

# 2

## Matrix Algebra

### INTRODUCTORY EXAMPLE

### Computer Models in Aircraft Design

To design the next generation of commercial and military aircraft, engineers at Boeing’s Phantom Works use 3D modeling and computational fluid dynamics (CFD). They study the airflow around a virtual airplane to answer important design questions before physical models are created. This has drastically reduced design cycle times and cost—and linear algebra plays a crucial role in the process.

The virtual airplane begins as a mathematical “wire-frame” model that exists only in computer memory and on graphics display terminals. (A model of a Boeing 777 is shown.) This mathematical model organizes and influences each step of the design and manufacture of the airplane—both the exterior and interior. The CFD analysis concerns the exterior surface.

Although the finished skin of a plane may seem smooth, the geometry of the surface is complicated. In addition to wings and a fuselage, an aircraft has nacelles, stabilizers, slats, flaps, and ailerons. The way air flows around these structures determines how the plane moves through the sky. Equations that describe the airflow are complicated, and they must account for engine intake, engine exhaust, and the wakes left by the wings of the plane. To study the airflow, engineers need a highly refined description of the plane’s surface.

A computer creates a model of the surface by first superimposing a three-dimensional grid of “boxes” on the



original wire-frame model. Boxes in this grid lie either completely inside or completely outside the plane, or they intersect the surface of the plane. The computer selects the boxes that intersect the surface and subdivides them, retaining only the smaller boxes that still intersect the surface. The subdividing process is repeated until the grid is extremely fine. A typical grid can include over 400,000 boxes.

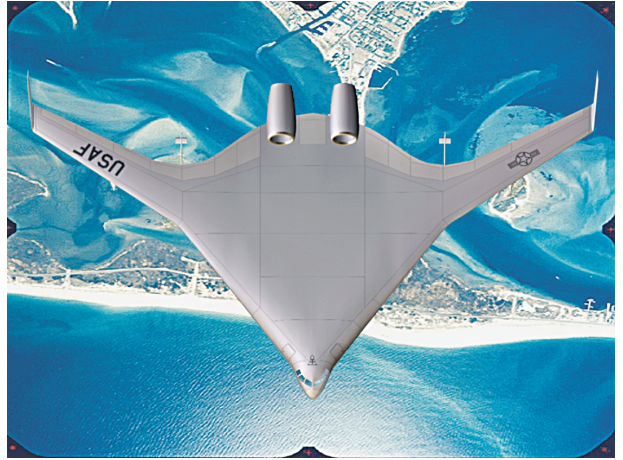
The process for finding the airflow around the plane involves repeatedly solving a system of linear equations  $A\mathbf{x} = \mathbf{b}$  that may involve up to 2 million equations and variables. The vector  $\mathbf{b}$  changes each time, based on data from the grid and solutions of previous equations. Using the fastest computers available commercially, a Phantom Works team can spend from a few hours to several days setting up and solving a single airflow problem. After the team analyzes the solution, they may make small changes to the airplane surface and begin the whole process again. Thousands of CFD runs may be required.

This chapter presents two important concepts that assist in the solution of such massive systems of equations:

- *Partitioned matrices:* A typical CFD system of equations has a “sparse” coefficient matrix with mostly zero entries. Grouping the variables correctly leads to a partitioned matrix with many zero blocks. Section 2.4 introduces such matrices and describes some of their applications.

- *Matrix factorizations:* Even when written with partitioned matrices, the system of equations is complicated. To further simplify the computations, the CFD software at Boeing uses what is called an LU factorization of the coefficient matrix. Section 2.5 discusses LU and other useful matrix factorizations. Further details about factorizations appear at several points later in the text.

To analyze a solution of an airflow system, engineers want to visualize the airflow over the surface of the plane. They use computer graphics, and linear algebra provides the engine for the graphics. The wire-frame model of the plane's surface is stored as data in many matrices. Once the image has been rendered on a computer screen, engineers can change its scale, zoom in or out of small regions, and rotate the image to see parts that may be hidden from view. Each of these operations is accomplished by appropriate



Modern CFD has revolutionized wing design. The Boeing Blended Wing Body is in design for the year 2020 or sooner.

matrix multiplications. Section 2.7 explains the basic ideas.

WEB

Our ability to analyze and solve equations will be greatly enhanced when we can perform algebraic operations with matrices. Furthermore, the definitions and theorems in this chapter provide some basic tools for handling the many applications of linear algebra that involve two or more matrices. For square matrices, the Invertible Matrix Theorem in Section 2.3 ties together most of the concepts treated earlier in the text. Sections 2.4 and 2.5 examine partitioned matrices and matrix factorizations, which appear in most modern uses of linear algebra. Sections 2.6 and 2.7 describe two interesting applications of matrix algebra, to economics and to computer graphics.

## 2.1 MATRIX OPERATIONS

If  $A$  is an  $m \times n$  matrix—that is, a matrix with  $m$  rows and  $n$  columns—then the scalar entry in the  $i$ th row and  $j$ th column of  $A$  is denoted by  $a_{ij}$  and is called the  $(i, j)$ -entry of  $A$ . See Fig. 1. For instance, the  $(3, 2)$ -entry is the number  $a_{32}$  in the third row, second column. Each column of  $A$  is a list of  $m$  real numbers, which identifies a vector in  $\mathbb{R}^m$ . Often, these columns are denoted by  $\mathbf{a}_1, \dots, \mathbf{a}_n$ , and the matrix  $A$  is written as

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

Observe that the number  $a_{ij}$  is the  $i$ th entry (from the top) of the  $j$ th column vector  $\mathbf{a}_j$ .

The **diagonal entries** in an  $m \times n$  matrix  $A = [a_{ij}]$  are  $a_{11}, a_{22}, a_{33}, \dots$ , and they form the **main diagonal** of  $A$ . A **diagonal matrix** is a square  $n \times n$  matrix whose nondiagonal entries are zero. An example is the  $n \times n$  identity matrix,  $I_n$ . An  $m \times n$  matrix whose entries are all zero is a **zero matrix** and is written as  $0$ . The size of a zero matrix is usually clear from the context.

$$\begin{array}{c} \text{Column} \\ j \\ \left[ \begin{array}{cccc} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ \text{Row } i & a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{array} \right] = A \\ \uparrow \qquad \qquad \qquad \uparrow \qquad \qquad \qquad \uparrow \\ \mathbf{a}_1 \qquad \qquad \qquad \mathbf{a}_j \qquad \qquad \qquad \mathbf{a}_n \end{array}$$

FIGURE 1 Matrix notation.

## Sums and Scalar Multiples

The arithmetic for vectors described earlier has a natural extension to matrices. We say that two matrices are **equal** if they have the same size (i.e., the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal. If  $A$  and  $B$  are  $m \times n$  matrices, then the **sum**  $A + B$  is the  $m \times n$  matrix whose columns are the sums of the corresponding columns in  $A$  and  $B$ . Since vector addition of the columns is done entrywise, each entry in  $A + B$  is the sum of the corresponding entries in  $A$  and  $B$ . The sum  $A + B$  is defined only when  $A$  and  $B$  are the same size.

**EXAMPLE 1** Let

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

Then

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

but  $A + C$  is not defined because  $A$  and  $C$  have different sizes. ■

If  $r$  is a scalar and  $A$  is a matrix, then the **scalar multiple**  $rA$  is the matrix whose columns are  $r$  times the corresponding columns in  $A$ . As with vectors,  $-A$  stands for  $(-1)A$ , and  $A - B$  is the same as  $A + (-1)B$ .

**EXAMPLE 2** If  $A$  and  $B$  are the matrices in Example 1, then

$$\begin{aligned} 2B &= 2 \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} \\ A - 2B &= \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix} \end{aligned}$$

It was unnecessary in Example 2 to compute  $A - 2B$  as  $A + (-1)2B$  because the usual rules of algebra apply to sums and scalar multiples of matrices, as the following theorem shows.

### THEOREM 1

Let  $A$ ,  $B$ , and  $C$  be matrices of the same size, and let  $r$  and  $s$  be scalars.

- |                                |                         |
|--------------------------------|-------------------------|
| a. $A + B = B + A$             | d. $r(A + B) = rA + rB$ |
| b. $(A + B) + C = A + (B + C)$ | e. $(r + s)A = rA + sA$ |
| c. $A + 0 = A$                 | f. $r(sA) = (rs)A$      |

Each equality in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because  $A$ ,  $B$ , and  $C$  are equal in size. The equality of columns follows immediately from analogous properties of vectors. For instance, if the  $j$ th columns of  $A$ ,  $B$ , and  $C$  are  $\mathbf{a}_j$ ,  $\mathbf{b}_j$ , and  $\mathbf{c}_j$ , respectively, then the  $j$ th columns of  $(A + B) + C$  and  $A + (B + C)$  are

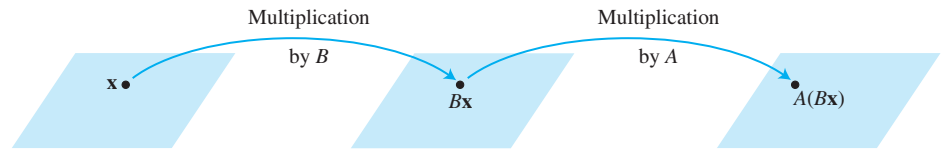
$$(\mathbf{a}_j + \mathbf{b}_j) + \mathbf{c}_j \quad \text{and} \quad \mathbf{a}_j + (\mathbf{b}_j + \mathbf{c}_j)$$

respectively. Since these two vector sums are equal for each  $j$ , property (b) is verified.

Because of the associative property of addition, we can simply write  $A + B + C$  for the sum, which can be computed either as  $(A + B) + C$  or as  $A + (B + C)$ . The same applies to sums of four or more matrices.

## Matrix Multiplication

When a matrix  $B$  multiplies a vector  $\mathbf{x}$ , it transforms  $\mathbf{x}$  into the vector  $B\mathbf{x}$ . If this vector is then multiplied in turn by a matrix  $A$ , the resulting vector is  $A(B\mathbf{x})$ . See Fig. 2.

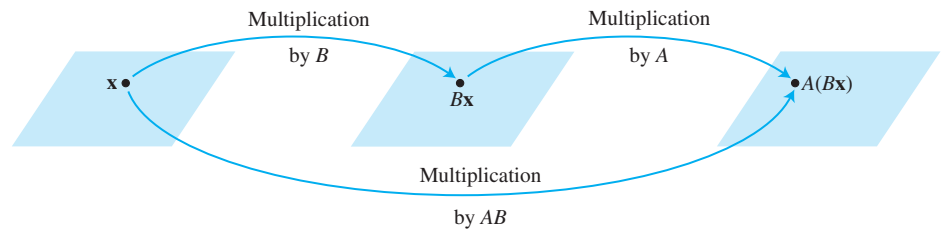


**FIGURE 2** Multiplication by  $B$  and then  $A$ .

Thus  $A(B\mathbf{x})$  is produced from  $\mathbf{x}$  by a *composition* of mappings—the linear transformations studied in Section 1.8. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by  $AB$ , so that

$$A(B\mathbf{x}) = (AB)\mathbf{x} \quad (1)$$

See Fig. 3.



**FIGURE 3** Multiplication by  $AB$ .

If  $A$  is  $m \times n$ ,  $B$  is  $n \times p$ , and  $\mathbf{x}$  is in  $\mathbb{R}^p$ , denote the columns of  $B$  by  $\mathbf{b}_1, \dots, \mathbf{b}_p$  and the entries in  $\mathbf{x}$  by  $x_1, \dots, x_p$ . Then

$$B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p$$

By the linearity of multiplication by  $A$ ,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1) + \cdots + A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1 + \cdots + x_pA\mathbf{b}_p \end{aligned}$$

The vector  $A(B\mathbf{x})$  is a linear combination of the vectors  $A\mathbf{b}_1, \dots, A\mathbf{b}_p$ , using the entries in  $\mathbf{x}$  as weights. In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]\mathbf{x}$$

Thus multiplication by  $[A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$  transforms  $\mathbf{x}$  into  $A(B\mathbf{x})$ . We have found the matrix we sought!

**DEFINITION**

If  $A$  is an  $m \times n$  matrix, and if  $B$  is an  $n \times p$  matrix with columns  $\mathbf{b}_1, \dots, \mathbf{b}_p$ , then the product  $AB$  is the  $m \times p$  matrix whose columns are  $A\mathbf{b}_1, \dots, A\mathbf{b}_p$ . That is,

$$AB = A[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p] = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$$

This definition makes equation (1) true for all  $\mathbf{x}$  in  $\mathbb{R}^p$ . Equation (1) proves that the composite mapping in Fig. 3 is a linear transformation and that its standard matrix is  $AB$ . *Multiplication of matrices corresponds to composition of linear transformations.*

**EXAMPLE 3** Compute  $AB$ , where  $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$  and  $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$ .

**SOLUTION** Write  $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$ , and compute:

$$\begin{aligned} A\mathbf{b}_1 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, & A\mathbf{b}_2 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix}, & A\mathbf{b}_3 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 11 \\ -1 \end{bmatrix} & &= \begin{bmatrix} 0 \\ 13 \end{bmatrix} & &= \begin{bmatrix} 21 \\ -9 \end{bmatrix} \end{aligned}$$

Then

$$AB = A[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$\begin{matrix} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{matrix}$

Notice that since the first column of  $AB$  is  $A\mathbf{b}_1$ , this column is a linear combination of the columns of  $A$  using the entries in  $\mathbf{b}_1$  as weights. A similar statement is true for each column of  $AB$ .

Each column of  $AB$  is a linear combination of the columns of  $A$  using weights from the corresponding column of  $B$ .

Obviously, the number of columns of  $A$  must match the number of rows in  $B$  in order for a linear combination such as  $A\mathbf{b}_1$  to be defined. Also, the definition of  $AB$  shows that  $AB$  has the same number of rows as  $A$  and the same number of columns as  $B$ .

**EXAMPLE 4** If  $A$  is a  $3 \times 5$  matrix and  $B$  is a  $5 \times 2$  matrix, what are the sizes of  $AB$  and  $BA$ , if they are defined?

**SOLUTION** Since  $A$  has 5 columns and  $B$  has 5 rows, the product  $AB$  is defined and is a  $3 \times 2$  matrix:

$$\begin{array}{ccc}
 A & B & AB \\
 \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} & = \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \\
 3 \times 5 & 5 \times 2 & 3 \times 2
 \end{array}$$

The product  $BA$  is *not* defined because the 2 columns of  $B$  do not match the 3 rows of  $A$ . ■

The definition of  $AB$  is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in  $AB$  when working small problems by hand.

**ROW-COLUMN RULE FOR COMPUTING  $AB$**

If the product  $AB$  is defined, then the entry in row  $i$  and column  $j$  of  $AB$  is the sum of the products of corresponding entries from row  $i$  of  $A$  and column  $j$  of  $B$ . If  $(AB)_{ij}$  denotes the  $(i, j)$ -entry in  $AB$ , and if  $A$  is an  $m \times n$  matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

To verify this rule, let  $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$ . Column  $j$  of  $AB$  is  $A\mathbf{b}_j$ , and we can compute  $A\mathbf{b}_j$  by the row-vector rule for computing  $A\mathbf{x}$  from Section 1.4. The  $i$ th entry in  $A\mathbf{b}_j$  is the sum of the products of corresponding entries from row  $i$  of  $A$  and the vector  $\mathbf{b}_j$ , which is precisely the computation described in the rule for computing the  $(i, j)$ -entry of  $AB$ .

**EXAMPLE 5** Use the row-column rule to compute two of the entries in  $AB$  for the matrices in Example 3. An inspection of the numbers involved will make it clear how the two methods for calculating  $AB$  produce the same matrix.

**SOLUTION** To find the entry in row 1 and column 3 of  $AB$ , consider row 1 of  $A$  and column 3 of  $B$ . Multiply corresponding entries and add the results, as shown below:

$$AB \rightarrow \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

For the entry in row 2 and column 2 of  $AB$ , use row 2 of  $A$  and column 2 of  $B$ :

$$\rightarrow \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$



**EXAMPLE 6** Find the entries in the second row of  $AB$ , where

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

**SOLUTION** By the row–column rule, the entries of the second row of  $AB$  come from row 2 of  $A$  (and the columns of  $B$ ):

$$\begin{array}{c} \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \begin{array}{c} \rightarrow \\ \rightarrow \end{array} \end{array} \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} \\ = \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

Notice that since Example 6 requested only the second row of  $AB$ , we could have written just the second row of  $A$  to the left of  $B$  and computed

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

This observation about rows of  $AB$  is true in general and follows from the row–column rule. Let  $\text{row}_i(A)$  denote the  $i$ th row of a matrix  $A$ . Then

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B \quad (2)$$

## Properties of Matrix Multiplication

The following theorem lists the standard properties of matrix multiplication. Recall that  $I_m$  represents the  $m \times m$  identity matrix and  $I_m \mathbf{x} = \mathbf{x}$  for all  $\mathbf{x}$  in  $\mathbb{R}^m$ .

### THEOREM 2

Let  $A$  be an  $m \times n$  matrix, and let  $B$  and  $C$  have sizes for which the indicated sums and products are defined.

- $A(BC) = (AB)C$  (associative law of multiplication)
- $A(B + C) = AB + AC$  (left distributive law)
- $(B + C)A = BA + CA$  (right distributive law)
- $r(AB) = (rA)B = A(rB)$   
for any scalar  $r$
- $I_m A = A = A I_n$  (identity for matrix multiplication)

**PROOF** Properties (b)–(e) are considered in the exercises. Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative. Here is another proof of (a) that rests on the “column definition” of the product of two matrices. Let

$$C = [\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_p]$$

By the definition of matrix multiplication,

$$\begin{aligned} BC &= [B\mathbf{c}_1 \ \cdots \ B\mathbf{c}_p] \\ A(BC) &= [A(B\mathbf{c}_1) \ \cdots \ A(B\mathbf{c}_p)] \end{aligned}$$

Recall from equation (1) that the definition of  $AB$  makes  $A(B\mathbf{x}) = (AB)\mathbf{x}$  for all  $\mathbf{x}$ , so

$$A(BC) = [(AB)\mathbf{c}_1 \ \cdots \ (AB)\mathbf{c}_p] = (AB)C \quad \blacksquare$$

The associative and distributive laws in Theorems 1 and 2 say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write  $ABC$  for the product, which can be computed either as  $A(BC)$  or as  $(AB)C$ .<sup>1</sup> Similarly, a product  $ABCD$  of four matrices can be computed as  $A(BCD)$  or  $(ABC)D$  or  $A(BC)D$ , and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because  $AB$  and  $BA$  are usually not the same. This is not surprising, because the columns of  $AB$  are linear combinations of the columns of  $A$ , whereas the columns of  $BA$  are constructed from the columns of  $B$ . The position of the factors in the product  $AB$  is emphasized by saying that  $A$  is *right-multiplied* by  $B$  or that  $B$  is *left-multiplied* by  $A$ . If  $AB = BA$ , we say that  $A$  and  $B$  **commute** with one another.

**EXAMPLE 7** Let  $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$  and  $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$ . Show that these matrices do not commute. That is, verify that  $AB \neq BA$ .

**SOLUTION**

$$\begin{aligned} AB &= \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix} \\ BA &= \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix} \end{aligned} \quad \blacksquare$$

Example 7 illustrates the first of the following list of important differences between matrix algebra and the ordinary algebra of real numbers. See Exercises 9–12 for examples of these situations.

**WARNINGS:**

1. In general,  $AB \neq BA$ .
2. The cancellation laws do *not* hold for matrix multiplication. That is, if  $AB = AC$ , then it is *not* true in general that  $B = C$ . (See Exercise 10.)
3. If a product  $AB$  is the zero matrix, you *cannot* conclude in general that either  $A = 0$  or  $B = 0$ . (See Exercise 12.)

## Powers of a Matrix

**WEB**

If  $A$  is an  $n \times n$  matrix and if  $k$  is a positive integer, then  $A^k$  denotes the product of  $k$

<sup>1</sup>When  $B$  is square and  $C$  has fewer columns than  $A$  has rows, it is more efficient to compute  $A(BC)$  than  $(AB)C$ .



copies of  $A$ :

$$A^k = \underbrace{A \cdots A}_k$$

If  $A$  is nonzero and if  $\mathbf{x}$  is in  $\mathbb{R}^n$ , then  $A^k\mathbf{x}$  is the result of left-multiplying  $\mathbf{x}$  by  $A$  repeatedly  $k$  times. If  $k = 0$ , then  $A^0\mathbf{x}$  should be  $\mathbf{x}$  itself. Thus  $A^0$  is interpreted as the identity matrix. Matrix powers are useful in both theory and applications (Sections 2.6, 4.9, and later in the text).

## The Transpose of a Matrix

Given an  $m \times n$  matrix  $A$ , the **transpose** of  $A$  is the  $n \times m$  matrix, denoted by  $A^T$ , whose columns are formed from the corresponding rows of  $A$ .

**EXAMPLE 8** Let

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

Then

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix} \quad \blacksquare$$

### THEOREM 3

Let  $A$  and  $B$  denote matrices whose sizes are appropriate for the following sums and products.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- For any scalar  $r$ ,  $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$

Proofs of (a)–(c) are straightforward and are omitted. For (d), see Exercise 33. Usually,  $(AB)^T$  is not equal to  $A^T B^T$ , even when  $A$  and  $B$  have sizes such that the product  $A^T B^T$  is defined.

The generalization of Theorem 3(d) to products of more than two factors can be stated in words as follows:

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

The exercises contain numerical examples that illustrate properties of transposes.

## NUMERICAL NOTES

1. The fastest way to obtain  $AB$  on a computer depends on the way in which the computer stores matrices in its memory. The standard high-performance algorithms, such as in LAPACK, calculate  $AB$  by columns, as in our definition of the product. (A version of LAPACK written in C++ calculates  $AB$  by rows.)
2. The definition of  $AB$  lends itself well to parallel processing on a computer. The columns of  $B$  are assigned individually or in groups to different processors, which independently and hence simultaneously compute the corresponding columns of  $AB$ .

## PRACTICE PROBLEMS

1. Since vectors in  $\mathbb{R}^n$  may be regarded as  $n \times 1$  matrices, the properties of transposes in Theorem 3 apply to vectors, too. Let

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

Compute  $(A\mathbf{x})^T$ ,  $\mathbf{x}^T A^T$ ,  $\mathbf{x}\mathbf{x}^T$ , and  $\mathbf{x}^T \mathbf{x}$ . Is  $A^T \mathbf{x}^T$  defined?

2. Let  $A$  be a  $4 \times 4$  matrix and let  $\mathbf{x}$  be a vector in  $\mathbb{R}^4$ . What is the fastest way to compute  $A^2 \mathbf{x}$ ? Count the multiplications.

## 2.1 EXERCISES

In Exercises 1 and 2, compute each matrix sum or product if it is defined. If an expression is undefined, explain why. Let

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

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

1.  $-2A$ ,  $B - 2A$ ,  $AC$ ,  $CD$
2.  $A + 3B$ ,  $2C - 3E$ ,  $DB$ ,  $EC$

In the rest of this exercise set and in those to follow, assume that each matrix expression is defined. That is, the sizes of the matrices (and vectors) involved “match” appropriately.

3. Let  $A = \begin{bmatrix} 2 & -5 \\ 3 & -2 \end{bmatrix}$ . Compute  $3I_2 - A$  and  $(3I_2)A$ .

4. Compute  $A - 5I_3$  and  $(5I_3)A$ , where

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

In Exercises 5 and 6, compute the product  $AB$  in two ways: (a) by the definition, where  $A\mathbf{b}_1$  and  $A\mathbf{b}_2$  are computed separately, and (b) by the row–column rule for computing  $AB$ .

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

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

7. If a matrix  $A$  is  $5 \times 3$  and the product  $AB$  is  $5 \times 7$ , what is the size of  $B$ ?
8. How many rows does  $B$  have if  $BC$  is a  $5 \times 4$  matrix?
9. Let  $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$  and  $B = \begin{bmatrix} 1 & 9 \\ -3 & k \end{bmatrix}$ . What value(s) of  $k$ , if any, will make  $AB = BA$ ?

10. Let  $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$ ,  $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$ , and  $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$ . Verify that  $AB = AC$  and yet  $B \neq C$ .

11. Let  $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$  and  $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ . Compute  $AD$  and  $DA$ . Explain how the columns or rows of  $A$  change when  $A$  is multiplied by  $D$  on the right or on the left. Find a  $3 \times 3$  matrix  $B$ , not the identity matrix or the zero matrix, such that  $AB = BA$ .

12. Let  $A = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}$ . Construct a  $2 \times 2$  matrix  $B$  such that  $AB$  is the zero matrix. Use two different nonzero columns for  $B$ .

13. Let  $\mathbf{r}_1, \dots, \mathbf{r}_p$  be vectors in  $\mathbb{R}^n$ , and let  $Q$  be an  $m \times n$  matrix. Write the matrix  $[Q\mathbf{r}_1 \ \cdots \ Q\mathbf{r}_p]$  as a *product* of two matrices (neither of which is an identity matrix).
14. Let  $U$  be the  $3 \times 2$  cost matrix described in Example 6 in Section 1.8. The first column of  $U$  lists the costs per dollar of output for manufacturing product  $B$ , and the second column lists the costs per dollar of output for product  $C$ . (The costs are categorized as materials, labor, and overhead.) Let  $\mathbf{q}_1$  be a vector in  $\mathbb{R}^2$  that lists the output (measured in dollars) of products  $B$  and  $C$  manufactured during the first quarter of the year, and let  $\mathbf{q}_2, \mathbf{q}_3$ , and  $\mathbf{q}_4$  be the analogous vectors that list the amounts of products  $B$  and  $C$  manufactured in the second, third, and fourth quarters, respectively. Give an economic description of the data in the matrix  $UQ$ , where  $Q = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3 \ \mathbf{q}_4]$ .

Exercises 15 and 16 concern arbitrary matrices  $A, B$ , and  $C$  for which the indicated sums and products are defined. Mark each statement True or False. Justify each answer.

15. a. If  $A$  and  $B$  are  $2 \times 2$  matrices with columns  $\mathbf{a}_1, \mathbf{a}_2$ , and  $\mathbf{b}_1, \mathbf{b}_2$ , respectively, then  $AB = [\mathbf{a}_1\mathbf{b}_1 \ \mathbf{a}_2\mathbf{b}_2]$ .  
 b. Each column of  $AB$  is a linear combination of the columns of  $B$  using weights from the corresponding column of  $A$ .  
 c.  $AB + AC = A(B + C)$   
 d.  $A^T + B^T = (A + B)^T$   
 e. The transpose of a product of matrices equals the product of their transposes in the same order.
16. a. The first row of  $AB$  is the first row of  $A$  multiplied on the right by  $B$ .  
 b. If  $A$  and  $B$  are  $3 \times 3$  matrices and  $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3]$ , then  $AB = [A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3]$ .  
 c. If  $A$  is an  $n \times n$  matrix, then  $(A^2)^T = (A^T)^2$   
 d.  $(ABC)^T = C^T A^T B^T$   
 e. The transpose of a sum of matrices equals the sum of their transposes.
17. If  $A = \begin{bmatrix} 1 & -3 \\ -3 & 5 \end{bmatrix}$  and  $AB = \begin{bmatrix} -3 & -11 \\ 1 & 17 \end{bmatrix}$ , determine the first and second columns of  $B$ .
18. Suppose the third column of  $B$  is all zeros. What can be said about the third column of  $AB$ ?
19. Suppose the third column of  $B$  is the sum of the first two columns. What can be said about the third column of  $AB$ ? Why?
20. Suppose the first two columns,  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , of  $B$  are equal. What can be said about the columns of  $AB$ ? Why?
21. Suppose the last column of  $AB$  is entirely zeros but  $B$  itself has no column of zeros. What can be said about the columns of  $A$ ?
22. Show that if the columns of  $B$  are linearly dependent, then so are the columns of  $AB$ .
23. Suppose  $CA = I_n$  (the  $n \times n$  identity matrix). Show that the equation  $A\mathbf{x} = \mathbf{0}$  has only the trivial solution. Explain why  $A$  cannot have more columns than rows.
24. Suppose  $A$  is a  $3 \times n$  matrix whose columns span  $\mathbb{R}^3$ . Explain how to construct an  $n \times 3$  matrix  $D$  such that  $AD = I_3$ .
25. Suppose  $A$  is an  $m \times n$  matrix and there exist  $n \times m$  matrices  $C$  and  $D$  such that  $CA = I_n$  and  $AD = I_m$ . Prove that  $m = n$  and  $C = D$ . [Hint: Think about the product  $CAD$ .]
26. Suppose  $AD = I_m$  (the  $m \times m$  identity matrix). Show that for any  $\mathbf{b}$  in  $\mathbb{R}^m$ , the equation  $A\mathbf{x} = \mathbf{b}$  has a solution. [Hint: Think about the equation  $AD\mathbf{b} = \mathbf{b}$ .] Explain why  $A$  cannot have more rows than columns.

In Exercises 27 and 28, view vectors in  $\mathbb{R}^n$  as  $n \times 1$  matrices. For  $\mathbf{u}$  and  $\mathbf{v}$  in  $\mathbb{R}^n$ , the matrix product  $\mathbf{u}^T \mathbf{v}$  is a  $1 \times 1$  matrix, called the **scalar product**, or **inner product**, of  $\mathbf{u}$  and  $\mathbf{v}$ . It is usually written as a single real number without brackets. The matrix product  $\mathbf{u}\mathbf{v}^T$  is an  $n \times n$  matrix, called the **outer product** of  $\mathbf{u}$  and  $\mathbf{v}$ . The products  $\mathbf{u}^T \mathbf{v}$  and  $\mathbf{u}\mathbf{v}^T$  will appear later in the text.

27. Let  $\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ -5 \end{bmatrix}$  and  $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ . Compute  $\mathbf{u}^T \mathbf{v}$ ,  $\mathbf{v}^T \mathbf{u}$ ,  $\mathbf{u}\mathbf{v}^T$ , and  $\mathbf{v}\mathbf{u}^T$ .

28. If  $\mathbf{u}$  and  $\mathbf{v}$  are in  $\mathbb{R}^n$ , how are  $\mathbf{u}^T \mathbf{v}$  and  $\mathbf{v}^T \mathbf{u}$  related? How are  $\mathbf{u}\mathbf{v}^T$  and  $\mathbf{v}\mathbf{u}^T$  related?

29. Prove Theorem 2(b) and 2(c). Use the row–column rule. The  $(i, j)$ -entry in  $A(B + C)$  can be written as

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

or

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

30. Prove Theorem 2(d). [Hint: The  $(i, j)$ -entry in  $(rA)B$  is  $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$ .]

31. Show that  $I_m A = A$  where  $A$  is an  $m \times n$  matrix. Assume  $I_m \mathbf{x} = \mathbf{x}$  for all  $\mathbf{x}$  in  $\mathbb{R}^m$ .

32. Show that  $A I_n = A$  when  $A$  is an  $m \times n$  matrix. [Hint: Use the (column) definition of  $A I_n$ .]

33. Prove Theorem 3(d). [Hint: Consider the  $j$ th row of  $(AB)^T$ .]

34. Give a formula for  $(AB\mathbf{x})^T$ , where  $\mathbf{x}$  is a vector and  $A$  and  $B$  are matrices of appropriate sizes.

35. [M] Read the documentation for your matrix program, and write the commands that will produce the following matrices (without keying in each entry of the matrix).

- A  $4 \times 5$  matrix of zeros
- A  $5 \times 3$  matrix of ones
- The  $5 \times 5$  identity matrix
- A  $4 \times 4$  diagonal matrix, with diagonal entries 3, 4, 2, 5

A useful way to test new ideas in matrix algebra, or to make conjectures, is to make calculations with matrices selected at random. Checking a property for a few matrices does not prove that the property holds in general, but it makes the property more believable. Also, if the property is actually false, making a few calculations may help to discover this.

36. [M] Write the command(s) that will create a  $5 \times 6$  matrix with random entries. In what range of numbers do the entries lie? Tell how to create a  $4 \times 4$  matrix with random integer entries between  $-9$  and  $9$ . [Hint: If  $x$  is a random number such that  $0 < x < 1$ , then  $-9.5 < 19(x - .5) < 9.5$ .]
37. [M] Construct random  $4 \times 4$  matrices  $A$  and  $B$  to test whether  $AB = BA$ . The best way to do this is to compute  $AB - BA$  and check whether this difference is the zero matrix. Then test  $AB - BA$  for three more pairs of random  $4 \times 4$  matrices. Report your conclusions.
38. [M] Construct a random  $5 \times 5$  matrix  $A$  and test whether  $(A + I)(A - I) = A^2 - I$ . The best way to do this is to compute  $(A + I)(A - I) - (A^2 - I)$  and verify that this difference is the zero matrix. Do this for three random matrices. Then test  $(A + B)(A - B) = A^2 - B^2$  the same

way for three pairs of random  $4 \times 4$  matrices. Report your conclusions.

39. [M] Use at least three pairs of random  $4 \times 4$  matrices  $A$  and  $B$  to test the equalities  $(A + B)^T = A^T + B^T$  and  $(AB)^T = B^T A^T$ , as well as  $(AB)^T = A^T B^T$ . (See Exercise 37.) Report your conclusions. [Note: Most matrix programs use  $A'$  for  $A^T$ .]

40. [M] Let

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

Compute  $S^k$  for  $k = 2, \dots, 6$ .

41. [M] Describe in words what happens when  $A^5, A^{10}, A^{20}$ , and  $A^{30}$  are computed for

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \\ 1/4 & 1/6 & 7/12 \end{bmatrix}$$

### SOLUTIONS TO PRACTICE PROBLEMS

1.  $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$ . So  $(A\mathbf{x})^T = [-4 \quad 2]$ . Also,

$$\mathbf{x}^T A^T = [5 \quad 3] \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = [-4 \quad 2].$$

The quantities  $(A\mathbf{x})^T$  and  $\mathbf{x}^T A^T$  are equal, by Theorem 3(d). Next,

$$\mathbf{xx}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix} [5 \quad 3] = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T \mathbf{x} = [5 \quad 3] \begin{bmatrix} 5 \\ 3 \end{bmatrix} = [25 + 9] = 34$$

A  $1 \times 1$  matrix such as  $\mathbf{x}^T \mathbf{x}$  is usually written without the brackets. Finally,  $A^T \mathbf{x}^T$  is not defined, because  $\mathbf{x}^T$  does not have two rows to match the two columns of  $A^T$ .

2. The fastest way to compute  $A^2 \mathbf{x}$  is to compute  $A(A\mathbf{x})$ . The product  $A\mathbf{x}$  requires 16 multiplications, 4 for each entry, and  $A(A\mathbf{x})$  requires 16 more. In contrast, the product  $A^2$  requires 64 multiplications, 4 for each of the 16 entries in  $A^2$ . After that,  $A^2 \mathbf{x}$  takes 16 more multiplications, for a total of 80.

## 2.2 THE INVERSE OF A MATRIX

Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers. This section investigates the matrix analogue of the reciprocal, or multiplicative inverse, of a nonzero number.