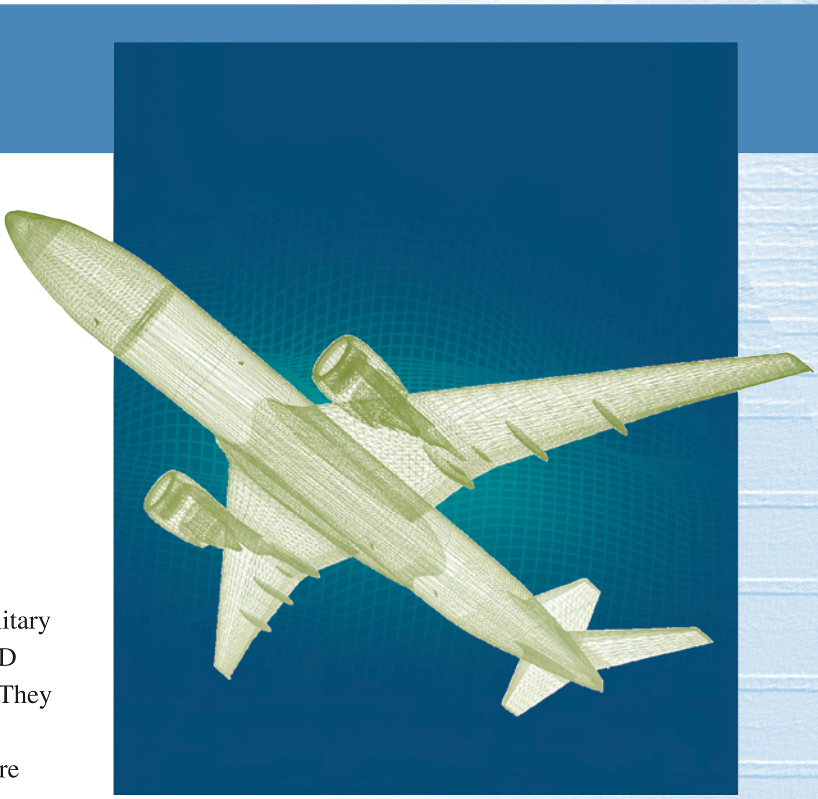
2

Matrix Algebra



WEB

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 matrix multiplications. Section 2.7 explains the basic ideas.

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

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 ith row and jth 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. The columns of A are vectors in $\mathbb{R}^m$ and are denoted by (boldface) $\mathbf{a}_1, \dots, \mathbf{a}_n$. We focus attention on these columns when we write

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

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

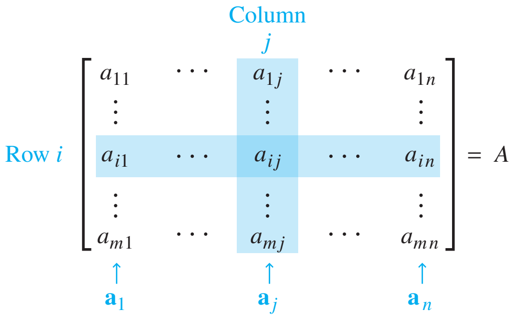


FIGURE 1 Matrix notation.

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 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.

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

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

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, we define $-A$ to mean $(-1)A$, and we write $A - B$ in place of $A + (-1)B$.

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

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

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

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 we see in the following theorem.

THEOREM 1

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

a. $A + B = B + A$

b. $(A + B) + C = A + (B + C)$

c. $A + 0 = A$

d. $r(A + B) = rA + rB$

e. $(r + s)A = rA + sA$

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 jth columns of A, B, and C are $\mathbf{a}_j$, $\mathbf{b}_j$, and $\mathbf{c}_j$, respectively, then the jth 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.

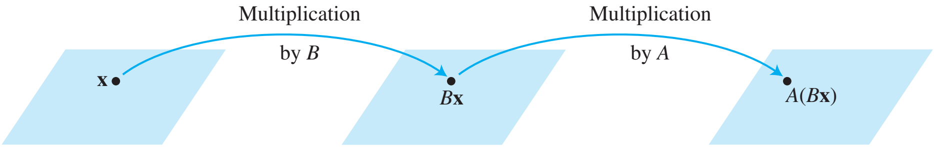


FIGURE 2 Multiplication by B and then A.

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

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

See Fig. 3.

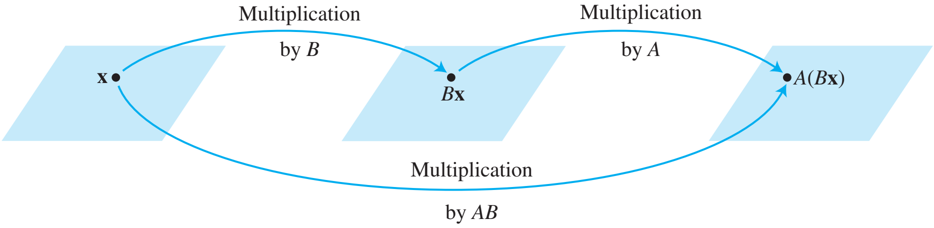


FIGURE 3 Multiplication by AB.

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. If we rewrite these vectors as the columns of a matrix, we have

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

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 (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:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

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

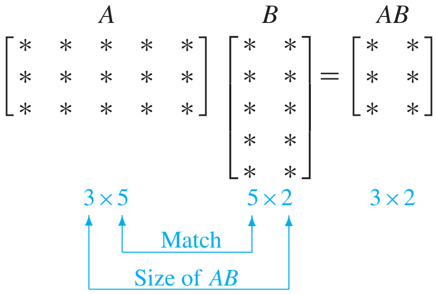
Notice that since the first column of AB is $A\mathbf{b}_1$, this column is a linear combination of the columns of A using the entries in $\mathbf{b}_1$ as weights. A similar statement is true for each column of AB.

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 3×5 matrix and B is a 5×2 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 3×2 matrix:



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 ith 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:

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

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

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

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

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

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):

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

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$

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 ith row of a matrix A. Then

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B \qquad (1)$$

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. $A(BC) = (AB)C$	(associative law of multiplication)
b. $A(B + C) = AB + AC$	(left distributive law)
c. $(B + C)A = BA + CA$	(right distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r	
e. $I_m A = A = AI_n$	(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,

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

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

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

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 $(AB)C$.[1] 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.

---

[1] When B is square and C has fewer columns than A has rows, it is more efficient to compute $A(BC)$ instead of $(AB)C$.

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

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

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix}$$

For emphasis, we include the remark about commutativity with the following list of important differences between matrix algebra and 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 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

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

Then

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

THEOREM 3

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

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^TA^T$

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

2. Let A be a 4×4 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. $A + 2B$, $3C - E$, CB, EB

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

3. Let $A = \begin{bmatrix} 4 & -1 \\ 5 & -2 \end{bmatrix}$. Compute $3I_2 - A$ and $(3I_2)A$.

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

$$A = \begin{bmatrix} 9 & -1 & 3 \\ -8 & 7 & -6 \\ -4 & 1 & 8 \end{bmatrix}.$$

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

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

7. If a matrix A is 5×3 and the product AB is 5×7, what is the size of B?

8. How many rows does B have if BC is a 3×4 matrix?

9. Let $A = \begin{bmatrix} 2 & 5 \\ -3 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & -5 \\ 3 & k \end{bmatrix}$. What value(s) of k, if any, will make $AB = BA$?

10. Let $A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 8 & 4 \\ 5 & 5 \end{bmatrix}$, and $C = \begin{bmatrix} 5 & -2 \\ 3 & 1 \end{bmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.

11. Let $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 5 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$. 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 3×3 matrix B, not the identity matrix or the zero matrix, such that $AB = BA$.

12. Let $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$. Construct a 2×2 matrix B such that AB is the zero matrix. Use two different nonzero columns for B.

13. Let $\mathbf{r}_1, \ldots, \mathbf{r}_p$ be vectors in $\mathbb{R}^n$, and let Q be an $m \times n$ matrix. Write the matrix $[\, Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_p \,]$ as a *product* of two matrices (neither of which is an identity matrix).

14. Let U be the 3×2 cost matrix described in Example 6 of 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 $Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3 \quad \mathbf{q}_4]$.

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 2×2 with columns $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}_1$, $\mathbf{b}_2$, respectively, then $AB = [\,\mathbf{a}_1\mathbf{b}_1 \quad \mathbf{a}_2\mathbf{b}_2\,]$.

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. If A and B are 3×3 and $B = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,]$, then $AB = [A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3]$.

b. The second row of AB is the second row of A multiplied on the right by B.

c. $(AB)C = (AC)B$

d. $(AB)^T = A^TB^T$

e. The transpose of a sum of matrices equals the sum of their transposes.

17. If $A = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$ and $AB = \begin{bmatrix} -1 & 2 & -1 \\ 6 & -9 & 3 \end{bmatrix}$, determine the first and second columns of B.

18. Suppose the first two columns, $\mathbf{b}_1$ and $\mathbf{b}_2$, of B are equal. What can you say about the columns of AB (if AB is defined)? Why?

19. Suppose the third column of B is the sum of the first two columns. What can you say about the third column of AB? Why?

20. Suppose the second column of B is all zeros. What can you say about the second column of AB?

21. Suppose the last column of AB is entirely zero but B itself has no column of zeros. What can you say 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 $AD = I_m$ (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.

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 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 $AD = I_3$.

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 1×1 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 $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \\ -4 \end{bmatrix}$ 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

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj}) \quad \text{or} \quad \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

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 $I_mA = A$ when A is an $m\times n$ matrix. You can assume $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

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

33. Prove Theorem 3(d). [*Hint:* Consider the jth 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. A 5×6 matrix of zeros

b. A 3×5 matrix of ones

c. The 6×6 identity matrix

d. A 5×5 diagonal matrix, with diagonal entries 3, 5, 7, 2, 4

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, you may discover this when you make a few calculations.

36. **[M]** Write the command(s) that will create a 6×4 matrix with random entries. In what range of numbers do the entries lie? Tell how to create a 3×3 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 a random 4×4 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 4×4 matrices. Report your conclusions.

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

39. **[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 $k = 2, \dots, 6$.

40. **[M]** Describe in words what happens when you compute A^5, A^{10}, A^{20}, and A^{30} for

$$A = \begin{bmatrix} 1/6 & 1/2 & 1/3 \\ 1/2 & 1/4 & 1/4 \\ 1/3 & 1/4 & 5/12 \end{bmatrix}$$

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 $(A\mathbf{x})^T = \begin{bmatrix} -4 & 2 \end{bmatrix}$. Also, $\mathbf{x}^TA^T = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 2 \end{bmatrix}$. The quantities $(A\mathbf{x})^T$ and $\mathbf{x}^TA^T$ are equal, as we expect from Theorem 3(d). Next,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

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

A 1×1 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.

Recall that the multiplicative inverse of a number such as 5 is 1/5 or 5^{-1}. This inverse satisfies the equations

$$5^{-1}\cdot 5 = 1 \quad \text{and} \quad 5\cdot 5^{-1} = 1$$

The matrix generalization requires *both* equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative. Furthermore, a full generalization is possible only if the matrices involved are square.[1]

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A. In fact, C is uniquely determined by A, because if B were another inverse of A, then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1}, so that

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

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

EXAMPLE 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

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

Thus $C = A^{-1}$.

Here is a simple formula for the inverse of a 2×2 matrix, along with a test to tell if the inverse exists.

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A, and we write

$$\det A = ad - bc$$

Theorem 4 says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

---

[1] One could say that an $m \times n$ matrix A is invertible if there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. However, these equations imply that A is square and $C = D$. Thus A is invertible as defined above. See Exercises 23–25 in Section 2.1.

EXAMPLE 2 Find the inverse of $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

Solution Since det $A = 3(6) - 4(5) = -2 \neq 0$, A is invertible, and

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}$$

Invertible matrices are indispensable in linear algebra—mainly for algebraic calculations and formula derivations, as in the next theorem. There are also occasions when an inverse matrix provides insight into a mathematical model of a real-life situation, as in Example 3, below.

THEOREM 5 If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

PROOF Take any $\mathbf{b}$ in $\mathbb{R}^n$. A solution exists because if $A^{-1}\mathbf{b}$ is substituted for $\mathbf{x}$, then $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. So $A^{-1}\mathbf{b}$ is a solution. To prove that the solution is unique, show that if $\mathbf{u}$ is any solution, then $\mathbf{u}$, in fact, must be $A^{-1}\mathbf{b}$. Indeed, if $A\mathbf{u} = \mathbf{b}$, we can multiply both sides by A^{-1} and obtain

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \qquad I\mathbf{u} = A^{-1}\mathbf{b}, \qquad \text{and} \qquad \mathbf{u} = A^{-1}\mathbf{b}$$

EXAMPLE 3 A horizontal elastic beam is supported at each end and is subjected to forces at points 1, 2, 3, as shown in Fig. 1. Let $\mathbf{f}$ in $\mathbb{R}^3$ list the forces at these points, and let $\mathbf{y}$ in $\mathbb{R}^3$ list the amounts of deflection (that is, movement) of the beam at the three points. Using Hooke's law from physics, it can be shown that

$$\mathbf{y} = D\mathbf{f}$$

where D is a *flexibility matrix*. Its inverse is called the *stiffness matrix*. Describe the physical significance of the columns of D and D^{-1}.

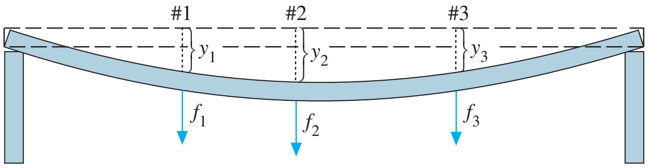


FIGURE 1 Deflection of an elastic beam.

Solution Write $I_3 = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3]$ and observe that

$$D = DI_3 = [D\mathbf{e}_1 \quad D\mathbf{e}_2 \quad D\mathbf{e}_3]$$

Interpret the vector $\mathbf{e}_1 = (1, 0, 0)$ as a unit force applied downward at point 1 on the beam (with zero force at the other two points). Then $D\mathbf{e}_1$, the first column of D, lists the beam deflections due to a unit force at point 1. Similar descriptions apply to the second and third columns of D.

To study the stiffness matrix D^{-1}, observe that the equation $\mathbf{f} = D^{-1}\mathbf{y}$ computes a force vector $\mathbf{f}$ when a deflection vector $\mathbf{y}$ is given. Write

$$D^{-1} = D^{-1}I_3 = [D^{-1}\mathbf{e}_1 \quad D^{-1}\mathbf{e}_2 \quad D^{-1}\mathbf{e}_3]$$

Now interpret $\mathbf{e}_1$ as a deflection vector. Then $D^{-1}\mathbf{e}_1$ lists the forces that create the deflection. That is, the first column of D^{-1} lists the forces that must be applied at the three points to produce a unit deflection at point 1 and zero deflections at the other points. Similarly, columns 2 and 3 of D^{-1} list the forces required to produce unit deflections at points 2 and 3, respectively. In each column, one or two of the forces must be negative (point upward) to produce a unit deflection at the desired point and zero deflections at the other two points. If the flexibility is measured, for example, in inches of deflection per pound of load, then the stiffness matrix entries are given in pounds of load per inch of deflection.

The formula of Theorem 5 is seldom used to solve an equation $A\mathbf{x} = \mathbf{b}$ numerically because row reduction of $[\,A \quad \mathbf{b}\,]$ is nearly always faster. (Row reduction is usually more accurate, too, when computations involve rounding off numbers.) One possible exception is the 2×2 case. In this case, mental computations to solve $A\mathbf{x} = \mathbf{b}$ are sometimes easier using the formula for A^{-1}, as in the next example.

EXAMPLE 4 Use the inverse of the matrix A in Example 2 to solve the system

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

Solution This system is equivalent to $A\mathbf{x} = \mathbf{b}$, so

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$

The next theorem provides three useful facts about invertible matrices.

THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and

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

b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

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

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,

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

PROOF To verify (a), we must find a matrix C such that

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

However, we already know that these equations are satisfied with A in place of C. Hence A^{-1} is invertible, and A is its inverse. Next, to prove (b), we compute:

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

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$. For (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. Similarly, $A^T(A^{-1})^T = I^T = I$. Hence A^T is invertible, and its inverse is $(A^{-1})^T$. ■

The following generalization of Theorem 6(b) is needed later.

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by *watching the row reduction of A to I.*

Elementary Matrices

An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix. The next example illustrates the three kinds of elementary matrices.

EXAMPLE 5 Let

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

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

Compute E_1A, E_2A, and E_3A, and describe how these products can be obtained by elementary row operations on A.

Solution We have

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \qquad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}$$

Addition of -4 times row 1 of A to row 3 produces E_1A. (This is a row replacement operation.) An interchange of rows 1 and 2 of A produces E_2A, and multiplication of row 3 of A by 5 produces E_3A.

Left-multiplication (that is, multiplication on the left) by E_1 in Example 5 has the same effect on any $3 \times n$ matrix. It adds -4 times row 1 to row 3. In particular, since $E_1 \cdot I = E_1$, we see that E_1 *itself* is produced by this same row operation on the identity. Thus Example 5 illustrates the following general fact about elementary matrices. See Exercises 27 and 28.

> If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Since row operations are reversible, as we showed in Section 1.1, elementary matrices are invertible, for if E is produced by a row operation on I, then there is another row operation of the same type that changes E back into I. Hence there is an elementary matrix F such that $FE = I$. Since E and F correspond to reverse operations, $EF = I$, too.

> Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

EXAMPLE 6 Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Solution To transform E_1 into I, add +4 times row 1 to row 3. The elementary matrix that does this is

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$

The following theorem provides the best way to "visualize" an invertible matrix, and the theorem leads immediately to a method for finding the inverse of a matrix.

THEOREM 7

> An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

PROOF Suppose that A is invertible. Then, since the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ (Theorem 5), A has a pivot position in every row (Theorem 4 in Section 1.4). Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n. That is, $A \sim I_n$.

Now suppose, conversely, that $A \sim I_n$. Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices $E_1, \dots, E_p$ such that

$$A \sim E_1 A \sim E_2(E_1 A) \sim \cdots \sim E_p(E_{p-1} \cdots E_1 A) = I_n$$

That is,

$$E_p \cdots E_1 A = I_n \tag{1}$$

Since the product $E_p \cdots E_1$ of invertible matrices is invertible, (1) leads to

$$\begin{aligned}(E_p \cdots E_1)^{-1}(E_p \cdots E_1)A &= (E_p \cdots E_1)^{-1} I_n \\ A &= (E_p \cdots E_1)^{-1}\end{aligned}$$

Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,

$$A^{-1} = [\,(E_p \cdots E_1)^{-1}\,]^{-1} = E_p \cdots E_1$$

Then $A^{-1} = E_p \cdots E_1 \cdot I_n$, which says that A^{-1} results from applying $E_1, \dots, E_p$ successively to I_n. This is the same sequence in (1) that reduced A to I_n. ■

An Algorithm for Finding A^{-1}

If we place A and I side-by-side to form an augmented matrix $[\,A \quad I\,]$, then row operations on this matrix produce identical operations on A and on I. By Theorem 7, either there are row operations that transform A to I_n and I_n to A^{-1} or else A is not invertible.

ALGORITHM FOR FINDING A^{-1}

Row reduce the augmented matrix $[\,A \quad I\,]$. If A is row equivalent to I, then $[\,A \quad I\,]$ is row equivalent to $[\,I \quad A^{-1}\,]$. Otherwise, A does not have an inverse.

EXAMPLE 7 Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

Solution

$$[\,A \quad I\,] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

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

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

$$\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

Since $A \sim I$, we conclude that A is invertible by Theorem 7, and

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

It is a good idea to check the final answer:

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It is not necessary to check that $A^{-1}A = I$ since A is invertible. ■

Another View of Matrix Inversion

Denote the columns of I_n by $\mathbf{e}_1, \dots, \mathbf{e}_n$. Then row reduction of $[\,A \quad I\,]$ to $[\,I \quad A^{-1}\,]$ can be viewed as the simultaneous solution of the n systems

$$A\mathbf{x} = \mathbf{e}_1, \quad A\mathbf{x} = \mathbf{e}_2, \quad \dots, \quad A\mathbf{x} = \mathbf{e}_n \tag{2}$$

where the "augmented columns" of these systems have all been placed next to A to form $[\,A \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n\,] = [\,A \quad I\,]$. The equation $AA^{-1} = I$ and the definition of matrix multiplication show that the columns of A^{-1} are precisely the solutions of the systems in (2). This observation is useful because some applied problems may require finding only one or two columns of A^{-1}. In this case, only the corresponding systems in (2) need be solved.

NUMERICAL NOTE

In practical work, A^{-1} is seldom computed, unless the entries of A^{-1} are needed. Computing both A^{-1} and $A^{-1}\mathbf{b}$ takes about three times as many arithmetic operations as solving $A\mathbf{x} = \mathbf{b}$ by row reduction, and row reduction may be more accurate.

CD Exploring Properties of Inverses

PRACTICE PROBLEMS

1. Use determinants to determine which of the following matrices are invertible.

a. $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ b. $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ c. $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. Find the inverse of the matrix $A = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$, if it exists.

2.2 EXERCISES

Find the inverses of the matrices in Exercises 1–4.

1. $\begin{bmatrix} 8 & 6 \\ 5 & 4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & 2 \\ 7 & 4 \end{bmatrix}$

3. $\begin{bmatrix} 8 & 5 \\ -7 & -5 \end{bmatrix}$

4. $\begin{bmatrix} 3 & -4 \\ 7 & -8 \end{bmatrix}$

5. Use the inverse found in Exercise 1 to solve the system

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. Use the inverse found in Exercise 3 to solve the system

$$\begin{aligned} 8x_1 + 5x_2 &= -9 \\ -7x_1 - 5x_2 &= 11 \end{aligned}$$

7. Let $A = \begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$, and $\mathbf{b}_4 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

a. Find A^{-1}, and use it to solve the four equations
$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \quad A\mathbf{x} = \mathbf{b}_4$

b. The four equations in part (a) can be solved by the *same* set of row operations, since the coefficient matrix is the same in each case. Solve the four equations in part (a) by row reducing the augmented matrix $[\,A \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \mathbf{b}_4\,]$.

8. Use matrix algebra to show that if A is invertible and D satisfies $AD = I$, then $D = A^{-1}$.

In Exercises 9 and 10, mark each statement True or False. Justify each answer.

9. a. In order for a matrix B to be the inverse of A, both equations $AB = I$ and $BA = I$ must be true.

b. If A and B are $n \times n$ and invertible, then $A^{-1}B^{-1}$ is the inverse of AB.

c. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $ab - cd \neq 0$, then A is invertible.

d. If A is an invertible $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for *each* $\mathbf{b}$ in $\mathbb{R}^n$.

e. Each elementary matrix is invertible.

10. a. A product of invertible $n \times n$ matrices is invertible, and the inverse of the product is the product of their inverses in the same order.

b. If A is invertible, then the inverse of A^{-1} is A itself.

c. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $ad = bc$, then A is not invertible.

d. If A can be row reduced to the identity matrix, then A must be invertible.

e. If A is invertible, then elementary row operations that reduce A to the identity I_n also reduce A^{-1} to I_n.

11. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Show that the equation $AX = B$ has a unique solution $A^{-1}B$.

12. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Explain why $A^{-1}B$ can be computed by row reduction:

If $[A \quad B] \sim \cdots \sim [I \quad X]$, then $X = A^{-1}B$.

If A is larger than 2×2, then row reduction of $[A \quad B]$ is much faster than computing both A^{-1} and $A^{-1}B$.

13. Suppose $AB = AC$, where B and C are $n \times p$ matrices and A is invertible. Show that $B = C$. Is this true, in general, when A is not invertible?

14. Suppose $(B - C)D = 0$, where B and C are $m \times n$ matrices and D is invertible. Show that $B = C$.

15. Suppose A, B, and C are invertible $n \times n$ matrices. Show that ABC is also invertible by producing a matrix D such that $(ABC)D = I$ and $D(ABC) = I$.

16. Suppose A and B are $n \times n$, B is invertible, and AB is invertible. Show that A is invertible. [*Hint:* Let $C = AB$, and solve this equation for A.]

17. Solve the equation $AB = BC$ for A, assuming that A, B, and C are square and B is invertible.

18. Suppose P is invertible and $A = PBP^{-1}$. Solve for B in terms of A.

19. If A, B, and C are $n \times n$ invertible matrices, does the equation $C^{-1}(A + X)B^{-1} = I_n$ have a solution, X? If so, find it.

20. Suppose A, B, and X are $n \times n$ matrices with A, X, and $A - AX$ invertible, and suppose

$$(A - AX)^{-1} = X^{-1}B \tag{3}$$

a. Explain why B is invertible.

b. Solve (3) for X. If you need to invert a matrix, explain why that matrix is invertible.

21. Explain why the columns of an $n \times n$ matrix A are linearly independent when A is invertible.

22. Explain why the columns of an $n \times n$ matrix A span $\mathbb{R}^n$ when A is invertible. [*Hint:* Review Theorem 4 in Section 1.4.]

23. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A has n pivot columns and A is row equivalent to I_n. By Theorem 7, this shows that A must be invertible. (This exercise and Exercise 24 will be cited in Section 2.3.)

24. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Explain why A must be invertible. [*Hint:* Is A row equivalent to I_n?]

Exercises 25 and 26 prove Theorem 4 for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. Show that if $ad - bc = 0$, then the equation $A\mathbf{x} = \mathbf{0}$ has more than one solution. Why does this imply that A is not invertible? [*Hint:* First, consider $a = b = 0$. Then, if a and b are not both zero, consider the vector $\mathbf{x} = \begin{bmatrix} -b \\ a \end{bmatrix}$.]

26. Show that if $ad - bc \neq 0$, the formula for A^{-1} works.

Exercises 27 and 28 prove special cases of the facts about elementary matrices stated in the box following Example 5. Here A is a 3×3 matrix and $I = I_3$. (A general proof would require slightly more notation.)

27. a. Use equation (1) from Section 2.1 to show that $\text{row}_i(A) = \text{row}_i(I) \cdot A$, for $i = 1, 2, 3$.

b. Show that if rows 1 and 2 of A are interchanged, then the result may be written as EA, where E is an elementary matrix formed by interchanging rows 1 and 2 of I.

c. Show that if row 3 of A is multiplied by 5, then the result may be written as EA, where E is formed by multiplying row 3 of I by 5.

28. Show that if row 3 of A is replaced by $\text{row}_3(A) - 4 \cdot \text{row}_1(A)$, the result is EA, where E is formed from I by replacing $\text{row}_3(I)$ by $\text{row}_3(I) - 4 \cdot \text{row}_1(I)$.

Find the inverses of the matrices in Exercises 29–32, if they exist. Use the algorithm introduced in this section.

29. $\begin{bmatrix} 1 & 2 \\ 4 & 7 \end{bmatrix}$

30. $\begin{bmatrix} 5 & 10 \\ 4 & 7 \end{bmatrix}$

31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$

32. $\begin{bmatrix} 1 & -2 & 1 \\ 4 & -7 & 3 \\ -2 & 6 & -4 \end{bmatrix}$

33. Use the algorithm from this section to find the inverses of $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Let A be the corresponding $n \times n$ matrix, and let B be its inverse. Guess the form of B, and then prove that $AB = I$ and $BA = I$.

34. Repeat the strategy of Exercise 33 to guess the inverse of $A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & & 0 \\ 1 & 2 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$. Prove that your guess is correct.

35. Let $A = \begin{bmatrix} -2 & -7 & -9 \\ 2 & 5 & 6 \\ 1 & 3 & 4 \end{bmatrix}$. Find the third column of A^{-1} without computing the other columns.

36. **[M]** Let $A = \begin{bmatrix} -25 & -9 & -27 \\ 546 & 180 & 537 \\ 154 & 50 & 149 \end{bmatrix}$. Find the second and third columns of A^{-1} without computing the first column.

37. Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. Construct a 2×3 matrix C (by trial and error) using only 1, -1, and 0 as entries, such that $CA = I_2$. Compute AC and note that $AC \neq I_3$.

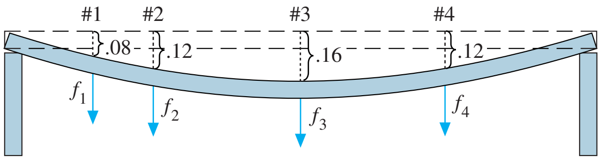
38. Let $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$. Construct a 4×2 matrix D using only 1 and 0 as entries, such that $AD = I_2$. Is it possible that $CA = I_4$ for some 4×2 matrix C? Why or why not?

39. Let $D = \begin{bmatrix} .005 & .002 & .001 \\ .002 & .004 & .002 \\ .001 & .002 & .005 \end{bmatrix}$ be a flexibility matrix, with flexibility measured in inches per pound. Suppose that forces of 30, 50, and 20 lb are applied at points 1, 2, and 3, respectively, in Fig. 1 of Example 3. Find the corresponding deflections.

40. **[M]** Compute the stiffness matrix D^{-1} for D in Exercise 39. List the forces needed to produce a deflection of .04 in. at point 3, with zero deflections at the other points.

41. **[M]** Let $D = \begin{bmatrix} .0040 & .0030 & .0010 & .0005 \\ .0030 & .0050 & .0030 & .0010 \\ .0010 & .0030 & .0050 & .0030 \\ .0005 & .0010 & .0030 & .0040 \end{bmatrix}$ be a flexibility matrix for an elastic beam with four points at which force is applied. Units are centimeters per newton of force.

Measurements at the four points show deflections of .08, .12, .16, and .12 cm. Determine the forces at the four points.



Deflection of elastic beam in Exercises 41 and 42.

42. **[M]** With D as in Exercise 41, determine the forces that produce a deflection of .24 cm at the second point on the beam, with zero deflections at the other three points. How is the answer related to the entries in D^{-1}? [*Hint:* First answer the question when the deflection is 1 cm at the second point.]

SOLUTIONS TO PRACTICE PROBLEMS

1. a. $\det\begin{bmatrix}3 & -9\\ 2 & 6\end{bmatrix} = 3\cdot 6 - (-9)\cdot 2 = 18 + 18 = 36$. The determinant is nonzero, so the matrix is invertible.

b. $\det\begin{bmatrix}4 & -9\\ 0 & 5\end{bmatrix} = 4\cdot 5 - (-9)\cdot 0 = 20 \neq 0$. The matrix is invertible.

c. $\det\begin{bmatrix}6 & -9\\ -4 & 6\end{bmatrix} = 6\cdot 6 - (-9)(-4) = 36 - 36 = 0$. The matrix is not invertible.

2. $$[\,A \quad I\,] \sim \begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ -1 & 5 & 6 & 0 & 1 & 0\\ 5 & -4 & 5 & 0 & 0 & 1\end{bmatrix}$$

$$\sim \begin{bmatrix}1 & -2 & -1 & 1 & 0 & 0\\ 0 & 3 & 5 & 1 & 1 & 0\\ 0 & 6 & 10 & -5 & 0 & 1\end{bmatrix}$$

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

We have obtained a matrix of the form $[\,B \quad D\,]$, where B is square and has a row of zeros. Further row operations will not transform B into I, so we stop. A does not have an inverse.

2.3 CHARACTERIZATIONS OF INVERTIBLE MATRICES

This section provides a review of most of the concepts introduced in Chapter 1, in relation to systems of n linear equations in n unknowns and to *square* matrices. The main result is Theorem 8.

THEOREM 8 **The Invertible Matrix Theorem**

Let A be a square $n \times n$ matrix. Then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix.
b. A is row equivalent to the $n \times n$ identity matrix.
c. A has n pivot positions.
d. The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
e. The columns of A form a linearly independent set.
f. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.
g. The equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.
h. The columns of A span $\mathbb{R}^n$.
i. The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ onto $\mathbb{R}^n$.
j. There is an $n \times n$ matrix C such that $CA = I$.
k. There is an $n \times n$ matrix D such that $AD = I$.
l. A^T is an invertible matrix.

First, we need some notation. If the truth of statement (a) always implies that statement (j) is true, we say that (a) *implies* (j) and write (a) $\Rightarrow$ (j). We will establish the "circle" of implications shown in Fig. 1. If any one of these five statements is true, then so are the others. Finally, we will link the remaining statements of the theorem to the statements in this circle.

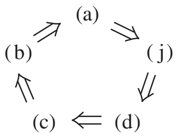
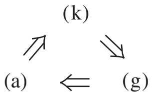


FIGURE 1

PROOF If (a) is true, then A^{-1} works for C in (j), so (a) $\Rightarrow$ (j). Next, (j) $\Rightarrow$ (d) by Exercise 23 in Section 2.1. (Turn back and read the exercise.) Also, (d) $\Rightarrow$ (c) by Exercise 23 in Section 2.2. If A is square and has n pivot positions, then the pivots must lie on the main diagonal, in which case, the reduced echelon form of A is I_n. Thus (c) $\Rightarrow$ (b). Also, (b) $\Rightarrow$ (a) by Theorem 7 in Section 2.2. This completes the circle in Fig. 1.

Next, (a) $\Rightarrow$ (k) because A^{-1} works for D. Also, (k) $\Rightarrow$ (g) by Exercise 24 in Section 2.1, and (g) $\Rightarrow$ (a) by Exercise 24 in Section 2.2. So (k) and (g) are linked to the circle. Further, (g), (h), and (i) are equivalent for any matrix, by Theorem 4 in Section 1.4 and Theorem 12(a) in Section 1.9. Thus, (h) and (i) are linked through (g) to the circle.

Since (d) is linked to the circle, so are (e) and (f), because (d), (e), and (f) are all equivalent for *any* matrix A. (See Section 1.7 and Theorem 12(b) in Section 1.9.) Finally, (a) $\Rightarrow$ (l) by Theorem 6(c) in Section 2.2, and (l) $\Rightarrow$ (a) by the same theorem with A and A^T interchanged. This completes the proof. ■



$$\text{(g)} \Leftrightarrow \text{(h)} \Leftrightarrow \text{(i)}$$

$$\text{(d)} \Leftrightarrow \text{(e)} \Leftrightarrow \text{(f)}$$

$$\text{(a)} \Leftrightarrow \text{(l)}$$

Because of Theorem 5 in Section 2.2, statement (g) in Theorem 8 could also be written as "The equation $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for each $\mathbf{b}$ in $\mathbb{R}^n$." This statement certainly implies (b) and hence implies that A is invertible.

The next fact follows from Theorem 8 and Exercise 8 in Section 2.2.

> Let A and B be square matrices. If $AB = I$, then A and B are both invertible, with $B = A^{-1}$ and $A = B^{-1}$.

The Invertible Matrix Theorem divides the set of all $n \times n$ matrices into two disjoint classes: the invertible (nonsingular) matrices, and the noninvertible (singular) matrices. Each statement in the theorem describes a property of every $n \times n$ invertible matrix. The *negation* of a statement in the theorem describes a property of every $n \times n$ singular matrix. For instance, an $n \times n$ singular matrix is *not* row equivalent to I_n, does *not* have n pivot positions, and has linearly *dependent* columns. Negations of other statements are considered in the exercises.

EXAMPLE 1 Use the Invertible Matrix Theorem to decide if A is invertible:

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

Solution

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

So A has three pivot positions and hence is invertible, by the Invertible Matrix Theorem, statement (c).

SG Expanded Table for the IMT 2–10

The power of the Invertible Matrix Theorem lies in the connections it provides between so many important concepts, such as linear independence of columns of a matrix A and the existence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$. It should be emphasized, however, that the Invertible Matrix Theorem *applies only to square matrices*. For example, if the columns of a 4×3 matrix are linearly independent, we cannot use the Invertible Matrix Theorem to conclude anything about the existence or nonexistence of solutions to equations of the form $A\mathbf{x} = \mathbf{b}$.

Invertible Linear Transformations

Recall from Section 2.1 that matrix multiplication corresponds to composition of linear transformations. When a matrix A is invertible, the equation $A^{-1}A\mathbf{x} = \mathbf{x}$ can be viewed as a statement about linear transformations. See Fig. 2.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **invertible** if there exists a function $S : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \tag{1}$$

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n \tag{2}$$

The next theorem shows that if such an S exists, it is unique and must be a linear transformation. We call S the **inverse** of T and write it as T^{-1}.

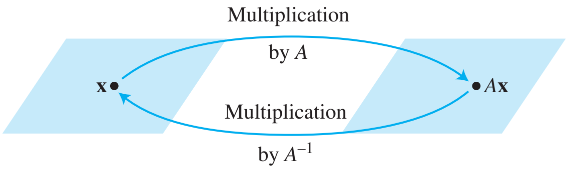


FIGURE 2 A^{-1} transforms $A\mathbf{x}$ back to $\mathbf{x}$.

THEOREM 9

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying (1) and (2).

PROOF Suppose that T is invertible. Then (2) shows that T is onto $\mathbb{R}^n$, for if $\mathbf{b}$ is in $\mathbb{R}^n$ and $\mathbf{x} = S(\mathbf{b})$, then $T(\mathbf{x}) = T(S(\mathbf{b})) = \mathbf{b}$, so each $\mathbf{b}$ is in the range of T. Thus A is invertible, by the Invertible Matrix Theorem, statement (i).

Conversely, suppose that A is invertible, and let $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Then, S is a linear transformation, and S obviously satisfies (1) and (2). For instance,

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = \mathbf{x}$$

Thus T is invertible. The proof that S is unique is outlined in Exercise 39. ■

EXAMPLE 2 What can you say about a one-to-one linear transformation T from $\mathbb{R}^n$ into $\mathbb{R}^n$?

Solution The columns of the standard matrix A of T are linearly independent (by Theorem 12 in Section 1.9). So A is invertible, by the Invertible Matrix Theorem, and T maps $\mathbb{R}^n$ *onto* $\mathbb{R}^n$. Also, T is invertible, by Theorem 9. ■

NUMERICAL NOTES

In practical work, you might occasionally encounter a "nearly singular" or **ill-conditioned** matrix—an invertible matrix that can become singular if some of its entries are changed ever so slightly. In this case, row reduction may produce fewer than n pivot positions, as a result of roundoff error. Also, roundoff error can sometimes make a singular matrix appear to be invertible.

WEB

Some matrix programs will compute a **condition number** for a square matrix. The larger the condition number, the closer the matrix is to being singular. The condition number of the identity matrix is 1. A singular matrix has an infinite condition number. In extreme cases, a matrix program may not be able to distinguish between a singular matrix and an ill-conditioned matrix.

Exercises 41–45 show that matrix computations can produce substantial error when a condition number is large.

PRACTICE PROBLEMS

1. Determine if $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ is invertible.

2. Suppose that for a certain $n \times n$ matrix A, statement (g) of the Invertible Matrix Theorem is *not* true. What can you say about equations of the form $A\mathbf{x} = \mathbf{b}$?

3. Suppose that A and B are $n \times n$ matrices and the equation $AB\mathbf{x} = \mathbf{0}$ has a nontrivial solution. What can you say about the matrix AB?

2.3 EXERCISES

Unless otherwise specified, assume that all matrices in these exercises are $n \times n$. Determine which of the matrices in Exercises 1–10 are invertible. Use as few calculations as possible. Justify your answers.

1. $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

2. $\begin{bmatrix} -4 & 6 \\ 6 & -9 \end{bmatrix}$

3. $\begin{bmatrix} 5 & 0 & 0 \\ -3 & -7 & 0 \\ 8 & 5 & -1 \end{bmatrix}$

4. $\begin{bmatrix} -7 & 0 & 4 \\ 3 & 0 & -1 \\ 2 & 0 & 9 \end{bmatrix}$

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

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

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

8. $\begin{bmatrix} 1 & 3 & 7 & 4 \\ 0 & 5 & 9 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 10 \end{bmatrix}$

9. **[M]** $\begin{bmatrix} 4 & 0 & -7 & -7 \\ -6 & 1 & 11 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 3 & -1 \end{bmatrix}$

10. **[M]** $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

In Exercises 11 and 12, the matrices are all $n \times n$. Each part of the exercises is an *implication* of the form "If ⟨ statement 1 ⟩, then ⟨ statement 2 ⟩." Mark an implication as True if the truth of ⟨ statement 2 ⟩ *always* follows whenever ⟨ statement 1 ⟩ happens to be true. An implication is False if there is an instance in which ⟨ statement 2 ⟩ is false but ⟨ statement 1 ⟩ is true. Justify each answer.

11. a. If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then A is row equivalent to the $n \times n$ identity matrix.

b. If the columns of A span $\mathbb{R}^n$, then the columns are linearly independent.

c. If A is an $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

d. If the equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, then A has fewer than n pivot positions.

e. If A^T is not invertible, then A is not invertible.

12. a. If there is an $n \times n$ matrix D such that $AD = I$, then there is also an $n \times n$ matrix C such that $CA = I$.

b. If the columns of A are linearly independent, then the columns of A span $\mathbb{R}^n$.

c. If the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$, then the solution is unique for each $\mathbf{b}$.

d. If the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ into $\mathbb{R}^n$, then A has n pivot positions.

e. If there is a $\mathbf{b}$ in $\mathbb{R}^n$ such that the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one.

13. An $m \times n$ **upper triangular matrix** is one whose entries *below* the main diagonal are 0's (as in Exercise 8). When is a square upper triangular matrix invertible? Justify your answer.

14. An $m \times n$ **lower triangular matrix** is one whose entries *above* the main diagonal are 0's (as in Exercise 3). When is a square lower triangular matrix invertible? Justify your answer.

15. Can a square matrix with two identical columns be invertible? Why or why not?

16. Is it possible for a 5×5 matrix to be invertible when its columns do not span $\mathbb{R}^5$? Why or why not?

17. If A is invertible, then the columns of A^{-1} are linearly independent. Explain why.

18. If C is 6×6 and the equation $C\mathbf{x} = \mathbf{v}$ is consistent for every $\mathbf{v}$ in $\mathbb{R}^6$, is it possible that for some $\mathbf{v}$, the equation $C\mathbf{x} = \mathbf{v}$ has more than one solution? Why or why not?

19. If the columns of a 7×7 matrix D are linearly independent, what can you say about solutions of $D\mathbf{x} = \mathbf{b}$? Why?

20. If $n \times n$ matrices E and F have the property that $EF = I$, then E and F commute. Explain why.

21. If the equation $G\mathbf{x} = \mathbf{y}$ has more than one solution for some $\mathbf{y}$ in $\mathbb{R}^n$, can the columns of G span $\mathbb{R}^n$? Why or why not?

22. If the equation $H\mathbf{x} = \mathbf{c}$ is inconsistent for some $\mathbf{c}$ in $\mathbb{R}^n$, what can you say about the equation $H\mathbf{x} = \mathbf{0}$? Why?

23. If an $n \times n$ matrix K cannot be row reduced to I_n, what can you say about the columns of K? Why?

24. If L is $n \times n$ and the equation $L\mathbf{x} = \mathbf{0}$ has the trivial solution, do the columns of L span $\mathbb{R}^n$? Why?

25. Verify the boxed statement preceding Example 1.

26. Explain why the columns of A^2 span $\mathbb{R}^n$ whenever the columns of A are linearly independent.

27. Show that if AB is invertible, so is A. You cannot use Theorem 6(b), because you cannot *assume* that A and B are invertible. [*Hint:* There is a matrix W such that $ABW = I$. Why?]

28. Show that if AB is invertible, so is B.

29. If A is an $n \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ has more than one solution for some $\mathbf{b}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one. What else can you say about this transformation? Justify your answer.

30. If A is an $n \times n$ matrix and the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one, what else can you say about this transformation? Justify your answer.

31. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Without using Theorems 5 or 8, explain why each equation $A\mathbf{x} = \mathbf{b}$ has in fact exactly one solution.

32. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Without using the Invertible Matrix Theorem, explain directly why the equation $A\mathbf{x} = \mathbf{b}$ must have a solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

In Exercises 33 and 34, T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$. Show that T is invertible and find a formula for T^{-1}.

33. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

34. $T(x_1, x_2) = (6x_1 - 8x_2, -5x_1 + 7x_2)$

35. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Explain why T is both one-to-one and onto $\mathbb{R}^n$. Use equations (1) and (2). Then give a second explanation using one or more theorems.

36. Let T be a linear transformation that maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Show that T^{-1} exists and maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Is T^{-1} also one-to-one?

37. Suppose T and U are linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $T(U(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Is it true that $U(T(x)) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$? Why or why not?

38. Suppose a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has the property that $T(\mathbf{u}) = T(\mathbf{v})$ for some pair of distinct vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. Can T map $\mathbb{R}^n$ onto $\mathbb{R}^n$? Why or why not?

39. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation, and let S and U be functions from $\mathbb{R}^n$ into $\mathbb{R}^n$ such that $S(T(\mathbf{x})) = \mathbf{x}$ and $U(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that $U(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$. This will show that T has a unique inverse, as asserted in Theorem 9. [*Hint:* Given any $\mathbf{v}$ in $\mathbb{R}^n$, we can write $\mathbf{v} = T(\mathbf{x})$ for some $\mathbf{x}$. Why? Compute $S(\mathbf{v})$ and $U(\mathbf{v})$.]

40. Suppose T and S satisfy the invertibility equations (1) and (2), where T is a linear transformation. Show directly that S is a linear transformation. [*Hint:* Given $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, let $\mathbf{x} = S(\mathbf{u})$, $\mathbf{y} = S(\mathbf{v})$. Then $T(\mathbf{x}) = \mathbf{u}$, $T(\mathbf{y}) = \mathbf{v}$. Why? Apply S to both sides of the equation $T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y})$. Also, consider $T(c\mathbf{x}) = cT(\mathbf{x})$.]

41. **[M]** Suppose an experiment leads to the following system of equations:

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.249 \\ 1.6x_1 + 1.1x_2 &= 6.843 \end{aligned} \tag{3}$$

a. Solve system (3), and then solve system (4), below, in which the data on the right have been rounded to two decimal places. In each case, find the *exact* solution.

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.25 \\ 1.6x_1 + 1.1x_2 &= 6.84 \end{aligned} \tag{4}$$

b. The entries in (4) differ from those in (3) by less than .05%. Find the percentage error when using the solution of (4) as an approximation for the solution of (3).

c. Use your matrix program to produce the condition number of the coefficient matrix in (3).

Exercises 42–44 show how to use the condition number of a matrix A to estimate the accuracy of a computed solution of $A\mathbf{x} = \mathbf{b}$. If the entries of A and $\mathbf{b}$ are accurate to about r significant digits and if the condition number of A is approximately 10^k (with k a positive integer), then the computed solution of $A\mathbf{x} = \mathbf{b}$ should usually be accurate to at least $r - k$ significant digits.

42. **[M]** Find the condition number of the matrix A in Exercise 9. Construct a random vector $\mathbf{x}$ in $\mathbb{R}^4$ and compute $\mathbf{b} = A\mathbf{x}$. Then use your matrix program to compute the solution $\mathbf{x}_1$ of $A\mathbf{x} = \mathbf{b}$. To how many digits do $\mathbf{x}$ and $\mathbf{x}_1$ agree? Find out the number of digits your matrix program stores accurately, and report how many digits of accuracy are lost when $\mathbf{x}_1$ is used in place of the exact solution $\mathbf{x}$.

43. **[M]** Repeat Exercise 42 for the matrix in Exercise 10.

44. **[M]** Solve an equation $A\mathbf{x} = \mathbf{b}$ for a suitable $\mathbf{b}$ to find the last column of the inverse of the *fifth-order Hilbert matrix*

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

How many digits in each entry of $\mathbf{x}$ do you expect to be correct? Explain. [*Note:* The exact solution is (630, −12600, 56700, −88200, 44100).]

45. **[M]** Some matrix programs, such as MATLAB, have a command to create Hilbert matrices of various sizes. If possible, use an inverse command to compute the inverse of a twelfth-order or larger Hilbert matrix, A. Compute AA^{-1}. Report what you find.

SG Mastering: Reviewing and Reflecting 2–13

SOLUTIONS TO PRACTICE PROBLEMS

1. The columns of A are obviously linearly dependent because columns 2 and 3 are multiples of column 1. Hence A cannot be invertible, by the Invertible Matrix Theorem.
2. If statement (g) is *not* true, then the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent for at least one $\mathbf{b}$ in $\mathbb{R}^n$.
3. Apply the Invertible Matrix Theorem to the matrix AB in place of A. Then statement (d) becomes: $AB\mathbf{x} = \mathbf{0}$ has only the trivial solution. This is not true. So AB is not invertible.

2.4 PARTITIONED MATRICES

A key feature of our work with matrices has been the ability to regard a matrix A as a list of column vectors rather than just a rectangular array of numbers. This point of view has been so useful that we wish to consider other **partitions** of A, indicated by horizontal and vertical dividing rules, as in Example 1 below. Partitioned matrices appear in most modern applications of linear algebra because the notation highlights essential structures in matrix analysis, as in the chapter introductory example on aircraft design. This section provides an opportunity to review matrix algebra and use the Invertible Matrix Theorem.

EXAMPLE 1 The matrix

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

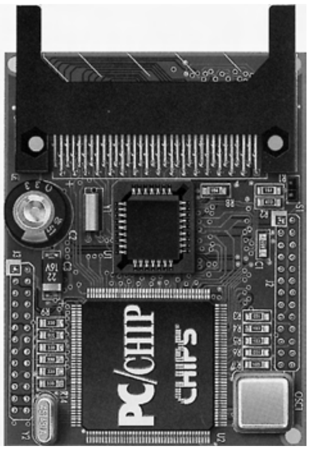
can also be written as the 2 × 3 **partitioned** (or **block**) **matrix**

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

whose entries are the *blocks* (or *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \qquad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \qquad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \qquad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \qquad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$



EXAMPLE 2 When a matrix A appears in a mathematical model of a physical system such as an electrical network, a transportation system, or a large corporation, it may be natural to regard A as a partitioned matrix. For instance, if a microcomputer circuit board consists mainly of three VLSI (very large-scale integrated) microchips, then the matrix for the circuit board might have the general form

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

The submatrices on the "diagonal" of A—namely, A_{11}, A_{22}, and A_{33}—concern the three VLSI chips, while the other submatrices depend on the interconnections among those microchips.

Addition and Scalar Multiplication

If matrices A and B are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum $A + B$. In this case, each block of $A + B$ is the (matrix) sum of the corresponding blocks of A and B. Multiplication of a partitioned matrix by a scalar is also computed block by block.

Multiplication of Partitioned Matrices

Partitioned matrices can be multiplied by the usual row–column rule as if the block entries were scalars, provided that for a product AB, the column partition of A matches the row partition of B.

EXAMPLE 3 Let

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

The 5 columns of A are partitioned into a set of 3 columns and then a set of 2 columns. The 5 rows of B are partitioned in the same way—into a set of 3 rows and then a set of 2 rows. We say that the partitions of A and B are **conformable** for **block multiplication**. It can be shown that the ordinary product AB can be written as

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{rr} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

It is important that each smaller product in the expression for AB is written with the submatrix from A on the left, since matrix multiplication is not commutative. For instance,

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

Hence the top block in AB is

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$

The row–column rule for multiplication of block matrices provides the most general way to regard the product of two matrices. Each of the following views of a product has already been described using simple partitions of matrices: (1) the definition of $A\mathbf{x}$ using the columns of A, (2) the column definition of AB, (3) the row–column rule for computing AB, and (4) the rows of AB as products of the rows of A and the matrix B. A fifth view of AB, again using partitions, follows in Theorem 10 below.

The calculations in the next example prepare the way for Theorem 10. Here $\text{col}_k(A)$ is the kth column of A, and $\text{row}_k(B)$ is the kth row of B.

EXAMPLE 4 Let $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Verify that

$$AB = \text{col}_1(A)\,\text{row}_1(B) + \text{col}_2(A)\,\text{row}_2(B) + \text{col}_3(A)\,\text{row}_3(B)$$

Solution Each term above is an *outer product*. (See Exercises 27 and 28 in Section 2.1.) By the row–column rule for computing a matrix product,

$$\text{col}_1(A)\,\text{row}_1(B) = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix}$$

$$\text{col}_2(A)\,\text{row}_2(B) = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix}$$

$$\text{col}_3(A)\,\text{row}_3(B) = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} e & f \end{bmatrix} = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix}$$

Thus

$$\sum_{k=1}^{3} \text{col}_k(A)\,\text{row}_k(B) = \begin{bmatrix} -3a + c + 2e & -3b + d + 2f \\ a - 4c + 5e & b - 4d + 5f \end{bmatrix}$$

This matrix is obviously AB. Notice that the (1, 1)-entry in AB is the sum of the (1, 1)-entries in the three outer products, the (1, 2)-entry in AB is the sum of the (1, 2)-entries in the three outer products, and so on.

THEOREM 10

Column–Row Expansion of AB

If A is $m \times n$ and B is $n \times p$, then

$$AB = [\,\text{col}_1(A) \quad \text{col}_2(A) \quad \cdots \quad \text{col}_n(A)\,] \begin{bmatrix} \text{row}_1(B) \\ \text{row}_2(B) \\ \vdots \\ \text{row}_n(B) \end{bmatrix} \tag{1}$$

$$= \text{col}_1(A)\,\text{row}_1(B) + \cdots + \text{col}_n(A)\,\text{row}_n(B)$$

PROOF For each row index i and column index j, the (i, j)-entry in $\text{col}_k(A)\,\text{row}_k(B)$ is the product of a_{ik} from $\text{col}_k(A)$ and b_{kj} from $\text{row}_k(B)$. Hence the (i, j)-entry in the sum shown in (1) is

$$\underset{(k=1)}{a_{i1}b_{1j}} \quad + \quad \underset{(k=2)}{a_{i2}b_{2j}} \quad + \quad \cdots \quad + \quad \underset{(k=n)}{a_{in}b_{nj}}$$

This sum is also the (i, j)-entry in AB, by the row–column rule.

Inverses of Partitioned Matrices

The next example illustrates calculations involving inverses and partitioned matrices.

EXAMPLE 5 A matrix of the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

is said to be *block upper triangular*. Assume that A_{11} is $p \times p$, A_{22} is $q \times q$, and A is invertible. Find a formula for A^{-1}.

Solution Denote A^{-1} by B and partition B so that

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \tag{2}$$

This matrix equation provides four equations that will lead to the unknown blocks $B_{11}, \ldots, B_{22}$. Compute the product on the left side of (2), and equate each entry with the corresponding block in the identity matrix on the right. That is, set

$$A_{11}B_{11} + A_{12}B_{21} = I_p \tag{3}$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \tag{4}$$

$$A_{22}B_{21} = 0 \tag{5}$$

$$A_{22}B_{22} = I_q \tag{6}$$

By itself, (6) does not say that A_{22} is invertible, because we do not yet know that $B_{22}A_{22} = I_q$. But, using the Invertible Matrix Theorem and the fact that A_{22} is square, we *can* conclude that A_{22} is invertible and $B_{22} = A_{22}^{-1}$. Now we can use (5) to find

$$B_{21} = A_{22}^{-1}0 = 0$$

so that (3) simplifies to

$$A_{11}B_{11} + 0 = I_p$$

This shows that A_{11} is invertible and $B_{11} = A_{11}^{-1}$. Finally, from (4),

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{and} \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$

A **block diagonal matrix** is a partitioned matrix with zero blocks off the main diagonal (of blocks). Such a matrix is invertible if and only if each block on the diagonal is invertible. See Exercises 13 and 14.

NUMERICAL NOTES

1. When matrices are too large to fit in a computer's high-speed memory, partitioning permits the computer to work with only two or three submatrices at a time. For instance, one linear programming research team simplified a problem by partitioning the matrix into 837 rows and 51 columns. The problem's solution took about 4 minutes on a Cray supercomputer.[1]
2. Some high-speed computers, particularly those with vector pipeline architecture, perform matrix calculations more efficiently when the algorithms use partitioned matrices.[2]
3. Professional software for high-performance numerical linear algebra, such as LAPACK, makes intensive use of partitioned matrix calculations.

---

[1] The solution time doesn't sound too impressive until you learn that each of the 51 block columns contained about 250,000 individual columns. The original problem had 837 equations and over 12,750,000 variables! Nearly 100 million of the more than 10 billion entries in the matrix were nonzero. See Robert E. Bixby et al., "Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods," *Operations Research*, 40, no. 5 (1992): 885–897.

[2] The importance of block matrix algorithms for computer calculations is described in *Matrix Computations*, 3rd ed., by Gene H. Golub and Charles F. van Loan (Baltimore: Johns Hopkins University Press, 1996).

The exercises that follow give practice with matrix algebra and illustrate typical calculations found in applications.

PRACTICE PROBLEMS

1. Show that $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible and find its inverse.
2. Compute X^TX, when X is partitioned as $\begin{bmatrix} X_1 & X_2 \end{bmatrix}$.

2.4 EXERCISES

In Exercises 1–9, assume that the matrices are partitioned conformably for block multiplication. Compute the products shown in Exercises 1–4.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$

4. $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

In Exercises 5–8, find formulas for X, Y, and Z in terms of A, B, and C, and justify your calculations. In some cases, you may need to make assumptions about the size of a matrix in order to produce a formula. [*Hint:* Compute the product on the left, and set it equal to the right side.]

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. Suppose A_{11} is an invertible matrix. Find matrices X and Y such that the product below has the form indicated. Also, compute B_{22}. [*Hint:* Compute the product on the left, and set it equal to the right side.]

$$\begin{bmatrix} I & 0 & 0 \\ X & I & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \\ 0 & B_{32} \end{bmatrix}$$

10. The inverse of $\begin{bmatrix} I & 0 & 0 \\ C & I & 0 \\ A & B & I \end{bmatrix}$ is $\begin{bmatrix} I & 0 & 0 \\ Z & I & 0 \\ X & Y & I \end{bmatrix}$. Find X, Y, and Z.

In Exercises 11 and 12, mark each statement True or False. Justify each answer.

11. a. If $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, with A_1 and A_2 the same sizes as B_1 and B_2, respectively, then $A + B = \begin{bmatrix} A_1 + B_1 & A_2 + B_2 \end{bmatrix}$.

b. If $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, then the partitions of A and B are conformable for block multiplication.

12. a. The definition of the matrix–vector product $A\mathbf{x}$ is a special case of block multiplication.

b. If A_1, A_2, B_1, and B_2 are $n \times n$ matrices, $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, and $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, then the product BA is defined, but AB is not.

13. Let $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$, where B and C are square. Show that A is invertible if and only if both B and C are invertible.

14. Show that the block upper triangular matrix A in Example 5 is invertible if and only if both A_{11} and A_{22} are invertible. [*Hint:* If A_{11} and A_{22} are invertible, the formula for A^{-1} given in Example 5 actually works as the inverse of A.] This fact about A is an important part of several computer algorithms that estimate eigenvalues of matrices. Eigenvalues are discussed in Chapter 5.

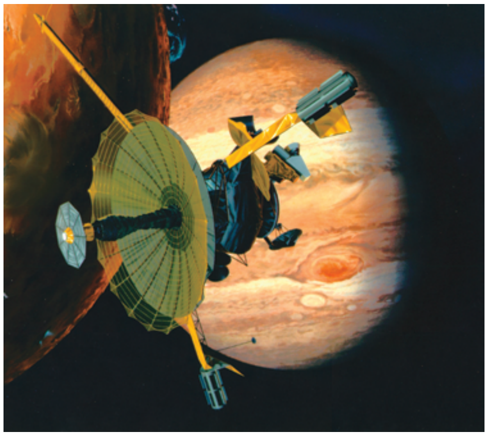
15. Suppose A_{11} is invertible. Find X and Y such that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \tag{7}$$

where $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$. The matrix S is called the **Schur complement** of A_{11}. Likewise, if A_{22} is invertible, the matrix $A_{11} - A_{12}A_{22}^{-1}A_{21}$ is called the Schur complement of A_{22}. Such expressions occur frequently in the theory of systems engineering, and elsewhere.

16. Suppose the block matrix A on the left side of (7) is invertible and A_{11} is invertible. Show that the Schur complement S of A_{11} is invertible. [*Hint:* The outside factors on the right side of (7) are always invertible. Verify this.] When A and A_{11} are both invertible, (7) leads to a formula for A^{-1}, using S^{-1}, A_{11}^{-1}, and the other entries in A.

17. When a deep space probe is launched, corrections may be necessary to place the probe on a precisely calculated trajectory. Radio telemetry provides a stream of vectors, $\mathbf{x}_1, \dots, \mathbf{x}_k$, giving information at different times about how the probe's position compares with its planned trajectory. Let X_k be the matrix $[\,\mathbf{x}_1 \ \cdots \ \mathbf{x}_k\,]$. The matrix $G_k = X_k X_k^T$ is computed as the radar data are analyzed. When $\mathbf{x}_{k+1}$ arrives, a new G_{k+1} must be computed. Since the data vectors arrive at high speed, the computational burden could be severe. But partitioned matrix multiplication helps tremendously. Compute the column–row expansions of G_k and G_{k+1}, and describe what must be computed in order to *update* G_k to form G_{k+1}.



The probe Galileo was launched October 18, 1989, and arrived near Jupiter in early December 1995.

18. Let X be an $m \times n$ data matrix such that X^TX is invertible, and let $M = I_m - X(X^TX)^{-1}X^T$. Add a column $\mathbf{x}_0$ to the data and form

$$W = [\,X \quad \mathbf{x}_0\,]$$

Compute W^TW. The (1, 1)-entry is X^TX. Show that the Schur complement (Exercise 15) of X^TX can be written in the form $\mathbf{x}_0^TM\mathbf{x}_0$. It can be shown that the quantity $(\mathbf{x}_0^TM\mathbf{x}_0)^{-1}$ is the (2, 2)-entry in $(W^TW)^{-1}$. This entry has a useful statistical interpretation, under appropriate hypotheses.

In the study of engineering control of physical systems, a standard set of differential equations is transformed by Laplace transforms into the following system of linear equations:

$$\begin{bmatrix} A - sI_n & B \\ C & I_m \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{8}$$

where A is $n \times n$, B is $n \times m$, C is $m \times n$, and s is a variable. The vector $\mathbf{u}$ in $\mathbb{R}^m$ is the "input" to the system, $\mathbf{y}$ in $\mathbb{R}^m$ is the "output," and $\mathbf{x}$ in $\mathbb{R}^n$ is the "state" vector. (Actually, the vectors $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{y}$ are functions of s, but we suppress this fact because it does not affect the algebraic calculations in Exercises 19 and 20.)

19. Assume $A - sI_n$ is invertible and view (8) as a system of two matrix equations. Solve the top equation for $\mathbf{x}$ and substitute into the bottom equation. The result is an equation of the form $W(s)\mathbf{u} = \mathbf{y}$, where $W(s)$ is a matrix that depends on s. $W(s)$ is called the *transfer function* of the system because it transforms the input $\mathbf{u}$ into the output $\mathbf{y}$. Find $W(s)$ and describe how it is related to the partitioned *system matrix* on the left side of (8). See Exercise 15.

20. Suppose the transfer function $W(s)$ in Exercise 19 is invertible for some s. It can be shown that the inverse transfer function $W(s)^{-1}$, which transforms outputs into inputs, is the Schur complement of $A - BC - sI_n$ for the matrix below. Find this Schur complement. See Exercise 15.

$$\begin{bmatrix} A - BC - sI_n & B \\ -C & I_m \end{bmatrix}$$

21. a. Verify that $A^2 = I$ when $A = \begin{bmatrix} 1 & 0 \\ 3 & -1 \end{bmatrix}$.

b. Use partitioned matrices to show that $M^2 = I$ when

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

22. Generalize the idea of Exercise 21(a) [not 21(b)] by constructing a 5×5 matrix $M = \begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$ such that $M^2 = I$. Make C a nonzero 2×3 matrix. Show that your construction works.

23. Use partitioned matrices to prove by induction that the product of two lower triangular matrices is also lower triangular. [*Hint*: A $(k+1) \times (k+1)$ matrix A_1 can be written in the form below, where a is a scalar, $\mathbf{v}$ is in $\mathbb{R}^k$, and A is a $k \times k$ lower triangular matrix. See the *Study Guide* for help with induction.]

$$A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}$$

SG The Principle of Induction 2–20

24. Use partitioned matrices to prove by induction that for $n = 2, 3, \ldots,$ the $n \times n$ matrix A shown below is invertible and B is its inverse.

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & & 0 \\ 1 & 1 & 1 & & 0 \\ \vdots & & & \ddots & \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & & 0 \\ -1 & 1 & 0 & & & 0 \\ 0 & -1 & 1 & & & 0 \\ \vdots & & \ddots & \ddots & & \\ 0 & & \cdots & & -1 & 1 \end{bmatrix}$$

For the induction step, assume A and B are $(k+1) \times (k+1)$ matrices, and partition A and B in a form similar to that displayed in Exercise 23.

25. Without using row reduction, find the inverse of

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 5 & 6 \end{bmatrix}$$

26. **[M]** For block operations, it may be necessary to access or enter submatrices of a large matrix. Describe the functions or commands of your matrix program that accomplish the following tasks. Suppose A is a 20×30 matrix.

a. Display the submatrix of A from rows 15 to 20 and columns 5 to 10.

b. Insert a 5×10 matrix B into A, beginning at row 10 and column 20.

c. Create a 50×50 matrix of the form $B = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$.

[*Note:* It may not be necessary to specify the zero blocks in B.]

27. **[M]** Suppose memory or size restrictions prevent your matrix program from working with matrices having more than 32 rows and 32 columns, and suppose some project involves 50×50 matrices A and B. Describe the commands or operations of your matrix program that accomplish the following tasks.

a. Compute $A + B$.

b. Compute AB.

c. Solve $A\mathbf{x} = \mathbf{b}$ for some vector $\mathbf{b}$ in $\mathbb{R}^{50}$, assuming that A can be partitioned into a 2×2 block matrix $[A_{ij}]$, with A_{11} an invertible 20×20 matrix, A_{22} an invertible 30×30 matrix, and A_{12} a zero matrix. [*Hint:* Describe appropriate smaller systems to solve, without using any matrix inverses.]

SOLUTIONS TO PRACTICE PROBLEMS

1. If $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible, its inverse has the form $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. We compute

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW + Y & AX + Z \end{bmatrix}$$

So W, X, Y, Z must satisfy $W = I$, $X = 0$, $AW + Y = 0$, and $AX + Z = I$. It follows that $Y = -A$ and $Z = I$. Hence

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

The product in the reverse order is also the identity, so the block matrix is invertible, and its inverse is $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (You could also appeal to the Invertible Matrix Theorem.)

2. $X^TX = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^TX_1 & X_1^TX_2 \\ X_2^TX_1 & X_2^TX_2 \end{bmatrix}$. The partitions of X^T and X are automatically conformable for block multiplication because the columns of X^T are the rows of X. This partition of X^TX is used in several computer algorithms for matrix computations.

2.5 MATRIX FACTORIZATIONS

A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices. Whereas matrix multiplication involves a *synthesis* of data (combining the effect of two or more linear transformations into a single matrix), matrix factorization is an *analysis* of data. In the language of computer science, the expression of A as a product amounts to a *preprocessing* of the data in A, organizing that data into two or more parts whose structures are more useful in some way, perhaps more accessible for computation.

Matrix factorizations and, later, factorizations of linear transformations will appear at a number of key points throughout the text. This section focuses on a factorization that lies at the heart of several important computer programs widely used in applications, such as the airflow problem described in the chapter introduction. Several other factorizations, to be studied later, are introduced in the exercises.

The LU Factorization

The LU factorization, described below, is motivated by the fairly common industrial and business problem of solving a sequence of equations, all with the same coefficient matrix:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{b}_p \tag{1}$$

See Exercise 32, for example. Also see Section 5.8, where the inverse power method is used to estimate eigenvalues of a matrix by solving equations like those in (1), one at a time.

When A is invertible, one could compute A^{-1} and then compute $A^{-1}\mathbf{b}_1$, $A^{-1}\mathbf{b}_2$, and so on. However, it is more efficient to solve the first equation in (1) by row reduction and obtain an LU factorization of A at the same time. Thereafter, the remaining equations in (1) are solved with the LU factorization.

At first, assume that A is an $m \times n$ matrix that can be row reduced to echelon form, *without row interchanges*. (Later, we will treat the general case.) Then A can be written in the form $A = LU$, where L is an $m \times m$ lower triangular matrix with 1's on the diagonal and U is an $m \times n$ echelon form of A. For instance, see Fig. 1. Such a factorization is called an **LU factorization** of A. The matrix L is invertible and is called a *unit* lower triangular matrix.

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}$$

FIGURE 1 An LU factorization.

Before studying how to construct L and U, we should look at why they are so useful. When $A = LU$, the equation $A\mathbf{x} = \mathbf{b}$ can be written as $L(U\mathbf{x}) = \mathbf{b}$. Writing $\mathbf{y}$ for $U\mathbf{x}$, we can find $\mathbf{x}$ by solving the *pair* of equations

$$\boxed{\begin{aligned} L\mathbf{y} &= \mathbf{b} \\ U\mathbf{x} &= \mathbf{y} \end{aligned}} \tag{2}$$

First solve $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, and then solve $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. See Fig. 2. Each equation is easy to solve because L and U are triangular.

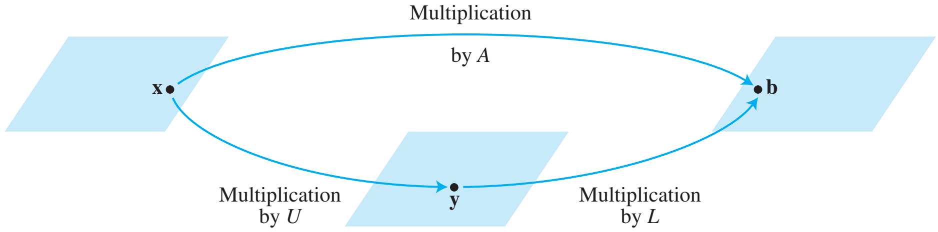


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

EXAMPLE 1 It can be verified that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

Use this LU factorization of A to solve $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

Solution The solution of $L\mathbf{y} = \mathbf{b}$ needs only 6 multiplications and 6 additions, because the arithmetic takes place only in column 5. (The zeros below each pivot in L are created automatically by the choice of row operations.)

$$\begin{bmatrix} L & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{y} \end{bmatrix}$$

Then, for $U\mathbf{x} = \mathbf{y}$, the "backward" phase of row reduction requires 4 divisions, 6 multiplications, and 6 additions. (For instance, creating the zeros in column 4 of $[\,U \quad \mathbf{y}\,]$ requires 1 division in row 4 and 3 multiplication–addition pairs to add multiples of row 4 to the rows above.)

$$\begin{bmatrix} U & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

To find $\mathbf{x}$ requires 28 arithmetic operations, or "flops" (floating point operations), excluding the cost of finding L and U. In contrast, row reduction of $[\,A \quad \mathbf{b}\,]$ to $[\,I \quad \mathbf{x}\,]$ takes 62 operations.

The computational efficiency of the LU factorization depends on knowing L and U. The next algorithm shows that the row reduction of A to an echelon form U amounts to an LU factorization because it produces L with essentially no extra work. After the first row reduction, L and U are available for solving additional equations whose coefficient matrix is A.

An LU Factorization Algorithm

Suppose A can be reduced to an echelon form U using only row replacements that add a multiple of one row to another row *below it*. In this case, there exist unit lower triangular elementary matrices $E_1, \dots, E_p$ such that

$$E_p \cdots E_1 A = U \tag{3}$$

Then

$$A = (E_p \cdots E_1)^{-1} U = LU$$

where

$$L = (E_p \cdots E_1)^{-1} \tag{4}$$

It can be shown that products and inverses of unit lower triangular matrices are also unit lower triangular. (For instance, see Exercise 19.) Thus L is unit lower triangular.

Note that the row operations in (3), which reduce A to U, also reduce the L in (4) to I, because $E_p \cdots E_1 L = (E_p \cdots E_1)(E_p \cdots E_1)^{-1} = I$. This observation is the key to *constructing* L.

Algorithm for an LU Factorization

1. Reduce A to an echelon form U by a sequence of row replacement operations, if possible.
2. Place entries in L such that the *same sequence of row operations* reduces L to I.

Step 1 is not always possible, but when it is, the argument above shows that an LU factorization exists. Example 2 will show how to implement step 2. By construction, L will satisfy

$$(E_p \cdots E_1)L = I$$

using the same $E_1, \ldots, E_p$ as in (3). Thus L will be invertible, by the Invertible Matrix Theorem, with $(E_p \cdots E_1) = L^{-1}$. From (3), $L^{-1}A = U$, and $A = LU$. So step 2 will produce an acceptable L.

EXAMPLE 2 Find an LU factorization of

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

Solution Since A has four rows, L should be 4×4. The first column of L is the first column of A divided by the top pivot entry:

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

Compare the first columns of A and L. *The row operations that create zeros in the first column of A will also create zeros in the first column of L.* We want this same correspondence of row operations to hold for the rest of L, so we watch a row reduction of A to an echelon form U:

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1 \tag{5}$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

The highlighted entries above determine the row reduction of A to U. At each pivot column, divide the highlighted entries by the pivot and place the result into L:

$$\begin{array}{cccc} \begin{bmatrix}2\\-4\\2\\-6\end{bmatrix} & \begin{bmatrix}3\\-9\\12\end{bmatrix} & \begin{bmatrix}2\\4\end{bmatrix} & \begin{bmatrix}5\end{bmatrix} \\ \div 2 & \div 3 & \div 2 & \div 5 \\ \downarrow & \downarrow & \downarrow & \downarrow \end{array}$$

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

An easy calculation verifies that this L and U satisfy $LU = A$.

In practical work, row interchanges are nearly always needed, because partial pivoting is used for high accuracy. (Recall that this procedure selects, among the possible choices for a pivot, an entry in the column having the largest absolute value.) To handle row interchanges, the LU factorization above can be modified easily to produce an L that is *permuted lower triangular*, in the sense that a rearrangement (called a permutation) of the rows of L can make L (unit) lower triangular. The resulting *permuted LU factorization* solves $A\mathbf{x} = \mathbf{b}$ in the same way as before, except that the reduction of $[\,L \quad \mathbf{b}\,]$ to $[\,I \quad \mathbf{y}\,]$ follows the order of the pivots in L from left to right, starting with the pivot in the first column. A reference to an "LU factorization" usually includes the possibility that L might be permuted lower triangular. For details, see the *Study Guide*.

SG Permuted LU Factorizations 2–24

Numerical Notes

The following operation counts apply to an $n \times n$ dense matrix A (with most entries nonzero) for n moderately large, say, $n \geq 30$.[1]

1. Computing an LU factorization of A takes about $2n^3/3$ flops (about the same as row reducing $[\,A \quad \mathbf{b}\,]$), whereas finding A^{-1} requires about $2n^3$ flops.
2. Solving $L\mathbf{y} = \mathbf{b}$ and $U\mathbf{x} = \mathbf{y}$ requires about $2n^2$ flops, because any $n \times n$ triangular system can be solved in about n^2 flops.
3. Multiplication of $\mathbf{b}$ by A^{-1} also requires about $2n^2$ flops, but the result may not be as accurate as that obtained from L and U (because of roundoff error when computing both A^{-1} and $A^{-1}\mathbf{b}$).
4. If A is sparse (with mostly zero entries), then L and U may be sparse, too, whereas A^{-1} is likely to be dense. In this case, a solution of $A\mathbf{x} = \mathbf{b}$ with an LU factorization is *much* faster than using A^{-1}. See Exercise 31.

CD Floating Point Operations

---

[1] See Section 3.8 in *Applied Linear Algebra*, 3rd ed., by Ben Noble and James W. Daniel (Englewood Cliffs, NJ: Prentice-Hall, 1988). Recall that for our purposes, a *flop* is +, −, ×, or ÷.

A Matrix Factorization in Electrical Engineering

Matrix factorization is intimately related to the problem of constructing an electrical network with specified properties. The following discussion gives just a glimpse of the connection between factorization and circuit design.

Suppose the box in Fig. 3 represents some sort of electric circuit, with an input and output. Record the input voltage and current by $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ (with voltage v in volts and current i in amps), and record the output voltage and current by $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$. Frequently, the transformation $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \mapsto \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ is linear. That is, there is a matrix A, called the *transfer matrix*, such that

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

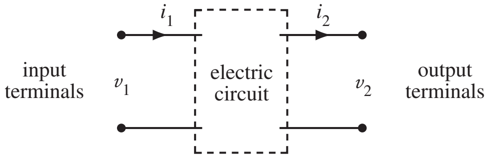


FIGURE 3 A circuit with input and output terminals.

Figure 4 shows a *ladder network*, where two circuits (there could be more) are connected in series, so that the output of one circuit becomes the input of the next circuit. The left circuit in Fig. 4 is called a *series circuit*, with resistance R_1 (in ohms).

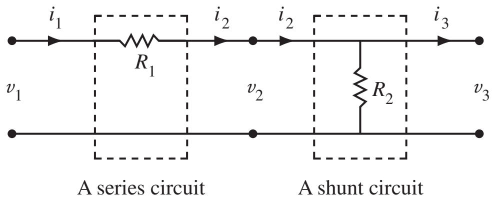


FIGURE 4 A ladder network.

The right circuit in Fig. 4 is a *shunt circuit*, with resistance R_2. Using Ohm's law and Kirchhoff's laws, one can show that the transfer matrices of the series and shunt circuits, respectively, are

$$\underset{\text{Transfer matrix of series circuit}}{\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix}} \quad \text{and} \quad \underset{\text{Transfer matrix of shunt circuit}}{\begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}}$$

EXAMPLE 3

a. Compute the transfer matrix of the ladder network in Fig. 4.

b. Design a ladder network whose transfer matrix is $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$.

Solution

a. Let A_1 and A_2 be the transfer matrices of the series and shunt circuits, respectively. Then an input vector $\mathbf{x}$ is transformed first into $A_1\mathbf{x}$ and then into $A_2(A_1\mathbf{x})$. The series connection of the circuits corresponds to composition of linear transformations, and the transfer matrix of the ladder network is (note the order)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} \tag{6}$$

b. We seek to factor the matrix $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$ into the product of transfer matrices, as in (6). So we look for R_1 and R_2 in Fig. 4 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$$

From the (1, 2)-entries, $R_1 = 8$ ohms, and from the (2, 1)-entries, $1/R_2 = .5$ ohm and $R_2 = 1/.5 = 2$ ohms. With these values, the network in Fig. 4 has the desired transfer matrix.

A network transfer matrix summarizes the input–output behavior (the design specifications) of the network without reference to the interior circuits. To physically build a network with specified properties, an engineer first determines if such a network can be constructed (or *realized*). Then the engineer tries to factor the transfer matrix into matrices corresponding to smaller circuits that perhaps are already manufactured and ready for assembly. In the common case of alternating current, the entries in the transfer matrix are usually rational complex-valued functions. (See Exercises 19 and 20 in Section 2.4 and Example 2 in Section 3.3.) A standard problem is to find a *minimal realization* that uses the smallest number of electrical components.

PRACTICE PROBLEM

Find an LU factorization of $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$. [*Note*: It will turn out that A has only three pivot columns, so the method of Example 2 will produce only the first three columns of L. The remaining two columns of L come from I_5.]

2.5 EXERCISES

In Exercises 1–6, solve the equation $A\mathbf{x} = \mathbf{b}$ by using the LU factorization given for A. In Exercises 1 and 2, also solve $A\mathbf{x} = \mathbf{b}$ by ordinary row reduction.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

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

2. $A = \begin{bmatrix} 4 & 3 & -5 \\ -4 & -5 & 7 \\ 8 & 6 & -8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

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

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

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

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

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

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

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

6. $A = \begin{bmatrix} 1 & 3 & 4 & 0 \\ -3 & -6 & -7 & 2 \\ 3 & 3 & 0 & -4 \\ -5 & -3 & 2 & 9 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

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

Find an LU factorization of the matrices in Exercises 7–16 (with L unit lower triangular). Note that MATLAB will usually produce a permuted LU factorization because it uses partial pivoting for numerical accuracy.

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

8. $\begin{bmatrix} 6 & 9 \\ 4 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -1 & 2 \\ -3 & -2 & 10 \\ 9 & -5 & 6 \end{bmatrix}$

10. $\begin{bmatrix} -5 & 3 & 4 \\ 10 & -8 & -9 \\ 15 & 1 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 3 & -6 & 3 \\ 6 & -7 & 2 \\ -1 & 7 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 2 & -4 & 2 \\ 1 & 5 & -4 \\ -6 & -2 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 3 & -5 & -3 \\ -1 & -5 & 8 & 4 \\ 4 & 2 & -5 & -7 \\ -2 & -4 & 7 & 5 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 4 & -1 & 5 \\ 3 & 7 & -2 & 9 \\ -2 & -3 & 1 & -4 \\ -1 & 6 & -1 & 7 \end{bmatrix}$

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

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

17. When A is invertible, MATLAB finds A^{-1} by factoring $A = LU$ (where L may be permuted lower triangular), inverting L and U, and then computing $U^{-1}L^{-1}$. Use this method to compute the inverse of A in Exercise 2. (Apply the algorithm of Section 2.2 to L and to U.)

18. Find A^{-1} as in Exercise 17, using A from Exercise 3.

19. Let A be a lower triangular $n \times n$ matrix with nonzero entries on the diagonal. Show that A is invertible and A^{-1} is lower triangular. [*Hint:* Explain why A can be changed into I using only row replacements and scaling. (Where are the pivots?) Also, explain why the row operations that reduce A to I change I into a lower triangular matrix.]

20. Let $A = LU$ be an LU factorization. Explain why A can be row reduced to U using only replacement operations. (This fact is the converse of what was proved in the text.)

21. Suppose $A = BC$, where B is invertible. Show that any sequence of row operations that reduces B to I also reduces A

to C. The converse is not true, since the zero matrix may be factored as $0 = B\cdot 0$.

Exercises 22–26 provide a glimpse of some widely used matrix factorizations, some of which are discussed later in the text.

22. (*Reduced LU Factorization*) With A as in the Practice Problem, find a 5×3 matrix B and a 3×4 matrix C such that $A = BC$. Generalize this idea to the case where A is $m\times n$, $A = LU$, and U has only three nonzero rows.

23. (*Rank Factorization*) Suppose an $m\times n$ matrix A admits a factorization $A = CD$ where C is $m\times 4$ and D is $4\times n$.

a. Show that A is the sum of four outer products. (See Section 2.4.)

b. Let $m = 400$ and $n = 100$. Explain why a computer programmer might prefer to store the data from A in the form of two matrices C and D.

24. (*QR Factorization*) Suppose $A = QR$, where Q and R are $n\times n$, R is invertible and upper triangular, and Q has the property that $Q^TQ = I$. Show that for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. What computations with Q and R will produce the solution?

WEB

25. (*Singular Value Decomposition*) Suppose $A = UDV^T$, where U and V are $n\times n$ matrices with the property that $U^TU = I$ and $V^TV = I$, and where D is a diagonal matrix with positive numbers $\sigma_1, \dots, \sigma_n$ on the diagonal. Show that A is invertible, and find a formula for A^{-1}.

26. (*Spectral Factorization*) Suppose a 3×3 matrix A admits a factorization as $A = PDP^{-1}$, where P is some invertible 3×3 matrix and D is the diagonal matrix

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

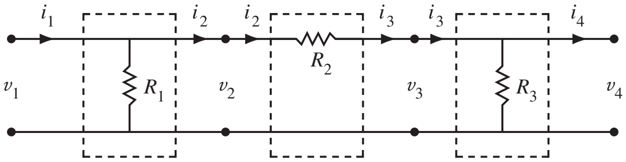
Show that this factorization is useful when computing high powers of A. Find fairly simple formulas for A^2, A^3, and A^k (k a positive integer), using P and the entries in D.

27. Design two different ladder networks that each output 9 volts and 4 amps when the input is 12 volts and 6 amps.

28. Show that if three shunt circuits (with resistances R_1, R_2, R_3) are connected in series, the resulting network has the same transfer matrix as a single shunt circuit. Find a formula for the resistance in that circuit.

29. a. Compute the transfer matrix of the network in the figure.

b. Let $A = \begin{bmatrix} 4/3 & -12 \\ -1/4 & 3 \end{bmatrix}$. Design a ladder network whose transfer matrix is A by finding a suitable matrix factorization of A.

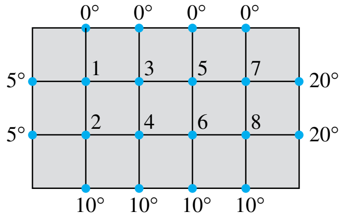


30. Find a different factorization of the A in Exercise 29, and thereby design a different ladder network whose transfer matrix is A.

31. **[M]** The solution to the steady-state heat flow problem for the plate in the figure is approximated by the solution to the equation $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (5, 15, 0, 10, 0, 10, 20, 30)$ and

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

WEB



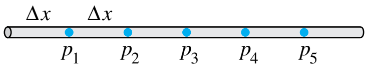
(Refer to Exercise 33 of Section 1.1.) The missing entries in A are zeros. The nonzero entries of A lie within a band along the main diagonal. Such *band matrices* occur in a variety of applications and often are extremely large (with thousands of rows and columns but relatively narrow bands).

a. Use the method of Example 2 to construct an LU factorization of A, and note that both factors are band matrices (with two nonzero diagonals below or above the main diagonal). Compute $LU - A$ to check your work.

b. Use the LU factorization to solve $A\mathbf{x} = \mathbf{b}$.

c. Obtain A^{-1} and note that A^{-1} is a dense matrix with no band structure. When A is large, L and U can be stored in much less space than A^{-1}. This fact is another reason for preferring the LU factorization of A to A^{-1} itself.

32. **[M]** The band matrix A shown below can be used to estimate the unsteady conduction of heat in a rod when the temperatures at points $p_1, \dots, p_5$ on the rod change with time.[2]



The constant C in the matrix depends on the physical nature of the rod, the distance Δx between the points on the rod, and the length of time Δt between successive temperature measurements. Suppose that for $k = 0, 1, 2, \dots$, a vector $\mathbf{t}_k$ in $\mathbb{R}^5$ lists the temperatures at time $k\Delta t$. If the two ends of the rod are maintained at 0°, then the temperature vectors satisfy the equation $A\mathbf{t}_{k+1} = \mathbf{t}_k$ $(k = 0, 1, \dots)$, where

$$A = \begin{bmatrix} (1+2C) & -C & & & \\ -C & (1+2C) & -C & & \\ & -C & (1+2C) & -C & \\ & & -C & (1+2C) & -C \\ & & & -C & (1+2C) \end{bmatrix}$$

a. Find the LU factorization of A when $C = 1$. A matrix such as A with three nonzero diagonals is called a *tridiagonal matrix*. The L and U factors are *bidiagonal matrices*.

b. Suppose $C = 1$ and $\mathbf{t}_0 = (10, 12, 12, 12, 10)$. Use the LU factorization of A to find the temperature distributions $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, and $\mathbf{t}_4$.

---

[2] See Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200–201.

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

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

Divide the entries in each highlighted column by the pivot at the top. The resulting columns form the first three columns in the lower half of L. This suffices to make row reduction of L to I correspond to reduction of A to U. Use the last two columns of I_5 to make L unit lower triangular.

$$\begin{array}{ccc} \begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix} & \begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix} & \begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix} \\ \div 2 & \div 3 & \div 5 \\ \downarrow & \downarrow & \downarrow \end{array}$$

$$\begin{bmatrix} 1 & & & & \\ 3 & 1 & & & \\ 1 & -1 & 1 & \cdots & \\ 2 & 2 & -1 & & \\ -3 & -3 & 2 & & \end{bmatrix}, \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

2.6 THE LEONTIEF INPUT–OUTPUT MODEL

Linear algebra played an essential role in the Nobel prize–winning work of Wassily Leontief, as mentioned at the beginning of Chapter 1. The economic model described in this section is the basis for more elaborate models used in many parts of the world.

Suppose a nation's economy is divided into n sectors that produce goods or services, and let $\mathbf{x}$ be a **production vector** in $\mathbb{R}^n$ that lists the output of each sector for one year. Also, suppose another part of the economy (called the *open sector*) does not produce goods or services but only consumes them, and let $\mathbf{d}$ be a **final demand vector** (or **bill of final demands)** that lists the value of the goods and services demanded from the various sectors by the nonproductive part of the economy. The vector $\mathbf{d}$ can represent consumer demand, government consumption, surplus production, exports, or other external demand.

As the various sectors produce goods to meet consumer demand, the producers themselves create additional **intermediate demand** for goods they need as inputs for their own production. The interrelations between the sectors are very complex, and the connection between the final demand and the production is unclear. Leontief asked if there is a production level $\mathbf{x}$ such that the amounts produced (or "supplied") will exactly balance the total demand for that production, so that

$$\left\{\begin{matrix}\text{amount}\\ \text{produced}\\ \mathbf{x}\end{matrix}\right\} = \left\{\begin{matrix}\text{intermediate}\\ \text{demand}\end{matrix}\right\} + \left\{\begin{matrix}\text{final}\\ \text{demand}\\ \mathbf{d}\end{matrix}\right\} \tag{1}$$

The basic assumption of Leontief's input–output model is that for each sector, there is a **unit consumption vector** in $\mathbb{R}^n$ that lists the inputs needed *per unit of output* of the sector. All input and output units are measured in millions of dollars, rather than in quantities such as tons or bushels. (Prices of goods and services are held constant.)

As a simple example, suppose the economy consists of three sectors—manufacturing, agriculture, and services—with unit consumption vectors $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$ shown in the table below:

Purchased from:	Inputs Consumed per Unit of Output		
	Manufacturing	Agriculture	Services
Manufacturing	.50	.40	.20
Agriculture	.20	.30	.10
Services	.10	.10	.30
	↑ $\mathbf{c}_1$	↑ $\mathbf{c}_2$	↑ $\mathbf{c}_3$

EXAMPLE 1 What amounts will be consumed by the manufacturing sector if it decides to produce 100 units?

Solution Compute

$$100\mathbf{c}_1 = 100\begin{bmatrix}.50\\ .20\\ .10\end{bmatrix} = \begin{bmatrix}50\\ 20\\ 10\end{bmatrix}$$

To produce 100 units, manufacturing will order (i.e., "demand") and consume 50 units from other parts of the manufacturing sector, 20 units from agriculture, and 10 units from services.

If manufacturing decides to produce x_1 units of output, then $x_1\mathbf{c}_1$ represents the *intermediate demands* of manufacturing, because the amounts in $x_1\mathbf{c}_1$ will be consumed in the process of creating the x_1 units of output. Likewise, if x_2 and x_3 denote the planned outputs of the agriculture and services sectors, $x_2\mathbf{c}_2$ and $x_3\mathbf{c}_3$ list their corresponding intermediate demands. The total intermediate demand from all three sectors is given by

$$\begin{aligned}\{\text{intermediate demand}\} &= x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + x_3\mathbf{c}_3\\ &= C\mathbf{x} \qquad (2)\end{aligned}$$

where C is the **consumption matrix** $[\,\mathbf{c}_1\quad \mathbf{c}_2\quad \mathbf{c}_3\,]$, namely,

$$C = \begin{bmatrix} .50 & .40 & .20\\ .20 & .30 & .10\\ .10 & .10 & .30\end{bmatrix} \qquad (3)$$

Equations (1) and (2) yield Leontief's model.

THE LEONTIEF INPUT–OUTPUT MODEL, OR PRODUCTION EQUATION

$$\underset{\substack{\text{Amount}\\ \text{produced}}}{\mathbf{x}} = \underset{\substack{\text{Intermediate}\\ \text{demand}}}{C\mathbf{x}} + \underset{\substack{\text{Final}\\ \text{demand}}}{\mathbf{d}} \qquad (4)$$

Writing $\mathbf{x}$ as $I\mathbf{x}$ and using matrix algebra, we can rewrite (4):

$$\begin{aligned} I\mathbf{x} - C\mathbf{x} &= \mathbf{d}\\ (I - C)\mathbf{x} &= \mathbf{d} \qquad (5)\end{aligned}$$

EXAMPLE 2 Consider the economy whose consumption matrix is given by (3). Suppose the final demand is 50 units for manufacturing, 30 units for agriculture, and 20 units for services. Find the production level $\mathbf{x}$ that will satisfy this demand.

Solution The coefficient matrix in (5) is

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

To solve (5), row reduce the augmented matrix

$$\begin{bmatrix} .5 & -.4 & -.2 & 50\\ -.2 & .7 & -.1 & 30\\ -.1 & -.1 & .7 & 20\end{bmatrix} \sim \begin{bmatrix} 5 & -4 & -2 & 500\\ -2 & 7 & -1 & 300\\ -1 & -1 & 7 & 200\end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 226\\ 0 & 1 & 0 & 119\\ 0 & 0 & 1 & 78\end{bmatrix}$$

The last column is rounded to the nearest whole unit. Manufacturing must produce approximately 226 units, agriculture 119 units, and services only 78 units.

If the matrix $I - C$ is invertible, then we can apply Theorem 5 in Section 2.2, with A replaced by $(I - C)$, and from the equation $(I - C)\mathbf{x} = \mathbf{d}$ obtain $\mathbf{x} = (I - C)^{-1}\mathbf{d}$. The theorem below shows that in most practical cases, $I - C$ *is* invertible and the production vector $\mathbf{x}$ is economically feasible, in the sense that the entries in $\mathbf{x}$ are nonnegative.

In the theorem, the term **column sum** denotes the sum of the entries in a column of a matrix. Under ordinary circumstances, the column sums of a consumption matrix are less than 1 because a sector should require less than one unit's worth of inputs to produce one unit of output.

THEOREM 11

Let C be the consumption matrix for an economy, and let $\mathbf{d}$ be the final demand. If C and $\mathbf{d}$ have nonnegative entries and if each column sum of C is less than 1, then $(I - C)^{-1}$ exists and the production vector

$$\mathbf{x} = (I - C)^{-1}\mathbf{d}$$

has nonnegative entries and is the unique solution of

$$\mathbf{x} = C\mathbf{x} + \mathbf{d}$$

The following discussion will suggest why the theorem is true and will lead to a new way to compute $(I - C)^{-1}$.

A Formula for $(I - C)^{-1}$

Imagine that the demand represented by $\mathbf{d}$ is presented to the various industries at the beginning of the year, and the industries respond by setting their production levels at $\mathbf{x} = \mathbf{d}$, which will exactly meet the final demand. As the industries prepare to produce $\mathbf{d}$, they send out orders for their raw materials and other inputs. This creates an intermediate demand of $C\mathbf{d}$ for inputs.

To meet the additional demand of $C\mathbf{d}$, the industries will need as additional inputs the amounts in $C(C\mathbf{d}) = C^2\mathbf{d}$. Of course, this creates a second round of intermediate demand, and when the industries decide to produce even more to meet this new demand, they create a third round of demand, namely, $C(C^2\mathbf{d}) = C^3\mathbf{d}$. And so it goes.

Theoretically, we can imagine this process continuing indefinitely, although in real life it would not take place in such a rigid sequence of events. We can diagram this hypothetical situation as follows:

	Demand That Must Be Met	Inputs Needed to Meet This Demand
Final demand	$\mathbf{d}$	$C\mathbf{d}$
Intermediate demand		
1st round	$C\mathbf{d}$	$C(C\mathbf{d}) = C^2\mathbf{d}$
2nd round	$C^2\mathbf{d}$	$C(C^2\mathbf{d}) = C^3\mathbf{d}$
3rd round	$C^3\mathbf{d}$	$C(C^3\mathbf{d}) = C^4\mathbf{d}$
	$\vdots$	$\vdots$

The production level **x** that will meet all of this demand is

$$\begin{aligned}\mathbf{x} &= \mathbf{d} + C\mathbf{d} + C^2\mathbf{d} + C^3\mathbf{d} + \cdots \\ &= (I + C + C^2 + C^3 + \cdots)\mathbf{d} \qquad (6)\end{aligned}$$

To make sense of (6), we use the following algebraic identity:

$$(I - C)(I + C + C^2 + \cdots + C^m) = I - C^{m+1} \qquad (7)$$

It can be shown that if the column sums in C are all strictly less than 1, then $I - C$ is invertible, C^m approaches the zero matrix as m gets arbitrarily large, and $I - C^{m+1} \to I$. (This fact is analogous to the fact that if a positive number t is less than 1, then $t^m \to 0$ as m increases.) Using (7), we write

$$(I - C)^{-1} \approx I + C + C^2 + C^3 + \cdots + C^m \qquad (8)$$

when the column sums of C are less than 1.

We interpret (8) as meaning that the right side can be made as close to $(I - C)^{-1}$ as desired by taking m sufficiently large.

In actual input–output models, powers of the consumption matrix approach the zero matrix rather quickly. So (8) really provides a practical way to compute $(I - C)^{-1}$. Likewise, for any **d**, the vectors $C^m\mathbf{d}$ approach the zero vector quickly, and (6) is a practical way to solve $(I - C)\mathbf{x} = \mathbf{d}$. If the entries in C and **d** are nonnegative, then (6) shows that the entries in **x** are nonnegative, too.

The Economic Importance of Entries in $(I - C)^{-1}$

The entries in $(I - C)^{-1}$ are significant because they can be used to predict how the production **x** will have to change when the final demand **d** changes. In fact, the entries in column j of $(I - C)^{-1}$ are the *increased* amounts the various sectors will have to produce in order to satisfy *an increase of 1 unit* in the final demand for output from sector j. See Exercise 8.

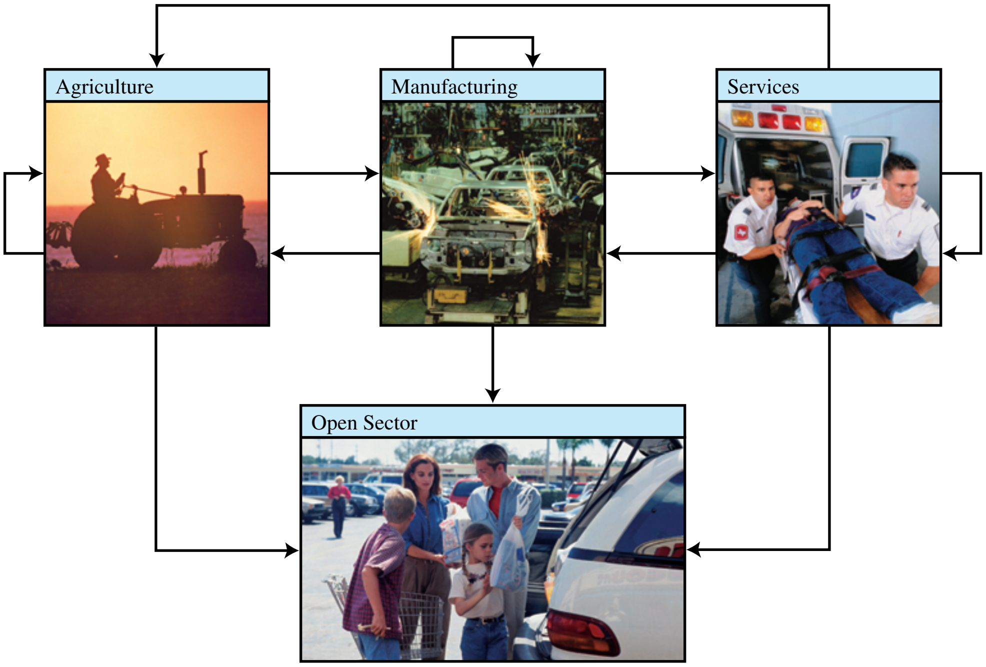
NUMERICAL NOTE

In any applied problem (not just in economics), an equation $A\mathbf{x} = \mathbf{b}$ can always be written as $(I - C)\mathbf{x} = \mathbf{b}$, with $C = I - A$. If the system is large and *sparse* (with mostly zero entries), it can happen that the column sums of the absolute values in C are less than 1. In this case, $C^m \to 0$. If C^m approaches zero quickly enough, (6) and (8) will provide practical formulas for solving $A\mathbf{x} = \mathbf{b}$ and finding A^{-1}.

PRACTICE PROBLEM

Suppose an economy has two sectors, goods and services. One unit of output from goods requires inputs of .2 unit from goods and .5 unit from services. One unit of output from services requires inputs of .4 unit from goods and .3 unit from services. There is a final demand of 20 units of goods and 30 units of services. Set up the Leontief input–output model for this situation.

2.6 EXERCISES



Exercises 1–4 refer to an economy that is divided into three sectors—manufacturing, agriculture, and services. For each unit of output, manufacturing requires .10 unit from other companies in that sector, .30 unit from agriculture, and .30 unit from services. For each unit of output, agriculture uses .20 unit of its own output, .60 unit from manufacturing, and .10 unit from services. For each unit of output, the services sector consumes .10 unit from services, .60 unit from manufacturing, but no agricultural products.

1. Construct the consumption matrix for this economy, and determine what intermediate demands are created if agriculture plans to produce 100 units.
2. Determine the production levels needed to satisfy a final demand of 18 units for agriculture, with no final demand for the other sectors. (Do not compute an inverse matrix.)
3. Determine the production levels needed to satisfy a final demand of 18 units for manufacturing, with no final demand for the other sectors. (Do not compute an inverse matrix.)
4. Determine the production levels needed to satisfy a final demand of 18 units for manufacturing, 18 units for agriculture, and 0 units for services.
5. Consider the production model $\mathbf{x} = C\mathbf{x} + \mathbf{d}$ for an economy with two sectors, where
$$C = \begin{bmatrix} .0 & .5 \\ .6 & .2 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 50 \\ 30 \end{bmatrix}$$
Use an inverse matrix to determine the production level necessary to satisfy the final demand.
6. Repeat Exercise 5 with $C = \begin{bmatrix} .1 & .6 \\ .5 & .2 \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} 18 \\ 11 \end{bmatrix}$.
7. Let C and $\mathbf{d}$ be as in Exercise 5.
   a. Determine the production level necessary to satisfy a final demand for 1 unit of output from sector 1.

b. Use an inverse matrix to determine the production level necessary to satisfy a final demand of $\begin{bmatrix} 51 \\ 30 \end{bmatrix}$.

c. Use the fact that $\begin{bmatrix} 51 \\ 30 \end{bmatrix} = \begin{bmatrix} 50 \\ 30 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ to explain how and why the answers to parts (a) and (b) and to Exercise 5 are related.

8. Let C be an $n \times n$ consumption matrix whose column sums are less than 1. Let $\mathbf{x}$ be the production vector that satisfies a final demand $\mathbf{d}$, and let $\Delta\mathbf{x}$ be a production vector that satisfies a different final demand $\Delta\mathbf{d}$.

a. Show that if the final demand changes from $\mathbf{d}$ to $\mathbf{d} + \Delta\mathbf{d}$, then the new production level must be $\mathbf{x} + \Delta\mathbf{x}$. Thus $\Delta\mathbf{x}$ gives the amounts by which production must *change* in order to accommodate the *change* $\Delta\mathbf{d}$ in demand.

b. Let $\Delta\mathbf{d}$ be the vector in $\mathbb{R}^n$ with 1 as the first entry and 0's elsewhere. Explain why the corresponding production $\Delta\mathbf{x}$ is the first column of $(I - C)^{-1}$. This shows that the first column of $(I - C)^{-1}$ gives the amounts the various sectors must produce to satisfy an increase of 1 unit in the final demand for output from sector 1.

9. Solve the Leontief production equation for an economy with three sectors, given that

$$C = \begin{bmatrix} .2 & .2 & .0 \\ .3 & .1 & .3 \\ .1 & .0 & .2 \end{bmatrix} \quad \text{and} \quad \mathbf{d} = \begin{bmatrix} 40 \\ 60 \\ 80 \end{bmatrix}$$

10. The consumption matrix C for the U.S. economy in 1972 has the property that *every entry* in the matrix $(I - C)^{-1}$ is nonzero (and positive).[1] What does that say about the effect of raising the demand for the output of just one sector of the economy?

11. The Leontief production equation, $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, is usually accompanied by a dual **price equation**,

$$\mathbf{p} = C^T\mathbf{p} + \mathbf{v}$$

where $\mathbf{p}$ is a **price vector** whose entries list the price per unit for each sector's output, and $\mathbf{v}$ is a **value added vector** whose entries list the value added per unit of output. (Value added includes wages, profit, depreciation, etc.) An important fact in economics is that the gross domestic product (GDP) can be expressed in two ways:

$$\{\text{gross domestic product}\} = \mathbf{p}^T\mathbf{d} = \mathbf{v}^T\mathbf{x}$$

Verify the second equality. [*Hint*: Compute $\mathbf{p}^T\mathbf{x}$ in two ways.]

12. Let C be a consumption matrix such that $C^m \to 0$ as $m \to \infty$, and for $m = 1, 2, \ldots$, let $D_m = I + C + \cdots + C^m$. Find a difference equation that relates D_m and D_{m+1} and thereby obtain an iterative procedure for computing formula (8) for $(I - C)^{-1}$.

13. **[M]** The consumption matrix C below is based on input–output data for the U.S. economy in 1958, with data for 81 sectors grouped into 7 larger sectors: (1) nonmetal household and personal products, (2) final metal products (such as motor vehicles), (3) basic metal products and mining, (4) basic nonmetal products and agriculture, (5) energy, (6) services, and (7) entertainment and miscellaneous products.[2] Find the production levels needed to satisfy the final demand $\mathbf{d}$. (Units are in millions of dollars.)

$$\begin{bmatrix} .1588 & .0064 & .0025 & .0304 & .0014 & .0083 & .1594 \\ .0057 & .2645 & .0436 & .0099 & .0083 & .0201 & .3413 \\ .0264 & .1506 & .3557 & .0139 & .0142 & .0070 & .0236 \\ .3299 & .0565 & .0495 & .3636 & .0204 & .0483 & .0649 \\ .0089 & .0081 & .0333 & .0295 & .3412 & .0237 & .0020 \\ .1190 & .0901 & .0996 & .1260 & .1722 & .2368 & .3369 \\ .0063 & .0126 & .0196 & .0098 & .0064 & .0132 & .0012 \end{bmatrix},$$

$$\mathbf{d} = \begin{bmatrix} 74{,}000 \\ 56{,}000 \\ 10{,}500 \\ 25{,}000 \\ 17{,}500 \\ 196{,}000 \\ 5{,}000 \end{bmatrix}$$

14. **[M]** The demand vector in Exercise 13 is reasonable for 1958 data, but Leontief's discussion of the economy in the reference cited there used a demand vector closer to 1964 data:

$\mathbf{d} = (99640, 75548, 14444, 33501, 23527, 263985, 6526)$

Find the production levels needed to satisfy this demand.

15. **[M]** Use equation (6) to solve the problem in Exercise 13. Set $\mathbf{x}^{(0)} = \mathbf{d}$, and for $k = 1, 2, \ldots$, compute $\mathbf{x}^{(k)} = \mathbf{d} + C\mathbf{x}^{(k-1)}$. How many steps are needed to obtain the answer in Exercise 13 to four significant figures?

---

[1] Wassily W. Leontief, "The World Economy of the Year 2000," *Scientific American*, September 1980, pp. 206–231.

[2] Wassily W. Leontief, "The Structure of the U.S. Economy," *Scientific American*, April 1965, pp. 30–32.

SOLUTION TO PRACTICE PROBLEM

The following data are given:

Purchased from:	Inputs Needed per Unit of Output Goods	Services	External Demand
Goods	.2	.4	20
Services	.5	.3	30

The Leontief input–output model is $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, where

$$C = \begin{bmatrix} .2 & .4 \\ .5 & .3 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 20 \\ 30 \end{bmatrix}$$

2.7 APPLICATIONS TO COMPUTER GRAPHICS

Computer graphics are images displayed or animated on a computer screen. Applications of computer graphics are widespread and growing rapidly. For instance, computer-aided design (CAD) is an integral part of many engineering processes, such as the aircraft design process described in the chapter introduction. The entertainment industry has made the most spectacular use of computer graphics—from the special effects in *The Matrix* to PlayStation 2 and the Xbox.

Most interactive computer software for business and industry makes use of computer graphics in the screen displays and for other functions, such as graphical display of data, desktop publishing, and slide production for commercial and educational presentations. Consequently, anyone studying a computer language invariably spends time learning how to use at least two-dimensional (2D) graphics.

This section examines some of the basic mathematics used to manipulate and display graphical images such as a wire-frame model of an airplane. Such an image (or picture) consists of a number of points, connecting lines or curves, and information about how to fill in closed regions bounded by the lines and curves. Often, curved lines are approximated by short straight-line segments, and a figure is defined mathematically by a list of points.

Among the simplest 2D graphics symbols are letters used for labels on the screen. Some letters are stored as wire-frame objects; others that have curved portions are stored with additional mathematical formulas for the curves.

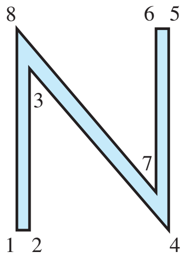


FIGURE 1
Regular N.

EXAMPLE 1 The capital letter N in Fig. 1 is determined by eight points, or *vertices*. The coordinates of the points can be stored in a data matrix, D.

$$\begin{array}{r|cccccccc|}
 & \multicolumn{8}{c}{\text{Vertex:}} \\
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
x\text{-coordinate} & 0 & .5 & .5 & 6 & 6 & 5.5 & 5.5 & 0 \\
y\text{-coordinate} & 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8
\end{array} = D$$

In addition to D, it is necessary to specify which vertices are connected by lines, but we omit this detail.

The main reason graphical objects are described by collections of straight-line segments is that the standard transformations in computer graphics map line segments onto other line segments. (For instance, see Exercise 27 in Section 1.8.) Once the vertices that describe an object have been transformed, their images can be connected with the appropriate straight lines to produce the complete image of the original object.

EXAMPLE 2 Given $A = \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix}$, describe the effect of the shear transformation $\mathbf{x} \mapsto A\mathbf{x}$ on the letter N in Example 1.

Solution By definition of matrix multiplication, the columns of the product AD contain the images of the vertices of the letter N.

$$AD = \begin{array}{c} \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 0 & .5 & 2.105 & 6 & 8 & 7.5 & 5.895 & 2 \\ 0 & 0 & 6.420 & 0 & 8 & 8 & 1.580 & 8 \end{bmatrix} \end{array}$$

The transformed vertices are plotted in Fig. 2, along with connecting line segments that correspond to those in the original figure.

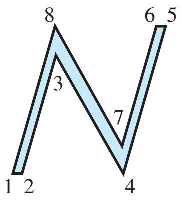


FIGURE 2
Slanted N.

The italic N in Fig. 2 looks a bit too wide. To compensate, we can shrink the width by a scale transformation.

EXAMPLE 3 Compute the matrix of the transformation that performs a shear transformation, as in Example 2, and then scales all x-coordinates by a factor of .75.

Solution The matrix that multiplies the x-coordinate of a point by .75 is

$$S = \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix}$$

So the matrix of the composite transformation is

$$\begin{aligned} SA &= \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} .75 & .1875 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

The result of this composite transformation is shown in Fig. 3.

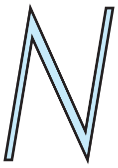
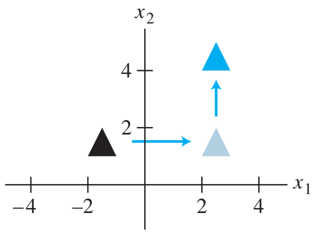


FIGURE 3
Composite transformation of N.

The mathematics of computer graphics is intimately connected with matrix multiplication. Unfortunately, translating an object on a screen does not correspond directly to matrix multiplication because translation is not a linear transformation. The standard way to avoid this difficulty is to introduce what are called *homogeneous coordinates*.

Homogeneous Coordinates

Each point (x, y) in $\mathbb{R}^2$ can be identified with the point $(x, y, 1)$ on the plane in $\mathbb{R}^3$ that lies one unit above the xy-plane. We say that (x, y) has *homogeneous coordinates* $(x, y, 1)$. For instance, the point $(0, 0)$ has homogeneous coordinates $(0, 0, 1)$. Homogeneous coordinates for points are not added or multiplied by scalars, but they can be transformed via multiplication by 3×3 matrices.



Translation by $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

EXAMPLE 4 A translation of the form $(x, y) \mapsto (x+h, y+k)$ is written in homogeneous coordinates as $(x, y, 1) \mapsto (x+h, y+k, 1)$. This transformation can be computed via matrix multiplication:

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+h \\ y+k \\ 1 \end{bmatrix}$$

EXAMPLE 5 Any linear transformation on $\mathbb{R}^2$ is represented with respect to homogeneous coordinates by a partitioned matrix of the form $\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$, where A is a 2×2 matrix. Typical examples are

$$\underset{\substack{\text{Counterclockwise}\\ \text{rotation about the}\\ \text{origin, angle } \varphi}}{\begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}}, \quad \underset{\substack{\text{Reflection}\\ \text{through } y = x}}{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}, \quad \underset{\substack{\text{Scale } x \text{ by } s\\ \text{and } y \text{ by } t}}{\begin{bmatrix} s & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

Composite Transformations

The movement of a figure on a computer screen often requires two or more basic transformations. The composition of such transformations corresponds to matrix multiplication when homogeneous coordinates are used.

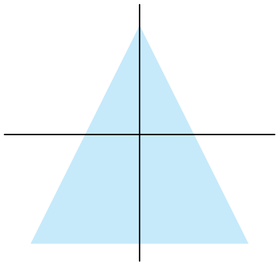
EXAMPLE 6 Find the 3×3 matrix that corresponds to the composite transformation of a scaling by .3, a rotation of 90°, and finally a translation that adds $(-.5, 2)$ to each point of a figure.

Solution If $\varphi = \pi/2$, then $\sin\varphi = 1$ and $\cos\varphi = 0$. From Examples 4 and 5, we have

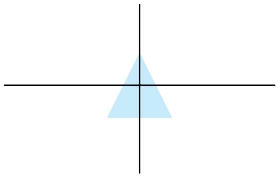
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \xrightarrow{\text{Scale}} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\xrightarrow{\text{Rotate}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

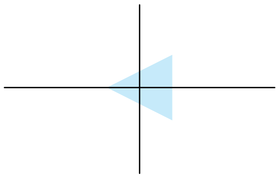
$$\xrightarrow{\text{Translate}} \begin{bmatrix} 1 & 0 & -.5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



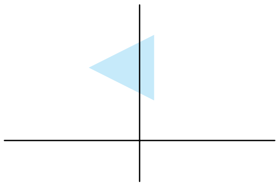
Original Figure



After Scaling



After Rotating



After Translating

The matrix for the composite transformation is

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

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

3D Computer Graphics

Some of the newest and most exciting work in computer graphics is connected with molecular modeling. With 3D (three-dimensional) graphics, a biologist can examine a simulated protein molecule and search for active sites that might accept a drug molecule. The biologist can rotate and translate an experimental drug and attempt to attach it to the protein. This ability to *visualize* potential chemical reactions is vital to modern drug and cancer research. In fact, advances in drug design depend to some extent upon progress in the ability of computer graphics to construct realistic simulations of molecules and their interactions.[1]

Current research in molecular modeling is focused on *virtual reality*, an environment in which a researcher can see and *feel* the drug molecule slide into the protein. In Fig. 4, such tactile feedback is provided by a force-displaying remote manipulator.

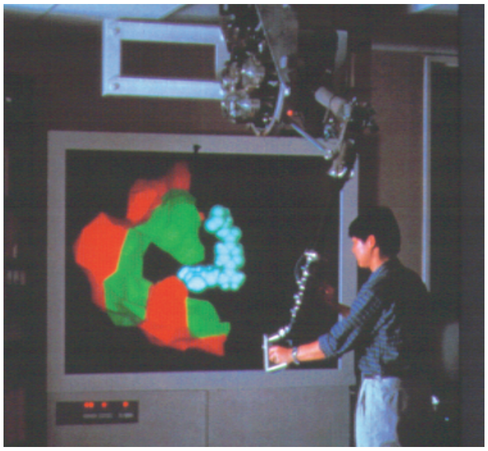


FIGURE 4 Molecular modeling in virtual reality. (Computer Science Department, University of North Carolina at Chapel Hill. Photo by Bo Strain.)

---

[1] Robert Pool, "Computing in Science," *Science* **256**, 3 April 1992, p. 45.

Another design for virtual reality involves a helmet and glove that detect head, hand, and finger movements. The helmet contains two tiny computer screens, one for each eye. Making this virtual environment more realistic is a challenge to engineers, scientists, and mathematicians. The mathematics we examine here barely opens the door to this interesting field of research.

Homogeneous 3D Coordinates

By analogy with the 2D case, we say that $(x, y, z, 1)$ are homogeneous coordinates for the point (x, y, z) in $\mathbb{R}^3$. In general, (X, Y, Z, H) are **homogeneous coordinates** for (x, y, z) if $H \neq 0$ and

$$x = \frac{X}{H}, \qquad y = \frac{Y}{H}, \quad \text{and} \quad z = \frac{Z}{H} \tag{1}$$

Each nonzero scalar multiple of $(x, y, z, 1)$ gives a set of homogeneous coordinates for (x, y, z). For instance, both $(10, -6, 14, 2)$ and $(-15, 9, -21, -3)$ are homogeneous coordinates for $(5, -3, 7)$.

The next example illustrates the transformations used in molecular modeling to move a drug into a protein molecule.

EXAMPLE 7 Give 4×4 matrices for the following transformations:

a. Rotation about the y-axis through an angle of 30°. (By convention, a positive angle is the counterclockwise direction when looking toward the origin from the positive half of the axis of rotation—in this case, the y-axis.)

b. Translation by the vector $\mathbf{p} = (-6, 4, 5)$.

Solution

a. First, construct the 3×3 matrix for the rotation. The vector $\mathbf{e}_1$ rotates down toward the negative z-axis, stopping at $(\cos 30°, 0, -\sin 30°) = (\sqrt{3}/2, 0, -.5)$. The vector $\mathbf{e}_2$ on the y-axis does not move, but $\mathbf{e}_3$ on the z-axis rotates down toward the positive x-axis, stopping at $(\sin 30°, 0, \cos 30°) = (.5, 0, \sqrt{3}/2)$. See Fig. 5. From Section 1.9, the standard matrix for this rotation is

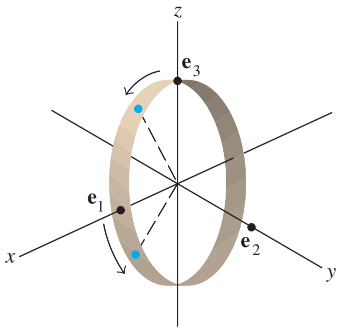


FIGURE 5

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

So the rotation matrix for homogeneous coordinates is

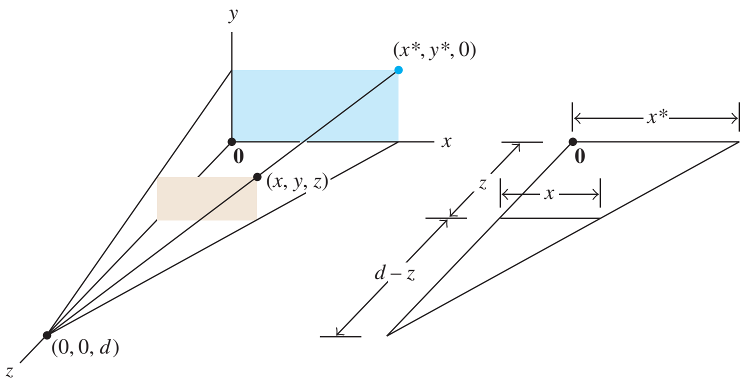
$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ -.5 & 0 & \sqrt{3}/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b. We want $(x, y, z, 1)$ to map to $(x - 6, y + 4, z + 5, 1)$. The matrix that does this is

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

Perspective Projections

A three-dimensional object is represented on the two-dimensional computer screen by projecting the object onto a *viewing plane*. (We ignore other important steps, such as selecting the portion of the viewing plane to display on the screen.) For simplicity, let the xy-plane represent the computer screen, and imagine that the eye of a viewer is along the positive z-axis, at a point $(0, 0, d)$. A *perspective projection* maps each point (x, y, z) onto an image point $(x^*, y^*, 0)$ so that the two points and the eye position, called the *center of projection*, are on a line. See Fig. 6(a).



(a) (b)

FIGURE 6 Perspective projection of (x, y, z) onto $(x^*, y^*, 0)$.

The triangle in the xz-plane in Fig. 6(a) is redrawn in part (b) showing the lengths of line segments. Similar triangles show that

$$\frac{x^*}{d} = \frac{x}{d - z} \quad \text{and} \quad x^* = \frac{dx}{d - z} = \frac{x}{1 - z/d}$$

Similarly,

$$y^* = \frac{y}{1 - z/d}$$

Using homogeneous coordinates, we can represent the perspective projection by a matrix, say, P. We want $(x, y, z, 1)$ to map into $\left(\frac{x}{1 - z/d}, \frac{y}{1 - z/d}, 0, 1\right)$. Scaling these coordinates by $1 - z/d$, we can also use $(x, y, 0, 1 - z/d)$ as homogeneous coordinates for the image. Now it is easy to display P. In fact,

$$P\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \\ 1 - z/d \end{bmatrix}$$

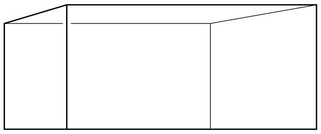
EXAMPLE 8 Let S be the box with vertices $(3, 1, 5)$, $(5, 1, 5)$, $(5, 0, 5)$, $(3, 0, 5)$, $(3, 1, 4)$, $(5, 1, 4)$, $(5, 0, 4)$, and $(3, 0, 4)$. Find the image of S under the perspective projection with center of projection at $(0, 0, 10)$.

Solution Let P be the projection matrix, and let D be the data matrix for S using homogeneous coordinates. The data matrix for the image of S is

$$PD = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/10 & 1 \end{bmatrix} \begin{array}{c} \text{Vertex:} \\ \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 5 & 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .5 & .5 & .5 & .5 & .6 & .6 & .6 & .6 \end{bmatrix}$$

To obtain $\mathbb{R}^3$ coordinates, use (1) and divide the top three entries in each column by the corresponding entry in the fourth row:



S under the perspective transformation.

$$\begin{array}{c} \text{Vertex:} \\ \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 6 & 10 & 10 & 6 & 5 & 8.3 & 8.3 & 5 \\ 2 & 2 & 0 & 0 & 1.7 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

This text's web site has some interesting applications of computer graphics, including a further discussion of perspective projections. One of the computer projects on the web site involves simple animation.

NUMERICAL NOTE

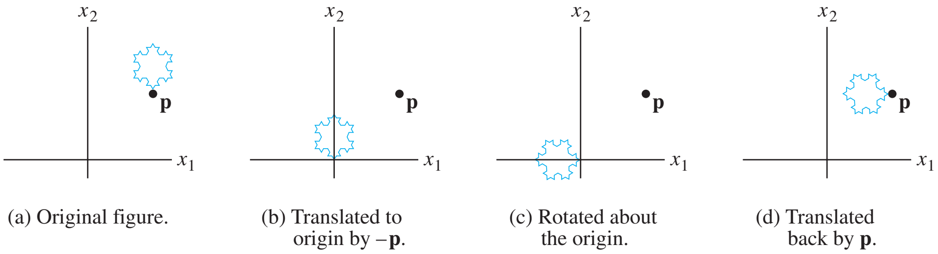
Continuous movement of graphical 3D objects requires intensive computation with 4×4 matrices, particularly when the surfaces are *rendered* to appear realistic, with texture and appropriate lighting. *High-end computer graphics boards* have 4×4 matrix operations and graphics algorithms embedded in their microchips and circuitry. Such boards can perform the billions of matrix multiplications per second needed for realistic color animation in 3D gaming programs.[2]

Further Reading

James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes, *Computer Graphics: Principles and Practice*, 3rd ed. (Boston, MA: Addison-Wesley, 2002), Chapters 5 and 6.

PRACTICE PROBLEM

Rotation of a figure about a point **p** in $\mathbb{R}^2$ is accomplished by first translating the figure by $-\mathbf{p}$, rotating about the origin, and then translating back by **p**. See Fig. 7. Construct the 3×3 matrix that rotates points $-30°$ about the point $(-2, 6)$, using homogeneous coordinates.



(a) Original figure. (b) Translated to origin by $-\mathbf{p}$. (c) Rotated about the origin. (d) Translated back by **p**.

FIGURE 7 Rotation of figure about point **p**.

2.7 EXERCISES

1. What 3×3 matrix will have the same effect on homogeneous coordinates for $\mathbb{R}^2$ that the shear matrix A has in Example 2?

2. Use matrix multiplication to find the image of the triangle with data matrix $D = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 2 & 3 \end{bmatrix}$ under the transformation that reflects points through the y-axis. Sketch both the original triangle and its image.

In Exercises 3–8, find the 3×3 matrices that produce the described composite 2D transformations, using homogeneous coordinates.

---

[2] See Jan Ozer, "High-Performance Graphics Boards," *PC Magazine* **19**, 1 September 2000, pp. 187–200. Also, "The Ultimate Upgrade Guide: Moving On Up," *PC Magazine* **21**, 29 January 2002, pp. 82–91.

3. Translate by (3, 1), and then rotate 45° about the origin.

4. Translate by (−2, 3), and then scale the x-coordinate by .8 and the y-coordinate by 1.2.

5. Reflect points through the x-axis, and then rotate 30° about the origin.

6. Rotate points 30°, and then reflect through the x-axis.

7. Rotate points through 60° about the point (6, 8).

8. Rotate points through 45° about the point (3, 7).

9. A 2×200 data matrix D contains the coordinates of 200 points. Compute the number of multiplications required to transform these points using two arbitrary 2×2 matrices A and B. Consider the two possibilities $A(BD)$ and $(AB)D$. Discuss the implications of your results for computer graphics calculations.

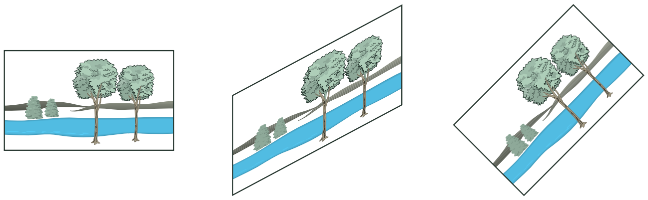
10. Consider the following geometric 2D transformations: D, a dilation (in which x-coordinates and y-coordinates are scaled by the same factor); R, a rotation; and T, a translation. Does D *commute* with R? That is, is $D(R(\mathbf{x})) = R(D(\mathbf{x}))$ for all $\mathbf{x}$ in $\mathbb{R}^2$? Does D commute with T? Does R commute with T?

11. A rotation on a computer screen is sometimes implemented as the product of two shear-and-scale transformations, which can speed up calculations that determine how a graphic image actually appears in terms of screen pixels. (The screen consists of rows and columns of small dots, called *pixels*.) The first transformation A_1 shears vertically and then compresses each column of pixels; the second transformation A_2 shears horizontally and then stretches each row of pixels. Let

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \sec\varphi & -\tan\varphi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Show that the composition of the two transformations is a rotation in $\mathbb{R}^2$.



12. A rotation in $\mathbb{R}^2$ usually requires four multiplications. Compute the product below, and show that the matrix for a rotation can be factored into three shear transformations (each of which requires only one multiplication).

$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ \sin\varphi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

13. The usual transformations on homogeneous coordinates for 2D computer graphics involve 3×3 matrices of the form

$$\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

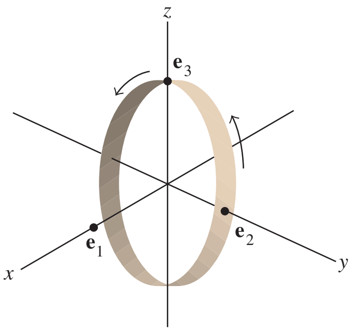
where A is a 2×2 matrix and $\mathbf{p}$ is in $\mathbb{R}^2$. Show that such a transformation amounts to a linear transformation on $\mathbb{R}^2$ followed by a translation. [*Hint:* Find an appropriate matrix factorization involving partitioned matrices.]

14. Show that the transformation in Exercise 7 is equivalent to a rotation about the origin followed by a translation by $\mathbf{p}$. Find $\mathbf{p}$.

15. What vector in $\mathbb{R}^3$ has homogeneous coordinates $(\frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, \frac{1}{24})$?

16. Are (1, −2, 3, 4) and (10, −20, 30, 40) homogeneous coordinates for the same point in $\mathbb{R}^3$? Why or why not?

17. Give the 4×4 matrix that rotates points in $\mathbb{R}^3$ about the x-axis through an angle of 60°. (See the figure.)



18. Give the 4×4 matrix that rotates points in $\mathbb{R}^3$ about the z-axis through an angle of −30°, and then translates by $\mathbf{p} = (5, -2, 1)$.

19. Let S be the triangle with vertices (4.2, 1.2, 4), (6, 4, 2), (2, 2, 6). Find the image of S under the perspective projection with center of projection at (0, 0, 10).

20. Let S be the triangle with vertices $(9, 3, -5)$, $(12, 8, 2)$, $(1.8, 2.7, 1)$. Find the image of S under the perspective projection with center of projection at $(0, 0, 10)$.

Exercises 21 and 22 concern the way in which color is specified for display in computer graphics. A color on a computer screen is encoded by three numbers (R, G, B) that list the amount of energy an electron gun must transmit to red, green, and blue phosphor dots on the computer screen. (A fourth number specifies the luminance or intensity of the color.)

21. **[M]** The actual color a viewer sees on a screen is influenced by the specific type and amount of phosphors on the screen. So each computer screen manufacturer must convert between the (R, G, B) data and an international CIE standard for color, which uses three primary colors, called X, Y, and Z. A typical conversion for short-persistence phosphors is

$$\begin{bmatrix} .61 & .29 & .150 \\ .35 & .59 & .063 \\ .04 & .12 & .787 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

A computer program will send a stream of color information to the screen, using standard CIE data (X, Y, Z). Find the equation that converts these data to the (R, G, B) data needed for the screen's electron gun.

22. **[M]** The signal broadcast by commercial television describes each color by a vector (Y, I, Q). If the screen is black and white, only the Y-coordinate is used. (This gives a better monochrome picture than using CIE data for colors.) The correspondence between *YIQ* and a "standard" *RGB* color is given by

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.528 & .311 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(A screen manufacturer would change the matrix entries to work for its *RGB* screens.) Find the equation that converts the *YIQ* data transmitted by the television station to the *RGB* data needed for the television screen.

SOLUTION TO PRACTICE PROBLEM

Assemble the matrices right-to-left for the three operations. Using $\mathbf{p} = (-2, 6)$, $\cos(-30°) = \sqrt{3}/2$, and $\sin(-30°) = -.5$, we have

$$\underbrace{\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Translate back by } p}\underbrace{\begin{bmatrix} \sqrt{3}/2 & 1/2 & 0 \\ -1/2 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Rotate around the origin}}\underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Translate by } -p}$$

$$= \begin{bmatrix} \sqrt{3}/2 & 1/2 & \sqrt{3} - 5 \\ -1/2 & \sqrt{3}/2 & -3\sqrt{3} + 5 \\ 0 & 0 & 1 \end{bmatrix}$$

2.8 SUBSPACES OF $\mathbb{R}^n$

This section focuses on important sets of vectors in $\mathbb{R}^n$ called *subspaces*. Often subspaces arise in connection with some matrix A, and they provide useful information about the equation $A\mathbf{x} = \mathbf{b}$. The concepts and terminology in this section will be used repeatedly throughout the rest of the book.[1]

---

[1] Sections 2.8 and 2.9 are included here to permit readers to postpone the study of most or all of the next two chapters and to skip directly to Chapter 5, if so desired. *Omit* these two sections if you plan to work through Chapter 4 before Chapter 5.

DEFINITION A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:

a. The zero vector is in H.

b. For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.

c. For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

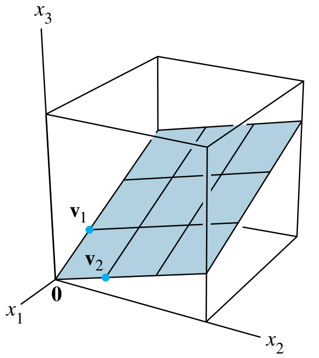


FIGURE 1
Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ as a plane through the origin.

In words, a subspace is *closed* under addition and scalar multiplication. As you will see in the next few examples, most sets of vectors discussed in Chapter 1 are subspaces. For instance, a plane through the origin is the standard way to visualize the subspace in Example 1. See Fig. 1.

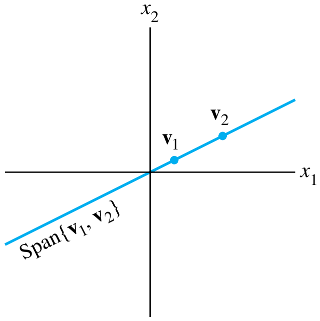
EXAMPLE 1 If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^n$ and $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, then H is a subspace of $\mathbb{R}^n$. To verify this statement, note that the zero vector is in H (because $0\mathbf{v} + 0\mathbf{u}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$). Now take two arbitrary vectors in H, say,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{and} \quad \mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Then

$$\mathbf{u} + \mathbf{v} = (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2$$

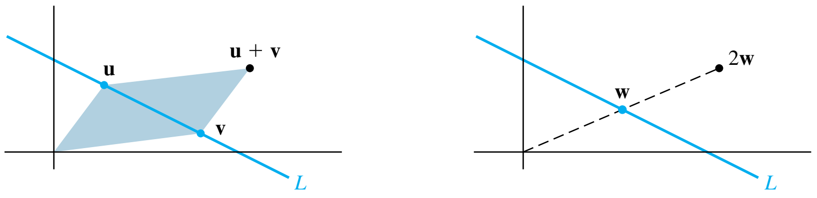
which shows that $\mathbf{u} + \mathbf{v}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ and hence is in H. Also, for any scalar c, the vector $c\mathbf{u}$ is in H, because $c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$.



$\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2 = k\mathbf{v}_1$.

If $\mathbf{v}_1$ is not zero and if $\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$, then $\mathbf{v}_1$ and $\mathbf{v}_2$ simply span a *line* through the origin. So a line through the origin is another example of a subspace.

EXAMPLE 2 A line L *not* through the origin is *not* a subspace, because it does not contain the origin, as required. Also, Fig. 2 shows that L is not closed under addition or scalar multiplication.



$\mathbf{u} + \mathbf{v}$ is not on L

$2\mathbf{w}$ is not on L

FIGURE 2

EXAMPLE 3 For $\mathbf{v}_1, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$, the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is a subspace of $\mathbb{R}^n$. The verification of this statement is similar to the argument given in

Example 1. We shall now refer to Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ as **the subspace spanned** (or **generated**) by $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Note that $\mathbb{R}^n$ is a subspace of itself because it has the three properties required for a subspace. Another special subspace is the set consisting of only the zero vector in $\mathbb{R}^n$. This set, called the **zero subspace**, also satisfies the conditions for a subspace.

Column Space and Null Space of a Matrix

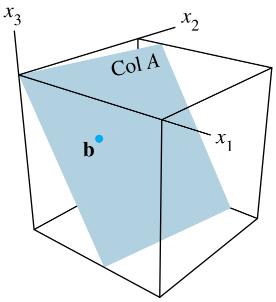
Subspaces of $\mathbb{R}^n$ usually occur in applications and theory in one of two ways. In both cases, the subspace can be related to a matrix.

DEFINITION

The **column space** of a matrix A is the set Col A of all linear combinations of the columns of A.

If $A = [\,\mathbf{a}_1 \;\; \cdots \;\; \mathbf{a}_n\,]$, with the columns in $\mathbb{R}^m$, then Col A is the same as Span $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. Example 3 shows that **the column space of an $m \times n$ matrix is a subspace of $\mathbb{R}^m$.**

EXAMPLE 4 Let $A = \begin{bmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix}$. Determine whether $\mathbf{b}$ is in the column space of A.



Solution The vector $\mathbf{b}$ is a linear combination of the columns of A if and only if $\mathbf{b}$ can be written as $A\mathbf{x}$ for some $\mathbf{x}$, that is, if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution. Row reducing the augmented matrix $[\,A \;\; \mathbf{b}\,]$,

$$\begin{bmatrix} 1 & -3 & -4 & 3 \\ -4 & 6 & -2 & 3 \\ -3 & 7 & 6 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & -2 & -6 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

we conclude that $A\mathbf{x} = \mathbf{b}$ is consistent and $\mathbf{b}$ is in Col A.

The solution of Example 4 shows that when a system of linear equations is written in the form $A\mathbf{x} = \mathbf{b}$, the column space of A is the set of all $\mathbf{b}$ for which the system has a solution.

DEFINITION

The **null space** of a matrix A is the set Nul A of all solutions to the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

When A has n columns, the solutions of $A\mathbf{x} = \mathbf{0}$ belong to $\mathbb{R}^n$, and the null space of A is a subset of $\mathbb{R}^n$. In fact, Nul A has the properties of a *subspace* of $\mathbb{R}^n$.

THEOREM 12 The null space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^n$. Equivalently, the set of all solutions to a system $A\mathbf{x} = \mathbf{0}$ of m homogeneous linear equations in n unknowns is a subspace of $\mathbb{R}^n$.

PROOF The zero vector is in Nul A (because $A\mathbf{0} = \mathbf{0}$). To show that Nul A satisfies the other two properties required for a subspace, take any $\mathbf{u}$ and $\mathbf{v}$ in Nul A. That is, suppose $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$. Then, by a property of matrix multiplication,

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Thus $\mathbf{u} + \mathbf{v}$ satisfies $A\mathbf{x} = \mathbf{0}$, and so $\mathbf{u} + \mathbf{v}$ is in Nul A. Also, for any scalar c, $A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0}$, which shows that $c\mathbf{u}$ is in Nul A. ■

To test whether a given vector $\mathbf{v}$ is in Nul A, just compute $A\mathbf{v}$ to see whether $A\mathbf{v}$ is the zero vector. Because Nul A is described by a condition that must be checked for each vector, we say that the null space is defined *implicitly*. In contrast, the column space is defined *explicitly*, because vectors in Col A can be constructed (by linear combinations) from the columns of A. To create an explicit description of Nul A, solve the equation $A\mathbf{x} = \mathbf{0}$ and write the solution in parametric vector form. (See Example 6, below.)[2]

Basis for a Subspace

Because a subspace typically contains an infinite number of vectors, some problems involving a subspace are handled best by working with a small finite set of vectors that span the subspace. The smaller the set, the better. It can be shown that the smallest possible spanning set must be linearly independent.

DEFINITION A **basis** for a subspace H of $\mathbb{R}^n$ is a linearly independent set in H that spans H.

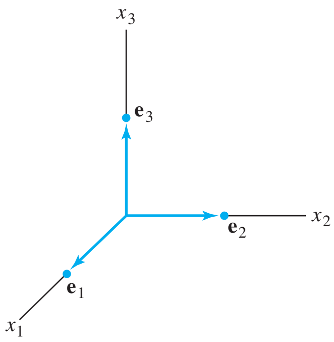


FIGURE 3
The standard basis for $\mathbb{R}^3$.

EXAMPLE 5 The columns of an invertible $n \times n$ matrix form a basis for all of $\mathbb{R}^n$ because they are linearly independent and span $\mathbb{R}^n$, by the Invertible Matrix Theorem. One such matrix is the $n \times n$ identity matrix. Its columns are denoted by $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

The set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is called the **standard basis** for $\mathbb{R}^n$. See Fig. 3. ■

---

[2] The contrast between Nul A and Col A is discussed further in Section 4.2.

The next example shows that the standard procedure for writing the solution set of $A\mathbf{x} = \mathbf{0}$ in parametric vector form actually identifies a basis for Nul A. This fact will be used throughout Chapter 5.

EXAMPLE 6 Find a basis for the null space of the matrix

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

Solution First, write the solution of $A\mathbf{x} = \mathbf{0}$ in parametric vector form:

$$A \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{array}{rcrcrcrcl} x_1 & - & 2x_2 & & & - & x_4 & + & 3x_5 = 0 \\ & & & & x_3 & + & 2x_4 & - & 2x_5 = 0 \\ & & & & & & & & 0 = 0 \end{array}$$

The general solution is $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, with x_2, x_4, and x_5 free.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \mathbf{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \tag{1}$$

Equation (1) shows that Nul A coincides with the set of all linear combinations of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. That is, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ generates Nul A. In fact, this construction of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ automatically makes them linearly independent, because (1) shows that $\mathbf{0} = x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ only if the weights x_2, x_4, and x_5 are all zero. (Examine entries 2, 4, and 5 in the vector $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$.) So $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a *basis* for Nul A.

Finding a basis for the column space of a matrix is actually less work than finding a basis for the null space. However, the method requires some explanation. Let's begin with a simple case.

EXAMPLE 7 Find a basis for the column space of the matrix

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

Solution Denote the columns of B by $\mathbf{b}_1, \ldots, \mathbf{b}_5$ and note that $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$ and $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. The fact that $\mathbf{b}_3$ and $\mathbf{b}_4$ are combinations of the pivot columns means that any combination of $\mathbf{b}_1, \ldots, \mathbf{b}_5$ is actually just a combination of $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_5$. Indeed, if $\mathbf{v}$ is any vector in Col B, say,

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 + c_5\mathbf{b}_5$$

then, substituting for $\mathbf{b}_3$ and $\mathbf{b}_4$, we can write $\mathbf{v}$ in the form

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3(-3\mathbf{b}_1 + 2\mathbf{b}_2) + c_4(5\mathbf{b}_1 - \mathbf{b}_2) + c_5\mathbf{b}_5$$

which is a linear combination of $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_5$. So $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ spans Col B. Also, $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_5$ are linearly independent, because they are columns from an identity matrix. So the pivot columns of B form a basis for Col B.

The matrix B in Example 7 is in reduced echelon form. To handle a general matrix A, recall that linear dependence relations among the columns of A can be expressed in the form $A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x}$. (If some columns are not involved in a particular dependence relation, then the corresponding entries in $\mathbf{x}$ are zero.) When A is row reduced to echelon form B, the columns are drastically changed, but the equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same set of solutions. That is, the columns of A have *exactly the same linear dependence relationships* as the columns of B.

EXAMPLE 8 It can be verified that the matrix

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_5\,] = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$$

is row equivalent to the matrix B in Example 7. Find a basis for Col A.

Solution From Example 7, the pivot columns of A are columns 1, 2, and 5. Also, $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$ and $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. Since row operations do not affect linear dependence relations among the columns of the matrix, we should have

$$\mathbf{a}_3 = -3\mathbf{a}_1 + 2\mathbf{a}_2 \quad \text{and} \quad \mathbf{a}_4 = 5\mathbf{a}_1 - \mathbf{a}_2$$

Check that this is true! By the argument in Example 7, $\mathbf{a}_3$ and $\mathbf{a}_4$ are not needed to generate the column space of A. Also, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ must be linearly independent, because any dependence relation among $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_5$ would imply the same dependence relation among $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_5$. Since $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ is linearly independent, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ is also linearly independent and hence is a basis for Col A.

The argument in Example 8 can be adapted to prove the following theorem.

THEOREM 13 The pivot columns of a matrix A form a basis for the column space of A.

Warning: Be careful to use *pivot columns of A itself* for the basis of Col A. The columns of an echelon form B are often not in the column space of A. (For instance, in Examples 7 and 8, the columns of B all have zeros in their last entries and cannot generate the columns of A.)

PRACTICE PROBLEMS

1. Let $A = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix}$. Is $\mathbf{u}$ in Nul A? Is $\mathbf{u}$ in Col A? Justify each answer.

2. Given $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, find a vector in Nul A and a vector in Col A.

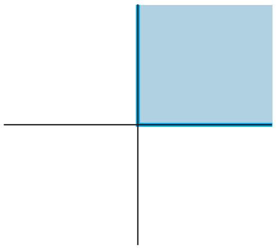
3. Suppose an $n \times n$ matrix A is invertible. What can you say about Col A? About Nul A?

SG Mastering: Subspace, Col A, Nul A, Basis 2–37

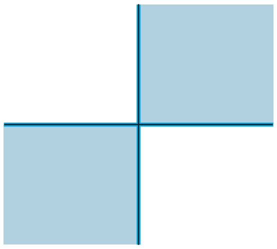
2.8 EXERCISES

Exercises 1–4 display sets in $\mathbb{R}^2$. Assume the sets include the bounding lines. In each case, give a specific reason why the set H is *not* a subspace of $\mathbb{R}^2$. (For instance, find two vectors in H whose sum is *not* in H, or find a vector in H with a scalar multiple that is not in H. Draw a picture.)

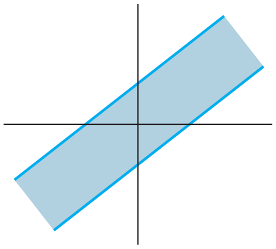
1.



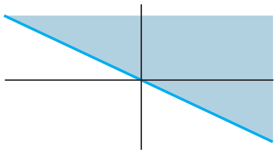
2.



3.



4.



5. Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \\ -5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ -5 \\ 8 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 8 \\ 2 \\ -9 \end{bmatrix}$. Determine if $\mathbf{w}$ is in the subspace of $\mathbb{R}^3$ generated by $\mathbf{v}_1$ and $\mathbf{v}_2$.

6. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ -7 \\ 9 \\ 7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ -8 \\ 6 \\ 5 \end{bmatrix}$, and $\mathbf{u} = \begin{bmatrix} -4 \\ 10 \\ -7 \\ -5 \end{bmatrix}$. Determine if $\mathbf{u}$ is in the subspace of $\mathbb{R}^4$ generated by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

7. Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -3 \\ 8 \\ -7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -4 \\ 6 \\ -7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 6 \\ -10 \\ 11 \end{bmatrix}$, and $A = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \,]$.

a. How many vectors are in $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

b. How many vectors are in Col A?

c. Is $\mathbf{p}$ in Col A? Why or why not?

8. Let $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ -6 \\ 3 \end{bmatrix}$, and $\mathbf{p} = \begin{bmatrix} 1 \\ 14 \\ -9 \end{bmatrix}$. Determine if $\mathbf{p}$ is in Col A, where $A = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \,]$.

9. With A and $\mathbf{p}$ as in Exercise 7, determine if $\mathbf{p}$ is in Nul A.

10. With $\mathbf{u} = (-2, 3, 1)$ and A as in Exercise 8, determine if $\mathbf{u}$ is in Nul A.

In Exercises 11 and 12, give integers p and q such that Nul A is a subspace of $\mathbb{R}^p$ and Col A is a subspace of $\mathbb{R}^q$.

11. $A = \begin{bmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -5 & -1 & 0 \\ 2 & 7 & 11 \end{bmatrix}$

13. For A as in Exercise 11, find a nonzero vector in Nul A and a nonzero vector in Col A.

14. For A as in Exercise 12, find a nonzero vector in Nul A and a nonzero vector in Col A.

Determine which sets in Exercises 15–20 are bases for $\mathbb{R}^2$ or $\mathbb{R}^3$. Justify each answer.

15. $\begin{bmatrix} 5 \\ -2 \end{bmatrix}, \begin{bmatrix} 10 \\ -3 \end{bmatrix}$

16. $\begin{bmatrix} -4 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

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

18. $\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -5 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

19. $\begin{bmatrix} 3 \\ -8 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -5 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ -6 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 7 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 8 \\ 9 \end{bmatrix}$

In Exercises 21 and 22, mark each statement True or False. Justify each answer.

21. a. A subspace of $\mathbb{R}^n$ is any set H such that (i) the zero vector is in H, (ii) $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} + \mathbf{v}$ are in H, and (iii) c is a scalar and $c\mathbf{u}$ is in H.

b. If $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is the same as the column space of the matrix $[\,\mathbf{v}_1 \ \cdots \ \mathbf{v}_p\,]$.

c. The set of all solutions of a system of m homogeneous equations in n unknowns is a subspace of $\mathbb{R}^m$.

d. The columns of an invertible $n \times n$ matrix form a basis for $\mathbb{R}^n$.

e. Row operations do not affect linear dependence relations among the columns of a matrix.

22. a. A subset H of $\mathbb{R}^n$ is a subspace if the zero vector is in H.

b. Given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ in $\mathbb{R}^n$, the set of all linear combinations of these vectors is a subspace of $\mathbb{R}^n$.

c. The null space of an $m \times n$ matrix is a subspace of $\mathbb{R}^n$.

d. The column space of a matrix A is the set of solutions of $A\mathbf{x} = \mathbf{b}$.

e. If B is an echelon form of a matrix A, then the pivot columns of B form a basis for Col A.

Exercises 23–26 display a matrix A and an echelon form of A. Find a basis for Col A and a basis for Nul A.

23. $A = \begin{bmatrix} 4 & 5 & 9 & -2 \\ 6 & 5 & 1 & 12 \\ 3 & 4 & 8 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 6 & -5 \\ 0 & 1 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

24. $A = \begin{bmatrix} -3 & 9 & -2 & -7 \\ 2 & -6 & 4 & 8 \\ 3 & -9 & -2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 6 & 9 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 1 & 4 & 8 & -3 & -7 \\ -1 & 2 & 7 & 3 & 4 \\ -2 & 2 & 9 & 5 & 5 \\ 3 & 6 & 9 & -5 & -2 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 4 & 8 & 0 & 5 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

26. $A = \begin{bmatrix} 3 & -1 & 7 & 3 & 9 \\ -2 & 2 & -2 & 7 & 5 \\ -5 & 9 & 3 & 3 & 4 \\ -2 & 6 & 6 & 3 & 7 \end{bmatrix}$

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

27. Construct a 3×3 matrix A and a nonzero vector $\mathbf{b}$ such that $\mathbf{b}$ is in Col A, but $\mathbf{b}$ is not the same as any one of the columns of A.

28. Construct a 3×3 matrix A and a vector $\mathbf{b}$ such that $\mathbf{b}$ is *not* in Col A.

29. Construct a nonzero 3×3 matrix A and a nonzero vector $\mathbf{b}$ such that $\mathbf{b}$ is in Nul A.

30. Suppose the columns of a matrix $A = [\mathbf{a}_1 \cdots \mathbf{a}_p]$ are linearly independent. Explain why $\{\mathbf{a}_1, \ldots, \mathbf{a}_p\}$ is a basis for Col A.

In Exercises 31–36, respond as comprehensively as possible, and justify your answer.

31. Suppose F is a 5×5 matrix whose column space is not equal to $\mathbb{R}^5$. What can you say about Nul F?

32. If R is a 6×6 matrix and Nul R is *not* the zero subspace, what can you say about Col R?

33. If Q is a 4×4 matrix and Col $Q = \mathbb{R}^4$, what can you say about solutions of equations of the form $Q\mathbf{x} = \mathbf{b}$ for $\mathbf{b}$ in $\mathbb{R}^4$?

34. If P is a 5×5 matrix and Nul P is the zero subspace, what can you say about solutions of equations of the form $P\mathbf{x} = \mathbf{b}$ for $\mathbf{b}$ in $\mathbb{R}^5$?

35. What can you say about Nul B when B is a 5×4 matrix with linearly independent columns?

36. What can you say about the shape of an $m \times n$ matrix A when the columns of A form a basis for $\mathbb{R}^m$?

[M] In Exercises 37 and 38, construct bases for the column space and the null space of the given matrix A. Justify your work.

37. $A = \begin{bmatrix} 3 & -5 & 0 & -1 & 3 \\ -7 & 9 & -4 & 9 & -11 \\ -5 & 7 & -2 & 5 & -7 \\ 3 & -7 & -3 & 4 & 0 \end{bmatrix}$

38. $A = \begin{bmatrix} 5 & 2 & 0 & -8 & -8 \\ 4 & 1 & 2 & -8 & -9 \\ 5 & 1 & 3 & 5 & 19 \\ -8 & -5 & 6 & 8 & 5 \end{bmatrix}$

CD Column Space and Null Space

CD A Basis for Col A

SOLUTIONS TO PRACTICE PROBLEMS

1. To determine whether $\mathbf{u}$ is in Nul A, simply compute

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

The result shows that $\mathbf{u}$ is in Nul A. Deciding whether $\mathbf{u}$ is in Col A requires more work. Reduce the augmented matrix $[A \quad \mathbf{u}]$ to echelon form to determine whether the equation $A\mathbf{x} = \mathbf{u}$ is consistent:

$$\begin{bmatrix} 1 & -1 & 5 & -7 \\ 2 & 0 & 7 & 3 \\ -3 & -5 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & -8 & 12 & -19 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

The equation $A\mathbf{x} = \mathbf{u}$ has no solution, so $\mathbf{u}$ is not in Col A.

2. In contrast to Practice Problem 1, finding a vector in Nul A requires more work than testing whether a specified vector is in Nul A. However, since A is already in reduced echelon form, the equation $A\mathbf{x} = \mathbf{0}$ shows that if $\mathbf{x} = (x_1, x_2, x_3)$, then $x_2 = 0$, $x_3 = 0$, and x_1 is a free variable. Thus, a basis for Nul A is $\mathbf{v} = (1, 0, 0)$. Finding just one vector in Col A is trivial, since each column of A is in Col A. In this particular case, the same vector $\mathbf{v}$ is in both Nul A *and* Col A. For most $n \times n$ matrices, the zero vector of $\mathbb{R}^n$ is the only vector in both Nul A and Col A.

3. If A is invertible, then the columns of A span $\mathbb{R}^n$, by the Invertible Matrix Theorem. By definition, the columns of any matrix always span the column space, so in this case Col A is all of $\mathbb{R}^n$. In symbols, Col $A = \mathbb{R}^n$. Also, since A is invertible, the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. This means that Nul A is the zero subspace. In symbols, Nul $A = \{\mathbf{0}\}$.

2.9 DIMENSION AND RANK

This section continues the discussion of subspaces and bases for subspaces, beginning with the concept of a coordinate system. The definition and example below should make a useful new term, *dimension,* seem quite natural, at least for subspaces of $\mathbb{R}^3$.

Coordinate Systems

The main reason for selecting a basis for a subspace H, instead of merely a spanning set, is that each vector in H can be written in only one way as a linear combination of the basis vectors. To see why, suppose $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ is a basis for H, and suppose a vector $\mathbf{x}$ in H can be generated in two ways, say,

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p \quad \text{and} \quad \mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_p\mathbf{b}_p \tag{1}$$

Then, subtracting gives

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_p - d_p)\mathbf{b}_p \tag{2}$$

Since $\mathcal{B}$ is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \le j \le p$, which shows that the two representations in (1) are actually the same.

DEFINITION

Suppose the set $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ is a basis for a subspace H. For each $\mathbf{x}$ in H, the **coordinates of x relative to the basis** $\mathcal{B}$ are the weights $c_1, \ldots, c_p$ such that $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p$, and the vector in $\mathbb{R}^p$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

is called the **coordinate vector of x (relative to** $\mathcal{B}$**)** or the $\mathcal{B}$**-coordinate vector of x.**[1]

EXAMPLE 1 Let $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Then $\mathcal{B}$ is a basis for $H = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ because $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. Determine if $\mathbf{x}$ is in H, and if it is, find the coordinate vector of $\mathbf{x}$ relative to $\mathcal{B}$.

Solution If $\mathbf{x}$ is in H, then the following vector equation is consistent:

$$c_1\begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

---

[1] It is important that the elements of $\mathcal{B}$ are numbered because the entries in $[\mathbf{x}]_{\mathcal{B}}$ depend on the order of the vectors in $\mathcal{B}$.

The scalars c_1 and c_2, if they exist, are the $\mathcal{B}$-coordinates of $\mathbf{x}$. Row operations show that

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

Thus $c_1 = 2$, $c_2 = 3$, and $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The basis $\mathcal{B}$ determines a "coordinate system" on H, which can be visualized by the grid shown in Fig. 1.

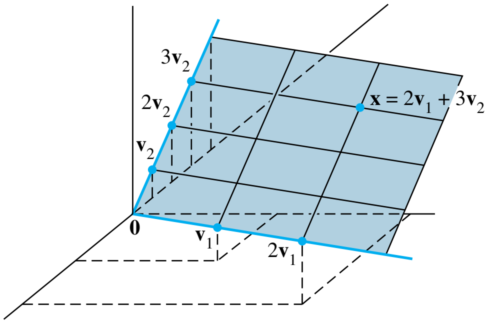


FIGURE 1 A coordinate system on a plane H in $\mathbb{R}^3$.

Notice that although points in H are also in $\mathbb{R}^3$, they are completely determined by their coordinate vectors, which belong to $\mathbb{R}^2$. The grid on the plane in Fig. 1 makes H "look" like $\mathbb{R}^2$. The correspondence $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$ is a one-to-one correspondence between H and $\mathbb{R}^2$ that preserves linear combinations. We call such a correspondence an *isomorphism*, and we say that H is *isomorphic* to $\mathbb{R}^2$.

In general, if $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for H, then the mapping $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is a one-to-one correspondence that makes H look and act the same as $\mathbb{R}^p$ (even though the vectors in H themselves may have more than p entries). (Section 4.4 has more details.)

The Dimension of a Subspace

It can be shown that if a subspace H has a basis of p vectors, then every basis of H must consist of exactly p vectors. (See Exercises 27 and 28.) Thus the following definition makes sense.

DEFINITION

The **dimension** of a nonzero subspace H, denoted by dim H, is the number of vectors in any basis for H. The dimension of the zero subspace $\{\mathbf{0}\}$ is defined to be zero.[2]

---

[2] The zero subspace has *no* basis (because the zero vector by itself forms a linearly dependent set).

The space $\mathbb{R}^n$ has dimension n. Every basis for $\mathbb{R}^n$ consists of n vectors. A plane through $\mathbf{0}$ in $\mathbb{R}^3$ is two-dimensional, and a line through $\mathbf{0}$ is one-dimensional.

EXAMPLE 2 Recall that the null space of the matrix A in Example 6 of Section 2.8 had a basis of 3 vectors. So the dimension of Nul A in this case is 3. Observe how each basis vector corresponds to a free variable in the equation $A\mathbf{x} = \mathbf{0}$. Our construction always produces a basis in this way. So, to find the dimension of Nul A, simply identify and count the number of free variables in $A\mathbf{x} = \mathbf{0}$.

DEFINITION The **rank** of a matrix A, denoted by rank A, is the dimension of the column space of A.

Since the pivot columns of A form a basis for Col A, the rank of A is just the number of pivot columns in A.

EXAMPLE 3 Determine the rank of the matrix

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

Solution Reduce A to echelon form:

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{array}{c} \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{ccccc} \uparrow & \uparrow & & \uparrow & \end{array} \end{array}$$

Pivot columns

The matrix A has 3 pivot columns, so rank $A = 3$.

The row reduction in Example 3 reveals that there are two free variables in $A\mathbf{x} = \mathbf{0}$, because two of the five columns of A are *not* pivot columns. (The nonpivot columns correspond to the free variables in $A\mathbf{x} = \mathbf{0}$.) Since the number of pivot columns plus the number of nonpivot columns is exactly the number of columns, the dimensions of Col A and Nul A have the following useful connection. (See the Rank Theorem in Section 4.6 for additional details.)

THEOREM 14 **The Rank Theorem**

If a matrix A has n columns, then rank A + dim Nul $A = n$.

The following theorem is important for applications and will be needed in Chapters 5 and 6. The theorem (proved in Section 4.5) is certainly plausible, if you think of a

p-dimensional subspace as isomorphic to $\mathbb{R}^p$. The Invertible Matrix Theorem shows that p vectors in $\mathbb{R}^p$ are linearly independent if and only if they also span $\mathbb{R}^p$.

THEOREM 15

The Basis Theorem

Let H be a p-dimensional subspace of $\mathbb{R}^n$. Any linearly independent set of exactly p elements in H is automatically a basis for H. Also, any set of p elements of H that spans H is automatically a basis for H.

Rank and the Invertible Matrix Theorem

The various vector space concepts associated with a matrix provide several more statements for the Invertible Matrix Theorem. They are presented below to follow the statements in the original theorem in Section 2.3.

THEOREM

The Invertible Matrix Theorem (continued)

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

m. The columns of A form a basis of $\mathbb{R}^n$.
n. $\text{Col } A = \mathbb{R}^n$
o. $\dim \text{Col } A = n$
p. $\text{rank } A = n$
q. $\text{Nul } A = \{\mathbf{0}\}$
r. $\dim \text{Nul } A = 0$

PROOF Statement (m) is logically equivalent to statements (e) and (h) regarding linear independence and spanning. The other five statements are linked to the earlier ones of the theorem by the following chain of almost trivial implications:

$$(g) \Rightarrow (n) \Rightarrow (o) \Rightarrow (p) \Rightarrow (r) \Rightarrow (q) \Rightarrow (d)$$

Statement (g), which says that the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$, implies (n), because Col A is precisely the set of all $\mathbf{b}$ such that the equation $A\mathbf{x} = \mathbf{b}$ is consistent. The implications (n) $\Rightarrow$ (o) $\Rightarrow$ (p) follow from the definitions of *dimension* and *rank*. If the rank of A is n, the number of columns of A, then $\dim \text{Nul } A = 0$, by the Rank Theorem, and so $\text{Nul } A = \{\mathbf{0}\}$. Thus (p) $\Rightarrow$ (r) $\Rightarrow$ (q). Also, (q) implies that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, which is statement (d). Since statements (d) and (g) are already known to be equivalent to the statement that A is invertible, the proof is complete. ■

SG Expanded Table for the IMT 2–39

NUMERICAL NOTES

Many algorithms discussed in this text are useful for understanding concepts and making simple computations by hand. However, the algorithms are often unsuitable for large-scale problems in real life.

Rank determination is a good example. It would seem easy to reduce a matrix to echelon form and count the pivots. But unless exact arithmetic is performed on a matrix whose entries are specified exactly, row operations can change the apparent rank of a matrix. For instance, if the value of x in the matrix $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ is not stored exactly as 7 in a computer, then the rank may be 1 or 2, depending on whether the computer treats $x - 7$ as zero.

In practical applications, the effective rank of a matrix A is often determined from the singular value decomposition of A, to be discussed in Section 7.4.

CD The `rank` command

PRACTICE PROBLEMS

1. Determine the dimension of the subspace H of $\mathbb{R}^3$ spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. (First, find a basis for H.)

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 6 \\ -7 \end{bmatrix}$$

2. Consider the basis $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ .2 \end{bmatrix}, \begin{bmatrix} .2 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{R}^2$. If $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, what is $\mathbf{x}$?

3. Could $\mathbb{R}^3$ possibly contain a four-dimensional subspace? Explain.

2.9 EXERCISES

In Exercises 1 and 2, find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ and the given basis $\mathcal{B}$. Illustrate your answer with a figure, as in the solution of Practice Problem 2.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

In Exercises 3–6, the vector $\mathbf{x}$ is in a subspace H with a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$.

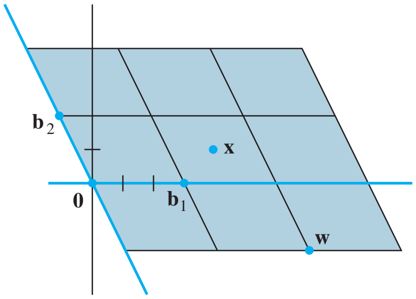
3. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$

4. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -7 \\ 5 \end{bmatrix}$

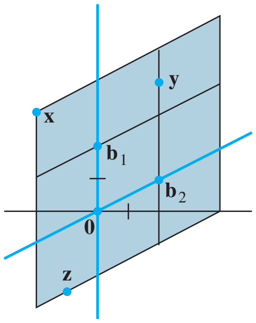
5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ -7 \\ 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 10 \\ -7 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 7 \\ 5 \\ -6 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 11 \\ 0 \\ 7 \end{bmatrix}$

7. Let $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use the figure to estimate $[\mathbf{w}]_{\mathcal{B}}$ and $[\mathbf{x}]_{\mathcal{B}}$. Confirm your estimate of $[\mathbf{x}]_{\mathcal{B}}$ by using it and $\{\mathbf{b}_1, \mathbf{b}_2\}$ to compute $\mathbf{x}$.



8. Let $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use the figure to estimate $[\mathbf{x}]_{\mathcal{B}}$, $[\mathbf{y}]_{\mathcal{B}}$, and $[\mathbf{z}]_{\mathcal{B}}$. Confirm your estimates of $[\mathbf{y}]_{\mathcal{B}}$ and $[\mathbf{z}]_{\mathcal{B}}$ by using them and $\{\mathbf{b}_1, \mathbf{b}_2\}$ to compute $\mathbf{y}$ and $\mathbf{z}$.



Exercises 9–12 display a matrix A and an echelon form of A. Find bases for Col A and Nul A, and then state the dimensions of these subspaces.

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

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

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

11. $A = \begin{bmatrix} 1 & 2 & -5 & 0 & -1 \\ 2 & 5 & -8 & 4 & 3 \\ -3 & -9 & 9 & -7 & -2 \\ 3 & 10 & -7 & 11 & 7 \end{bmatrix}$

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

12. $A = \begin{bmatrix} 1 & 2 & -4 & 3 & 3 \\ 5 & 10 & -9 & -7 & 8 \\ 4 & 8 & -9 & -2 & 7 \\ -2 & -4 & 5 & 0 & -6 \end{bmatrix}$

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

In Exercises 13 and 14, find a basis for the subspace spanned by the given vectors. What is the dimension of the subspace?

13. $\begin{bmatrix} 1 \\ -3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \\ -6 \\ 12 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -3 \\ 7 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -6 \\ 8 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ -7 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ -8 \\ 9 \\ -5 \end{bmatrix}$

15. Suppose a 3×5 matrix A has three pivot columns. Is Col $A = \mathbb{R}^3$? Is Nul $A = \mathbb{R}^2$? Explain your answers.

16. Suppose a 4×7 matrix A has three pivot columns. Is Col $A = \mathbb{R}^3$? What is the dimension of Nul A? Explain your answers.

In Exercises 17 and 18, mark each statement True or False. Justify each answer. Here A is an $m \times n$ matrix.

17. a. If $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is a basis for a subspace H and if $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, then $c_1, \ldots, c_p$ are the coordinates of $\mathbf{x}$ relative to the basis $\mathcal{B}$.

b. Each line in $\mathbb{R}^n$ is a one-dimensional subspace of $\mathbb{R}^n$.

c. The dimension of Col A is the number of pivot columns of A.

d. The dimensions of Col A and Nul A add up to the number of columns of A.

   e. If a set of p vectors spans a p-dimensional subspace H of $\mathbb{R}^n$, then these vectors form a basis for H.

18. a. If $\mathcal{B}$ is a basis for a subspace H, then each vector in H can be written in only one way as a linear combination of the vectors in $\mathcal{B}$.

   b. If $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a basis for a subspace H of $\mathbb{R}^n$, then the correspondence $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ makes H look and act the same as $\mathbb{R}^p$.

   c. The dimension of Nul A is the number of variables in the equation $A\mathbf{x} = \mathbf{0}$.

   d. The dimension of the column space of A is rank A.

   e. If H is a p-dimensional subspace of $\mathbb{R}^n$, then a linearly independent set of p vectors in H is a basis for H.

In Exercises 19–24, justify each answer or construction.

19. If the subspace of all solutions of $A\mathbf{x} = \mathbf{0}$ has a basis consisting of three vectors and if A is a 5×7 matrix, what is the rank of A?

20. What is the rank of a 4×5 matrix whose null space is three-dimensional?

21. If the rank of a 7×6 matrix A is 4, what is the dimension of the solution space of $A\mathbf{x} = \mathbf{0}$?

22. Show that a set $\{\mathbf{v}_1, \dots, \mathbf{v}_5\}$ in $\mathbb{R}^n$ is linearly dependent if $\dim \mathrm{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_5\} = 4$.

23. If possible, construct a 3×4 matrix A such that dim Nul $A = 2$ and dim Col $A = 2$.

24. Construct a 4×3 matrix with rank 1.

25. Let A be an $n \times p$ matrix whose column space is p-dimensional. Explain why the columns of A must be linearly independent.

26. Suppose columns 1, 3, 5, and 6 of a matrix A are linearly independent (but are not necessarily pivot columns) and the rank of A is 4. Explain why the four columns mentioned must be a basis for the column space of A.

27. Suppose vectors $\mathbf{b}_1, \dots, \mathbf{b}_p$ span a subspace W, and let $\{\mathbf{a}_1, \dots, \mathbf{a}_q\}$ be any set in W containing more than p vectors. Fill in the details of the following argument to show that $\{\mathbf{a}_1, \dots, \mathbf{a}_q\}$ must be linearly dependent. First, let $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$ and $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_q]$.

   a. Explain why for each vector $\mathbf{a}_j$, there exists a vector $\mathbf{c}_j$ in $\mathbb{R}^p$ such that $\mathbf{a}_j = B\mathbf{c}_j$.

   b. Let $C = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_q]$. Explain why there is a nonzero vector $\mathbf{u}$ such that $C\mathbf{u} = \mathbf{0}$.

   c. Use B and C to show that $A\mathbf{u} = \mathbf{0}$. This shows that the columns of A are linearly dependent.

28. Use Exercise 27 to show that if $\mathcal{A}$ and $\mathcal{B}$ are bases for a subspace W of $\mathbb{R}^n$, then $\mathcal{A}$ cannot contain more vectors than $\mathcal{B}$, and, conversely, $\mathcal{B}$ cannot contain more vectors than $\mathcal{A}$.

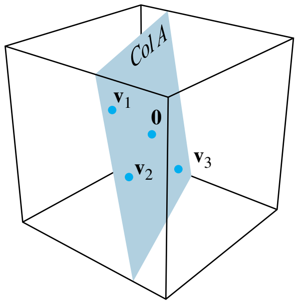
29. **[M]** Let $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Show that $\mathbf{x}$ is in H, and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, when

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

30. **[M]** Let $H = \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Show that $\mathcal{B}$ is a basis for H and $\mathbf{x}$ is in H, and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, when

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

SG Mastering: Dimension and Rank 2–41

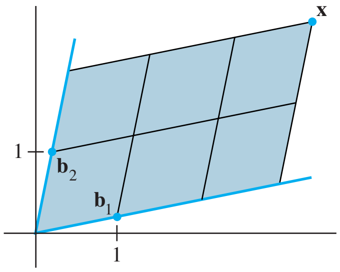


SOLUTIONS TO PRACTICE PROBLEMS

1. Construct $A = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \mathbf{v}_3]$ so that the subspace spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is the column space of A. A basis for this space is provided by the pivot columns of A.

$$A = \begin{bmatrix} 2 & 3 & -1 \\ -8 & -7 & 6 \\ 6 & -1 & -7 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & -10 & -4 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

The first two columns of A are pivot columns and form a basis for H. Thus $\dim H = 2$.



2. If $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, then **x** is formed from a linear combination of the basis vectors using weights 3 and 2:

$$\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ .2 \end{bmatrix} + 2\begin{bmatrix} .2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 2.6 \end{bmatrix}$$

The basis $\{\mathbf{b}_1, \mathbf{b}_2\}$ determines a *coordinate system* for $\mathbb{R}^2$, illustrated by the grid in the figure. Note how **x** is 3 units in the $\mathbf{b}_1$-direction and 2 units in the $\mathbf{b}_2$-direction.

3. A four-dimensional subspace would contain a basis of four linearly independent vectors. This is impossible inside $\mathbb{R}^3$. Since any linearly independent set in $\mathbb{R}^3$ has no more than three vectors, any subspace of $\mathbb{R}^3$ has dimension no more than 3. The space $\mathbb{R}^3$ itself is the only three-dimensional subspace of $\mathbb{R}^3$. Other subspaces of $\mathbb{R}^3$ have dimension 2, 1, or 0.

CHAPTER 2 SUPPLEMENTARY EXERCISES

1. Assume that the matrices mentioned in the statements below have appropriate sizes. Mark each statement True or False. Justify each answer.

a. If A and B are $m \times n$, then both AB^T and A^TB are defined.

b. If $AB = C$ and C has 2 columns, then A has 2 columns.

c. Left-multiplying a matrix B by a diagonal matrix A, with nonzero entries on the diagonal, scales the rows of B.

d. If $BC = BD$, then $C = D$.

e. If $AC = 0$, then either $A = 0$ or $C = 0$.

f. If A and B are $n \times n$, then $(A + B)(A - B) = A^2 - B^2$.

g. An elementary $n \times n$ matrix has either n or $n + 1$ nonzero entries.

h. The transpose of an elementary matrix is an elementary matrix.

i. An elementary matrix must be square.

j. Every square matrix is a product of elementary matrices.

k. If A is a 3×3 matrix with three pivot positions, there exist elementary matrices $E_1, \ldots, E_p$ such that $E_p \cdots E_1 A = I$.

l. If $AB = I$, then A is invertible.

m. If A and B are square and invertible, then AB is invertible, and $(AB)^{-1} = A^{-1}B^{-1}$.

n. If $AB = BA$ and if A is invertible, then $A^{-1}B = BA^{-1}$.

o. If A is invertible and if $r \neq 0$, then $(rA)^{-1} = rA^{-1}$.

p. If A is a 3×3 matrix and the equation $A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has a unique solution, then A is invertible.

2. Find the matrix C whose inverse is $C^{-1} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$.

3. Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Show that $A^3 = 0$. Use matrix algebra to compute the product $(I - A)(I + A + A^2)$.

4. Suppose $A^n = 0$ for some $n > 1$. Find an inverse for $I - A$.

5. Suppose an $n \times n$ matrix A satisfies the equation $A^2 - 2A + I = 0$. Show that $A^3 = 3A - 2I$ and $A^4 = 4A - 3I$.

6. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. These are *Pauli spin matrices* used in the study of electron spin in quantum mechanics. Show that $A^2 = I$, $B^2 = I$, and $AB = -BA$. Matrices such that $AB = -BA$ are said to *anticommute*.

7. Let $A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 4 & 11 \\ 1 & 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 5 \\ 1 & 5 \\ 3 & 4 \end{bmatrix}$. Compute $A^{-1}B$ without computing A^{-1}. [*Hint*: $A^{-1}B$ is the solution of the equation $AX = B$.]

8. Find a matrix A such that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ into $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectively. [*Hint*: Write a matrix equation involving A, and solve for A.]

9. Suppose $AB = \begin{bmatrix} 5 & 4 \\ -2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$. Find A.

10. Suppose A is invertible. Explain why A^TA is also invertible. Then show that $A^{-1} = (A^TA)^{-1}A^T$.

11. Let $x_1, \dots, x_n$ be fixed numbers. The matrix below, called a *Vandermonde matrix,* occurs in applications such as signal processing, error-correcting codes, and polynomial interpolation.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Given $\mathbf{y} = (y_1, \dots, y_n)$ in $\mathbb{R}^n$, suppose $\mathbf{c} = (c_0, \dots, c_{n-1})$ in $\mathbb{R}^n$ satisfies $V\mathbf{c} = \mathbf{y}$, and define the polynomial

$$p(t) = c_0 + c_1t + c_2t^2 + \cdots + c_{n-1}t^{n-1}$$

a. Show that $p(x_1) = y_1, \dots, p(x_n) = y_n$. We call $p(t)$ an *interpolating polynomial for the points* $(x_1, y_1), \dots, (x_n, y_n)$ because the graph of $p(t)$ passes through the points.

b. Suppose $x_1, \dots, x_n$ are distinct numbers. Show that the columns of V are linearly independent. [*Hint*: How many zeros can a polynomial of degree $n-1$ have?]

c. Prove: "If $x_1, \dots, x_n$ are distinct numbers, and $y_1, \dots, y_n$ are arbitrary numbers, then there is an interpolating polynomial of degree $\le n-1$ for $(x_1, y_1), \dots, (x_n, y_n)$."

12. Let $A = LU$, where L is an invertible lower triangular matrix and U is upper triangular. Explain why the first column of A is a multiple of the first column of L. How is the second column of A related to the columns of L?

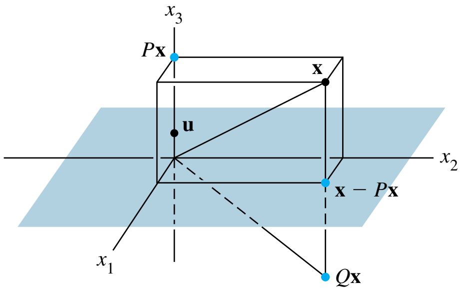
13. Given $\mathbf{u}$ in $\mathbb{R}^n$ with $\mathbf{u}^T\mathbf{u} = 1$, let $P = \mathbf{u}\mathbf{u}^T$ (an outer product) and $Q = I - 2P$. Justify statements (a), (b), and (c).

a. $P^2 = P$ b. $P^T = P$ c. $Q^2 = I$

The transformation $\mathbf{x} \mapsto P\mathbf{x}$ is called a *projection,* and $\mathbf{x} \mapsto Q\mathbf{x}$ is called a *Householder reflection.* Such reflections are used in computer programs to create multiple zeros in a vector (usually a column of a matrix).

CD Lagrange Interpolation

14. Let $\mathbf{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$. Determine P and Q as in Exercise 13, and compute $P\mathbf{x}$ and $Q\mathbf{x}$. The figure shows that $Q\mathbf{x}$ is the reflection of $\mathbf{x}$ through the x_1x_2-plane.



A Householder reflection through the plane $x_3 = 0$.

15. Suppose $C = E_3E_2E_1B$, where E_1, E_2, E_3 are elementary matrices. Explain why C is row equivalent to B.

16. Let A be an $n \times n$ singular matrix. Describe how to construct an $n \times n$ nonzero matrix B such that $AB = 0$.

17. Let A be a 6×4 matrix and B a 4×6 matrix. Show that the 6×6 matrix AB cannot be invertible.

18. Suppose A is a 5×3 matrix and there exists a 3×5 matrix C such that $CA = I_3$. Suppose further that for some given $\mathbf{b}$ in $\mathbb{R}^5$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution. Show that this solution is unique.

19. **[M]** Certain dynamical systems can be studied by examining powers of a matrix, such as those below. Determine what happens to A^k and B^k as k increases (for example, try $k = 2, \dots, 16$). Try to identify what is special about A and B. Investigate large powers of other matrices of this type, and make a conjecture about such matrices.

$$A = \begin{bmatrix} .4 & .2 & .3 \\ .3 & .6 & .3 \\ .3 & .2 & .4 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & .2 & .3 \\ .1 & .6 & .3 \\ .9 & .2 & .4 \end{bmatrix}$$

20. **[M]** Let A_n be the $n \times n$ matrix with 0's on the main diagonal and 1's elsewhere. Compute A_n^{-1} for $n = 4, 5,$ and 6, and make a conjecture about the general form of A_n^{-1} for larger values of n.