Pearson New International Edition

Elementary Linear Algebra
A Matrix Approach

L. Spence A. Insel S. Friedberg

Second Edition

Pearson New International Edition

Elementary Linear Algebra
A Matrix Approach
L. Spence A. Insel S. Friedberg
Second Edition

PEARSON

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

PEARSON

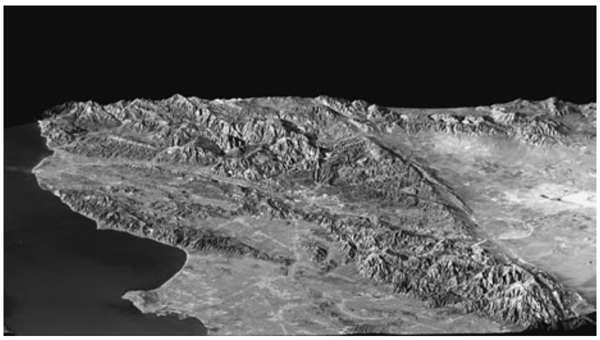
ISBN 10: 1-292-02503-4
ISBN 13: 978-1-292-02503-2

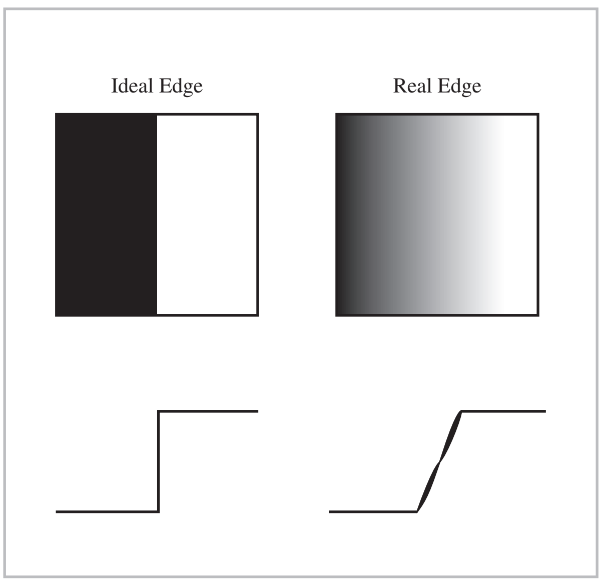
British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

1 INTRODUCTION





For computers to process digital images, whether satellite photos or x-rays, there is a need to recognize the edges of objects. Image edges, which are rapid changes or discontinuities in image intensity, reflect a boundary between dissimilar regions in an image and thus are important basic characteristics of an image. They often indicate the physical extent of objects in the image or a boundary between light and shadow on a single surface or other regions of interest.

The lowermost two figures at the left indicate the changes in image intensity of the ideal and real edges above, when moving from right to left. We see that real intensities can change rapidly, but not instantaneously. In principle, the edge may be found by looking for very large changes over small distances.

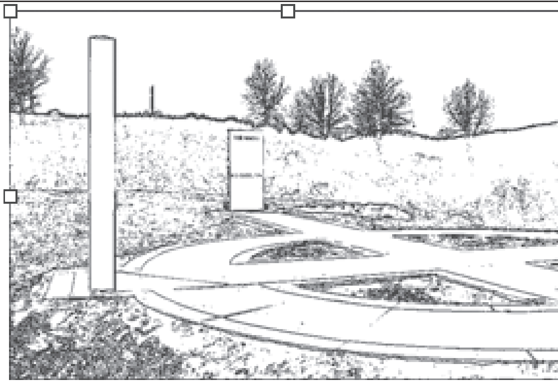
However, a digital image is discrete rather than continuous: it is a matrix of nonnegative entries that provide numerical descriptions of the shades of gray for the pixels in the image, where the entries vary from 0 for a white pixel to 1 for a black pixel. An analysis must be done using the discrete analog of the derivative to measure the rate of change of image intensity in two directions.

From Chapter 1 of *Elementary Linear Algebra*, Second Edition. Lawrence E. Spence, Arnold J. Insel, Stephen H. Friedberg.

The *Sobel* matrices, $S_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$ and $S_2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$ provide a method for measuring these intensity changes. Apply the Sobel matrices S_1 and S_2 in turn to the 3x3 subimage centered on each pixel in the original image. The results are the changes of intensity near the pixel in the horizontal and the vertical directions, respectively. The ordered pair of numbers that are obtained is a vector in the plane that provides the direction and magnitude of the intensity change at the pixel. This vector may be thought of as the discrete analog of the gradient vector of a function of two variables studied in calculus.

Replace each of the original pixel values by the lengths of these vectors, and choose an appropriate threshold value. The final image, called the *thresholded image*, is obtained by changing to black every pixel for which the length of the vector is greater than the threshold value, and changing to white all the other pixels. (See the images below.)

Original Image



Thresholded Image

Notice how the edges are emphasized in the thresholded image. In regions where image intensity is constant, these vectors have length zero, and hence the corresponding regions appear white in the thresholded image. Likewise, a rapid change in image intensity, which occurs at an edge of an object, results in a relatively dark colored boundary in the thresholded image.

CHAPTER 1

MATRICES, VECTORS, AND SYSTEMS OF LINEAR EQUATIONS

The most common use of linear algebra is to solve systems of linear equations, which arise in applications to such diverse disciplines as physics, biology, economics, engineering, and sociology. In this chapter, we describe the most efficient algorithm for solving systems of linear equations, *Gaussian elimination.* This algorithm, or some variation of it, is used by most mathematics software (such as MATLAB).

We can write systems of linear equations compactly, using arrays called *matrices* and *vectors*. More importantly, the arithmetic properties of these arrays enable us to compute solutions of such systems or to determine if no solutions exist. This chapter begins by developing the basic properties of matrices and vectors. In Sections 1.3 and 1.4, we begin our study of systems of linear equations. In Sections 1.6 and 1.7, we introduce two other important concepts of vectors, namely, generating sets and linear independence, which provide information about the existence and uniqueness of solutions of a system of linear equations.

1.1 MATRICES AND VECTORS

Many types of numerical data are best displayed in two-dimensional arrays, such as tables.

For example, suppose that a company owns two bookstores, each of which sells newspapers, magazines, and books. Assume that the sales (in hundreds of dollars) of the two bookstores for the months of July and August are represented by the following tables:

	July	
Store	1	2
Newspapers	6	8
Magazines	15	20
Books	45	64

and

	August	
Store	1	2
Newspapers	7	9
Magazines	18	31
Books	52	68

The first column of the July table shows that store 1 sold \$1500 worth of magazines and \$4500 worth of books during July. We can represent the information on July sales more simply as

$$\begin{bmatrix} 6 & 8 \\ 15 & 20 \\ 45 & 64 \end{bmatrix}.$$

Such a rectangular array of real numbers is called a *matrix*.[1] It is customary to refer to real numbers as **scalars** (originally from the word *scale*) when working with a matrix. We denote the set of real numbers by $\mathcal{R}$.

Definitions A **matrix** (*plural, matrices*) is a rectangular array of scalars. If the matrix has m rows and n columns, we say that the **size** of the matrix is $\boldsymbol{m}$ **by** $\boldsymbol{n}$, written $m \times n$. The matrix is **square** if $m = n$. The scalar in the ith row and jth column is called the **(i, j)-entry** of the matrix.

If A is a matrix, we denote its (i, j)-entry by a_{ij}. We say that two matrices A and B are **equal** if they have the same size and have equal corresponding entries; that is, $a_{ij} = b_{ij}$ for all i and j. Symbolically, we write $A = B$.

In our bookstore example, the July and August sales are contained in the matrices

$$B = \begin{bmatrix} 6 & 8 \\ 15 & 20 \\ 45 & 64 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 7 & 9 \\ 18 & 31 \\ 52 & 68 \end{bmatrix}.$$

Note that $b_{12} = 8$ and $c_{12} = 9$, so $B \neq C$. Both B and C are 3×2 matrices. Because of the context in which these matrices arise, they are called *inventory matrices.*

Other examples of matrices are

$$\begin{bmatrix} \frac{2}{3} & -4 & 0 \\ \pi & 1 & 6 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 8 \\ 4 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} -2 & 0 & 1 & 1 \end{bmatrix}.$$

The first matrix has size 2×3, the second has size 3×1, and the third has size 1×4.

Practice Problem 1 ► Let $A = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$.

(a) What is the $(1, 2)$-entry of A?

(b) What is a_{22}? ◄

Sometimes we are interested in only a part of the information contained in a matrix. For example, suppose that we are interested in only magazine and book sales in July. Then the relevant information is contained in the last two rows of B; that is, in the matrix E defined by

$$E = \begin{bmatrix} 15 & 20 \\ 45 & 64 \end{bmatrix}.$$

E is called a *submatrix* of B. In general, a **submatrix** of a matrix M is obtained by deleting from M entire rows, entire columns, or both. It is permissible, when forming a submatrix of M, to delete none of the rows or none of the columns of M. As another example, if we delete the first row and the second column of B, we obtain the submatrix

$$\begin{bmatrix} 15 \\ 45 \end{bmatrix}.$$

---

[1] James Joseph Sylvester (1814–1897) coined the term *matrix* in the 1850s.

MATRIX SUMS AND SCALAR MULTIPLICATION

Matrices are more than convenient devices for storing information. Their usefulness lies in their *arithmetic*. As an example, suppose that we want to know the total numbers of newspapers, magazines, and books sold by both stores during July and August. It is natural to form one matrix whose entries are the sum of the corresponding entries of the matrices B and C, namely,

$$\begin{array}{r} \text{Store} \\ \text{Newspapers} \\ \text{Magazines} \\ \text{Books} \end{array} \begin{array}{c} \begin{array}{cc} 1 & \;\;2 \end{array} \\ \begin{bmatrix} 13 & 17 \\ 33 & 51 \\ 97 & 132 \end{bmatrix} \end{array}.$$

If A and B are $m \times n$ matrices, the **sum** of A and B, denoted by $A + B$, is the $m \times n$ matrix obtained by adding the corresponding entries of A and B; that is, $A + B$ is the $m \times n$ matrix whose (i, j)-entry is $a_{ij} + b_{ij}$. Notice that the matrices A and B must have the same size for their sum to be defined.

Suppose that in our bookstore example, July sales were to double in all categories. Then the new matrix of July sales would be

$$\begin{bmatrix} 12 & 16 \\ 30 & 40 \\ 90 & 128 \end{bmatrix}.$$

We denote this matrix by $2B$.

Let A be an $m \times n$ matrix and c be a scalar. The **scalar multiple** cA is the $m \times n$ matrix whose entries are c times the corresponding entries of A; that is, cA is the $m \times n$ matrix whose (i, j)-entry is ca_{ij}. Note that $1A = A$. We denote the matrix $(-1)A$ by $-A$ and the matrix $0A$ by O. We call the $m \times n$ matrix O in which each entry is 0 the $m \times n$ **zero matrix**.

Example 1 Compute the matrices $A + B$, $3A$, $-A$, and $3A + 4B$, where

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

Solution We have

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

and

$$3A + 4B = \begin{bmatrix} 9 & 12 & 6 \\ 6 & -9 & 0 \end{bmatrix} + \begin{bmatrix} -16 & 4 & 0 \\ 20 & -24 & 4 \end{bmatrix} = \begin{bmatrix} -7 & 16 & 6 \\ 26 & -33 & 4 \end{bmatrix}.$$

Just as we have defined addition of matrices, we can also define **subtraction**. For any matrices A and B of the same size, we define $A - B$ to be the matrix obtained by subtracting each entry of B from the corresponding entry of A. Thus the (i, j)-entry of $A - B$ is $a_{ij} - b_{ij}$. Notice that $A - A = O$ for all matrices A.

If, as in Example 1, we have

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

then

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

Practice Problem 2 ▶ Let $A = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 4 \end{bmatrix}$. Compute the following matrices:

(a) $A - B$

(b) $2A$

(c) $A + 3B$ ◀

We have now defined the operations of matrix addition and scalar multiplication. The power of linear algebra lies in the natural relations between these operations, which are described in our first theorem.

THEOREM 1.1

(Properties of Matrix Addition and Scalar Multiplication) Let A, B, and C be $m \times n$ matrices, and let s and t be any scalars. Then

(a) $A + B = B + A$. (commutative law of matrix addition)

(b) $(A + B) + C = A + (B + C)$. (associative law of matrix addition)

(c) $A + O = A$.

(d) $A + (-A) = O$.

(e) $(st)A = s(tA)$.

(f) $s(A + B) = sA + sB$.

(g) $(s + t)A = sA + tA$.

PROOF We prove parts (b) and (f). The rest are left as exercises.

(b) The matrices on each side of the equation are $m \times n$ matrices. We must show that each entry of $(A + B) + C$ is the same as the corresponding entry of $A + (B + C)$. Consider the (i, j)-entries. Because of the definition of matrix addition, the (i, j)-entry of $(A + B) + C$ is the sum of the (i, j)-entry of $A + B$, which is $a_{ij} + b_{ij}$, and the (i, j)-entry of C, which is c_{ij}. Therefore this sum equals $(a_{ij} + b_{ij}) + c_{ij}$. Similarly, the (i, j)-entry of $A + (B + C)$ is $a_{ij} + (b_{ij} + c_{ij})$. Because the associative law holds for addition of scalars, $(a_{ij} + b_{ij}) + c_{ij} = a_{ij} + (b_{ij} + c_{ij})$. Therefore the (i, j)-entry of $(A + B) + C$ equals the (i, j)-entry of $A + (B + C)$, proving (b).

(f) The matrices on each side of the equation are $m \times n$ matrices. As in the proof of (b), we consider the (i, j)-entries of each matrix. The (i, j)-entry of $s(A + B)$ is defined to be the product of s and the (i, j)-entry of $A + B$, which is $a_{ij} + b_{ij}$. This product equals $s(a_{ij} + b_{ij})$. The (i, j)-entry of $sA + sB$ is the sum of the (i, j)-entry of sA, which is sa_{ij}, and the (i, j)-entry of sB, which is sb_{ij}. This sum is $sa_{ij} + sb_{ij}$. Since $s(a_{ij} + b_{ij}) = sa_{ij} + sb_{ij}$, (f) is proved. ■

Because of the associative law of matrix addition, sums of three or more matrices can be written unambiguously without parentheses. Thus we may write $A + B + C$ instead of either $(A + B) + C$ or $A + (B + C)$.

MATRIX TRANSPOSES

In the bookstore example, we could have recorded the information about July sales in the following form:

Store	Newspapers	Magazines	Books
1	6	15	45
2	8	20	64

This representation produces the matrix

$$\begin{bmatrix} 6 & 15 & 45 \\ 8 & 20 & 64 \end{bmatrix}.$$

Compare this with

$$B = \begin{bmatrix} 6 & 8 \\ 15 & 20 \\ 45 & 64 \end{bmatrix}.$$

The rows of the first matrix are the columns of B, and the columns of the first matrix are the rows of B. This new matrix is called the *transpose* of B. In general, the **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix denoted by A^T whose (i, j)-entry is the (j, i)-entry of A.

The matrix C in our bookstore example and its transpose are

$$C = \begin{bmatrix} 7 & 9 \\ 18 & 31 \\ 52 & 68 \end{bmatrix} \quad \text{and} \quad C^T = \begin{bmatrix} 7 & 18 & 52 \\ 9 & 31 & 68 \end{bmatrix}.$$

Practice Problem 3 ▶ Let $A = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 4 \end{bmatrix}$. Compute the following matrices:

(a) A^T

(b) $(3B)^T$

(c) $(A + B)^T$ ◀

The following theorem shows that the transpose preserves the operations of matrix addition and scalar multiplication:

THEOREM 1.2

(Properties of the Transpose) Let A and B be $m \times n$ matrices, and let s be any scalar. Then

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

(b) $(sA)^T = sA^T$.

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

PROOF We prove part (a). The rest are left as exercises.

(a) The matrices on each side of the equation are $n \times m$ matrices. So we show that the (i, j)-entry of $(A + B)^T$ equals the (i, j)-entry of $A^T + B^T$. By the definition of transpose, the (i, j)-entry of $(A + B)^T$ equals the (j, i)-entry of $A + B$, which is $a_{ji} + b_{ji}$. On the other hand, the (i, j)-entry of $A^T + B^T$ equals the sum of the (i, j)-entry of A^T and the (i, j)-entry of B^T, that is, $a_{ji} + b_{ji}$. Because the (i, j)-entries of $(A + B)^T$ and $A^T + B^T$ are equal, (a) is proved. ■

VECTORS

A matrix that has exactly one row is called a **row vector**, and a matrix that has exactly one column is called a **column vector**. The term *vector* is used to refer to either a row vector or a column vector. The entries of a vector are called **components**. In this book, we normally work with column vectors, and we denote the set of all column vectors with n components by $\mathcal{R}^n$.

We write vectors as boldface lower case letters such as **u** and **v**, and denote the ith component of the vector **u** by u_i. For example, if $\mathbf{u} = \begin{bmatrix} 2 \\ -4 \\ 7 \end{bmatrix}$, then $u_2 = -4$. Occasionally, we identify a vector **u** in $\mathcal{R}^n$ with an n-tuple, $(u_1, u_2, \ldots, u_n)$.

Because vectors are special types of matrices, we can add them and multiply them by scalars. In this context, we call the two arithmetic operations on vectors **vector addition** and **scalar multiplication**. These operations satisfy the properties listed in Theorem 1.1. In particular, the vector in $\mathcal{R}^n$ with all zero components is denoted by **0** and is called the **zero vector**. It satisfies $\mathbf{u} + \mathbf{0} = \mathbf{u}$ and $0\mathbf{u} = \mathbf{0}$ for every **u** in $\mathcal{R}^n$.

Example 2

Let $\mathbf{u} = \begin{bmatrix} 2 \\ -4 \\ 7 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$. Then

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ 7 \end{bmatrix}, \qquad \mathbf{u} - \mathbf{v} = \begin{bmatrix} -3 \\ -7 \\ 7 \end{bmatrix}, \qquad \text{and} \qquad 5\mathbf{v} = \begin{bmatrix} 25 \\ 15 \\ 0 \end{bmatrix}.$$

For a given matrix, it is often advantageous to consider its rows and columns as vectors. For example, for the matrix $\begin{bmatrix} 2 & 4 & 3 \\ 0 & 1 & -2 \end{bmatrix}$, the **rows** are $\begin{bmatrix} 2 & 4 & 3 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & -2 \end{bmatrix}$, and the **columns** are $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

Because the columns of a matrix play a more important role than the rows, we introduce a special notation. When a capital letter denotes a matrix, we use the corresponding lower case letter in boldface with a subscript j to represent the jth column of that matrix. So if A is an $m \times n$ matrix, its jth column is

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

GEOMETRY OF VECTORS

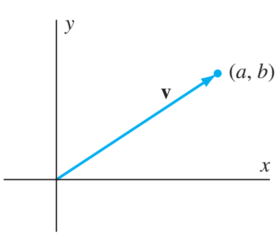


Figure 1.1 A vector in $\mathcal{R}^2$

For many applications,[2] it is useful to represent vectors geometrically as directed line segments, or arrows. For example, if $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ is a vector in $\mathcal{R}^2$, we can represent **v** as an arrow from the origin to the point (a, b) in the xy-plane, as shown in Figure 1.1.

[2] The importance of vectors in physics was recognized late in the nineteenth century. The algebra of vectors, developed by Oliver Heaviside (1850–1925) and Josiah Willard Gibbs (1839–1903), won out over the algebra of quaternions to become the language of physicists.

Example 3

Velocity Vectors A boat cruises in still water toward the northeast at 20 miles per hour. The velocity $\mathbf{u}$ of the boat is a vector that points in the direction of the boat's motion, and whose length is 20, the boat's speed. If the positive y-axis represents north and the positive x-axis represents east, the boat's direction makes an angle of 45° with the x-axis. (See Figure 1.2.) We can compute the components of $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ by using trigonometry:

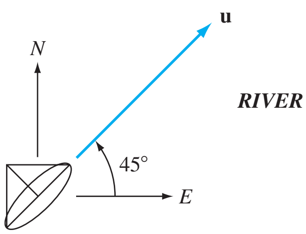


Figure 1.2

$$u_1 = 20\cos 45^\circ = 10\sqrt{2} \qquad \text{and} \qquad u_2 = 20\sin 45^\circ = 10\sqrt{2}.$$

Therefore $\mathbf{u} = \begin{bmatrix} 10\sqrt{2} \\ 10\sqrt{2} \end{bmatrix}$, where the units are in miles per hour.

VECTOR ADDITION AND THE PARALLELOGRAM LAW

We can represent vector addition graphically, using arrows, by a result called the *parallelogram law*.[3] To add nonzero vectors $\mathbf{u}$ and $\mathbf{v}$, first form a parallelogram with adjacent sides $\mathbf{u}$ and $\mathbf{v}$. Then the sum $\mathbf{u} + \mathbf{v}$ is the arrow along the diagonal of the parallelogram as shown in Figure 1.3.

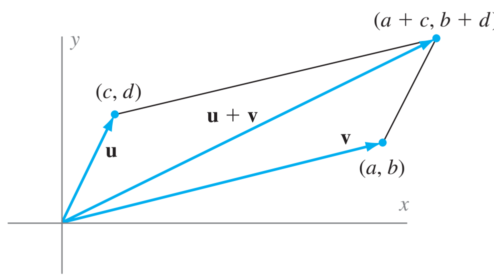


Figure 1.3 The parallelogram law of vector addition

Velocities can be combined by adding vectors that represent them.

Example 4

Imagine that the boat from the previous example is now cruising on a river, which flows to the east at 7 miles per hour. As before, the bow of the boat points toward the northeast, and its speed relative to the water is 20 miles per hour. In this case, the vector $\mathbf{u} = \begin{bmatrix} 10\sqrt{2} \\ 10\sqrt{2} \end{bmatrix}$, which we calculated in the previous example, represents the boat's velocity (in miles per hour) relative to the river. To find the velocity of the boat relative to the shore, we must add a vector $\mathbf{v}$, representing the velocity of the river, to the vector $\mathbf{u}$. Since the river flows toward the east at 7 miles per hour, its velocity vector is $\mathbf{v} = \begin{bmatrix} 7 \\ 0 \end{bmatrix}$. We can represent the sum of the vectors $\mathbf{u}$ and $\mathbf{v}$ by using the parallelogram law, as shown in Figure 1.4. The velocity of the boat relative to the shore (in miles per hour) is the vector

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 10\sqrt{2} + 7 \\ 10\sqrt{2} \end{bmatrix}.$$

[3] A justification of the parallelogram law by Heron of Alexandria (first century C.E.) appears in his *Mechanics*.

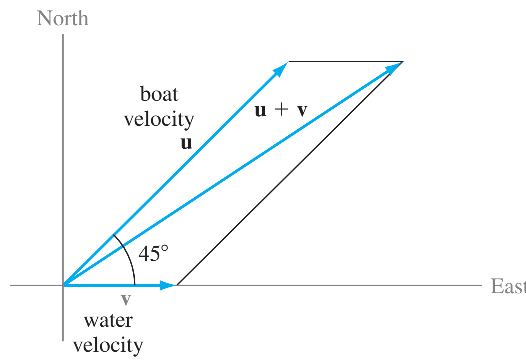


Figure 1.4

To find the speed of the boat, we use the Pythagorean theorem, which tells us that the length of a vector with endpoint (p, q) is $\sqrt{p^2 + q^2}$. Using the fact that the components of $\mathbf{u} + \mathbf{v}$ are $p = 10\sqrt{2} + 7$ and $q = 10\sqrt{2}$, respectively, it follows that the speed of the boat is

$$\sqrt{p^2 + q^2} \approx 25.44 \text{ mph.}$$

SCALAR MULTIPLICATION

We can also represent scalar multiplication graphically, using arrows. If $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ is a vector and c is a positive scalar, the scalar multiple $c\mathbf{v}$ is a vector that points in the same direction as $\mathbf{v}$, and whose length is c times the length of $\mathbf{v}$. This is shown in Figure 1.5(a). If c is negative, $c\mathbf{v}$ points in the opposite direction from $\mathbf{v}$, and has length $|c|$ times the length of $\mathbf{v}$. This is shown in Figure 1.5(b). We call two vectors **parallel** if one of them is a scalar multiple of the other.

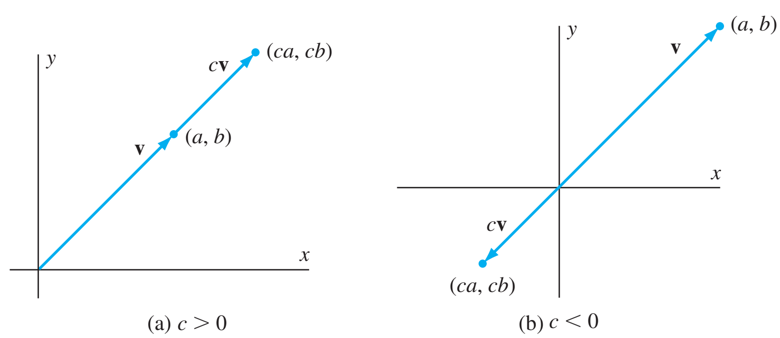


Figure 1.5 Scalar multiplication of vectors

VECTORS IN $\mathcal{R}^3$

If we identify $\mathcal{R}^3$ as the set of all ordered triples, then the same geometric ideas that hold in $\mathcal{R}^2$ are also true in $\mathcal{R}^3$. We may depict a vector $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathcal{R}^3$ as an arrow emanating from the origin of the xyz-coordinate system, with the point (a, b, c) as its

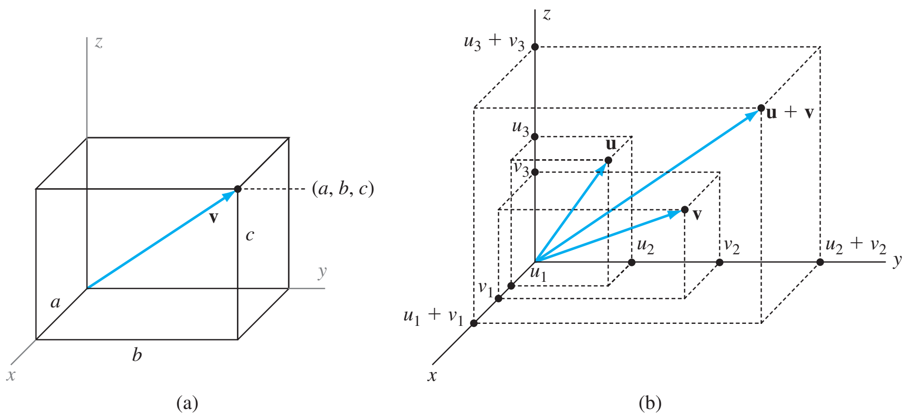


Figure 1.6 Vectors in $\mathcal{R}^3$

endpoint. (See Figure 1.6(a).) As is the case in $\mathcal{R}^2$, we can view two nonzero vectors in $\mathcal{R}^3$ as adjacent sides of a parallelogram, and we can represent their addition by using the parallelogram law. (See Figure 1.6(b).) In real life, motion takes place in 3-dimensional space, and we can depict quantities such as velocities and forces as vectors in $\mathcal{R}^3$.

EXERCISES

In Exercises 1–12, compute the indicated matrices, where

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

1. $4A$
2. $-A$
3. $4A - 2B$
4. $3A + 2B$
5. $(2B)^T$
6. $A^T + 2B^T$
7. $A + B$
8. $(A + 2B)^T$
9. A^T
10. $A - B$
11. $-(B^T)$
12. $(-B)^T$

In Exercises 13–24, compute the indicated matrices, if possible, where

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

13. $-A$
14. $3B$
15. $(-2)A$
16. $(2B)^T$
17. $A - B$
18. $A - B^T$
19. $A^T - B$
20. $3A + 2B^T$
21. $(A + B)^T$
22. $(4A)^T$
23. $B - A^T$
24. $(B^T - A)^T$

In Exercises 25–28, assume that $A = \begin{bmatrix} 3 & -2 \\ 0 & 1.6 \\ 2\pi & 5 \end{bmatrix}$.

25. Determine a_{12}.
26. Determine a_{21}.
27. Determine $\mathbf{a}_1$.
28. Determine $\mathbf{a}_2$.

In Exercises 29–32, assume that $C = \begin{bmatrix} 2 & -3 & 0.4 \\ 2e & 12 & 0 \end{bmatrix}$.

29. Determine $\mathbf{c}_1$.
30. Determine $\mathbf{c}_3$.
31. Determine the first row of C.
32. Determine the second row of C.

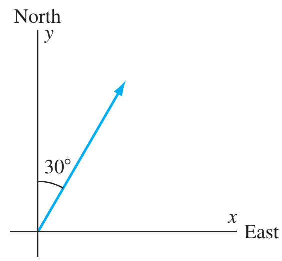


Figure 1.7 A view of the airplane from above

33. An airplane is flying with a ground speed of 300 mph at an angle of 30° east of due north. (See Figure 1.7.) In addition, the airplane is climbing at a rate of 10 mph. Determine the vector in $\mathcal{R}^3$ that represents the velocity (in mph) of the airplane.
34. A swimmer is swimming northeast at 2 mph in still water.
    (a) Give the velocity of the swimmer. Include a sketch.
    (b) A current in a northerly direction at 1 mph affects the velocity of the swimmer. Give the new velocity and speed of the swimmer. Include a sketch.
35. A pilot keeps her airplane pointed in a northeastward direction while maintaining an airspeed (speed relative to the surrounding air) of 300 mph. A wind from the west blows eastward at 50 mph.

(a) Find the velocity (in mph) of the airplane relative to the ground.

(b) What is the speed (in mph) of the airplane relative to the ground?

36. Suppose that in a medical study of 20 people, for each i, $1 \le i \le 20$, the 3×1 vector $\mathbf{u}_i$ is defined so that its components respectively represent the blood pressure, pulse rate, and cholesterol reading of the ith person. Provide an interpretation of the vector $\frac{1}{20}(\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_{20})$.

T&F *In Exercises 37–56, determine whether the statements are true or false.*

37. Matrices must be of the same size for their sum to be defined.
38. The transpose of a sum of two matrices is the sum of the transposed matrices.
39. Every vector is a matrix.
40. A scalar multiple of the zero matrix is the zero scalar.
41. The transpose of a matrix is a matrix of the same size.
42. A submatrix of a matrix may be a vector.
43. If B is a 3×4 matrix, then its rows are 4×1 vectors.
44. The $(3, 4)$-entry of a matrix lies in column 3 and row 4.
45. In a zero matrix, every entry is 0.
46. An $m \times n$ matrix has $m + n$ entries.
47. If $\mathbf{v}$ and $\mathbf{w}$ are vectors such that $\mathbf{v} = -3\mathbf{w}$, then $\mathbf{v}$ and $\mathbf{w}$ are parallel.
48. If A and B are any $m \times n$ matrices, then
$$A - B = A + (-1)B.$$
49. The (i, j)-entry of A^T equals the (j, i)-entry of A.
50. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 4 & 0 \end{bmatrix}$, then $A = B$.
51. In any matrix A, the sum of the entries of $3A$ equals three times the sum of the entries of A.
52. Matrix addition is commutative.
53. Matrix addition is associative.
54. For any $m \times n$ matrices A and B and any scalars c and d, $(cA + dB)^T = cA^T + dB^T$.
55. If A is a matrix, then cA is the same size as A for every scalar c.
56. If A is a matrix for which the sum $A + A^T$ is defined, then A is a square matrix.

---

57. Let A and B be matrices of the same size.
(a) Prove that the jth column of $A + B$ is $\mathbf{a}_j + \mathbf{b}_j$.
(b) Prove that for any scalar c, the jth column of cA is $c\mathbf{a}_j$.
58. For any $m \times n$ matrix A, prove that $0A = O$, the $m \times n$ zero matrix.
59. For any $m \times n$ matrix A, prove that $1A = A$.
60. Prove Theorem 1.1(a).
61. Prove Theorem 1.1(c).
62. Prove Theorem 1.1(d).
63. Prove Theorem 1.1(e).
64. Prove Theorem 1.1(g).
65. Prove Theorem 1.2(b).
66. Prove Theorem 1.2(c).

A square matrix A is called a **diagonal matrix** *if* $a_{ij} = 0$ *whenever* $i \ne j$. *Exercises 67–70 are concerned with diagonal matrices.*

67. Prove that a square zero matrix is a diagonal matrix.
68. Prove that if B is a diagonal matrix, then cB is a diagonal matrix for any scalar c.
69. Prove that if B is a diagonal matrix, then B^T is a diagonal matrix.
70. Prove that if B and C are diagonal matrices of the same size, then $B + C$ is a diagonal matrix.

A (square) matrix A is said to be **symmetric** *if* $A = A^T$. *Exercises 71–78 are concerned with symmetric matrices.*

71. Give examples of 2×2 and 3×3 symmetric matrices.
72. Prove that the (i, j)-entry of a symmetric matrix equals the (j, i)-entry.
73. Prove that a square zero matrix is symmetric.
74. Prove that if B is a symmetric matrix, then so is cB for any scalar c.
75. Prove that if B is a square matrix, then $B + B^T$ is symmetric.
76. Prove that if B and C are $n \times n$ symmetric matrices, then so is $B + C$.
77. Is a square submatrix of a symmetric matrix necessarily a symmetric matrix? Justify your answer.
78. Prove that a diagonal matrix is symmetric.

A (square) matrix A is called **skew-symmetric** *if* $A^T = -A$. *Exercises 79–81 are concerned with skew-symmetric matrices.*

79. What must be true about the (i, i)-entries of a skew-symmetric matrix? Justify your answer.
80. Give an example of a nonzero 2×2 skew-symmetric matrix B. Now show that every 2×2 skew-symmetric matrix is a scalar multiple of B.
81. Show that every 3×3 matrix can be written as the sum of a symmetric matrix and a skew-symmetric matrix.
82.[4] The **trace** of an $n \times n$ matrix A, written trace(A), is defined to be the sum
$$\text{trace}(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$
Prove that, for any $n \times n$ matrices A and B and scalar c, the following statements are true:
(a) $\text{trace}(A + B) = \text{trace}(A) + \text{trace}(B)$.
(b) $\text{trace}(cA) = c \cdot \text{trace}(A)$.
(c) $\text{trace}(A^T) = \text{trace}(A)$.
83. *Probability vectors* are vectors whose components are nonnegative and have a sum of 1. Show that if $\mathbf{p}$ and $\mathbf{q}$ are probability vectors and a and b are nonnegative scalars with $a + b = 1$, then $a\mathbf{p} + b\mathbf{q}$ is a probability vector.

[4] This exercise is used in Sections 2.2, 7.1, and 7.5 (on pages 115, 495, and 533, respectively).

In the following exercise, use either a calculator with matrix capabilities or computer software such as MATLAB to solve the problem:

84. Consider the matrices

$$A = \begin{bmatrix} 1.3 & 2.1 & -3.3 & 6.0 \\ 5.2 & 2.3 & -1.1 & 3.4 \\ 3.2 & -2.6 & 1.1 & -4.0 \\ 0.8 & -1.3 & -12.1 & 5.7 \\ -1.4 & 3.2 & 0.7 & 4.4 \end{bmatrix}$$

and

$$B = \begin{bmatrix} 2.6 & -1.3 & 0.7 & -4.4 \\ 2.2 & -2.6 & 1.3 & -3.2 \\ 7.1 & 1.5 & -8.3 & 4.6 \\ -0.9 & -1.2 & 2.4 & 5.9 \\ 3.3 & -0.9 & 1.4 & 6.2 \end{bmatrix}.$$

(a) Compute $A + 2B$.
(b) Compute $A - B$.
(c) Compute $A^T + B^T$.

SOLUTIONS TO THE PRACTICE PROBLEMS

1. (a) The $(1, 2)$-entry of A is 2.
   (b) The $(2, 2)$-entry of A is 3.

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

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

   (c) $$A + 3B = \begin{bmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \end{bmatrix} + 3\begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 4 \end{bmatrix}$$
   $$= \begin{bmatrix} 2 & -1 & 1 \\ 3 & 0 & -2 \end{bmatrix} + \begin{bmatrix} 3 & 9 & 0 \\ 6 & -3 & 12 \end{bmatrix}$$
   $$= \begin{bmatrix} 5 & 8 & 1 \\ 9 & -3 & 10 \end{bmatrix}$$

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

   (b) $$(3B)^T = \begin{bmatrix} 3 & 9 & 0 \\ 6 & -3 & 12 \end{bmatrix}^T = \begin{bmatrix} 3 & 6 \\ 9 & -3 \\ 0 & 12 \end{bmatrix}$$

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

1.2 LINEAR COMBINATIONS, MATRIX–VECTOR PRODUCTS, AND SPECIAL MATRICES

In this section, we explore some applications involving matrix operations and introduce the product of a matrix and a vector.

Suppose that 20 students are enrolled in a linear algebra course, in which two tests, a quiz, and a final exam are given. Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{20} \end{bmatrix}$, where u_i denotes the score of the ith student on the first test. Likewise, define vectors $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{z}$ similarly for the second test, quiz, and final exam, respectively. Assume that the instructor computes a student's course average by counting each test score twice as much as a quiz score, and the final exam score three times as much as a test score. Thus the *weights* for the tests, quiz, and final exam score are, respectively, 2/11, 2/11, 1/11, 6/11 (the weights must sum to one). Now consider the vector

$$\mathbf{y} = \frac{2}{11}\mathbf{u} + \frac{2}{11}\mathbf{v} + \frac{1}{11}\mathbf{w} + \frac{6}{11}\mathbf{z}.$$

The first component y_1 represents the first student's course average, the second component y_2 represents the second student's course average, and so on. Notice that $\mathbf{y}$ is a sum of scalar multiples of $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{z}$. This form of vector sum is so important that it merits its own definition.

Definitions A **linear combination** of vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ is a vector of the form

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k,$$

where $c_1, c_2, \ldots, c_k$ are scalars. These scalars are called the **coefficients** of the linear combination.

Note that a linear combination of one vector is simply a scalar multiple of that vector.

In the previous example, the vector **y** of the students' course averages is a linear combination of the vectors **u**, **v**, **w**, and **z**. The coefficients are the weights. Indeed, any weighted average produces a linear combination of the scores.

Notice that

$$\begin{bmatrix}2\\8\end{bmatrix} = (-3)\begin{bmatrix}1\\1\end{bmatrix} + 4\begin{bmatrix}1\\3\end{bmatrix} + 1\begin{bmatrix}1\\-1\end{bmatrix}.$$

Thus $\begin{bmatrix}2\\8\end{bmatrix}$ is a linear combination of $\begin{bmatrix}1\\1\end{bmatrix}$, $\begin{bmatrix}1\\3\end{bmatrix}$, and $\begin{bmatrix}1\\-1\end{bmatrix}$, with coefficients -3, 4, and 1. We can also write

$$\begin{bmatrix}2\\8\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} + 2\begin{bmatrix}1\\3\end{bmatrix} - 1\begin{bmatrix}1\\-1\end{bmatrix}.$$

This equation also expresses $\begin{bmatrix}2\\8\end{bmatrix}$ as a linear combination of $\begin{bmatrix}1\\1\end{bmatrix}$, $\begin{bmatrix}1\\3\end{bmatrix}$, and $\begin{bmatrix}1\\-1\end{bmatrix}$, but now the coefficients are 1, 2, and -1. So the set of coefficients that express one vector as a linear combination of the others need not be unique.

Example 1

(a) Determine whether $\begin{bmatrix}4\\-1\end{bmatrix}$ is a linear combination of $\begin{bmatrix}2\\3\end{bmatrix}$ and $\begin{bmatrix}3\\1\end{bmatrix}$.

(b) Determine whether $\begin{bmatrix}-4\\-2\end{bmatrix}$ is a linear combination of $\begin{bmatrix}6\\3\end{bmatrix}$ and $\begin{bmatrix}2\\1\end{bmatrix}$.

(c) Determine whether $\begin{bmatrix}3\\4\end{bmatrix}$ is a linear combination of $\begin{bmatrix}3\\2\end{bmatrix}$ and $\begin{bmatrix}6\\4\end{bmatrix}$.

Solution (a) We seek scalars x_1 and x_2 such that

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

That is, we seek a solution of the system of equations

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

Because these equations represent nonparallel lines in the plane, there is exactly one solution, namely, $x_1 = -1$ and $x_2 = 2$. Therefore $\begin{bmatrix}4\\-1\end{bmatrix}$ is a (unique) linear combination of the vectors $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, namely,

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

(See Figure 1.8.)

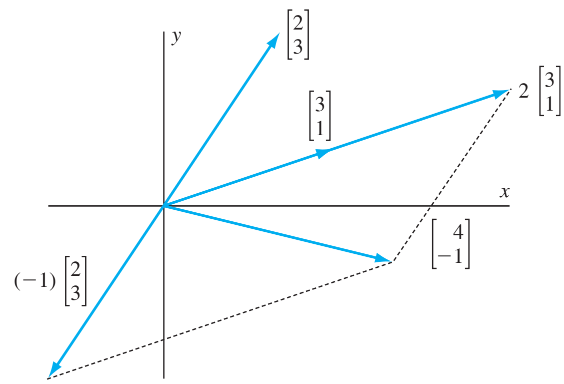


Figure 1.8 The vector $\begin{bmatrix} 4 \\ -1 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

(b) To determine whether $\begin{bmatrix} -4 \\ -2 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 6 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, we perform a similar computation and produce the set of equations

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

Since the first equation is twice the second, we need only solve $3x_1 + x_2 = -2$. This equation represents a line in the plane, and the coordinates of any point on the line give a solution. For example, we can let $x_1 = -2$ and $x_2 = 4$. In this case, we have

$$\begin{bmatrix} -4 \\ -2 \end{bmatrix} = (-2)\begin{bmatrix} 6 \\ 3 \end{bmatrix} + 4\begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

There are infinitely many solutions. (See Figure 1.9.)

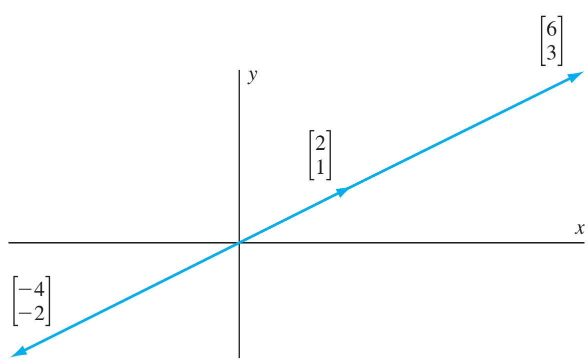


Figure 1.9 The vector $\begin{bmatrix} -4 \\ -2 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 6 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

(c) To determine if $\begin{bmatrix}3\\4\end{bmatrix}$ is a linear combination of $\begin{bmatrix}3\\2\end{bmatrix}$ and $\begin{bmatrix}6\\4\end{bmatrix}$, we must solve the system of equations

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

If we add $-\frac{2}{3}$ times the first equation to the second, we obtain $0 = 2$, an equation with no solutions. Indeed, the two original equations represent parallel lines in the plane, so the original system has no solutions. We conclude that $\begin{bmatrix}3\\4\end{bmatrix}$ is not a linear combination of $\begin{bmatrix}3\\2\end{bmatrix}$ and $\begin{bmatrix}6\\4\end{bmatrix}$. (See Figure 1.10.)

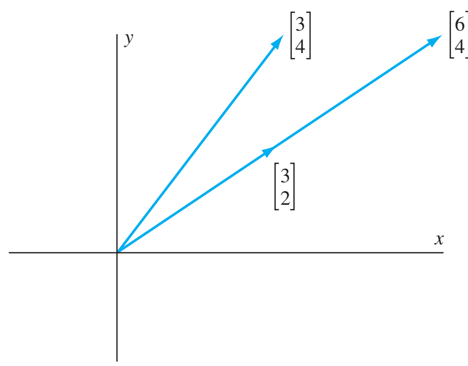


Figure 1.10 The vector $\begin{bmatrix}3\\4\end{bmatrix}$ is *not* a linear combination of $\begin{bmatrix}3\\2\end{bmatrix}$ and $\begin{bmatrix}6\\4\end{bmatrix}$.

Example 2 Given vectors $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, show that the sum of any two linear combinations of these vectors is also a linear combination of these vectors.

Solution Suppose that $\mathbf{w}$ and $\mathbf{z}$ are linear combinations of $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$. Then we may write

$$\mathbf{w} = a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 \qquad \text{and} \qquad \mathbf{z} = a'\mathbf{u}_1 + b'\mathbf{u}_2 + c'\mathbf{u}_3,$$

where a, b, c, a', b', c' are scalars. So

$$\mathbf{w} + \mathbf{z} = (a + a')\mathbf{u}_1 + (b + b')\mathbf{u}_2 + (c + c')\mathbf{u}_3,$$

which is also a linear combination of $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$.

STANDARD VECTORS

We can write any vector $\begin{bmatrix}a\\b\end{bmatrix}$ in $\mathcal{R}^2$ as a linear combination of the two vectors $\begin{bmatrix}1\\0\end{bmatrix}$ and $\begin{bmatrix}0\\1\end{bmatrix}$ as follows:

$$\begin{bmatrix}a\\b\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix}$$

The vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are called the *standard vectors* of $\mathcal{R}^2$. Similarly, we can write any vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ in $\mathcal{R}^3$ as a linear combination of the vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ as follows:

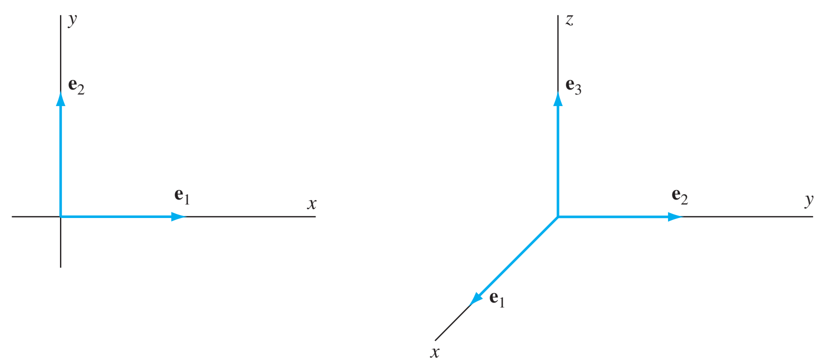
$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The vectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are called the *standard vectors* of $\mathcal{R}^3$.

In general, we define the **standard vectors** of $\mathcal{R}^n$ by

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

(See Figure 1.11.)



The standard vectors of $\mathcal{R}^2$ The standard vectors of $\mathcal{R}^3$

Figure 1.11

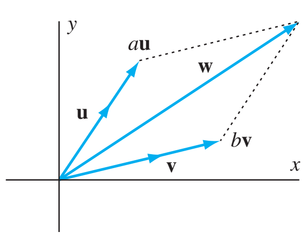


Figure 1.12 The vector **w** is a linear combination of the nonparallel vectors **u** and **v**.

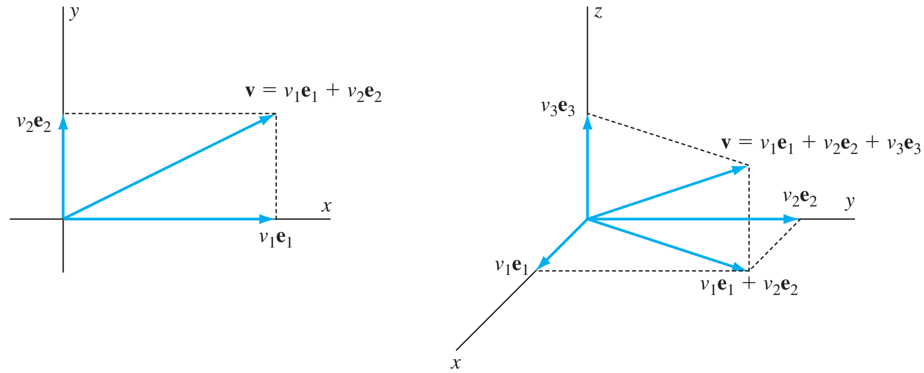
From the preceding equations, it is easy to see that every vector in $\mathcal{R}^n$ is a linear combination of the standard vectors of $\mathcal{R}^n$. In fact, for any vector $\mathbf{v}$ in $\mathcal{R}^n$,

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

(See Figure 1.13.)

Now let $\mathbf{u}$ and $\mathbf{v}$ be nonparallel vectors, and let $\mathbf{w}$ be any vector in $\mathcal{R}^2$. Begin with the endpoint of $\mathbf{w}$ and create a parallelogram with sides $a\mathbf{u}$ and $b\mathbf{v}$, so that $\mathbf{w}$ is its diagonal. It follows that $\mathbf{w} = a\mathbf{u} + b\mathbf{v}$; that is, $\mathbf{w}$ is a linear combination of the vectors $\mathbf{u}$ and $\mathbf{v}$. (See Figure 1.12.) More generally, the following statement is true:

> If $\mathbf{u}$ and $\mathbf{v}$ are any nonparallel vectors in $\mathcal{R}^2$, then every vector in $\mathcal{R}^2$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.



The vector **v** is a linear combination of standard vectors in $\mathcal{R}^2$.

The vector **v** is a linear combination of standard vectors in $\mathcal{R}^3$.

Figure 1.13

Practice Problem 1 ▶ Let $\mathbf{w} = \begin{bmatrix} -1 \\ 10 \end{bmatrix}$ and $\mathcal{S} = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$.

(a) Without doing any calculations, explain why **w** can be written as a linear combination of the vectors in $\mathcal{S}$.

(b) Express **w** as a linear combination of the vectors in $\mathcal{S}$. ◀

Suppose that a garden supply store sells three mixtures of grass seed. The deluxe mixture is 80% bluegrass and 20% rye, the standard mixture is 60% bluegrass and 40% rye, and the economy mixture is 40% bluegrass and 60% rye. One way to record this information is with the following 2×3 matrix:

$$B = \begin{array}{c} \begin{matrix} \text{deluxe} & \text{standard} & \text{economy} \end{matrix} \\ \begin{bmatrix} .80 & .60 & .40 \\ .20 & .40 & .60 \end{bmatrix} \end{array} \begin{array}{l} \text{bluegrass} \\ \text{rye} \end{array}$$

A customer wants to purchase a blend of grass seed containing 5 lb of bluegrass and 3 lb of rye. There are two natural questions that arise:

1. Is it possible to combine the three mixtures of seed into a blend that has exactly the desired amounts of bluegrass and rye, with no surplus of either?
2. If so, how much of each mixture should the store clerk add to the blend?

Let x_1, x_2, and x_3 denote the number of pounds of deluxe, standard, and economy mixtures, respectively, to be used in the blend. Then we have

$$\begin{aligned} .80x_1 + .60x_2 + .40x_3 &= 5 \\ .20x_1 + .40x_2 + .60x_3 &= 3. \end{aligned}$$

This is a *system of two linear equations in three unknowns.* Finding a solution of this system is equivalent to answering our second question. The technique for solving general systems is explored in great detail in Sections 1.3 and 1.4.

Using matrix notation, we may rewrite these equations in the form

$$\begin{bmatrix} .80x_1 + .60x_2 + .40x_3 \\ .20x_1 + .40x_2 + .60x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Now we use matrix operations to rewrite this matrix equation, using the columns of B, as

$$x_1\begin{bmatrix}.80\\.20\end{bmatrix}+x_2\begin{bmatrix}.60\\.40\end{bmatrix}+x_3\begin{bmatrix}.40\\.60\end{bmatrix}=\begin{bmatrix}5\\3\end{bmatrix}.$$

Thus we can rephrase the first question as follows: Is $\begin{bmatrix}5\\3\end{bmatrix}$ a linear combination of the columns $\begin{bmatrix}.80\\.20\end{bmatrix}$, $\begin{bmatrix}.60\\.40\end{bmatrix}$, and $\begin{bmatrix}.40\\.60\end{bmatrix}$ of B? The result in the box on page 17 provides an affirmative answer. Because no two of the three vectors are parallel, $\begin{bmatrix}5\\3\end{bmatrix}$ is a linear combination of any pair of these vectors.

MATRIX–VECTOR PRODUCTS

A convenient way to represent systems of linear equations is by *matrix–vector products*. For the preceding example, we represent the variables by the vector $\mathbf{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$ and define the *matrix–vector product* $B\mathbf{x}$ to be the linear combination

$$B\mathbf{x}=\begin{bmatrix}.80 & .60 & .40\\.20 & .40 & .60\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_1\begin{bmatrix}.80\\.20\end{bmatrix}+x_2\begin{bmatrix}.60\\.40\end{bmatrix}+x_3\begin{bmatrix}.40\\.60\end{bmatrix}.$$

This definition provides another way to state the first question in the preceding example: Does the vector $\begin{bmatrix}5\\3\end{bmatrix}$ equal $B\mathbf{x}$ for some vector $\mathbf{x}$? Notice that for the matrix–vector product to make sense, the number of columns of B must equal the number of components in $\mathbf{x}$. The general definition of a matrix–vector product is given next.

Definition Let A be an $m \times n$ matrix and $\mathbf{v}$ be an $n \times 1$ vector. We define the **matrix–vector product** of A and $\mathbf{v}$, denoted by $A\mathbf{v}$, to be the linear combination of the columns of A whose coefficients are the corresponding components of $\mathbf{v}$. That is,

$$A\mathbf{v}=v_1\mathbf{a}_1+v_2\mathbf{a}_2+\cdots+v_n\mathbf{a}_n.$$

As we have noted, for $A\mathbf{v}$ to exist, the number of columns of A must equal the number of components of $\mathbf{v}$. For example, suppose that

$$A=\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix} \quad \text{and} \quad \mathbf{v}=\begin{bmatrix}7\\8\end{bmatrix}.$$

Notice that A has two columns and $\mathbf{v}$ has two components. Then

$$A\mathbf{v}=\begin{bmatrix}1 & 2\\3 & 4\\5 & 6\end{bmatrix}\begin{bmatrix}7\\8\end{bmatrix}=7\begin{bmatrix}1\\3\\5\end{bmatrix}+8\begin{bmatrix}2\\4\\6\end{bmatrix}=\begin{bmatrix}7\\21\\35\end{bmatrix}+\begin{bmatrix}16\\32\\48\end{bmatrix}=\begin{bmatrix}23\\53\\83\end{bmatrix}.$$

Returning to the preceding garden supply store example, suppose that the store has 140 lb of seed in stock: 60 lb of the deluxe mixture, 50 lb of the standard mixture, and 30 lb of the economy mixture. We let $\mathbf{v} = \begin{bmatrix} 60 \\ 50 \\ 30 \end{bmatrix}$ represent this information. Now the matrix–vector product

$$B\mathbf{v} = \begin{bmatrix} .80 & .60 & .40 \\ .20 & .40 & .60 \end{bmatrix} \begin{bmatrix} 60 \\ 50 \\ 30 \end{bmatrix}$$

$$= 60 \begin{bmatrix} .80 \\ .20 \end{bmatrix} + 50 \begin{bmatrix} .60 \\ .40 \end{bmatrix} + 30 \begin{bmatrix} .40 \\ .60 \end{bmatrix}$$

$$= \overset{\text{seed (lb)}}{\begin{bmatrix} 90 \\ 50 \end{bmatrix}} \begin{matrix} \text{bluegrass} \\ \text{rye} \end{matrix}$$

gives the number of pounds of each type of seed contained in the 140 pounds of seed that the garden supply store has in stock. For example, there are 90 pounds of bluegrass because $90 = .80(60) + .60(50) + .40(30)$.

There is another approach to computing the matrix–vector product that relies more on the entries of A than on its columns. Consider the following example:

$$A\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$= v_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + v_3 \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \end{bmatrix}$$

Notice that the first component of the vector $A\mathbf{v}$ is the sum of products of the corresponding entries of the first row of A and the components of $\mathbf{v}$. Likewise, the second component of $A\mathbf{v}$ is the sum of products of the corresponding entries of the second row of A and the components of $\mathbf{v}$. With this approach to computing a matrix–vector product, we can omit the intermediate step in the preceding illustration. For example, suppose

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

Then

$$A\mathbf{v} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (2)(-1) + (3)(1) + (1)(3) \\ (1)(-1) + (-2)(1) + (3)(3) \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$