7x_2 + 4x_3 &= b_3 \end{aligned}$

19. $\begin{aligned} x_1 - x_2 + 3x_3 + 2x_1 &= b_1 \\ -2x_1 + x_2 + 5x_3 + x_1 &= b_2 \\ -3x_1 + 2x_2 + 2x_3 - x_1 &= b_3 \\ 4x_1 - 3x_2 + x_3 + 3x_1 &= b_4 \end{aligned}$

20. Consider the matrices

$$A=\begin{bmatrix}2&1&2\\2&2&-2\\3&1&1\end{bmatrix}\quad\text{and}\quad \mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

(a) Show that the equation $A\mathbf{x}=\mathbf{x}$ can be rewritten as $(A-I)\mathbf{x}=\mathbf{0}$ and use this result to solve $A\mathbf{x}=\mathbf{x}$ for $\mathbf{x}$.
(b) Solve $A\mathbf{x}=4\mathbf{x}$.

21. Solve the following matrix equation for X.

$$\begin{bmatrix}1&-1&1\\2&3&0\\0&2&-1\end{bmatrix}X=\begin{bmatrix}2&-1&5&7&8\\4&0&-3&0&1\\3&5&-7&2&1\end{bmatrix}$$

22. In each part determine whether the homogeneous system has a nontrivial solution (without using pencil and paper); then state whether the given matrix is invertible.

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

(b)
$$\begin{aligned}5x_1+x_2+4x_3+x_4&=0\\2x_3-x_4&=0\\x_3+x_4&=0\\7x_4&=0\end{aligned}\qquad\begin{bmatrix}5&1&4&1\\0&0&2&-1\\0&0&1&1\\0&0&0&7\end{bmatrix}$$

23. Let $A\mathbf{x}=\mathbf{0}$ be a homogeneous system of n linear equations in n unknowns that has only the trivial solution. Show that if k is any positive integer, then the system $A^k\mathbf{x}=\mathbf{0}$ also has only the trivial solution.

24. Let $A\mathbf{x}=\mathbf{0}$ be a homogeneous system of n linear equations in n unknowns, and let Q be an invertible $n\times n$ matrix. Show that $A\mathbf{x}=\mathbf{0}$ has just the trivial solution if and only if $(QA)\mathbf{x}=\mathbf{0}$ has just the trivial solution.

25. Let $A\mathbf{x}=\mathbf{b}$ be any consistent system of linear equations, and let $\mathbf{x}_1$ be a fixed solution. Show that every solution to the system can be written in the form $\mathbf{x}=\mathbf{x}_1+\mathbf{x}_0$, where $\mathbf{x}_0$ is a solution to $A\mathbf{x}=\mathbf{0}$. Show also that every matrix of this form is a solution.

26. Use part *(a)* of Theorem 1.6.3 to prove part *(b)*.

1.7 DIAGONAL, TRIANGULAR, AND SYMMETRIC MATRICES

In this section we shall consider certain classes of matrices that have special forms. The matrices that we study in this section are among the most important in linear algebra, and they will arise in many different settings throughout this text.

DIAGONAL MATRICES

A square matrix in which all of the entries off the main diagonal are zero is called a ***diagonal matrix***; some examples are

$$\begin{bmatrix}2&0\\0&-5\end{bmatrix}\quad\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\quad\begin{bmatrix}6&0&0&0\\0&-4&0&0\\0&0&0&0\\0&0&0&8\end{bmatrix}$$

A general $n\times n$ diagonal matrix D can be written as

$$D=\begin{bmatrix}d_1&0&\cdots&0\\0&d_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&d_n\end{bmatrix}\tag{1}$$

A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero; in this case the inverse of (1) is

$$D^{-1}=\begin{bmatrix}1/d_1&0&\cdots&0\\0&1/d_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&1/d_n\end{bmatrix}$$

The reader should verify that $DD^{-1}=D^{-1}D=I$.

Powers of diagonal matrices are easy to compute; we leave it for the reader to verify that if D is the diagonal matrix (1) and k is a positive integer, then

$$D^k=\begin{bmatrix}d_1^{\,k}&0&\cdots&0\\0&d_2^{\,k}&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&d_n^{\,k}\end{bmatrix}$$

Example 1 If

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

then

$$A^{-1}=\begin{bmatrix}1&0&0\\0&-\frac{1}{3}&0\\0&0&\frac{1}{2}\end{bmatrix}\quad A^5=\begin{bmatrix}1&0&0\\0&-243&0\\0&0&32\end{bmatrix}\quad A^{-5}=\begin{bmatrix}1&0&0\\0&-\frac{1}{243}&0\\0&0&\frac{1}{32}\end{bmatrix}$$

Matrix products that involve diagonal factors are especially easy to compute. For example,

$$\begin{bmatrix}d_1&0&0\\0&d_2&0\\0&0&d_3\end{bmatrix}\begin{bmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\end{bmatrix}=\begin{bmatrix}d_1a_{11}&d_1a_{12}&d_1a_{13}&d_1a_{14}\\d_2a_{21}&d_2a_{22}&d_2a_{23}&d_2a_{24}\\d_3a_{31}&d_3a_{32}&d_3a_{33}&d_3a_{34}\end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} = \begin{bmatrix} d_1a_{11} & d_2a_{12} & d_3a_{13} \\ d_1a_{21} & d_2a_{22} & d_3a_{23} \\ d_1a_{31} & d_2a_{32} & d_3a_{33} \\ d_1a_{41} & d_2a_{42} & d_3a_{43} \end{bmatrix}$$

In words, *to multiply a matrix A on the left by a diagonal matrix D, one can multiply successive rows of A by the successive diagonal entries of D, and to multiply A on the right by D one can multiply successive columns of A by the successive diagonal entries of D.*

TRIANGULAR MATRICES

A square matrix in which all the entries above the main diagonal are zero is called ***lower triangular***, and a square matrix in which all the entries below the main diagonal are zero is called ***upper triangular***. A matrix that is either upper triangular or lower triangular is called ***triangular***.

Example 2

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

A general 4×4 upper triangular matrix

A general 4×4 lower triangular matrix

REMARK. Observe that diagonal matrices are both upper triangular and lower triangular since they have zeros below and above the main diagonal. Observe also that a *square* matrix in row-echelon form is upper triangular since it has zeros below the main diagonal.

The following are four useful characterizations of triangular matrices. The reader will find it instructive to verify that the matrices in Example 2 have the stated properties.

- A square matrix $A = [a_{ij}]$ is upper triangular if and only if the ith row starts with at least $i - 1$ zeros.
- A square matrix $A = [a_{ij}]$ is lower triangular if and only if the jth column starts with at least $j - 1$ zeros.
- A square matrix $A = [a_{ij}]$ is upper triangular if and only if $a_{ij} = 0$ for $i > j$.
- A square matrix $A = [a_{ij}]$ is lower triangular if and only if $a_{ij} = 0$ for $i < j$.

The following theorem lists some of the basic properties of triangular matrices.

Theorem 1.7.1

(*a*) *The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.*

(*b*) *The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.*

(*c*) *A triangular matrix is invertible if and only if its diagonal entries are all nonzero.*

(*d*) *The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.*

Part (*a*) is evident from the fact that transposing a square matrix can be accomplished by reflecting the entries about the main diagonal; we omit the formal proof. We will prove (*b*), but we will defer the proofs of (*c*) and (*d*) to the next chapter, where we will have the tools to prove those results more efficiently.

Proof (*b*). We will prove the result for lower triangular matrices; the proof for upper triangular matrices is similar. Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be lower triangular $n \times n$ matrices, and let $C = [c_{ij}]$ be the product $C = AB$. From the remark preceding this theorem, we can prove that C is lower triangular by showing that $c_{ij} = 0$ for $i < j$. But from the definition of matrix multiplication,

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

If we assume that $i < j$, then the terms in this expression can be grouped as follows:

$$c_{ij} = \underbrace{a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ij-1}b_{j-1j}}_{\substack{\text{Terms in which the row} \\ \text{number of } b \text{ is less than the} \\ \text{column number of } b}} + \underbrace{a_{ij}b_{jj} + \cdots + a_{in}b_{nj}}_{\substack{\text{Terms in which the row} \\ \text{number of } a \text{ is less than} \\ \text{the column number of } a}}$$

In the first grouping all of the b factors are zero since B is lower triangular, and in the second grouping all of the a factors are zero since A is lower triangular. Thus, $c_{ij} = 0$, which is what we wanted to prove.

Example 3 Consider the upper triangular matrices

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

The matrix A is invertible, since its diagonal entries are nonzero, but the matrix B is not. We leave it for the reader to calculate the inverse of A by the method of Section 1.5 and show that

$$A^{-1} = \begin{bmatrix} 1 & -\frac{3}{2} & \frac{7}{5} \\ 0 & \frac{1}{2} & -\frac{2}{5} \\ 0 & 0 & \frac{1}{5} \end{bmatrix}$$

This inverse is upper triangular, as guaranteed by part (d) of Theorem 1.7.1. We also leave it for the reader to check that the product AB is

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

This product is upper triangular, as guaranteed by part (b) of Theorem 1.7.1.

SYMMETRIC MATRICES

A square matrix A is called ***symmetric*** if $A = A^T$.

Example 4 The following matrices are symmetric, since each is equal to its own transpose (verify).

$$\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix} \quad \begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix} \quad \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$

It is easy to recognize symmetric matrices by inspection: The entries on the main diagonal may be arbitrary, but "mirror images" of entries across the main diagonal must be equal (Figure 1).

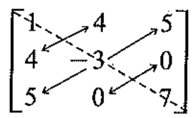


Figure 1

This follows from the fact that transposing a square matrix can be accomplished by interchanging entries that are symmetrically positioned about the main diagonal. Expressed in terms of the individual entries, a matrix $A = [a_{ij}]$ is symmetric if and only if $a_{ij} = a_{ji}$ for all values of i and j. As illustrated in Example 4, all diagonal matrices are symmetric.

The following theorem lists the main algebraic properties of symmetric matrices. The proofs are direct consequences of Theorem 1.4.9 and are left as exercises.

Theorem 1.7.2. *If A and B are symmetric matrices with the same size, and if k is any scalar, then:*

(a) A^T is symmetric.
(b) $A + B$ and $A - B$ are symmetric.
(c) kA is symmetric.

REMARK. It is not true, in general, that the product of symmetric matrices is symmetric. To see why this is so, let A and B be symmetric matrices with the same size. Then from part (d) of Theorem 1.4.9 and the symmetry we have

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

Since AB and BA are not usually equal, it follows that AB will not usually be symmetric. However, in the special case where $AB = BA$, the product AB will be symmetric. If A and B are matrices such that $AB = BA$, then we say that A and B ***commute***. In summary: *The product of two symmetric matrices is symmetric if and only if the matrices commute.*

Example 5 The first of the following equations shows a product of symmetric matrices that *is not* symmetric, and the second shows a product of symmetric matrices that *is* symmetric. We conclude that the factors in the first equation do not commute, but those in the second equation do. We leave it for the reader to verify that this is so.

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

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

In general, a symmetric matrix need not be invertible; for example, a square zero matrix is symmetric, but not invertible. However, if a symmetric matrix is invertible, then that inverse is also symmetric.

Theorem 1.7.3. *If A is an invertible symmetric matrix, then A^{-1} is symmetric.*

Proof. Assume that A is symmetric and invertible. From Theorem 1.4.10 and the fact that $A = A^T$ we have

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

which proves that A^{-1} is symmetric.

MATRICES OF THE FORM AA^T AND A^TA

Matrix products of the form AA^T and A^TA arise in a variety of applications. If A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so the products AA^T and A^TA are both square matrices; the matrix AA^T has size $m \times m$ and the matrix A^TA has size $n \times n$. Such products are always symmetric since

$$(AA^T)^T = (A^T)^T A^T = AA^T \quad \text{and} \quad (A^TA)^T = A^T(A^T)^T = A^TA$$

Example 6 Let A be the 2×3 matrix

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

Then

$$A^TA = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix}\begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 10 & -2 & -11 \\ -2 & 4 & -8 \\ -11 & -8 & 41 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} 21 & -17 \\ -17 & 34 \end{bmatrix}$$

Observe that A^TA and AA^T are symmetric as expected.

Later in this text, we will obtain general conditions on A under which AA^T and A^TA are invertible. However, in the special case where A is *square* we have the following result.

Theorem 1.7.4. *If A is an invertible matrix, then AA^T and A^TA are also invertible.*

Proof. Since A is invertible, so is A^T by Theorem 1.4.10. Thus, AA^T and A^TA are invertible, since they are the products of invertible matrices.

EXERCISE SET 1.7

1. Determine whether the matrix is invertible; if so, find the inverse by inspection.

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

2. Compute the product by inspection.

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

3. Find A^2, A^{-2}, and A^{-k} by inspection.

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

4. Which of the following matrices are symmetric?

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

5. By inspection, determine whether the given triangular matrix is invertible.

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

6. Find all values of a, b, and c for which A is symmetric.

$$A = \begin{bmatrix} 2 & a-2b+2c & 2a+b+c \\ 3 & 5 & a+c \\ 0 & -2 & 7 \end{bmatrix}$$

7. Find all values of a and b for which A and B are both not invertible.

$$A = \begin{bmatrix} a+b-1 & 0 \\ 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 0 \\ 0 & 2a-3b-7 \end{bmatrix}$$

8. Use the given equation to determine by inspection whether the matrices on the left commute.

(a) $\begin{bmatrix} 1 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -5 \\ -10 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 3 & 1 \end{bmatrix}$

9. Show that A and B commute if $a - d = 7b$.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

10. Find a diagonal matrix A that satisfies

(a) $A^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ (b) $A^{-2} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

11. (a) Factor A into the form $A = BD$, where D is a diagonal matrix.

$$A = \begin{bmatrix} 3a_{11} & 5a_{12} & 7a_{13} \\ 3a_{21} & 5a_{22} & 7a_{23} \\ 3a_{31} & 5a_{32} & 7a_{33} \end{bmatrix}$$

(b) Is your factorization the only one possible? Explain.

12. Verify Theorem 1.7.1*b* for the product AB, where

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

13. Verify Theorem 1.7.1*d* for the matrices A and B in Exercise 12.

14. Verify Theorem 1.7.3 for the given matrix A.

(a) $A = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 1 & -7 \\ 3 & -7 & 4 \end{bmatrix}$

15. Let A be a symmetric matrix.
(a) Show that A^2 is symmetric.
(b) Show that $2A^2 - 3A + I$ is symmetric.

16. Let A be a symmetric matrix.
(a) Show that A^k is symmetric if k is any nonnegative integer.
(b) If $p(x)$ is a polynomial, is $p(A)$ necessarily symmetric? Explain.

17. Let A be an upper triangular matrix and let $p(x)$ be a polynomial. Is $p(A)$ necessarily upper triangular? Explain.

18. Prove: If $A^TA = A$, then A is symmetric and $A = A^2$.

19. What is the maximum number of distinct entries that an $n \times n$ symmetric matrix can have?

20. Let $A = [a_{ij}]$ be an $n \times n$ matrix. Determine whether A is symmetric.
(a) $a_{ij} = i^2 + j^2$ (b) $a_{ij} = i^2 - j^2$
(c) $a_{ij} = 2i + 2j$ (d) $a_{ij} = 2i^2 + 2j^3$

21. Based on your experience with Exercise 20, devise a general test that can be applied to a formula for a_{ij} to determine whether $A = [a_{ij}]$ is symmetric.

22. A square matrix A is called ***skew-symmetric*** if $A^T = -A$. Prove:
(a) If A is an invertible skew-symmetric matrix, then A^{-1} is skew-symmetric.
(b) If A and B are skew-symmetric, then so are A^T, $A+B$, $A-B$, and kA for any scalar k.
(c) Every square matrix can be expressed as the sum of a symmetric matrix and a skew-symmetric matrix.

23. We showed in the text that the product of symmetric matrices is symmetric if and only if the matrices commute. Is the product of commuting skew-symmetric matrices skew-symmetric? Explain. [*Note.* See Exercise 22 for terminology.]

24. If the $n \times n$ matrix A can be expressed as $A = LU$, where L is a lower triangular matrix and U is an upper triangular matrix, then the linear system $A\mathbf{x} = \mathbf{b}$ can be expressed as $LU\mathbf{x} = \mathbf{b}$ and can be solved in two steps:

Step 1. Let $U\mathbf{x} = \mathbf{y}$, so that $LU\mathbf{x} = \mathbf{b}$ can be expressed as $L\mathbf{y} = \mathbf{b}$. Solve this system.
Step 2. Solve the system $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$.

In each part use this two-step method to solve the given system.

(a) $$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 3 & 0 \\ 2 & 4 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

(b) $$\begin{bmatrix} 2 & 0 & 0 \\ 4 & 1 & 0 \\ -3 & -2 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 & 2 \\ 0 & 4 & 1 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -5 \\ 2 \end{bmatrix}$$

SUPPLEMENTARY EXERCISES

1. Use Gauss–Jordan elimination to solve for x' and y' in terms of x and y.

$$x = \tfrac{3}{5}x' - \tfrac{4}{5}y'$$
$$y = \tfrac{4}{5}x' + \tfrac{3}{5}y'$$

2. Use Gauss–Jordan elimination to solve for x' and y' in terms of x and y.

$$x = x'\cos\theta - y'\sin\theta$$
$$y = x'\sin\theta + y'\cos\theta$$

3. Find a homogeneous linear system with two equations that are not multiples of one another and such that

$$x_1 = 1,\quad x_2 = -1,\quad x_3 = 1,\quad x_4 = 2$$

and

$$x_1 = 2,\quad x_2 = 0,\quad x_3 = 3,\quad x_4 = -1$$

are solutions of the system.

4. A box containing pennies, nickels, and dimes has 13 coins with a total value of 83 cents. How many coins of each type are in the box?

5. Find positive integers that satisfy

$$\begin{aligned} x + y + z &= 9 \\ x + 5y + 10z &= 44 \end{aligned}$$

6. For which value(s) of a does the following system have zero, one, infinitely many solutions?

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

7. Let

$$\begin{bmatrix} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{bmatrix}$$

be the augmented matrix for a linear system. For what values of a and b does the system have
(a) a unique solution,
(b) a one-parameter solution,
(c) a two-parameter solution,
(d) no solution?

8. Solve for x, y, and z.

$$\begin{aligned} xy - 2\sqrt{y} + 3zy &= 8 \\ 2xy - 3\sqrt{y} + 2zy &= 7 \\ -xy + \sqrt{y} + 2zy &= 4 \end{aligned}$$

9. Find a matrix K such that $AKB = C$ given that

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

10. How should the coefficients a, b, and c be chosen so that the system

$$\begin{aligned} ax + by - 3z &= -3 \\ -2x - by + cz &= -1 \\ ax + 3y - cz &= -3 \end{aligned}$$

has the solution $x = 1$, $y = -1$, and $z = 2$?

11. In each part solve the matrix equation for X.

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

(b) $$X\begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -5 & -1 & 0 \\ 6 & -3 & 7 \end{bmatrix}$$

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

12. (a) Express the equations

$$\begin{aligned} y_1 &= x_1 - x_2 + x_3 \\ y_2 &= 3x_1 + x_2 - 4x_3 \\ y_3 &= -2x_1 - 2x_2 + 3x_3 \end{aligned} \qquad \text{and} \qquad \begin{aligned} z_1 &= 4y_1 - y_2 + y_3 \\ z_2 &= -3y_1 + 5y_2 - y_3 \end{aligned}$$

in the matrix forms $Y = AX$ and $Z = BY$. Then use these to obtain a direct relationship $Z = CX$ between Z and X.
(b) Use the equation $Z = CX$ obtained in (a) to express z_1 and z_2 in terms of x_1, x_2, and x_3.
(c) Check the result in (b) by directly substituting the equations for y_1, y_2, and y_3 into the equations for z_1 and z_2 and then simplifying.

13. If A is $m \times n$ and B is $n \times p$, how many multiplication operations and how many addition operations are needed to calculate the matrix product AB?

14. Let A be a square matrix.
(a) Show that $(I - A)^{-1} = I + A + A^2 + A^3$ if $A^4 = 0$.
(b) Show that $(I - A)^{-1} = I + A + A^2 + \cdots + A^n$ if $A^{n+1} = 0$.

15. Find values of a, b, and c so that the graph of the polynomial $p(x) = ax^2 + bx + c$ passes through the points $(1, 2)$, $(-1, 6)$, and $(2, 3)$.

16. ***(For readers who have studied calculus.)*** Find values of a, b, and c so that the graph of the polynomial $p(x) = ax^2 + bx + c$ passes through the point $(-1, 0)$ and has a horizontal tangent at $(2, -9)$.

17. Let J_n be the $n \times n$ matrix each of whose entries is 1. Show that

$$(I - J_n)^{-1} = I - \frac{1}{n-1} J_n$$

18. Show that if a square matrix A satisfies $A^3 + 4A^2 - 2A + 7I = 0$, then so does A^T.

19. Prove: If B is invertible, then $AB^{-1} = B^{-1}A$ if and only if $AB = BA$.

20. Prove: If A is invertible, then $A + B$ and $I + BA^{-1}$ are both invertible or both not invertible.

21. Prove that if A and B are $n \times n$ matrices, then
(a) $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$ (b) $\operatorname{tr}(kA) = k \operatorname{tr}(A)$ (c) $\operatorname{tr}(A^T) = \operatorname{tr}(A)$ (d) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

22. Use Exercise 21 to show that there are no square matrices A and B such that

$$AB - BA = I$$

23. Prove: If A is an $m \times n$ matrix and B is the $n \times 1$ matrix each of whose entries is $1/n$, then

$$AB = \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \vdots \\ \bar{r}_m \end{bmatrix}$$

where $\bar{r}_i$ is the average of the entries in the ith row of A.

24. ***(For readers who have studied calculus.)*** If the entries of the matrix

$$C = \begin{bmatrix} c_{11}(x) & c_{12}(x) & \cdots & c_{1n}(x) \\ c_{21}(x) & c_{22}(x) & \cdots & c_{2n}(x) \\ \vdots & \vdots & & \vdots \\ c_{m1}(x) & c_{m2}(x) & \cdots & c_{mn}(x) \end{bmatrix}$$

are differentiable functions of x, then we define

$$\frac{dC}{dx} = \begin{bmatrix} c'_{11}(x) & c'_{12}(x) & \cdots & c'_{1n}(x) \\ c'_{21}(x) & c'_{22}(x) & \cdots & c'_{2n}(x) \\ \vdots & \vdots & & \vdots \\ c'_{m1}(x) & c'_{m2}(x) & \cdots & c'_{mn}(x) \end{bmatrix}$$

Show that if the entries in A and B are differentiable functions of x and the sizes of the matrices are such that the stated operations can be performed, then

(a) $\frac{d}{dx}(kA) = k\frac{dA}{dx}$ (b) $\frac{d}{dx}(A + B) = \frac{dA}{dx} + \frac{dB}{dx}$ (c) $\frac{d}{dx}(AB) = \frac{dA}{dx}B + A\frac{dB}{dx}$

25. ***(For readers who have studied calculus.)*** Use part (c) of Exercise 24 to show that

$$\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1}$$

State all the assumptions you make in obtaining this formula.

26. Find the values of A, B, and C that will make the equation

$$\frac{x^2 + x - 2}{(3x - 1)(x^2 + 1)} = \frac{A}{3x - 1} + \frac{Bx + C}{x^2 + 1}$$

an identity. [*Hint.* Multiply through by $(3x - 1)(x^2 + 1)$ and equate the corresponding coefficients of the polynomials on each side of the resulting equation.]

27. If P is an $n \times 1$ matrix such that $P^TP = 1$, then $H = I - 2PP^T$ is called the corresponding ***Householder matrix*** (named after the American mathematician A. S. Householder).
(a) Verify that $P^TP = 1$ if $P^T = [\frac{3}{4}\ \frac{1}{6}\ \frac{1}{4}\ \frac{5}{12}\ \frac{5}{12}]$ and compute the corresponding Householder matrix.
(b) Prove that if H is any Householder matrix, then $H = H^T$ and $H^TH = I$.
(c) Verify that the Householder matrix found in part (a) satisfies the conditions proved in part (b).

28. Assuming that the stated inverses exist, prove the following equalities.
(a) $(C^{-1} + D^{-1})^{-1} = C(C + D)^{-1}D$ (b) $(I + CD)^{-1}C = C(I + DC)^{-1}$
(c) $(C + DD^T)^{-1}D = C^{-1}D(I + D^TC^{-1}D)^{-1}$

29. (a) Show that if $a \neq b$, then

$$a^n + a^{n-1}b + a^{n-2}b^2 + \cdots + ab^{n-1} + b^n = \frac{a^{n+1} - b^{n+1}}{a - b}$$

(b) Use the result in part (a) to find A^n if

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

[*Note.* This exercise is based on a problem by John M. Johnson, *The Mathematics Teacher*, Vol. 85, No. 9, 1992.]