Linear Algebra

Instructor: Asst. Prof. Dr. Praphan Pavarangkoon
Office: Room no. 418-5, 4th Floor
Email: praphan@it.kmitl.ac.th

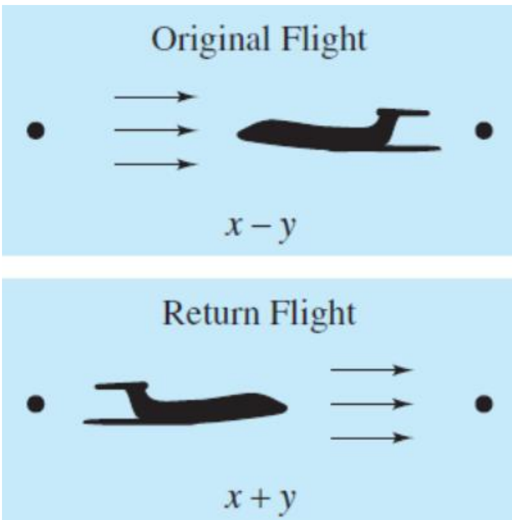
Outline

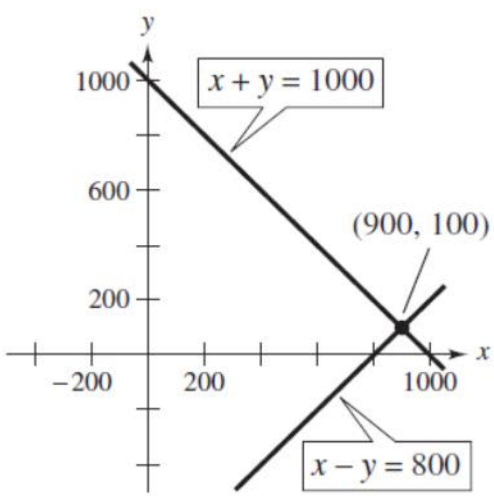
- Linear Algebra
- Matrices and determinants
- System of linear equations and its solutions
- Vector space
- Linear transformations
- Orthogonal matrix

What is Linear Algebra?

- The most fundamental theme of linear algebra is the theory of systems of linear equations.
- You have probably encountered small systems of linear equations in your previous mathematics courses.
- For example, suppose you travel on an airplane between two cities that are 5000 kilometers apart.
- If the trip one way against a headwind takes 6.25 hours and the return trip the same day in the direction of the wind takes only 5 hours, can you find the ground speed of the plane and the speed of the wind, assuming that both remain constant?

What is Linear Algebra? (cont.)





The lines intersect at (900, 100).

- If you let x represent the speed of the plane and y the speed of the wind, then the following system models the problem.

$$6.25(x - y) = 5000$$
$$5(x + y) = 5000$$

- This system of two equations and two unknowns simplifies to

$$x - y = 800$$
$$x + y = 1000.$$

- and the solution is $x = 900$ kilometers per hour and $y = 100$ kilometers per hour.
- Geometrically, this system represents two lines in the xy-plane.
- You can see in the figure that these lines intersect at the point (900, 100) which verifies the answer that was obtained.

What is Linear Algebra? (cont.)

- Solving systems of linear equations is one of the most important applications of linear algebra.
- It has been argued that the majority of all mathematical problems encountered in scientific and industrial applications involve solving a linear system at some point.
- Linear applications arise in such diverse areas as engineering, chemistry, economics, business, ecology, biology, and psychology.

Linear Algebra

- Linear algebra is a fairly extensive subject that covers vectors and matrices, determinants, systems of linear equations, vector spaces and linear transformations, eigenvalue problems, and other topics.
- As an area of study it has a broad appeal in that it has many applications in engineering, physics, geometry, computer science, economics, and other areas.
- It also contributes to a deeper understanding of mathematics itself.

Definition of a Matrix

- A matrix is a rectangular array of numbers or functions which we will enclose in brackets.
- The plural of matrix is matrices.

If m and n are positive integers, then an $m \times n$ **matrix** is a rectangular array

$$\underbrace{\left.\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}\right\} m \text{ rows}}_{n \text{ columns}}$$

in which each **entry,** a_{ij}, of the matrix is a number. An $m \times n$ matrix (read "m by n") has m **rows** (horizontal lines) and n **columns** (vertical lines).

Examples of Matrices

Each matrix has the indicated size.

(a) Size: 1×1

$$\begin{bmatrix} 2 \end{bmatrix}$$

(b) Size: 2×2

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

(c) Size: 1×4

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

(d) Size: 3×2

$$\begin{bmatrix} e & \pi \\ 2 & \sqrt{2} \\ -7 & 4 \end{bmatrix}$$

Definition of a Matrix (cont.)

- One very common use of matrices is to represent systems of linear equations.
- The matrix derived from the coefficients and constant terms of a system of linear equations is called the **augmented matrix** of the system.
- The matrix containing only the coefficients of the system is called the **coefficient matrix** of the system. Here is an example.

System

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

Augmented Matrix

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

Coefficient Matrix

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

Operations with Matrices

- Matrix Addition
- Scalar Multiplication
- Matrix Multiplication

Matrix Addition

- You can **add** two matrices (of the same size) by adding their corresponding entries.

If $A = [a_{ij}]$ and $B = [b_{ij}]$ are matrices of size $m \times n$, then their **sum** is the $m \times n$ matrix given by

$$A + B = [a_{ij} + b_{ij}].$$

The sum of two matrices of different sizes is undefined.

Addition of Matrices

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

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

(c) $\begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix} + \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(d) The sum of

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

is undefined.

Scalar Multiplication

- When working with matrices, real numbers are referred to as **scalars.**
- You can multiply a matrix *A* by a scalar *c* by multiplying each entry in *A* by *c*.

If $A = [a_{ij}]$ is an $m \times n$ matrix and c is a scalar, then the **scalar multiple** of A by c is the $m \times n$ matrix given by

$$cA = [ca_{ij}].$$

You can use $-A$ to represent the scalar product $(-1)A$. If A and B are of the same size, $A - B$ represents the sum of A and $(-1)B$. That is,

$$A - B = A + (-1)B. \qquad \text{Subtraction of matrices}$$

Scalar Multiplication and Matrix Subtraction

For the matrices

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

find (a) $3A$, (b) $-B$, and (c) $3A - B$.

SOLUTION

(a) $$3A = 3\begin{bmatrix} 1 & 2 & 4 \\ -3 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 3(1) & 3(2) & 3(4) \\ 3(-3) & 3(0) & 3(-1) \\ 3(2) & 3(1) & 3(2) \end{bmatrix} = \begin{bmatrix} 3 & 6 & 12 \\ -9 & 0 & -3 \\ 6 & 3 & 6 \end{bmatrix}$$

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

(c) $$3A - B = \begin{bmatrix} 3 & 6 & 12 \\ -9 & 0 & -3 \\ 6 & 3 & 6 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 1 & -4 & 3 \\ -1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 6 & 12 \\ -10 & 4 & -6 \\ 7 & 0 & 4 \end{bmatrix}$$

Matrix Multiplication

If $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the **product** AB is an $m \times p$ matrix

$$AB = [c_{ij}]$$

where

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdot \cdot \cdot + a_{in}b_{nj}.$$

- This definition means that the entry in the *i*th row and the *j*th column of the product *AB* is obtained by multiplying the entries in the *i*th row of *A* by the corresponding entries in the *j*th column of *B* and then adding the results.

Finding the Product of Two Matrices

Find the product AB, where

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

SOLUTION First note that the product AB is defined because A has size 3×2 and B has size 2×2. Moreover, the product AB has size 3×2 and will take the form

$$\begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}.$$

To find c_{11} (the entry in the first row and first column of the product), multiply corresponding entries in the first row of A and the first column of B. That is,

$$c_{11} = (-1)(-3) + (3)(-4) = -9$$

$$\begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} -9 & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}.$$

Finding the Product of Two Matrices (cont.)

Similarly, to find c_{12}, multiply corresponding entries in the first row of A and the second column of B to obtain

$c_{12} = (-1)(2) + (3)(1) = 1$

$$\begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} -9 & 1 \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}.$$

Continuing this pattern produces the results shown below.

$$\begin{aligned} c_{21} &= (4)(-3) + (-2)(-4) = -4 \\ c_{22} &= (4)(2) + (-2)(1) = 6 \\ c_{31} &= (5)(-3) + (0)(-4) = -15 \\ c_{32} &= (5)(2) + (0)(1) = 10 \end{aligned}$$

The product is

$$AB = \begin{bmatrix} -1 & 3 \\ 4 & -2 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} -3 & 2 \\ -4 & 1 \end{bmatrix} = \begin{bmatrix} -9 & 1 \\ -4 & 6 \\ -15 & 10 \end{bmatrix}.$$

Matrix Multiplication (cont.)

- Be sure you understand that for the product of two matrices to be defined, the number of columns of the first matrix must equal the number of rows of the second matrix. That is,

$$\underset{m \times n}{A} \qquad \underset{n \times p}{B} \qquad = \qquad \underset{m \times p}{AB}.$$

equal

size of AB

- So, the product *BA* is not defined for matrices such as *A* and *B* in the previous example.

Matrix Multiplication (cont.)

- The general pattern for matrix multiplication is as follows. To obtain the element in the ith row and the jth column of the product AB, use the ith row of A and the jth column of B.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2p} \\ b_{31} & b_{32} & \cdots & b_{3j} & \cdots & b_{3p} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1j} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2j} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{i1} & c_{i2} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mj} & \cdots & c_{mp} \end{bmatrix}$$

$$a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = c_{ij}$$

Matrix Multiplication

(a) $$\underbrace{\begin{bmatrix} 1 & 0 & 3 \\ 2 & -1 & -2 \end{bmatrix}}_{2 \times 3} \underbrace{\begin{bmatrix} -2 & 4 & 2 \\ 1 & 0 & 0 \\ -1 & 1 & -1 \end{bmatrix}}_{3 \times 3} = \underbrace{\begin{bmatrix} -5 & 7 & -1 \\ -3 & 6 & 6 \end{bmatrix}}_{2 \times 3}$$

(b) $$\underbrace{\begin{bmatrix} 3 & 4 \\ -2 & 5 \end{bmatrix}}_{2 \times 2} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{2 \times 2} = \underbrace{\begin{bmatrix} 3 & 4 \\ -2 & 5 \end{bmatrix}}_{2 \times 2}$$

(c) $$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}}_{2 \times 2} \underbrace{\begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}}_{2 \times 2} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{2 \times 2}$$

(d) $$\underbrace{\begin{bmatrix} 1 & -2 & -3 \end{bmatrix}}_{1 \times 3} \underbrace{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}}_{3 \times 1} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{1 \times 1}$$

(e) $$\underbrace{\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}}_{3 \times 1} \underbrace{\begin{bmatrix} 1 & -2 & -3 \end{bmatrix}}_{1 \times 3} = \underbrace{\begin{bmatrix} 2 & -4 & -6 \\ -1 & 2 & 3 \\ 1 & -2 & -3 \end{bmatrix}}_{3 \times 3}$$

Properties of Matrix Operations

- This section begins to develop the algebra of matrices.
- You will see that this algebra shares many (but not all) of the properties of the algebra of real numbers.

Properties of Matrix Addition and Scalar Multiplication

If A, B, and C are $m \times n$ matrices and c and d are scalars, then the following properties are true.

1. $A + B = B + A$	**Commutative property of addition**
2. $A + (B + C) = (A + B) + C$	**Associative property of addition**
3. $(cd)A = c(dA)$	**Associative property of multiplication**
4. $1A = A$	**Multiplicative identity**
5. $c(A + B) = cA + cB$	**Distributive property**
6. $(c + d)A = cA + dA$	**Distributive property**

Properties of Matrix Multiplication

If A, B, and C are matrices (with sizes such that the given matrix products are defined) and c is a scalar, then the following properties are true.

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $c(AB) = (cA)B = A(cB)$

Applications of Matrix Operations

- Representing user–product interaction data as matrices
- Applying matrix operations in computer graphics and image processing
- Building collaborative filtering models for recommender systems

Properties of the Identity Matrix

- You will now look at a special type of square matrix that has 1’s on the main diagonal and 0’s elsewhere.

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

For instance, if $n = 1, 2,$ or 3, we have

$$\underset{1 \times 1}{I_1 = [1]}, \qquad \underset{2 \times 2}{I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}, \qquad \underset{3 \times 3}{I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}.$$

- When the order of the matrix is understood to be n, you may denote I_n simply as I.

Properties of the Identity Matrix (cont.)

If A is a matrix of size $m \times n$, then the following properties are true.

1. $AI_n = A$
2. $I_mA = A$

As a special case of this theorem, note that if A is a *square* matrix of order n, then

$$AI_n = I_nA = A.$$

The Transpose of a Matrix

The **transpose** of a matrix is formed by writing its columns as rows. For instance, if A is the $m \times n$ matrix shown by

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix},$$

Size: $m \times n$

then the transpose, denoted by A^T, is the $n \times m$ matrix below

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \cdots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \cdots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \cdots & a_{m3} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \cdots & a_{mn} \end{bmatrix}.$$

Size: $n \times m$

The Transpose of a Matrix (cont.)

Find the transpose of each matrix.

(a) $A = \begin{bmatrix} 2 \\ 8 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ (c) $C = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (d) $D = \begin{bmatrix} 0 & 1 \\ 2 & 4 \\ 1 & -1 \end{bmatrix}$

SOLUTION (a) $A^T = [2 \quad 8]$ (b) $B^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ (c) $C^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(d) $D^T = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 4 & -1 \end{bmatrix}$

- Note that the square matrix in part (c) of the above example is equal to its transpose.
- Such a matrix is called **symmetric.**
- A matrix A is symmetric if $A = A^T$. From this definition it is clear that a symmetric matrix must be square. Also, if $A = [a_{ij}]$ is a symmetric matrix, then $a_{ij} = a_{ji}$ for all $i \neq j$.

Properties of Transposes

If A and B are matrices (with sizes such that the given matrix operations are defined) and c is a scalar, then the following properties are true.

1. $(A^T)^T = A$ — Transpose of a transpose
2. $(A + B)^T = A^T + B^T$ — Transpose of a sum
3. $(cA)^T = c(A^T)$ — Transpose of a scalar multiple
4. $(AB)^T = B^T A^T$ — Transpose of a product

The Product of a Matrix and Its Transpose

For the matrix

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

find the product AA^T and show that it is symmetric.

SOLUTION Because

$$AA^T = \begin{bmatrix} 1 & 3 \\ 0 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 3 & -2 & -1 \end{bmatrix} = \begin{bmatrix} 10 & -6 & -5 \\ -6 & 4 & 2 \\ -5 & 2 & 5 \end{bmatrix}$$

it follows that $AA^T = (AA^T)^T$, so AA^T is symmetric.

The Inverse of a Matrix

An $n \times n$ matrix A is **invertible** (or **nonsingular**) if there exists an $n \times n$ matrix B such that

$$AB = BA = I_n$$

where I_n is the identity matrix of order n. The matrix B is called the (multiplicative) **inverse** of A. A matrix that does not have an inverse is called **noninvertible** (or **singular**).

Finding the Inverse of a Matrix

Find the inverse of the matrix

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

SOLUTION To find the inverse of A, try to solve the matrix equation $AX = I$ for X.

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x_{11} + 4x_{21} & x_{12} + 4x_{22} \\ -x_{11} - 3x_{21} & -x_{12} - 3x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Now, by equating corresponding entries, you obtain the two systems of linear equations shown below.

$$\begin{aligned} x_{11} + 4x_{21} &= 1 \\ -x_{11} - 3x_{21} &= 0 \end{aligned} \qquad \begin{aligned} x_{12} + 4x_{22} &= 0 \\ -x_{12} - 3x_{22} &= 1 \end{aligned}$$

Solving the first system, you find that the first column of X is $x_{11} = -3$ and $x_{21} = 1$. Similarly, solving the second system, you find that the second column of X is $x_{12} = -4$ and $x_{22} = 1$. The inverse of A is

$$X = A^{-1} = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}.$$

Try using matrix multiplication to check this result.

Properties of Inverse Matrices

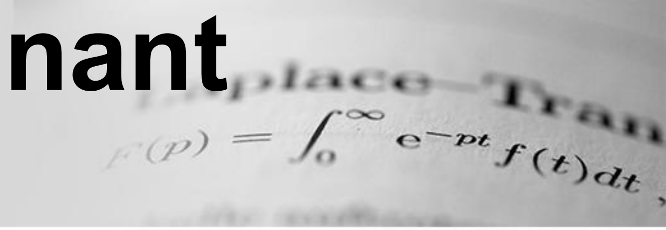
If A is an invertible matrix, k is a positive integer, and c is a scalar not equal to zero, then A^{-1}, A^k, cA, and A^T are invertible and the following are true.

1. $(A^{-1})^{-1} = A$
2. $(A^k)^{-1} = A^{-1}A^{-1} \cdot \cdot \cdot A^{-1} = (A^{-1})^k$
3. $(cA)^{-1} = \dfrac{1}{c}A^{-1}, \; c \neq 0$
4. $(A^T)^{-1} = (A^{-1})^T$

The Determinant of a Matrix

- Every *square* matrix can be associated with a real number called its determinant.
- Historically, the use of determinants arose from the recognition of special patterns that occur in the solutions of systems of linear equations.

Definition of the Determinant of a 2 × 2 Matrix



The **determinant** of the matrix

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

is given by

$$\det(A) = |A| = a_{11}a_{22} - a_{21}a_{12}.$$

REMARK: In this text, $\det(A)$ and $|A|$ are used interchangeably to represent the determinant of a matrix. Vertical bars are also used to denote the absolute value of a real number; the context will show which use is intended. Furthermore, it is common practice to delete the matrix brackets and write

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \quad \text{instead of} \quad \left| \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right|.$$

A convenient method for remembering the formula for the determinant of a 2×2 matrix is shown in the diagram below.

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

The determinant is the difference of the products of the two diagonals of the matrix. Note that the order is important, as demonstrated above.

The Determinant of a Matrix of Order 2

Find the determinant of each matrix.

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

SOLUTION

(a) $|A| = \begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix} = 2(2) - 1(-3) = 4 + 3 = 7$

(b) $|B| = \begin{vmatrix} 2 & 1 \\ 4 & 2 \end{vmatrix} = 2(2) - 4(1) = 4 - 4 = 0$

(c) $|C| = \begin{vmatrix} 0 & 3 \\ 2 & 4 \end{vmatrix} = 0(4) - 2(3) = 0 - 6 = -6$

Definitions of Minors and Cofactors of a Matrix

If A is a square matrix, then the **minor** M_{ij} of the element a_{ij} is the determinant of the matrix obtained by deleting the ith row and jth column of A. The **cofactor** C_{ij} is given by

$$C_{ij} = (-1)^{i+j} M_{ij}.$$

For example, if A is a 3×3 matrix, then the minors and cofactors of a_{21} and a_{22} are as shown in the diagram below.

Minor of a_{21}

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad M_{21} = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

Delete row 2 and column 1.

Cofactor of a_{21}

$$\begin{aligned} C_{21} &= (-1)^{2+1} M_{21} \\ &= -M_{21} \end{aligned}$$

Minor of a_{22}

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad M_{22} = \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

Delete row 2 and column 2.

Cofactor of a_{22}

$$\begin{aligned} C_{22} &= (-1)^{2+2} M_{22} \\ &= M_{22} \end{aligned}$$

Definitions of Minors and Cofactors of a Matrix (cont.)

As you can see, the minors and cofactors of a matrix can differ only in sign. To obtain the cofactors of a matrix, first find the minors and then apply the checkerboard pattern of $+$'s and $-$'s shown below.

Sign Pattern for Cofactors

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$$

3×3 *matrix*

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

4×4 *matrix*

$$\begin{bmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

$n \times n$ *matrix*

Note that *odd* positions (where $i + j$ is odd) have negative signs, and even positions (where $i + j$ is even) have positive signs.

Find the Minors and Cofactors of a Matrix

Find all the minors and cofactors of

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

SOLUTION To find the minor M_{11}, delete the first row and first column of A and evaluate the determinant of the resulting matrix.

$$\begin{bmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & 0 & 1 \end{bmatrix}, \quad M_{11} = \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} = -1(1) - 0(2) = -1$$

Similarly, to find M_{12}, delete the first row and second column.

$$\begin{bmatrix} 0 & 2 & 1 \\ 3 & -1 & 2 \\ 4 & 0 & 1 \end{bmatrix}, \quad M_{12} = \begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = 3(1) - 4(2) = -5$$

Continuing this pattern, you obtain

$$\begin{array}{lll} M_{11} = -1 & M_{12} = -5 & M_{13} = 4 \\ M_{21} = 2 & M_{22} = -4 & M_{23} = -8 \\ M_{31} = 5 & M_{32} = -3 & M_{33} = -6. \end{array}$$

Now, to find the cofactors, combine the checkerboard pattern of signs with these minors to obtain

$$\begin{array}{lll} C_{11} = -1 & C_{12} = 5 & C_{13} = 4 \\ C_{21} = -2 & C_{22} = -4 & C_{23} = 8 \\ C_{31} = 5 & C_{32} = 3 & C_{33} = -6. \end{array}$$

Definition of the Determinant of a Matrix

If A is a square matrix (of order 2 or greater), then the determinant of A is the sum of the entries in the first row of A multiplied by their cofactors. That is,

$$\det(A) = |A| = \sum_{j=1}^{n} a_{1j}C_{1j} = a_{11}C_{11} + a_{12}C_{12} + \cdot\cdot\cdot + a_{1n}C_{1n}.$$

The Determinant of a Matrix of Order 3

Find the determinant of

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

SOLUTION This matrix is the same as the one in Example 2. There you found the cofactors of the entries in the first row to be

$$C_{11} = -1, \qquad C_{12} = 5, \qquad C_{13} = 4.$$

By the definition of a determinant, you have

$$\begin{aligned} |A| &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} && \text{First row expansion} \\ &= 0(-1) + 2(5) + 1(4) = 14. \end{aligned}$$

Although the determinant is defined as an expansion by the cofactors in the first row, it can be shown that the determinant can be evaluated by expanding by *any* row or column. For instance, you could expand the 3×3 matrix in Example 3 by the second row to obtain

$$\begin{aligned} |A| &= a_{21}C_{21} + a_{22}C_{22} + a_{23}C_{23} && \text{Second row expansion} \\ &= 3(-2) + (-1)(-4) + 2(8) = 14 \end{aligned}$$

or by the first column to obtain

$$\begin{aligned} |A| &= a_{11}C_{11} + a_{21}C_{21} + a_{31}C_{31} && \text{First column expansion} \\ &= 0(-1) + 3(-2) + 4(5) = 14. \end{aligned}$$

Expansion by Cofactors

Let A be a square matrix of order n. Then the determinant of A is given by

$$\det(A) = |A| = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdot\cdot\cdot + a_{in}C_{in}$$

or

$$\det(A) = |A| = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdot\cdot\cdot + a_{nj}C_{nj}.$$

The Determinant of a Matrix of Order 4

Find the determinant of

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

SOLUTION By inspecting this matrix, you can see that three of the entries in the third column are zeros. You can eliminate some of the work in the expansion by using the third column.

$$|A| = 3(C_{13}) + 0(C_{23}) + 0(C_{33}) + 0(C_{43})$$

Because C_{23}, C_{33}, and C_{43} have zero coefficients, you need only find the cofactor C_{13}. To do this, delete the first row and third column of A and evaluate the determinant of the resulting matrix.

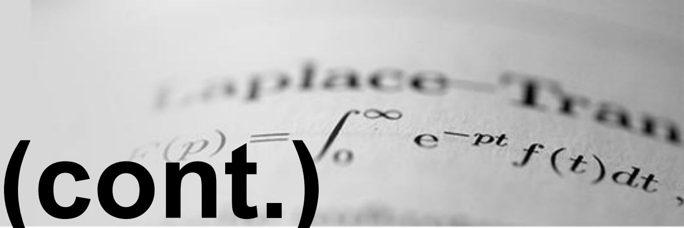
$$C_{13} = (-1)^{1+3}\begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 4 & -2 \end{vmatrix} = \begin{vmatrix} -1 & 1 & 2 \\ 0 & 2 & 3 \\ 3 & 4 & -2 \end{vmatrix}$$

Expanding by cofactors in the second row yields

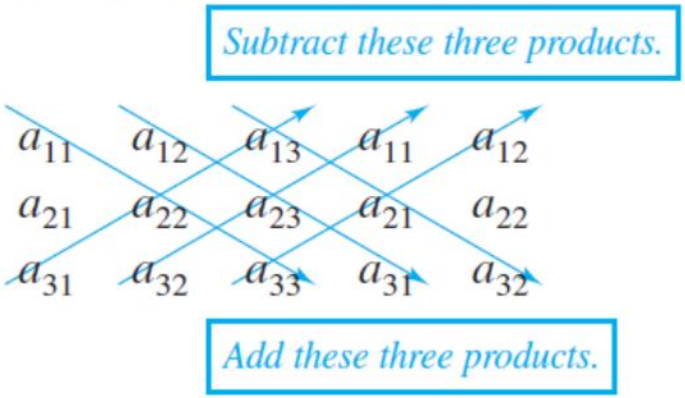
$$\begin{aligned} C_{13} &= (0)(-1)^{2+1}\begin{vmatrix} 1 & 2 \\ 4 & -2 \end{vmatrix} + (2)(-1)^{2+2}\begin{vmatrix} -1 & 2 \\ 3 & -2 \end{vmatrix} + (3)(-1)^{2+3}\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} \\ &= 0 + 2(1)(-4) + 3(-1)(-7) = 13. \end{aligned}$$

You obtain $|A| = 3(13) = 39$.

Definition of the Determinant of a Matrix (cont.)



There is an alternative method commonly used for evaluating the determinant of a 3×3 matrix A. To apply this method, copy the first and second columns of A to form fourth and fifth columns. The determinant of A is then obtained by adding (or subtracting) the products of the six diagonals, as shown in the following diagram.



Try confirming that the determinant of A is

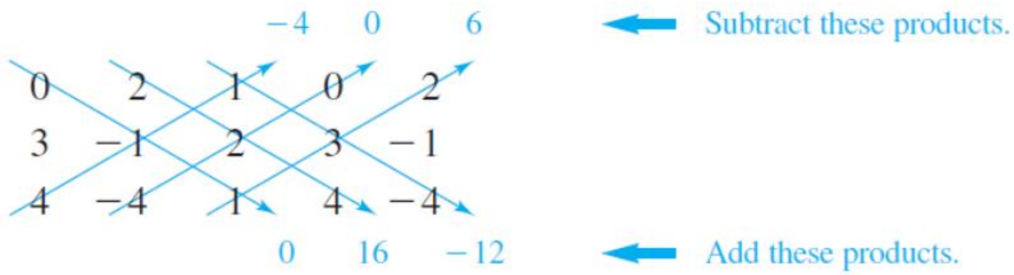
$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}.$$

The Determinant of a Matrix of Order 3

Find the determinant of

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

SOLUTION Begin by recopying the first two columns and then computing the six diagonal products as follows.



Now, by adding the lower three products and subtracting the upper three products, you can find the determinant of A to be $|A| = 0 + 16 + (-12) - (-4) - 0 - 6 = 2$.

Conditions That Yield a Zero Determinant

If A is a square matrix and any one of the following conditions is true, then $\det(A) = 0$.

1. An entire row (or an entire column) consists of zeros.
2. Two rows (or columns) are equal.
3. One row (or column) is a multiple of another row (or column).

$$\begin{vmatrix} 0 & 0 & 0 \\ 2 & 4 & -5 \\ 3 & -5 & 2 \end{vmatrix} = 0, \qquad \begin{vmatrix} 1 & -2 & 4 \\ 0 & 1 & 2 \\ 1 & -2 & 4 \end{vmatrix} = 0, \qquad \begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & -6 \\ -2 & 0 & 6 \end{vmatrix} = 0.$$

The first row has all zeros.

The first and third rows are the same.

The third column is a multiple of the first column.

Properties of Determinants

If A and B are square matrices of order n, then

$$\det(AB) = \det(A)\det(B).$$

If A is an $n \times n$ matrix and c is a scalar, then the determinant of cA is given by

$$\det(cA) = c^n \det(A).$$

A square matrix A is invertible (nonsingular) if and only if

$$\det(A) \neq 0.$$

If A is invertible, then

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

If A is a square matrix, then

$$\det(A) = \det(A^T).$$

The Adjoint of a Matrix

- If A is a square matrix, then the **matrix of cofactors** of A has the form

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

- The transpose of this matrix is called the adjoint of A and is denoted by adj(A). That is,

$$\operatorname{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$

Finding the Adjoint of a Square Matrix

Find the adjoint of

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

SOLUTION The cofactor C_{11} is given by

$$\begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{bmatrix} \longrightarrow C_{11} = (-1)^2 \begin{vmatrix} -2 & 1 \\ 0 & -2 \end{vmatrix} = 4.$$

Continuing this process produces the following matrix of cofactors of A.

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

The transpose of this matrix is the adjoint of A. That is,

$$\text{adj}(A) = \begin{bmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{bmatrix}.$$

The Inverse of a Matrix Given by Its Adjoint

If A is an $n \times n$ invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A).$$

Using the Adjoint of a Matrix to Find Its Inverse

Use the adjoint of

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

to find A^{-1}.

SOLUTION The determinant of this matrix is 3. Using the adjoint of A (found in Example 1), you can find the inverse of A to be

$$A^{-1} = \frac{1}{|A|}\operatorname{adj}(A) = \tfrac{1}{3}\begin{bmatrix} 4 & 6 & 7 \\ 1 & 0 & 1 \\ 2 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & 2 & \frac{7}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & 1 & \frac{2}{3} \end{bmatrix}.$$

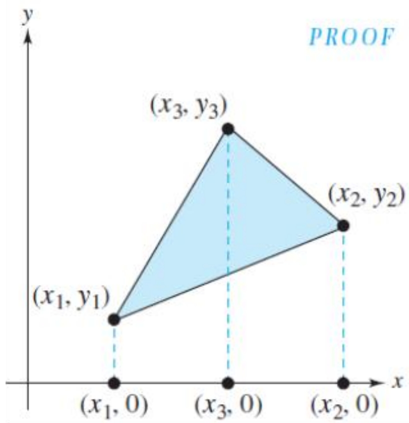
You can check to see that this matrix is the inverse of A by multiplying to obtain

$$AA^{-1} = \begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 1 \\ 1 & 0 & -2 \end{bmatrix}\begin{bmatrix} \frac{4}{3} & 2 & \frac{7}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{2}{3} & 1 & \frac{2}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Applications of Determinants

- Area, Volume, and Equations of Lines and Planes
- Two-Point Form of the Equation of a Line
- Volume of a Tetrahedron

Area, Volume, and Equations of Lines and Planes



The area of the triangle whose vertices are (x_1, y_1), (x_2, y_2), and (x_3, y_3) is given by

$$\text{Area} = \pm\frac{1}{2}\det\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix},$$

where the sign $(\pm)$ is chosen to yield a positive area.

Finding the Area of a Triangle

Find the area of the triangle whose vertices are (1, 0), (2, 2), and (4, 3).

SOLUTION It is not necessary to know the relative positions of the three vertices. Simply evaluate the determinant

$$\frac{1}{2}\begin{vmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 4 & 3 & 1 \end{vmatrix} = -\frac{3}{2}$$

and conclude that the area of the triangle is $\frac{3}{2}$.

Two-Point Form of the Equation of a Line

An equation of the line passing through the distinct points (x_1, y_1) and (x_2, y_2) is given by

$$\det\begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = 0.$$

Finding an Equation of the Line Passing Through Two Points

Find an equation of the line passing through the points $(2, 4)$ and $(-1, 3)$.

SOLUTION Applying the determinant formula for the equation of a line passing through two points produces

$$\begin{vmatrix} x & y & 1 \\ 2 & 4 & 1 \\ -1 & 3 & 1 \end{vmatrix} = 0.$$

To evaluate this determinant, expand by cofactors along the top row to obtain

$$x\begin{vmatrix} 4 & 1 \\ 3 & 1 \end{vmatrix} - y\begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} + 1\begin{vmatrix} 2 & 4 \\ -1 & 3 \end{vmatrix} = 0$$

$$x - 3y + 10 = 0.$$

An equation of the line is $x - 3y = -10$.

Volume of a Tetrahedron

- เตตระฮีดรอน ตรงกับภาษาอังกฤษว่า Tetrahedron (อ่านว่า tet-tra-he-dron) หมายถึงโครงสร้างปิระมิดฐานสามเหลี่ยม ซึ่งประกอบด้วย ด้าน (face) สามเหลี่ยม 4 ด้าน มุม (vertice) 4 มุม ขอบ (edge) 6 ขอบ เตตระฮีดรอนเป็นรูปทรงที่จัดอยู่ในประเภทพลาโทนิก (platonic)

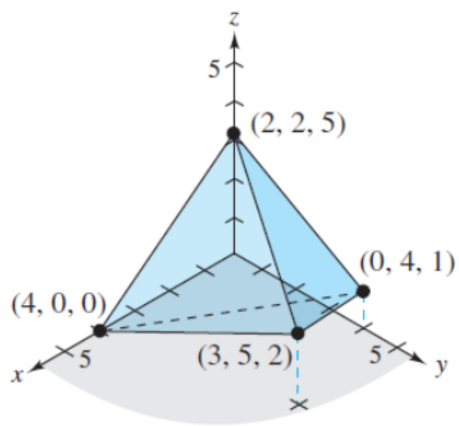
The volume of the tetrahedron whose vertices are (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), and (x_4, y_4, z_4) is given by

$$\text{Volume} = \pm\frac{1}{6}\det\begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{bmatrix},$$

where the sign $(\pm)$ is chosen to yield a positive volume.

Finding the Volume of a Tetrahedron

Find the volume of the tetrahedron whose vertices are (0, 4, 1), (4, 0, 0), (3, 5, 2), and (2, 2, 5)



SOLUTION Using the determinant formula for volume produces

$$\frac{1}{6}\begin{vmatrix} 0 & 4 & 1 & 1 \\ 4 & 0 & 0 & 1 \\ 3 & 5 & 2 & 1 \\ 2 & 2 & 5 & 1 \end{vmatrix} = \tfrac{1}{6}(-72) = -12.$$

The volume of the tetrahedron is 12.

Applications of Determinants and Inverses

- Solving linear regression analytically using matrix inversion
- Checking solvability and singularity of linear systems
- Ensuring numerical stability in algorithmic computation

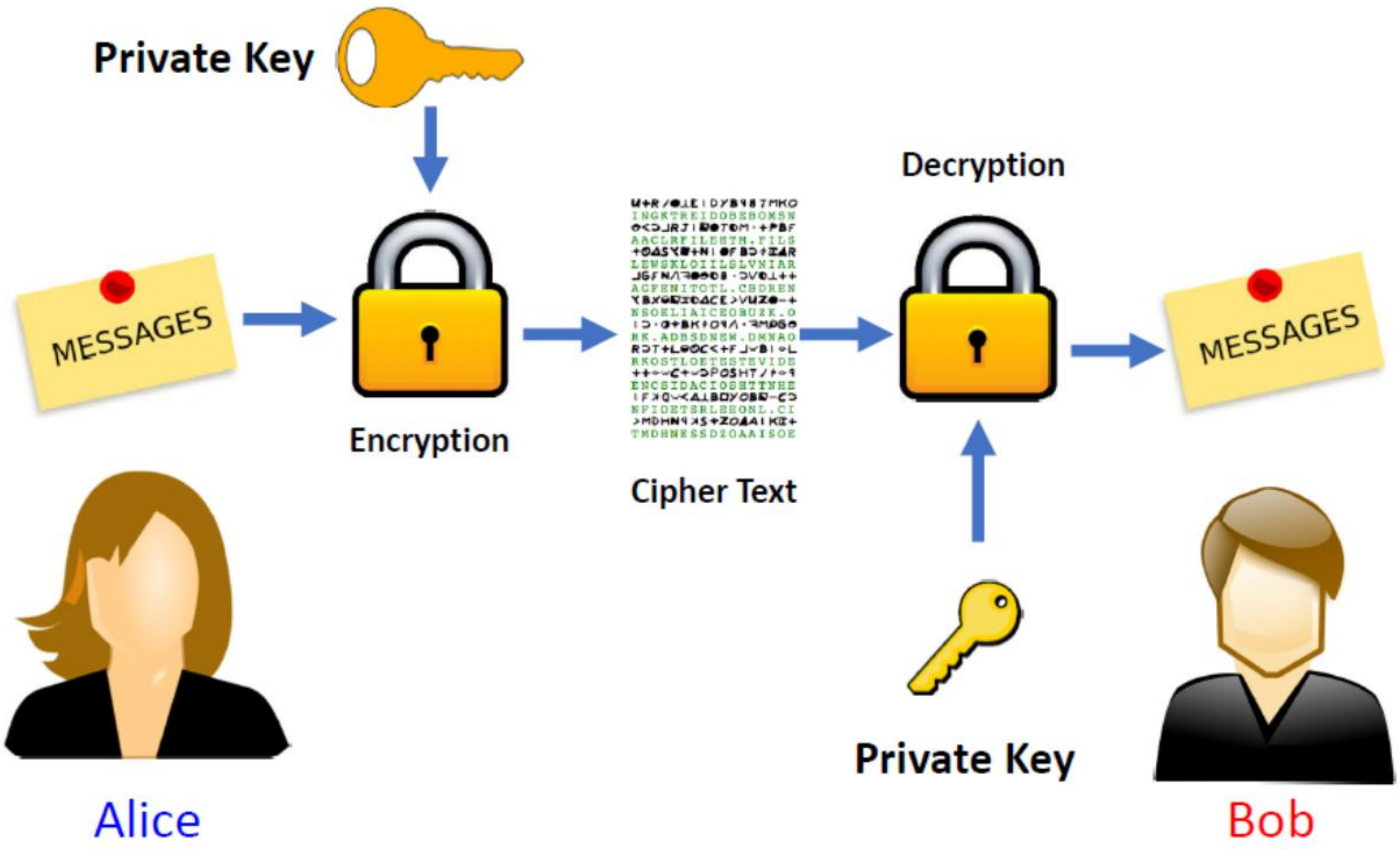
Applications of Matrix Cryptography

- Encoding and decoding data with matrix transformations
- Teaching symmetric cryptography using algebraic concepts
- Exploring matrix-based encryption methods in cybersecurity

Cryptography

- A **cryptogram** is a message written according to a secret code (the Greek word *kryptos* means “hidden”). This section describes a method of using matrix multiplication to **encode** and **decode** messages.

Symmetric Cryptography



Cryptography (cont.)

- Begin by assigning a number to each letter in the alphabet (with 0 assigned to a blank space), as follows.

0 = _	14 = N
1 = A	15 = O
2 = B	16 = P
3 = C	17 = Q
4 = D	18 = R
5 = E	19 = S
6 = F	20 = T
7 = G	21 = U
8 = H	22 = V
9 = I	23 = W
10 = J	24 = X
11 = K	25 = Y
12 = L	26 = Z
13 = M	

- Then the message is converted to numbers and partitioned into **uncoded row matrices**, each having *n* entries.

Forming Uncoded Row Matrices

Write the uncoded row matrices of size 1×3 for the message MEET ME MONDAY.

SOLUTION Partitioning the message (including blank spaces, but ignoring punctuation) into groups of three produces the following uncoded row matrices.

$$\begin{array}{ccccccccccccccc} [13 & 5 & 5] & [20 & 0 & 13] & [5 & 0 & 13] & [15 & 14 & 4] & [1 & 25 & 0] \\ M & E & E & T & \_ & M & E & \_ & M & O & N & D & A & Y & \_ \end{array}$$

Note that a blank space is used to fill out the last uncoded row matrix.

Cryptography (cont.)

- To **encode** a message, choose an $n \times n$ invertible matrix A and multiply the uncoded row matrices (on the right) by A to obtain **coded row matrices.**

Encoding a Message

Use the matrix

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

to encode the message MEET ME MONDAY.

SOLUTION The coded row matrices are obtained by multiplying each of the uncoded row matrices found in Example 4 by the matrix A, as follows.

Uncoded Row Matrix	*Encoding Matrix A*	*Coded Row Matrix*

$$[13 \quad 5 \quad 5]\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & 3 \\ 1 & -1 & -4 \end{bmatrix} = [13 \quad -26 \quad 21]$$

$$[20 \quad 0 \quad 13]\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & 3 \\ 1 & -1 & -4 \end{bmatrix} = [33 \quad -53 \quad -12]$$

$$[5 \quad 0 \quad 13]\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & 3 \\ 1 & -1 & -4 \end{bmatrix} = [18 \quad -23 \quad -42]$$

$$[15 \quad 14 \quad 4]\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & 3 \\ 1 & -1 & -4 \end{bmatrix} = [5 \quad -20 \quad 56]$$

$$[1 \quad 25 \quad 0]\begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & 3 \\ 1 & -1 & -4 \end{bmatrix} = [-24 \quad 23 \quad 77]$$

The sequence of coded row matrices is

$$[13 \quad -26 \quad 21][33 \quad -53 \quad -12][18 \quad -23 \quad -42][5 \quad -20 \quad 56][-24 \quad 23 \quad 77].$$

Finally, removing the brackets produces the cryptogram below.

$$13 \quad -26 \quad 21 \quad 33 \quad -53 \quad -12 \quad 18 \quad -23 \quad -42 \quad 5 \quad -20 \quad 56 \quad -24 \quad 23 \quad 77$$

Cryptography (cont.)

- For those who do not know the matrix *A*, decoding the cryptogram found in the previous example is difficult.
- But for an authorized receiver who knows the matrix *A*, decoding is simple.
- The receiver need only multiply the coded row matrices by A^{-1} to retrieve the uncoded row matrices.
- In other words, if

  $$X = [x_1 \ \ x_2 \ \cdot \ \cdot \ \cdot \ \ x_n]$$

  is an uncoded $1 \times n$ matrix, then $Y = XA$ is the corresponding encoded matrix.
- The receiver of the encoded matrix can decode *Y* by multiplying on the right by A^{-1} to obtain

  $$YA^{-1} = (XA)A^{-1} = X.$$

Decoding a Message

Use the inverse of the matrix

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

to decode the cryptogram

$$13 \quad -26 \quad 21 \quad 33 \quad -53 \quad -12 \quad 18 \quad -23 \quad -42 \quad 5 \quad -20 \quad 56 \quad -24 \quad 23 \quad 77.$$

SOLUTION Begin by using Gauss-Jordan elimination to find A^{-1}.

$$\underbrace{\left[\begin{array}{rrr|rrr} 1 & -2 & 2 & 1 & 0 & 0 \\ -1 & 1 & 3 & 0 & 1 & 0 \\ 1 & -1 & -4 & 0 & 0 & 1 \end{array}\right]}_{[A \,\vdots\, I]} \longrightarrow \underbrace{\left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -1 & -10 & -8 \\ 0 & 1 & 0 & -1 & -6 & -5 \\ 0 & 0 & 1 & 0 & -1 & -1 \end{array}\right]}_{[I \,\vdots\, A^{-1}]}$$

Now, to decode the message, partition the message into groups of three to form the coded row matrices

$$[13 \ -26 \quad 21][33 \ -53 \ -12][18 \ -23 \ -42][5 \ -20 \quad 56][-24 \quad 23 \quad 77].$$

Decoding a Message (cont.)

To obtain the decoded row matrices, multiply each coded row matrix by A^{-1} (on the right).

$$\begin{array}{ccc} \textit{Coded Row Matrix} & \textit{Decoding Matrix } A^{-1} & \textit{Decoded Row Matrix} \end{array}$$

$$[13 \quad -26 \quad 21]\begin{bmatrix} -1 & -10 & -8 \\ -1 & -6 & -5 \\ 0 & -1 & -1 \end{bmatrix} = [13 \quad 5 \quad 5]$$

$$[33 \quad -53 \quad -12]\begin{bmatrix} -1 & -10 & -8 \\ -1 & -6 & -5 \\ 0 & -1 & -1 \end{bmatrix} = [20 \quad 0 \quad 13]$$

$$[18 \quad -23 \quad -42]\begin{bmatrix} -1 & -10 & -8 \\ -1 & -6 & -5 \\ 0 & -1 & -1 \end{bmatrix} = [5 \quad 0 \quad 13]$$

$$[5 \quad -20 \quad 56]\begin{bmatrix} -1 & -10 & -8 \\ -1 & -6 & -5 \\ 0 & -1 & -1 \end{bmatrix} = [15 \quad 14 \quad 4]$$

$$[-24 \quad 23 \quad 77]\begin{bmatrix} -1 & -10 & -8 \\ -1 & -6 & -5 \\ 0 & -1 & -1 \end{bmatrix} = [1 \quad 25 \quad 0]$$

The sequence of decoded row matrices is

$$[13 \quad 5 \quad 5][20 \quad 0 \quad 13][5 \quad 0 \quad 13][15 \quad 14 \quad 4][1 \quad 25 \quad 0]$$

and the message is

13	5	5	20	0	13	5	0	13	15	14	4	1	25	0.
M	*E*	*E*	*T*	__	*M*	*E*	__	*M*	*O*	*N*	*D*	*A*	*Y*	__

Introduction to Systems of Linear Equations

- Linear algebra is a branch of mathematics rich in theory and applications.
- Because linear algebra arose from the study of systems of linear equations, you shall begin with linear equations.

Linear Equations in n Variables

A **linear equation in n variables** $x_1, x_2, x_3, \ldots, x_n$ has the form

$$a_1x_1 + a_2x_2 + a_3x_3 + \cdots + a_nx_n = b.$$

The **coefficients** $a_1, a_2, a_3, \ldots, a_n$ are real numbers, and the **constant term** b is a real number. The number a_1 is the **leading coefficient,** and x_1 is the **leading variable.**

Examples of Linear Equations and Nonlinear Equations

Each equation is linear.

(a) $3x + 2y = 7$

(b) $\frac{1}{2}x + y - \pi z = \sqrt{2}$

(c) $x_1 - 2x_2 + 10x_3 + x_4 = 0$

(d) $\left(\sin \frac{\pi}{2}\right)x_1 - 4x_2 = e^2$

Each equation is not linear.

(a) $xy + z = 2$

(b) $e^x - 2y = 4$

(c) $\sin x_1 + 2x_2 - 3x_3 = 0$

(d) $\frac{1}{x} + \frac{1}{y} = 4$

Systems of Linear Equations

A **system of *m* linear equations in *n* variables** is a set of *m* equations, each of which is linear in the same *n* variables:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}
$$

Systems of Two Equations in Two Variables

Solve each system of linear equations, and graph each system as a pair of straight lines.

(a) $x + y = 3$ $\quad$ (b) $x + y = 3$ $\quad$ (c) $x + y = 3$

$\quad x - y = -1$ $\quad$ $2x + 2y = 6$ $\quad$ $x + y = 1$

SOLUTION

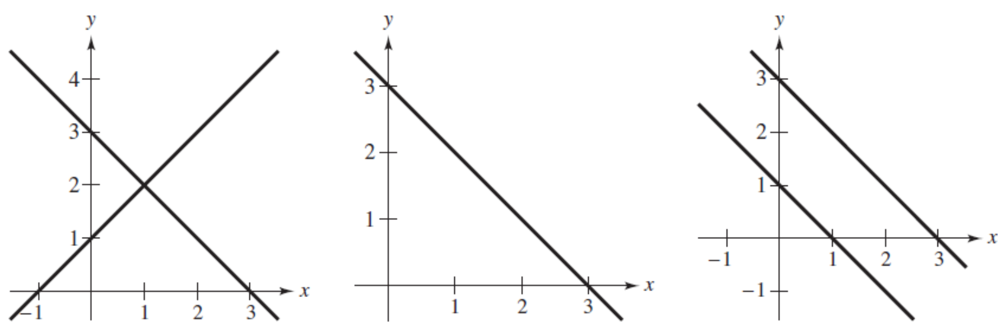
(a) This system has exactly one solution, $x = 1$ and $y = 2$. The solution can be obtained by adding the two equations to give $2x = 2$, which implies $x = 1$ and so $y = 2$. The graph of this system is represented by two *intersecting* lines, as shown in Figure 1.1(a).

(b) This system has an infinite number of solutions because the second equation is the result of multiplying both sides of the first equation by 2. A parametric representation of the solution set is shown as

$$x = 3 - t, \qquad y = t, \qquad t \text{ is any real number.}$$

The graph of this system is represented by two *coincident* lines, as shown in Figure 1.1(b).

(c) This system has no solution because it is impossible for the sum of two numbers to be 3 and 1 simultaneously. The graph of this system is represented by two *parallel* lines, as shown in Figure 1.1(c).



(a) Two intersecting lines: $x + y = 3$, $x - y = -1$

(b) Two coincident lines: $x + y = 3$, $2x + 2y = 6$

(c) Two parallel lines: $x + y = 3$, $x + y = 1$

Figure 1.1

Number of Solutions of a System of Linear Equations

For a system of linear equations in n variables, precisely one of the following is true.

1. The system has exactly one solution (consistent system).
2. The system has an infinite number of solutions (consistent system).
3. The system has no solution (inconsistent system).

Solving a System of Linear Equations

Which system is easier to solve algebraically?

$$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned} \qquad\qquad \begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ z &= 2 \end{aligned}$$

The system on the right is clearly easier to solve. This system is in **row-echelon form,** which means that it follows a stair-step pattern and has leading coefficients of 1. To solve such a system, use a procedure called **back-substitution.**

Systems of Equations

If A is an invertible matrix, then the system of linear equations $A\mathbf{x} = \mathbf{b}$ has a unique solution given by

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

PROOF Because A is nonsingular, the steps shown below are valid.

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ A^{-1}A\mathbf{x} &= A^{-1}\mathbf{b} \\ I\mathbf{x} &= A^{-1}\mathbf{b} \\ \mathbf{x} &= A^{-1}\mathbf{b} \end{aligned}$$

This solution is unique because if $\mathbf{x}_1$ and $\mathbf{x}_2$ were two solutions, you could apply the cancellation property to the equation $A\mathbf{x}_1 = \mathbf{b} = A\mathbf{x}_2$ to conclude that $\mathbf{x}_1 = \mathbf{x}_2$.

If C is an invertible matrix, then the following properties hold.

1. If $AC = BC$, then $A = B$. **Right cancellation property**
2. If $CA = CB$, then $A = B$. **Left cancellation property**

Solving a System of Equations Using an Inverse Matrix

Use an inverse matrix to solve each system.

(a) $\begin{aligned} 2x + 3y + z &= -1 \\ 3x + 3y + z &= 1 \\ 2x + 4y + z &= -2 \end{aligned}$

(b) $\begin{aligned} 2x + 3y + z &= 4 \\ 3x + 3y + z &= 8 \\ 2x + 4y + z &= 5 \end{aligned}$

(c) $\begin{aligned} 2x + 3y + z &= 0 \\ 3x + 3y + z &= 0 \\ 2x + 4y + z &= 0 \end{aligned}$

SOLUTION First note that the coefficient matrix for each system is

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

Using Gauss-Jordan elimination, you can find A^{-1} to be

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

To solve each system, use matrix multiplication, as follows.

(a) $\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$

The solution is $x = 2$, $y = -1$, and $z = -2$.

(b) $\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix}\begin{bmatrix} 4 \\ 8 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ -7 \end{bmatrix}$

The solution is $x = 4$, $y = 1$, and $z = -7$.

(c) $\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

The solution is trivial: $x = 0$, $y = 0$, and $z = 0$.

Cramer's Rule

Cramer's Rule, named after Gabriel Cramer (1704–1752), is a formula that uses determinants to solve a system of n linear equations in n variables. This rule can be applied only to systems of linear equations that have unique solutions.

To see how Cramer's Rule arises, look at the solution of a general system involving two linear equations in two variables.

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

Multiplying the first equation by $-a_{21}$ and the second by a_{11} and adding the results produces

$$\begin{array}{rcl} -a_{21}a_{11}x_1 - a_{21}a_{12}x_2 &=& -a_{21}b_1 \\ a_{11}a_{21}x_1 + a_{11}a_{22}x_2 &=& a_{11}b_2 \\ \hline (a_{11}a_{22} - a_{21}a_{12})x_2 &=& a_{11}b_2 - a_{21}b_1. \end{array}$$

Solving for x_2 (provided that $a_{11}a_{22} - a_{21}a_{12} \neq 0$) produces

$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}.$$

In a similar way, you can solve for x_1 to obtain

$$x_1 = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{21}a_{12}}.$$

Cramer's Rule (cont.)

Finally, recognizing that the numerators and denominators of both x_1 and x_2 can be represented as determinants, you have

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad a_{11}a_{22} - a_{21}a_{12} \neq 0.$$

The denominator for both x_1 and x_2 is simply the determinant of the coefficient matrix A. The determinant forming the numerator of x_1 can be obtained from A by replacing its first column by the column representing the constants of the system. The determinant forming the numerator of x_2 can be obtained in a similar way. These two determinants are denoted by $|A_1|$ and $|A_2|$, as follows.

$$|A_1| = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} \quad \text{and} \quad |A_2| = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

You have $x_1 = \dfrac{|A_1|}{|A|}$ and $x_2 = \dfrac{|A_2|}{|A|}$. This determinant form of the solution is called **Cramer's Rule.**

Using Cramer's Rule

Use Cramer's Rule to solve the system of linear equations.

$$\begin{aligned} 4x_1 - 2x_2 &= 10 \\ 3x_1 - 5x_2 &= 11 \end{aligned}$$

SOLUTION First find the determinant of the coefficient matrix.

$$|A| = \begin{vmatrix} 4 & -2 \\ 3 & -5 \end{vmatrix} = -14$$

Because $|A| \neq 0$, you know the system has a unique solution, and applying Cramer's Rule produces

$$x_1 = \frac{|A_1|}{|A|} = \frac{\begin{vmatrix} 10 & -2 \\ 11 & -5 \end{vmatrix}}{-14} = \frac{-28}{-14} = 2$$

$$x_2 = \frac{|A_2|}{|A|} = \frac{\begin{vmatrix} 4 & 10 \\ 3 & 11 \end{vmatrix}}{-14} = \frac{14}{-14} = -1.$$

The solution is $x_1 = 2$ and $x_2 = -1$.

Cramer's Rule (cont.)

If a system of n linear equations in n variables has a coefficient matrix with a nonzero determinant $|A|$, then the solution of the system is given by

$$x_1 = \frac{\det(A_1)}{\det(A)}, \qquad x_2 = \frac{\det(A_2)}{\det(A)}, \qquad \ldots, \qquad x_n = \frac{\det(A_n)}{\det(A)},$$

where the ith column of A_i is the column of constants in the system of equations.

PROOF Let the system be represented by $AX = B$. Because $|A|$ is nonzero, you can write

$$X = A^{-1}B = \frac{1}{|A|}\operatorname{adj}(A)B = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

If the entries of B are $b_1, b_2, \ldots, b_n$, then x_i is

$$x_i = \frac{1}{|A|}(b_1C_{1i} + b_2C_{2i} + \cdots + b_nC_{ni}),$$

but the sum (in parentheses) is precisely the cofactor expansion of A_i, which means that $x_i = |A_i|/|A|$, and the proof is complete.

Using Cramer's Rule

Use Cramer's Rule to solve the system of linear equations for x.

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

SOLUTION The determinant of the coefficient matrix is

$$|A| = \begin{vmatrix} -1 & 2 & -3 \\ 2 & 0 & 1 \\ 3 & -4 & 4 \end{vmatrix} = 10.$$

Because $|A| \neq 0$, you know the solution is unique, and Cramer's Rule can be applied to solve for x, as follows.

$$x = \frac{\begin{vmatrix} 1 & 2 & -3 \\ 0 & 0 & 1 \\ 2 & -4 & 4 \end{vmatrix}}{10} = \frac{(1)(-1)^5\begin{vmatrix} 1 & 2 \\ 2 & -4 \end{vmatrix}}{10} = \frac{(1)(-1)(-8)}{10} = \frac{4}{5}$$

Gaussian Elimination and Gauss-Jordan Elimination

- Gaussian elimination was introduced as a procedure for solving a system of linear equations.

Elementary Row Operations

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

- Rewriting a system of linear equations in row-echelon form usually involves a chain of equivalent systems, each of which is obtained by using one of the three basic operations. This process is called **Gaussian elimination**, after the German mathematician Carl Friedrich Gauss (1777–1855).

Elementary Row Operations (cont.)

(a) Interchange the first and second rows.

Original Matrix	*New Row-Equivalent Matrix*	*Notation*
$\begin{bmatrix} 0 & 1 & 3 & 4 \\ -1 & 2 & 0 & 3 \\ 2 & -3 & 4 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 2 & 0 & 3 \\ 0 & 1 & 3 & 4 \\ 2 & -3 & 4 & 1 \end{bmatrix}$	$R_1 \leftrightarrow R_2$

(b) Multiply the first row by $\frac{1}{2}$ to produce a new first row.

Original Matrix	*New Row-Equivalent Matrix*	*Notation*
$\begin{bmatrix} 2 & -4 & 6 & -2 \\ 1 & 3 & -3 & 0 \\ 5 & -2 & 1 & 2 \end{bmatrix}$	$\begin{bmatrix} 1 & -2 & 3 & -1 \\ 1 & 3 & -3 & 0 \\ 5 & -2 & 1 & 2 \end{bmatrix}$	$\left(\frac{1}{2}\right)R_1 \rightarrow R_1$

(c) Add -2 times the first row to the third row to produce a new third row.

Original Matrix	*New Row-Equivalent Matrix*	*Notation*
$\begin{bmatrix} 1 & 2 & -4 & 3 \\ 0 & 3 & -2 & -1 \\ 2 & 1 & 5 & -2 \end{bmatrix}$	$\begin{bmatrix} 1 & 2 & -4 & 3 \\ 0 & 3 & -2 & -1 \\ 0 & -3 & 13 & -8 \end{bmatrix}$	$R_3 + (-2)R_1 \rightarrow R_3$

Using Elementary Row Operations to Solve a System

Linear System	*Associated Augmented Matrix*	
$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$	
Add the first equation to the second equation.	Add the first row to the second row to produce a new second row.	
$\begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ 2x - 5y + 5z &= 17 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 2 & -5 & 5 & 17 \end{bmatrix}$	$R_2 + R_1 \to R_2$
Add -2 times the first equation to the third equation.	Add -2 times the first row to the third row to produce a new third row.	
$\begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ -y - z &= -1 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{bmatrix}$	$R_3 + (-2)R_1 \to R_3$
Add the second equation to the third equation.	Add the second row to the third row to produce a new third row.	
$\begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ 2z &= 4 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$	$R_3 + R_2 \to R_3$
Multiply the third equation by $\frac{1}{2}$.	Multiply the third row by $\frac{1}{2}$ to produce a new third row.	
$\begin{aligned} x - 2y + 3z &= 9 \\ y + 3z &= 5 \\ z &= 2 \end{aligned}$	$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$	$\left(\frac{1}{2}\right)R_3 \to R_3$

Row-Echelon Form of a Matrix

A matrix in **row-echelon form** has the following properties.

1. All rows consisting entirely of zeros occur at the bottom of the matrix.
2. For each row that does not consist entirely of zeros, the first nonzero entry is 1 (called a **leading 1**).
3. For two successive (nonzero) rows, the leading 1 in the higher row is farther to the left than the leading 1 in the lower row.

REMARK: A matrix in row-echelon form is in **reduced row-echelon form** if every column that has a leading 1 has zeros in every position above and below its leading 1.

Row-Echelon Form

The matrices below are in row-echelon form.

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

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

The matrices shown in parts (b) and (d) are in *reduced* row-echelon form. The matrices listed below are not in row-echelon form.

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

Gaussian Elimination with Back-Substitution

1. Write the augmented matrix of the system of linear equations.
2. Use elementary row operations to rewrite the augmented matrix in row-echelon form.
3. Write the system of linear equations corresponding to the matrix in row-echelon form, and use back-substitution to find the solution.

Gaussian elimination with back-substitution works well as an algorithmic method for solving systems of linear equations. For this algorithm, the order in which the elementary row operations are performed is important. Move from *left to right by columns*, changing all entries directly below the leading 1's to zeros.

Gaussian Elimination with Back-Substitution (cont.)

Solve the system.

$$\begin{aligned} x_2 + x_3 - 2x_4 &= -3 \\ x_1 + 2x_2 - x_3 \quad\quad &= 2 \\ 2x_1 + 4x_2 + x_3 - 3x_4 &= -2 \\ x_1 - 4x_2 - 7x_3 - x_4 &= -19 \end{aligned}$$

SOLUTION The augmented matrix for this system is

$$\begin{bmatrix} 0 & 1 & 1 & -2 & -3 \\ 1 & 2 & -1 & 0 & 2 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix}.$$

Obtain a leading 1 in the upper left corner and zeros elsewhere in the first column.

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix}$$

The first two rows are interchanged. $R_1 \leftrightarrow R_2$

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 1 & -4 & -7 & -1 & -19 \end{bmatrix}$$

Adding -2 times the first row to the third row produces a new third row. $R_3 + (-2)R_1 \rightarrow R_3$

Gaussian Elimination with Back-Substitution (cont.)

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & -6 & -6 & -1 & -21 \end{bmatrix}$$

Adding -1 times the first row to the fourth row produces a new fourth row. $R_4 + (-1)R_1 \rightarrow R_4$

Now that the first column is in the desired form, you should change the second column as shown below.

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & 0 & 0 & -13 & -39 \end{bmatrix}$$

Adding 6 times the second row to the fourth row produces a new fourth row. $R_4 + (6)R_2 \rightarrow R_4$

To write the third column in proper form, multiply the third row by $\frac{1}{3}$.

$$\begin{bmatrix} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & -13 & -39 \end{bmatrix}$$

Multiplying the third row by $\frac{1}{3}$ produces a new third row. $\left(\frac{1}{3}\right)R_3 \rightarrow R_3$

Similarly, to write the fourth column in proper form, you should multiply the fourth row by $-\frac{1}{13}$.

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

Multiplying the fourth row by $-\frac{1}{13}$ produces a new fourth row. $\left(-\frac{1}{13}\right)R_4 \rightarrow R_4$

The matrix is now in row-echelon form, and the corresponding system of linear equations is as shown below.

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

Using back-substitution, you can determine that the solution is

$$x_1 = -1, \quad x_2 = 2, \quad x_3 = 1, \quad x_4 = 3.$$

Gauss-Jordan Elimination

With Gaussian elimination, you apply elementary row operations to a matrix to obtain a (row-equivalent) row-echelon form. A second method of elimination, called **Gauss-Jordan elimination** after Carl Gauss and Wilhelm Jordan (1842–1899), continues the reduction process until a *reduced* row-echelon form is obtained. This procedure is demonstrated in the next example.

Gauss-Jordan Elimination (cont.)

Use Gauss-Jordan elimination to solve the system.

$$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y \quad\quad &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$$

SOLUTION In Example 3, Gaussian elimination was used to obtain the row-echelon form

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

Now, rather than using back-substitution, apply elementary row operations until you obtain a matrix in reduced row-echelon form. To do this, you must produce zeros above each of the leading 1's, as follows.

$$\begin{bmatrix} 1 & 0 & 9 & 19 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad R_1 + (2)R_2 \rightarrow R_1$$

$$\begin{bmatrix} 1 & 0 & 9 & 19 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad R_2 + (-3)R_3 \rightarrow R_2$$

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \qquad R_1 + (-9)R_3 \rightarrow R_1$$

Now, converting back to a system of linear equations, you have

$$\begin{aligned} x \quad\quad &= 1 \\ y \quad &= -1 \\ z &= 2. \end{aligned}$$

Applications of Systems of Linear Equations

- Systems of linear equations arise in a wide variety of applications and are one of the central themes in linear algebra.

Polynomial Curve Fitting

Suppose a collection of data is represented by n points in the xy-plane,

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

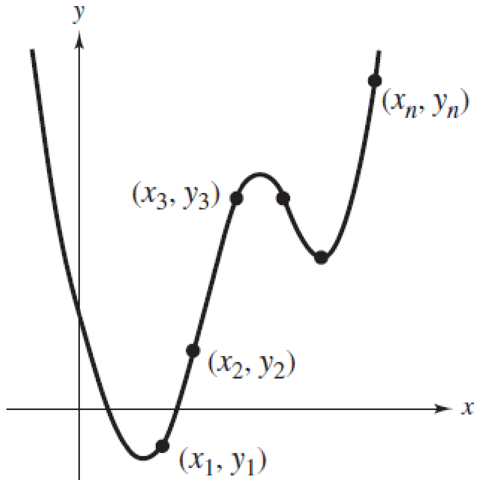
and you are asked to find a polynomial function of degree $n - 1$

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1}$$

whose graph passes through the specified points. This procedure is called **polynomial curve fitting.** If all x-coordinates of the points are distinct, then there is precisely one polynomial function of degree $n - 1$ (or less) that fits the n points, as shown in Figure 1.4.

To solve for the n coefficients of $p(x)$, substitute each of the n points into the polynomial function and obtain n linear equations in n variables $a_0, a_1, a_2, \ldots, a_{n-1}$.

$$\begin{aligned}
a_0 + a_1x_1 + a_2x_1^2 + \cdots + a_{n-1}x_1^{n-1} &= y_1 \\
a_0 + a_1x_2 + a_2x_2^2 + \cdots + a_{n-1}x_2^{n-1} &= y_2 \\
&\vdots \\
a_0 + a_1x_n + a_2x_n^2 + \cdots + a_{n-1}x_n^{n-1} &= y_n
\end{aligned}$$



Polynomial Curve Fitting

Polynomial Curve Fitting (cont.)

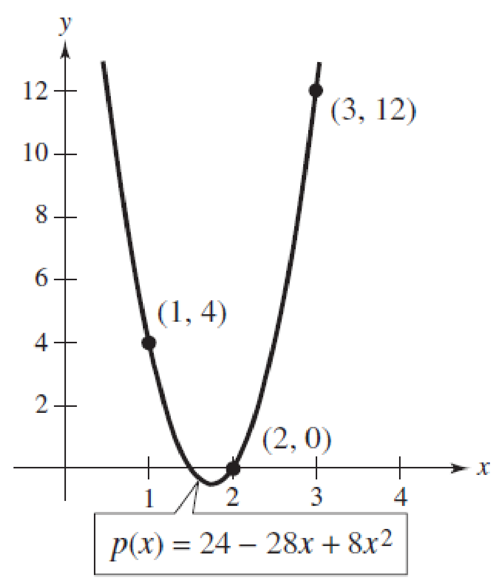
Determine the polynomial $p(x) = a_0 + a_1x + a_2x^2$ whose graph passes through the points $(1, 4)$, $(2, 0)$, and $(3, 12)$.

SOLUTION Substituting $x = 1$, 2, and 3 into $p(x)$ and equating the results to the respective y-values produces the system of linear equations in the variables a_0, a_1, and a_2 shown below.

$$\begin{aligned}
p(1) &= a_0 + a_1(1) + a_2(1)^2 = a_0 + a_1 + a_2 = 4\\
p(2) &= a_0 + a_1(2) + a_2(2)^2 = a_0 + 2a_1 + 4a_2 = 0\\
p(3) &= a_0 + a_1(3) + a_2(3)^2 = a_0 + 3a_1 + 9a_2 = 12
\end{aligned}$$

The solution of this system is $a_0 = 24$, $a_1 = -28$, and $a_2 = 8$, so the polynomial function is

$$p(x) = 24 - 28x + 8x^2.$$



Polynomial Curve Fitting (cont.)

Find a polynomial that fits the points $(-2, 3)$, $(-1, 5)$, $(0, 1)$, $(1, 4)$, and $(2, 10)$.

SOLUTION Because you are provided with five points, choose a fourth-degree polynomial function

$$p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4.$$

Substituting the given points into $p(x)$ produces the system of linear equations listed below.

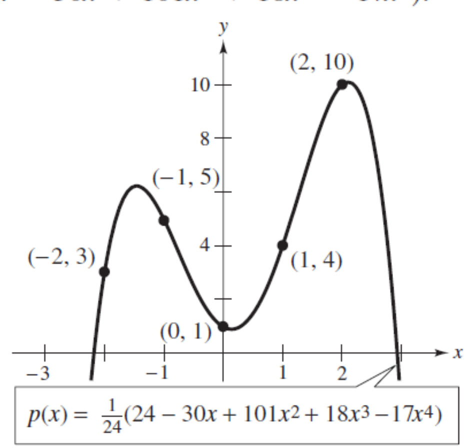
$$\begin{aligned}
a_0 - 2a_1 + 4a_2 - 8a_3 + 16a_4 &= 3\\
a_0 - a_1 + a_2 - a_3 + a_4 &= 5\\
a_0 &= 1\\
a_0 + a_1 + a_2 + a_3 + a_4 &= 4\\
a_0 + 2a_1 + 4a_2 + 8a_3 + 16a_4 &= 10
\end{aligned}$$

The solution of these equations is

$$a_0 = 1, \quad a_1 = -\tfrac{30}{24}, \quad a_2 = \tfrac{101}{24}, \quad a_3 = \tfrac{18}{24}, \quad a_4 = -\tfrac{17}{24}$$

which means the polynomial function is

$$\begin{aligned}
p(x) &= 1 - \tfrac{30}{24}x + \tfrac{101}{24}x^2 + \tfrac{18}{24}x^3 - \tfrac{17}{24}x^4\\
&= \tfrac{1}{24}(24 - 30x + 101x^2 + 18x^3 - 17x^4).
\end{aligned}$$



An Application of Curve Fitting

Find a polynomial that relates the periods of the first three planets to their mean distances from the sun, as shown in Table 1.1. Then test the accuracy of the fit by using the polynomial to calculate the period of Mars. (Distance is measured in astronomical units, and period is measured in years.) (Source: *CRC Handbook of Chemistry and Physics*)

TABLE 1.1

Planet	*Mercury*	*Venus*	*Earth*	*Mars*	*Jupiter*	*Saturn*
Mean Distance	0.387	0.723	1.0	1.523	5.203	9.541
Period	0.241	0.615	1.0	1.881	11.861	29.457

SOLUTION Begin by fitting a quadratic polynomial function

$$p(x) = a_0 + a_1x + a_2x^2$$

to the points (0.387, 0.241), (0.723, 0.615), and (1, 1). The system of linear equations obtained by substituting these points into $p(x)$ is

$$\begin{aligned} a_0 + 0.387a_1 + (0.387)^2a_2 &= 0.241 \\ a_0 + 0.723a_1 + (0.723)^2a_2 &= 0.615 \\ a_0 + \quad a_1 + \quad a_2 &= 1. \end{aligned}$$

The approximate solution of the system is

$$a_0 \approx -0.0634, \qquad a_1 \approx 0.6119, \qquad a_2 \approx 0.4515$$

which means that the polynomial function can be approximated by

$$p(x) = -0.0634 + 0.6119x + 0.4515x^2.$$

Using $p(x)$ to evaluate the period of Mars produces

$$p(1.523) \approx 1.916 \text{ years.}$$

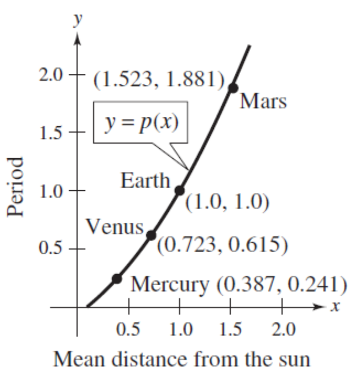


Figure 1.7

This estimate is compared graphically with the actual period of Mars in Figure 1.7. Note that the actual period (from Table 1.1) is 1.881 years.

Applications of Linear Systems

- Modeling supply and demand in logistics and finance
- Solving optimization problems in business operations
- Performing linear and multiple regression in analytics

Scalars and Vectors

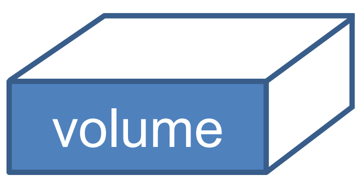
- A scalar quantity has only magnitude.
- A vector quantity has both magnitude and direction.

Scalar Quantities

length, area, volume, speed, mass, density, pressure, temperature, energy, entropy, work, power

Vector Quantities

displacement, velocity, acceleration, momentum, force, lift, drag, thrust, weight



Scalar

24

Vector

$\begin{bmatrix} 2 & -8 & 7 \end{bmatrix}$

row or column $\begin{bmatrix} 2 \\ -8 \\ 7 \end{bmatrix}$

Matrix

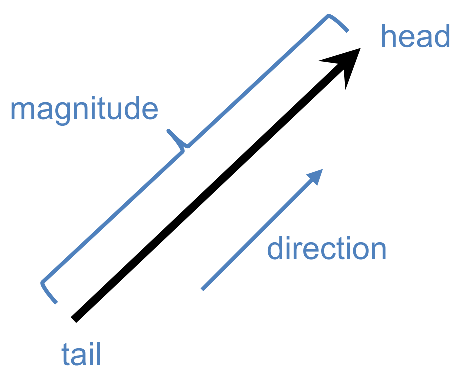
$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}$

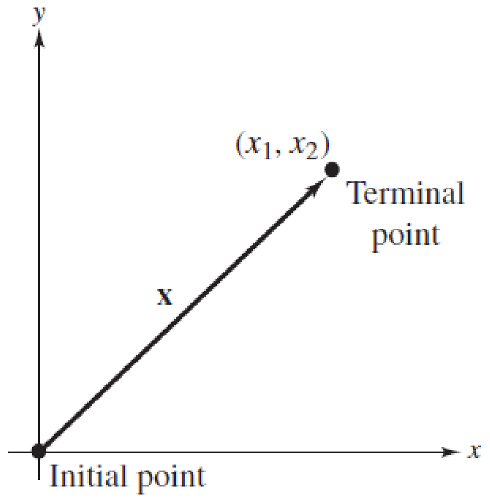
row(s) × column(s)

Vectors in the Plane

- A **vector in the plane** is represented geometrically by a **directed line segment** whose **initial point** is the origin and whose **terminal point** is the point (x_1, x_2).
- This vector is represented by the same ordered pair used to represent its terminal point. That is,

$$\mathbf{x} = (x_1, x_2)$$





Vectors in the Plane (cont.)

- The coordinates x_1 and x_2 are called the **components** of the vector $\mathbf{x}$.
- Two vectors in the plane $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ are **equal** if and only if $u_1 = v_1$ and $u_2 = v_2$.

Vectors in the Plane (cont.)

Use a directed line segment to represent each vector in the plane.
(a) $\mathbf{u} = (2, 3)$ (b) $\mathbf{v} = (-1, 2)$

SOLUTION To represent each vector, draw a directed line segment from the origin to the indicated terminal point, as shown in Figure 4.2.

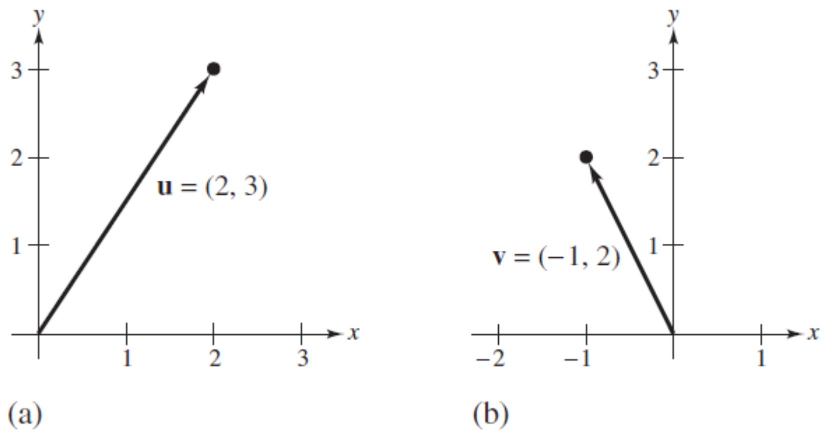


Figure 4.2

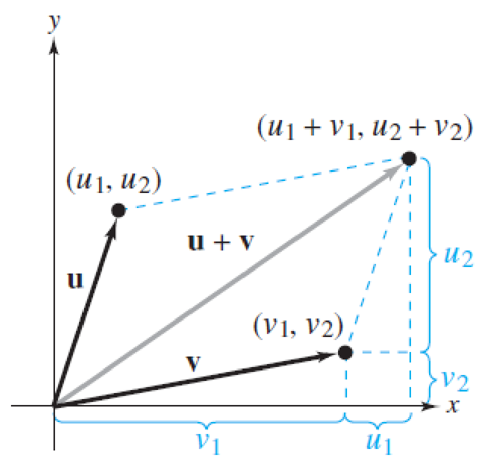
Operations with Vectors in the Plane

Vector Addition

- To add two vectors in the plane, add their corresponding components. That is, the **sum** of $\mathbf{u}$ and $\mathbf{v}$ is the vector

$$\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2).$$

- Geometrically, the sum of two vectors in the plane is represented as the diagonal of a parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as its adjacent sides.



Vector Addition

Adding Two Vectors in the Plane

Find the sum of the vectors.

(a) $\mathbf{u} = (1, 4)$, $\mathbf{v} = (2, -2)$ (b) $\mathbf{u} = (3, -2)$, $\mathbf{v} = (-3, 2)$ (c) $\mathbf{u} = (2, 1)$, $\mathbf{v} = (0, 0)$

SOLUTION

(a) $\mathbf{u} + \mathbf{v} = (1, 4) + (2, -2) = (3, 2)$

(b) $\mathbf{u} + \mathbf{v} = (3, -2) + (-3, 2) = (0, 0)$

(c) $\mathbf{u} + \mathbf{v} = (2, 1) + (0, 0) = (2, 1)$

Figure 4.4 gives the graphical representation of each sum.

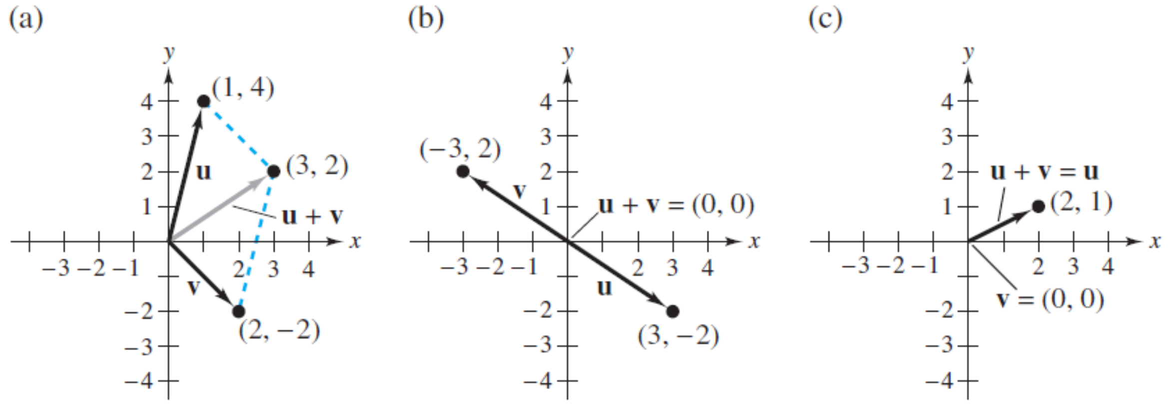


Figure 4.4

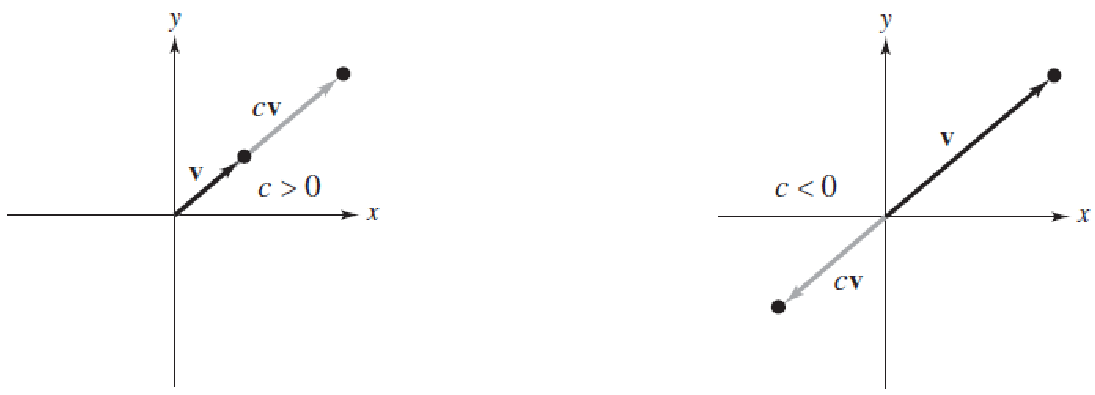
Operations with Vectors in the Plane (cont.)

Scalar Multiplication

- To multiply a vector $\mathbf{v}$ by a scalar c, multiply each of the components of $\mathbf{v}$ by c. That is,

$$c\mathbf{v} = c(v_1, v_2) = (cv_1, cv_2).$$

- The word *scalar* is used to mean a real number.
- In general, for a scalar c, the vector $c\mathbf{v}$ will be $|c|$ times as long as $\mathbf{v}$. If c is positive, then $c\mathbf{v}$ and $\mathbf{v}$ have the same direction, and if c is negative, then $c\mathbf{v}$ and $\mathbf{v}$ have opposite directions.



Operations with Vectors in the Plane (cont.)

Provided with $\mathbf{v} = (-2, 5)$ and $\mathbf{u} = (3, 4)$, find each vector.

(a) $\frac{1}{2}\mathbf{v}$ (b) $\mathbf{u} - \mathbf{v}$ (c) $\frac{1}{2}\mathbf{v} + \mathbf{u}$

SOLUTION (a) Because $\mathbf{v} = (-2, 5)$, you have

$$\frac{1}{2}\mathbf{v} = \left(\frac{1}{2}(-2), \frac{1}{2}(5)\right) = \left(-1, \frac{5}{2}\right).$$

(b) By the definition of vector subtraction, you have

$$\mathbf{u} - \mathbf{v} = (3 - (-2), 4 - 5) = (5, -1).$$

(c) Using the result of part(a), you have

$$\frac{1}{2}\mathbf{v} + \mathbf{u} = \left(-1, \frac{5}{2}\right) + (3, 4) = \left(-1 + 3, \frac{5}{2} + 4\right) = \left(2, \frac{13}{2}\right).$$

Figure 4.6 gives a graphical representation of these vector operations.

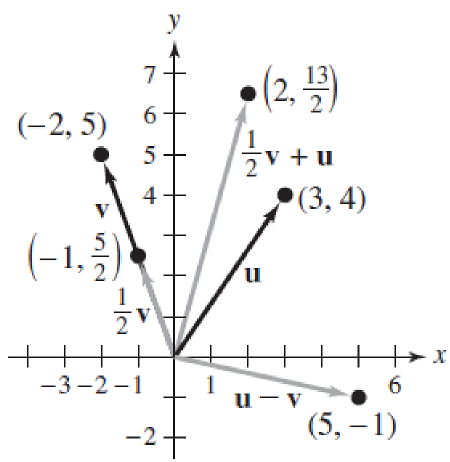


Figure 4.6

Properties of Vector Addition and Scalar Multiplication in the Plane

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in the plane, and let c and d be scalars.

1. $\mathbf{u} + \mathbf{v}$ is a vector in the plane.	Closure under addition
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Commutative property of addition
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	Associative property of addition
4. $\mathbf{u} + \mathbf{0} = \mathbf{u}$	Additive identity property
5. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	Additive inverse property
6. $c\mathbf{u}$ is a vector in the plane.	Closure under scalar multiplication
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	Distributive property
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	Distributive property
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$	Associative property of multiplication
10. $1(\mathbf{u}) = \mathbf{u}$	Multiplicative identity property

Vectors in R^n

- A vector in n-space is represented by an **ordered n-tuple.**
- The set of all n-tuples is called **n-space** and is denoted by R^n.

$$\begin{aligned}
R^1 &= \text{1-space} = \text{set of all real numbers} \\
R^2 &= \text{2-space} = \text{set of all ordered pairs of real numbers} \\
R^3 &= \text{3-space} = \text{set of all ordered triples of real numbers} \\
R^4 &= \text{4-space} = \text{set of all ordered quadruples of real numbers} \\
&\vdots \\
R^n &= n\text{-space} = \text{set of all ordered } n\text{-tuples of real numbers}
\end{aligned}$$

- The practice of using an ordered pair to represent either a point or a vector in R^2 continues in R^n.
- That is, an n-tuple $(x_1, x_2, x_3, \ldots, x_n)$ can be viewed as a **point** in R^n with the x_i's as its coordinates or as a **vector**

$$\mathbf{x} = (x_1, x_2, x_3, \ldots, x_n) \qquad \text{Vector in } R^n$$

with the x_i's as its components.
- As with vectors in the plane, two vectors in R^n are **equal** if and only if corresponding components are equal.

Vector Addition and Scalar Multiplication in R^n

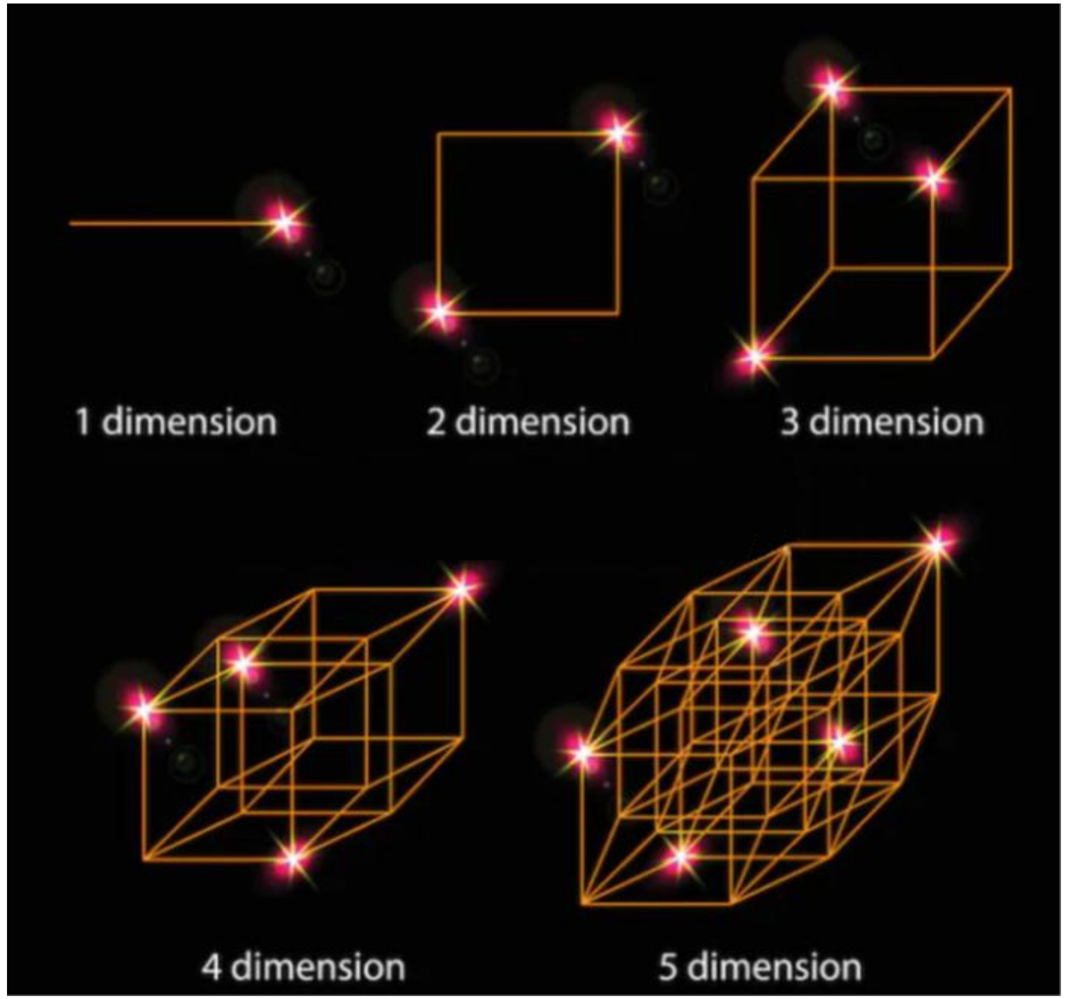
Let $\mathbf{u} = (u_1, u_2, u_3, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, v_3, \ldots, v_n)$ be vectors in R^n and let c be a real number. Then the sum of $\mathbf{u}$ and $\mathbf{v}$ is defined as the vector

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3, \ldots, u_n + v_n),$$

and the **scalar multiple** of $\mathbf{u}$ by c is defined as the vector

$$c\mathbf{u} = (cu_1, cu_2, cu_3, \ldots, cu_n).$$

Vectors in R^n (cont.)



Vector Operations in R^3

Provided that $\mathbf{u} = (-1, 0, 1)$ and $\mathbf{v} = (2, -1, 5)$ in R^3, find each vector.

(a) $\mathbf{u} + \mathbf{v}$ (b) $2\mathbf{u}$ (c) $\mathbf{v} - 2\mathbf{u}$

SOLUTION (a) To add two vectors, add their corresponding components, as follows.

$$\mathbf{u} + \mathbf{v} = (-1, 0, 1) + (2, -1, 5) = (1, -1, 6)$$

(b) To multiply a vector by a scalar, multiply each component by the scalar, as follows.

$$2\mathbf{u} = 2(-1, 0, 1) = (-2, 0, 2)$$

(c) Using the result of part (b), you have

$$\mathbf{v} - 2\mathbf{u} = (2, -1, 5) - (-2, 0, 2) = (4, -1, 3).$$

Figure 4.7 gives a graphical representation of these vector operations in R^3.

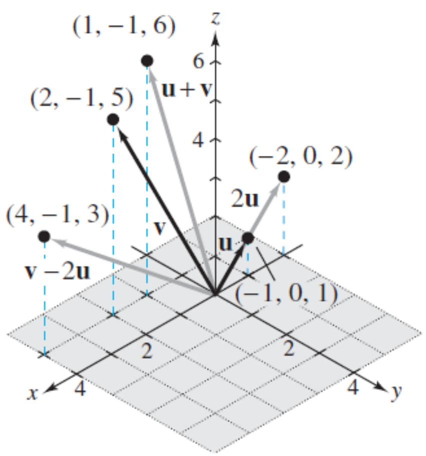


Figure 4.7

Properties of Vector Addition and Scalar Multiplication in R^n

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in R^n, and let c and d be scalars.

1. $\mathbf{u} + \mathbf{v}$ is a vector in R^n.	Closure under addition
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Commutative property of addition
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	Associative property addition
4. $\mathbf{u} + \mathbf{0} = \mathbf{u}$	Additive identity property
5. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	Additive inverse property
6. $c\mathbf{u}$ is a vector in R^n.	Closure under scalar multiplication
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	Distributive property
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	Distributive property
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$	Associative property of multiplication
10. $1(\mathbf{u}) = \mathbf{u}$	Multiplicative identity property

Vector Operations in R^4

Let $\mathbf{u} = (2, -1, 5, 0)$, $\mathbf{v} = (4, 3, 1, -1)$, and $\mathbf{w} = (-6, 2, 0, 3)$ be vectors in R^4. Solve for $\mathbf{x}$.

(a) $\mathbf{x} = 2\mathbf{u} - (\mathbf{v} + 3\mathbf{w})$ (b) $3(\mathbf{x} + \mathbf{w}) = 2\mathbf{u} - \mathbf{v} + \mathbf{x}$

SOLUTION (a) Using the properties listed in Theorem 4.2, you have

$$\begin{aligned}\mathbf{x} &= 2\mathbf{u} - (\mathbf{v} + 3\mathbf{w})\\ &= 2\mathbf{u} - \mathbf{v} - 3\mathbf{w}\\ &= (4, -2, 10, 0) - (4, 3, 1, -1) - (-18, 6, 0, 9)\\ &= (4 - 4 + 18, -2 - 3 - 6, 10 - 1 - 0, 0 + 1 - 9)\\ &= (18, -11, 9, -8).\end{aligned}$$

(b) Begin by solving for $\mathbf{x}$ as follows.

$$\begin{aligned}3(\mathbf{x} + \mathbf{w}) &= 2\mathbf{u} - \mathbf{v} + \mathbf{x}\\ 3\mathbf{x} + 3\mathbf{w} &= 2\mathbf{u} - \mathbf{v} + \mathbf{x}\\ 3\mathbf{x} - \mathbf{x} &= 2\mathbf{u} - \mathbf{v} - 3\mathbf{w}\\ 2\mathbf{x} &= 2\mathbf{u} - \mathbf{v} - 3\mathbf{w}\\ \mathbf{x} &= \tfrac{1}{2}(2\mathbf{u} - \mathbf{v} - 3\mathbf{w})\end{aligned}$$

Using the result of part (a) produces

$$\begin{aligned}\mathbf{x} &= \tfrac{1}{2}(18, -11, 9, -8)\\ &= \left(9, -\tfrac{11}{2}, \tfrac{9}{2}, -4\right).\end{aligned}$$

Properties of Additive Identity and Additive Inverse

The zero vector $\mathbf{0}$ in R^n is called the **additive identity** in R^n. Similarly, the vector $-\mathbf{v}$ is called the **additive inverse** of $\mathbf{v}$. The theorem below summarizes several important properties of the additive identity and additive inverse in R^n.

Let $\mathbf{v}$ be a vector in R^n, and let c be a scalar. Then the following properties are true.

1. The additive identity is unique. That is, if $\mathbf{v} + \mathbf{u} = \mathbf{v}$, then $\mathbf{u} = \mathbf{0}$.
2. The additive inverse of $\mathbf{v}$ is unique. That is, if $\mathbf{v} + \mathbf{u} = \mathbf{0}$, then $\mathbf{u} = -\mathbf{v}$.
3. $0\mathbf{v} = \mathbf{0}$
4. $c\mathbf{0} = \mathbf{0}$
5. If $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.
6. $-(-\mathbf{v}) = \mathbf{v}$

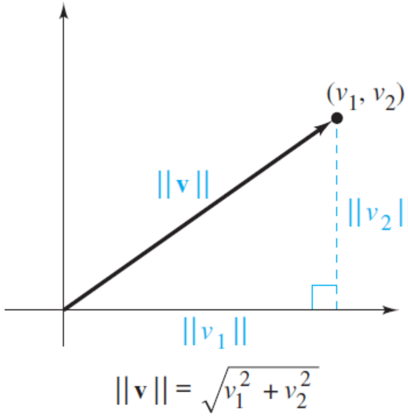
Length and Dot Product in R^n

- You will begin by reviewing the definition of the length of a vector in R^2.
- If $\mathbf{v} = (v_1, v_2)$ is a vector in the plane, then the **length,** or **magnitude,** of denoted by $\|\mathbf{v}\|$, is defined as

$$\|\mathbf{v}\| = \sqrt{\boldsymbol{v}_1^2 + \boldsymbol{v}_2^2}.$$

- This definition corresponds to the usual notion of length in Euclidean geometry.
- That is, the vector is thought of as the hypotenuse of a right triangle whose sides have lengths of $|v_1|$ and $|v_2|$.
- Applying the Pythagorean Theorem produces

$$\|\mathbf{v}\|^2 = |v_1|^2 + |v_2|^2 = v_1^2 + v_2^2.$$



Length of a Vector in R^n

The **length,** or **magnitude,** of a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in R^n is given by

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

REMARK: The length of a vector is also called its **norm.** If $\|\mathbf{v}\| = 1$, then the vector $\mathbf{v}$ is called a **unit vector.**

This definition shows that the length of a vector cannot be negative. That is, $\|\mathbf{v}\| \geq 0$. Moreover, $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v}$ is the zero vector $\mathbf{0}$.

The Length of a Vector in R^n

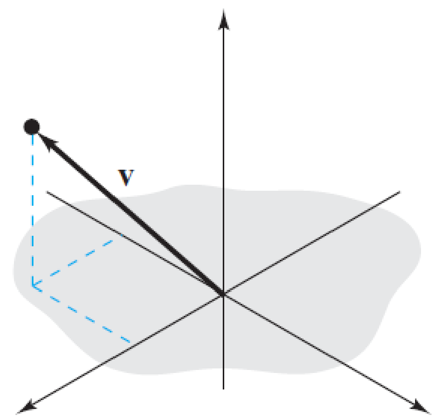
(a) In R^5, the length of $\mathbf{v} = (0, -2, 1, 4, -2)$ is

$$\|\mathbf{v}\| = \sqrt{0^2 + (-2)^2 + 1^2 + 4^2 + (-2)^2} = \sqrt{25} = 5.$$

(b) In R^3, the length of $\mathbf{v} = \left(2/\sqrt{17}, -2/\sqrt{17}, 3/\sqrt{17}\right)$ is

$$\|\mathbf{v}\| = \sqrt{\left(\frac{2}{\sqrt{17}}\right)^2 + \left(-\frac{2}{\sqrt{17}}\right)^2 + \left(\frac{3}{\sqrt{17}}\right)^2} = \sqrt{\frac{17}{17}} = 1.$$

Because its length is 1, $\mathbf{v}$ is a unit vector, as shown in Figure 5.2.



$$\mathbf{v} = \left(\frac{2}{\sqrt{17}}, -\frac{2}{\sqrt{17}}, \frac{3}{\sqrt{17}}\right)$$

Figure 5.2

Length of a Scalar Multiple

Let $\mathbf{v}$ be a vector in R^n and let c be a scalar. Then

$$\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|,$$

where $|c|$ is the absolute value of c.

PROOF Because $c\mathbf{v} = (cv_1, cv_2, \ldots, cv_n)$, it follows that

$$\begin{aligned}
\|c\mathbf{v}\| &= \|(cv_1, cv_2, \ldots, cv_n)\| \\
&= \sqrt{(cv_1)^2 + (cv_2)^2 + \cdots + (cv_n)^2} \\
&= \sqrt{c^2(v_1^2 + v_2^2 + \cdots + v_n^2)} \\
&= |c|\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \\
&= |c|\,\|\mathbf{v}\|.
\end{aligned}$$

Unit Vector in the Direction of v

If $\mathbf{v}$ is a nonzero vector in R^n, then the vector

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

has length 1 and has the same direction as $\mathbf{v}$. This vector $\mathbf{u}$ is called the **unit vector in the direction of v**.

Finding a Unit Vector

Find the unit vector in the direction of $\mathbf{v} = (3, -1, 2)$, and verify that this vector has length 1.

SOLUTION The unit vector in the direction of $\mathbf{v}$ is

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{(3, -1, 2)}{\sqrt{3^2 + (-1)^2 + 2^2}} = \frac{1}{\sqrt{14}}(3, -1, 2) = \left(\frac{3}{\sqrt{14}}, -\frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}\right),$$

which is a unit vector because

$$\sqrt{\left(\frac{3}{\sqrt{14}}\right)^2 + \left(-\frac{1}{\sqrt{14}}\right)^2 + \left(\frac{2}{\sqrt{14}}\right)^2} = \sqrt{\frac{14}{14}} = 1.$$

(See Figure 5.3.)

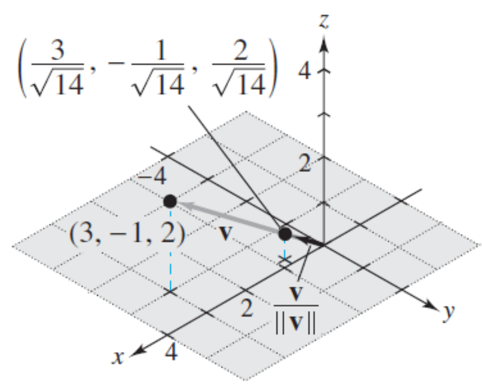


Figure 5.3

Distance Between Two Vectors in R^n

The **distance between two vectors u** and **v** in R^n is

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|.$$

You can easily verify the three properties of distance listed below.

1. $d(\mathbf{u}, \mathbf{v}) \geq 0$
2. $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$.
3. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$

Finding the Distance Between Two Vectors

The distance between $\mathbf{u} = (0, 2, 2)$ and $\mathbf{v} = (2, 0, 1)$ is

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \|(0 - 2, 2 - 0, 2 - 1)\| \\ &= \sqrt{(-2)^2 + 2^2 + 1^2} = 3. \end{aligned}$$

Dot Product in R^n

The **dot product** of $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ is the *scalar* quantity

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

REMARK: Notice that the dot product of two vectors is a scalar, not another vector.

Finding the Dot Product of Two Vectors

The dot product of $\mathbf{u} = (1, 2, 0, -3)$ and $\mathbf{v} = (3, -2, 4, 2)$ is

$$\mathbf{u} \cdot \mathbf{v} = (1)(3) + (2)(-2) + (0)(4) + (-3)(2) = -7.$$

Properties of the Dot Product

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in R^n and c is a scalar, then the following properties are true.

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
3. $c(\mathbf{u} \cdot \mathbf{v}) = (c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v})$
4. $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$
5. $\mathbf{v} \cdot \mathbf{v} \geq 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

PROOF The proofs of these properties follow easily from the definition of dot product. For example, to prove the first property, you can write

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \cdot \cdot \cdot + u_nv_n \\ &= v_1u_1 + v_2u_2 + \cdot \cdot \cdot + v_nu_n \\ &= \mathbf{v} \cdot \mathbf{u}. \end{aligned}$$

Finding Dot Products

Given $\mathbf{u} = (2, -2)$, $\mathbf{v} = (5, 8)$, and $\mathbf{w} = (-4, 3)$, find

(a) $\mathbf{u} \cdot \mathbf{v}$. (b) $(\mathbf{u} \cdot \mathbf{v})\mathbf{w}$. (c) $\mathbf{u} \cdot (2\mathbf{v})$. (d) $\|\mathbf{w}\|^2$. (e) $\mathbf{u} \cdot (\mathbf{v} - 2\mathbf{w})$.

SOLUTION (a) By definition, you have

$$\mathbf{u} \cdot \mathbf{v} = 2(5) + (-2)(8) = -6.$$

(b) Using the result in part (a), you have

$$(\mathbf{u} \cdot \mathbf{v})\mathbf{w} = -6\mathbf{w} = -6(-4, 3) = (24, -18).$$

(c) By Property 3 of Theorem 5.3, you have

$$\mathbf{u} \cdot (2\mathbf{v}) = 2(\mathbf{u} \cdot \mathbf{v}) = 2(-6) = -12.$$

(d) By Property 4 of Theorem 5.3, you have

$$\|\mathbf{w}\|^2 = \mathbf{w} \cdot \mathbf{w} = (-4)(-4) + (3)(3) = 25.$$

(e) Because $2\mathbf{w} = (-8, 6)$, you have

$$\mathbf{v} - 2\mathbf{w} = (5 - (-8), 8 - 6) = (13, 2).$$

Consequently,

$$\mathbf{u} \cdot (\mathbf{v} - 2\mathbf{w}) = 2(13) + (-2)(2) = 26 - 4 = 22.$$

Using Properties of the Dot Product

Provided with two vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n such that $\mathbf{u} \cdot \mathbf{u} = 39$, $\mathbf{u} \cdot \mathbf{v} = -3$, and $\mathbf{v} \cdot \mathbf{v} = 79$, evaluate $(\mathbf{u} + 2\mathbf{v}) \cdot (3\mathbf{u} + \mathbf{v})$.

SOLUTION Using Theorem 5.3, rewrite the dot product as

$$\begin{aligned}
(\mathbf{u} + 2\mathbf{v}) \cdot (3\mathbf{u} + \mathbf{v}) &= \mathbf{u} \cdot (3\mathbf{u} + \mathbf{v}) + (2\mathbf{v}) \cdot (3\mathbf{u} + \mathbf{v}) \\
&= \mathbf{u} \cdot (3\mathbf{u}) + \mathbf{u} \cdot \mathbf{v} + (2\mathbf{v}) \cdot (3\mathbf{u}) + (2\mathbf{v}) \cdot \mathbf{v} \\
&= 3(\mathbf{u} \cdot \mathbf{u}) + \mathbf{u} \cdot \mathbf{v} + 6(\mathbf{v} \cdot \mathbf{u}) + 2(\mathbf{v} \cdot \mathbf{v}) \\
&= 3(\mathbf{u} \cdot \mathbf{u}) + 7(\mathbf{u} \cdot \mathbf{v}) + 2(\mathbf{v} \cdot \mathbf{v}) \\
&= 3(39) + 7(-3) + 2(79) = 254.
\end{aligned}$$

The Angle Between Two Vectors in R^n

The **angle** θ between two nonzero vectors in R^n is given by

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}, \quad 0 \leq \theta \leq \pi.$$

REMARK: The angle between the zero vector and another vector is not defined.

Finding the Angle Between Two Vectors

The angle between $\mathbf{u} = (-4, 0, 2, -2)$ and $\mathbf{v} = (2, 0, -1, 1)$ is

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|} = \frac{-12}{\sqrt{24}\sqrt{6}} = -\frac{12}{\sqrt{144}} = -1.$$

Consequently, $\theta = \pi$. It makes sense that $\mathbf{u}$ and $\mathbf{v}$ should have opposite directions, because $\mathbf{u} = -2\mathbf{v}$.

θ	0^0 0^0	30^0 $\frac{\pi}{6}$	45^0 $\frac{\pi}{4}$	60^0 $\frac{\pi}{3}$	90^0 $\frac{\pi}{2}$	180^0 π	270^0 $\frac{3\pi}{2}$	360^0 2π
sin θ	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
cos θ	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0	-1	0	1
tan θ	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	$N.D.$	0	$N.D.$	0
cosec θ	$N.D.$	2	$\sqrt{2}$	$\frac{2}{\sqrt{3}}$	1	$N.D.$	-1	$N.D$
sec θ	1	$\frac{2}{\sqrt{3}}$	$\sqrt{2}$	2	$N.D.$	-1	$N.D.$	1
cot θ	$N.D.$	$\sqrt{3}$	1	$\frac{1}{\sqrt{3}}$	0	$N.D.$	0	$N.D$

The Angle Between Two Vectors in R^n (cont.)

Note that because $\|\mathbf{u}\|$ and $\|\mathbf{v}\|$ are always positive, $\mathbf{u} \cdot \mathbf{v}$ and $\cos\theta$ will always have the same sign. Moreover, because the cosine is positive in the first quadrant and negative in the second quadrant, the sign of the dot product of two vectors can be used to determine whether the angle between them is acute or obtuse, as shown in Figure 5.6.

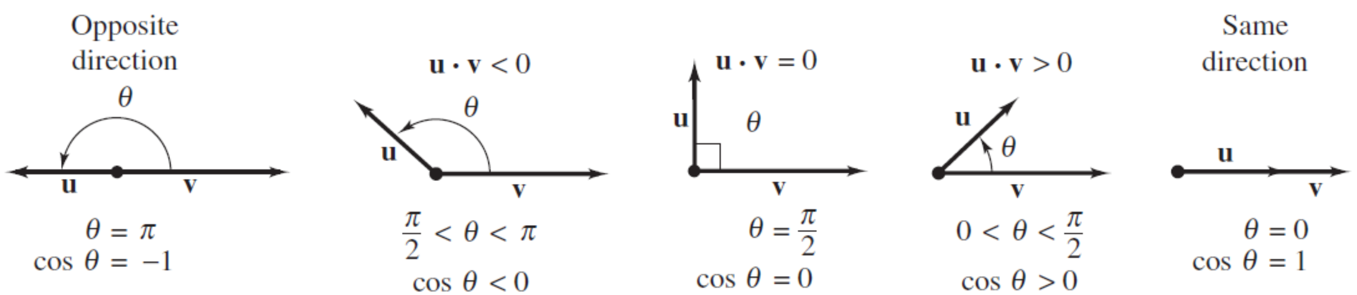


Figure 5.6

Orthogonal Vectors

Two vectors **u** and **v** in R^n are **orthogonal** if

$$\mathbf{u} \cdot \mathbf{v} = 0.$$

REMARK: Even though the angle between the zero vector and another vector is not defined, it is convenient to extend the definition of orthogonality to include the zero vector. In other words, the vector **0** is said to be orthogonal to every vector.

Orthogonal Vectors in R^n

(a) The vectors $\mathbf{u} = (1, 0, 0)$ and $\mathbf{v} = (0, 1, 0)$ are orthogonal because

$$\mathbf{u} \cdot \mathbf{v} = (1)(0) + (0)(1) + (0)(0) = 0.$$

(b) The vectors $\mathbf{u} = (3, 2, -1, 4)$ and $\mathbf{v} = (1, -1, 1, 0)$ are orthogonal because

$$\mathbf{u} \cdot \mathbf{v} = (3)(1) + (2)(-1) + (-1)(1) + (4)(0) = 0.$$

Finding Orthogonal Vectors

Determine all vectors in R^2 that are orthogonal to $\mathbf{u} = (4, 2)$.

SOLUTION Let $\mathbf{v} = (v_1, v_2)$ be orthogonal to $\mathbf{u}$. Then

$$\mathbf{u} \cdot \mathbf{v} = (4, 2) \cdot (v_1, v_2) = 4v_1 + 2v_2 = 0,$$

which implies that $2v_2 = -4v_1$ and $v_2 = -2v_1$. So, every vector that is orthogonal to $(4, 2)$ is of the form

$$\mathbf{v} = (t, -2t) = t(1, -2),$$

where t is a real number. (See Figure 5.7.)

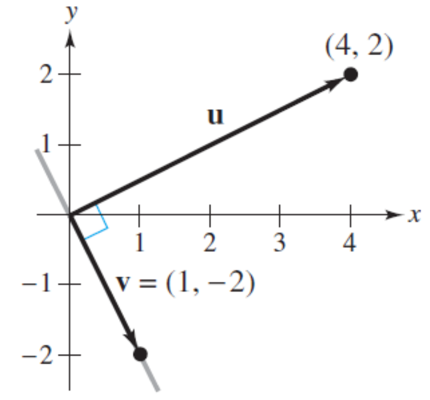


Figure 5.7

Applications of Vectors and Dot Products

- Measuring document similarity using cosine similarity
- Working with word embeddings in NLP and sentiment analysis
- Modeling directional quantities in business and physics

The Cross Product of Two Vectors in Space

Many problems in linear algebra involve finding a vector orthogonal to each vector in a set. Here you will look at a vector product that yields a vector in R^3 orthogonal to two vectors. This vector product is called the **cross product,** and it is most conveniently defined and calculated with vectors written in standard unit vector form.

$$\mathbf{v} = (v_1, v_2, v_3) = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$$

Let $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ and $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ be vectors in R^3. The **cross product** of $\mathbf{u}$ and $\mathbf{v}$ is the vector

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2)\mathbf{i} - (u_1v_3 - u_3v_1)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}.$$

REMARK: The cross product is defined only for vectors in R^3. The cross product of two vectors in R^2 or of vectors in R^n, $n > 3$, is not defined here.

A convenient way to remember the formula for the cross product $\mathbf{u} \times \mathbf{v}$ is to use the determinant form shown below.

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \begin{array}{l} \\ \leftarrow \text{Components of } \mathbf{u} \\ \leftarrow \text{Components of } \mathbf{v} \end{array}$$

Finding the Cross Product of Two Vectors

Provided that $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$ and $\mathbf{v} = 3\mathbf{i} + \mathbf{j} - 2\mathbf{k}$, find

(a) $\mathbf{u} \times \mathbf{v}$. (b) $\mathbf{v} \times \mathbf{u}$. (c) $\mathbf{v} \times \mathbf{v}$.

SOLUTION

(a)
$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -2 & 1 \\ 3 & 1 & -2 \end{vmatrix}$$
$$= \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & 1 \\ 3 & -2 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & -2 \\ 3 & 1 \end{vmatrix}\mathbf{k}$$
$$= 3\mathbf{i} + 5\mathbf{j} + 7\mathbf{k}$$

(b)
$$\mathbf{v} \times \mathbf{u} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & -2 \\ 1 & -2 & 1 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & -2 \\ -2 & 1 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 3 & -2 \\ 1 & 1 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 3 & 1 \\ 1 & -2 \end{vmatrix}\mathbf{k}$$
$$= -3\mathbf{i} - 5\mathbf{j} - 7\mathbf{k}$$

Note that this result is the negative of that in part (a).

(c)
$$\mathbf{v} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & 1 & -2 \\ 3 & 1 & -2 \end{vmatrix}$$
$$= \begin{vmatrix} 1 & -2 \\ 1 & -2 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 3 & -2 \\ 3 & -2 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 3 & 1 \\ 3 & 1 \end{vmatrix}\mathbf{k}$$
$$= 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k} = \mathbf{0}$$

Algebraic Properties of the Cross Product

If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in R^3 and c is a scalar, then the following properties are true.

1. $\mathbf{u} \times \mathbf{v} = -(\mathbf{v} \times \mathbf{u})$
2. $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) + (\mathbf{u} \times \mathbf{w})$
3. $c(\mathbf{u} \times \mathbf{v}) = c\mathbf{u} \times \mathbf{v} = \mathbf{u} \times c\mathbf{v}$
4. $\mathbf{u} \times \mathbf{0} = \mathbf{0} \times \mathbf{u} = \mathbf{0}$
5. $\mathbf{u} \times \mathbf{u} = \mathbf{0}$
6. $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$

PROOF The proof of the first property is shown here. The proofs of the other properties are left to you. (See Exercises 40–44.) Let $\mathbf{u}$ and $\mathbf{v}$ be

$$\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$$

and

$$\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}.$$

Then $\mathbf{u} \times \mathbf{v}$ is

$$\begin{aligned}
\mathbf{u} \times \mathbf{v} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} \\
&= (u_2v_3 - u_3v_2)\mathbf{i} - (u_1v_3 - u_3v_1)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k},
\end{aligned}$$

and $\mathbf{v} \times \mathbf{u}$ is

$$\begin{aligned}
\mathbf{v} \times \mathbf{u} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ v_1 & v_2 & v_3 \\ u_1 & u_2 & u_3 \end{vmatrix} \\
&= (v_2u_3 - v_3u_2)\mathbf{i} - (v_1u_3 - v_3u_1)\mathbf{j} + (v_1u_2 - v_2u_1)\mathbf{k} \\
&= -(u_2v_3 - u_3v_2)\mathbf{i} + (u_1v_3 - u_3v_1)\mathbf{j} - (u_1v_2 - u_2v_1)\mathbf{k} \\
&= -(\mathbf{v} \times \mathbf{u}).
\end{aligned}$$

Geometric Properties of the Cross Product

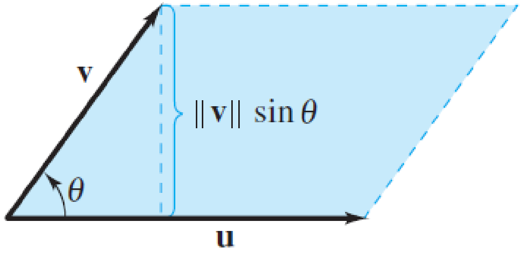
If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in R^3, then the following properties are true.

1. $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.
2. The angle θ between $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\,\|\mathbf{v}\| \sin \theta.$$

3. $\mathbf{u}$ and $\mathbf{v}$ are parallel if and only if $\mathbf{u} \times \mathbf{v} = \mathbf{0}$.
4. The parallelogram having $\mathbf{u}$ and $\mathbf{v}$ as adjacent sides has an area of $\|\mathbf{u} \times \mathbf{v}\|$.

$$\text{Area} = \overbrace{\|\mathbf{u}\|}^{\text{Base}}\,\overbrace{\|\mathbf{v}\| \sin \theta}^{\text{Height}} = \|\mathbf{u} \times \mathbf{v}\|.$$



Finding a Vector Orthogonal to Two Given Vectors

Find a unit vector orthogonal to both

$$\mathbf{u} = \mathbf{i} - 4\mathbf{j} + \mathbf{k}$$

and

$$\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}.$$

SOLUTION From Property 1 of Theorem 5.18, you know that the cross product

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & -4 & 1 \\ 2 & 3 & 0 \end{vmatrix}$$
$$= -3\mathbf{i} + 2\mathbf{j} + 11\mathbf{k}$$

is orthogonal to both **u** and **v**, as shown in Figure 5.30. Then, by dividing by the length of $\mathbf{u} \times \mathbf{v}$,

$$\|\mathbf{u} \times \mathbf{v}\| = \sqrt{(-3)^2 + 2^2 + 11^2}$$
$$= \sqrt{134},$$

you obtain the unit vector

$$\frac{\mathbf{u} \times \mathbf{v}}{\|\mathbf{u} \times \mathbf{v}\|} = -\frac{3}{\sqrt{134}}\mathbf{i} + \frac{2}{\sqrt{134}}\mathbf{j} + \frac{11}{\sqrt{134}}\mathbf{k},$$

which is orthogonal to both **u** and **v**, as follows.

$$\left(-\frac{3}{\sqrt{134}}, \frac{2}{\sqrt{134}}, \frac{11}{\sqrt{134}}\right) \cdot (1, -4, 1) = 0$$

$$\left(-\frac{3}{\sqrt{134}}, \frac{2}{\sqrt{134}}, \frac{11}{\sqrt{134}}\right) \cdot (2, 3, 0) = 0$$

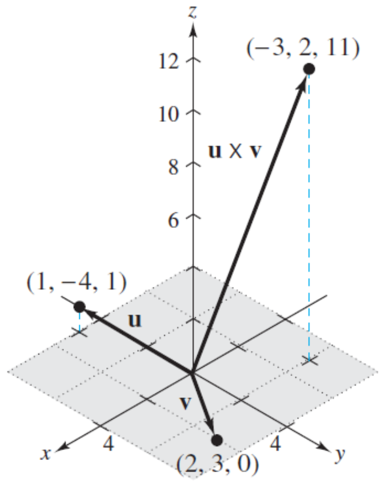


Figure 5.30

Finding the Area of a Parallelogram

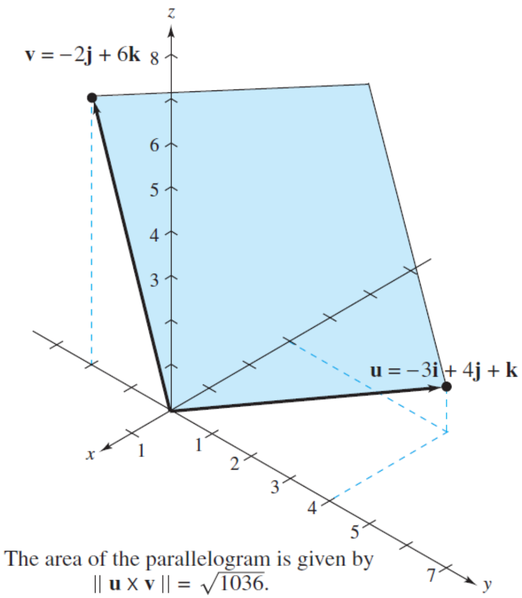
Find the area of the parallelogram that has

$$\mathbf{u} = -3\mathbf{i} + 4\mathbf{j} + \mathbf{k}$$

and

$$\mathbf{v} = -2\mathbf{j} + 6\mathbf{k}$$

as adjacent sides, as shown in Figure 5.31.



The area of the parallelogram is given by $\|\mathbf{u} \times \mathbf{v}\| = \sqrt{1036}$.

Figure 5.31

SOLUTION From Property 4 of Theorem 5.18, you know that the area of this parallelogram is $\|\mathbf{u} \times \mathbf{v}\|$. Because

$$\mathbf{u} \times \mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 4 & 1 \\ 0 & -2 & 6 \end{vmatrix} = 26\mathbf{i} + 18\mathbf{j} + 6\mathbf{k},$$

the area of the parallelogram is

$$\|\mathbf{u} \times \mathbf{v}\| = \sqrt{26^2 + 18^2 + 6^2} = \sqrt{1036} \approx 32.19.$$

Applications of Cross Product

- Calculating surface normals in 3D graphics
- Modeling torque and rotation in robotics and simulations
- Computing area and spatial relationships between vectors

สัจพจน์ (Axiom)

- หมายถึง ข้อความที่ยอมรับว่าเป็นจริงโดยไม่ต้องพิสูจน์ ซึ่งตรงข้ามกับคำว่า “ทฤษฎีบท” ซึ่งจะถูกยอมรับว่าเป็นจริงได้ก็ต่อเมื่อมีการพิสูจน์
- ดังนั้นสัจพจน์จึงถูกใช้เป็นจุดเริ่มต้นในการพิสูจน์ทางคณิตศาสตร์ และทฤษฎีบททุกอัน จะต้องอนุมาน (inference) มายังสัจพจน์ได้ เช่น
  - เราสามารถลากเส้นตรงผ่านจุดสองจุดได้
  - เราสามารถขยายส่วนของเส้นตรงไปเป็นเส้นตรงได้เส้นเดียว
  - มุมฉากทุกมุมย่อมเท่ากัน
  - สมการที่ถูกบวกด้วยค่าเท่ากันทั้งสองข้างก็ยังเป็นสมการอยู่

ที่มา: https://th.wikipedia.org/wiki/%E0%B8%AA%E0%B8%B1%E0%B8%88%E0%B8%9E%E0%B8%88%E0%B8%99%E0%B9%8C

เซต (Set)

- หมายถึง กลุ่มของสิ่งใดสิ่งหนึ่ง
- เซต = {สมาชิกของเซตตัวที่ 1, สมาชิกของเซตตัวที่ 2, ... }
- สมาชิกของเซต จะอยู่ภายใต้เครื่องหมาย ปีกกา { }

  เช่น เซต A คือของจำนวนคู่ที่เป็นบวก จะได้ว่า

  $$A = \{2, 4, 6, \dots\}$$

  การเป็นสมาชิกของเซต ใช้เครื่องหมาย ∈ เช่น
- เซตที่เป็นที่รู้จัก เช่น
  - เซตของจำนวนจริง แทนด้วย R
  - เซตของคู่ลำดับในระนาบ แทนด้วย R^2
  - เซตของจำนวนเต็มบวก แทนด้วย I^+
  - เซตของเมทริกซ์ ขนาด $m \times n$ แทนด้วย $M_{m\times n}$

Vector Spaces

- *Any* set that satisfies these properties (or **axioms**) is called a **vector space**, and the objects in the set are called **vectors**.

Let V be a set on which two operations (**vector addition** and **scalar multiplication**) are defined. If the listed axioms are satisfied for every $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and every scalar (real number) c and d, then V is called a **vector space.**

Addition:

1. $\mathbf{u} + \mathbf{v}$ is in V.	**Closure under addition**
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	**Commutative property**
3. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$	**Associative property**
4. V has a **zero vector** $\mathbf{0}$ such that for every $\mathbf{u}$ in V, $\mathbf{u} + \mathbf{0} = \mathbf{u}$.	**Additive identity**
5. For every $\mathbf{u}$ in V, there is a vector in V denoted by $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.	**Additive inverse**

Scalar Multiplication:

6. $c\mathbf{u}$ is in V.	**Closure under scalar multiplication**
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	**Distributive property**
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	**Distributive property**
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$	**Associative property**
10. $1(\mathbf{u}) = \mathbf{u}$	**Scalar identity**

R^2 with the Standard Operations Is a Vector Space

The set of all ordered pairs of real numbers R^2 with the standard operations is a vector space. To verify this, look back at Theorem 4.1. Vectors in this space have the form

$$\mathbf{v} = (v_1, v_2).$$

Let **u**, **v**, and **w** be vectors in the plane, and let c and d be scalars.

1.	$\mathbf{u} + \mathbf{v}$ is a vector in the plane.	Closure under addition
2.	$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Commutative property of addition
3.	$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	Associative property of addition
4.	$\mathbf{u} + \mathbf{0} = \mathbf{u}$	Additive identity property
5.	$\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	Additive inverse property
6.	$c\mathbf{u}$ is a vector in the plane.	Closure under scalar multiplication
7.	$c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	Distributive property
8.	$(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	Distributive property
9.	$c(d\mathbf{u}) = (cd)\mathbf{u}$	Associative property of multiplication
10.	$1(\mathbf{u}) = \mathbf{u}$	Multiplicative identity property

R^n with the Standard Operations Is a Vector Space

The set of all ordered n-tuples of real numbers R^n with the standard operations is a vector space. This is verified by Theorem 4.2. Vectors in this space are of the form

$$\mathbf{v} = (v_1, v_2, v_3, \ldots, v_n).$$

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in R^n, and let c and d be scalars.

1. $\mathbf{u} + \mathbf{v}$ is a vector in R^n.	Closure under addition
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$	Commutative property of addition
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$	Associative property addition
4. $\mathbf{u} + \mathbf{0} = \mathbf{u}$	Additive identity property
5. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$	Additive inverse property
6. $c\mathbf{u}$ is a vector in R^n.	Closure under scalar multiplication
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$	Distributive property
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$	Distributive property
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$	Associative property of multiplication
10. $1(\mathbf{u}) = \mathbf{u}$	Multiplicative identity property

The Vector Space of All 2 × 3 Matrices

Show that the set of all 2×3 matrices with the operations of matrix addition and scalar multiplication is a vector space.

SOLUTION If A and B are 2×3 matrices and c is a scalar, then $A + B$ and cA are also 2×3 matrices. The set is, therefore, closed under matrix addition and scalar multiplication. Moreover, the other eight vector space axioms follow directly from Theorems 2.1 and 2.2 (see Section 2.2). You can conclude that the set is a vector space. Vectors in this space have the form

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

The Vector Space of All Polynomials of Degree 2 or Less

Let P_2 be the set of all polynomials of the form

$$p(x) = a_2x^2 + a_1x + a_0,$$

where $a_0, a_1,$ and a_2 are real numbers. The *sum* of two polynomials $p(x) = a_2x^2 + a_1x + a_0$ and $q(x) = b_2x^2 + b_1x + b_0$ is defined in the usual way by

$$p(x) + q(x) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0),$$

and the *scalar multiple* of $p(x)$ by the scalar c is defined by

$$cp(x) = ca_2x^2 + ca_1x + ca_0.$$

Show that P_2 is a vector space.

SOLUTION Verification of each of the ten vector space axioms is a straightforward application of the properties of real numbers. For instance, because the set of real numbers is closed under addition, it follows that $a_2 + b_2$, $a_1 + b_1$, and $a_0 + b_0$ are real numbers, and

$$p(x) + q(x) = (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)$$

is in the set P_2 because it is a polynomial of degree 2 or less. P_2 is closed under addition. Similarly, you can use the fact that the set of real numbers is closed under multiplication to show that P_2 is closed under scalar multiplication. To verify the commutative axiom of addition, write

$$\begin{aligned} p(x) + q(x) &= (a_2x^2 + a_1x + a_0) + (b_2x^2 + b_1x + b_0) \\ &= (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0) \\ &= (b_2 + a_2)x^2 + (b_1 + a_1)x + (b_0 + a_0) \\ &= (b_2x^2 + b_1x + b_0) + (a_2x^2 + a_1x + a_0) \\ &= q(x) + p(x). \end{aligned}$$

Can you see where the commutative property of addition of real numbers was used? The zero vector in this space is the zero polynomial given by $\mathbf{0}(x) = 0x^2 + 0x + 0$, for all x. Try verifying the other vector space axioms. You may then conclude that P_2 is a vector space.

Properties of Scalar Multiplication

Let $\mathbf{v}$ be any element of a vector space V, and let c be any scalar. Then the following properties are true.

1. $0\mathbf{v} = \mathbf{0}$
2. $c\mathbf{0} = \mathbf{0}$
3. If $c\mathbf{v} = \mathbf{0}$, then $c = 0$ or $\mathbf{v} = \mathbf{0}$.
4. $(-1)\mathbf{v} = -\mathbf{v}$

PROOF To prove these properties, you are restricted to using the ten vector space axioms. For instance, to prove the second property, note from axiom 4 that $\mathbf{0} = \mathbf{0} + \mathbf{0}$. This allows you to write the steps below.

$$
\begin{aligned}
c\mathbf{0} &= c(\mathbf{0} + \mathbf{0}) && \text{Additive identity} \\
c\mathbf{0} &= c\mathbf{0} + c\mathbf{0} && \text{Left distributive property} \\
c\mathbf{0} + (-c\mathbf{0}) &= (c\mathbf{0} + c\mathbf{0}) + (-c\mathbf{0}) && \text{Add } -c\mathbf{0} \text{ to both sides.} \\
c\mathbf{0} + (-c\mathbf{0}) &= c\mathbf{0} + [c\mathbf{0} + (-c\mathbf{0})] && \text{Associative property} \\
\mathbf{0} &= c\mathbf{0} + \mathbf{0} && \text{Additive inverse} \\
\mathbf{0} &= c\mathbf{0} && \text{Additive identity}
\end{aligned}
$$

To prove the third property, suppose that $c\mathbf{v} = \mathbf{0}$. To show that this implies either $c = 0$ or $\mathbf{v} = \mathbf{0}$, assume that $c \neq 0$. (If $c = 0$, you have nothing more to prove.) Now, because $c \neq 0$, you can use the reciprocal $1/c$ to show that $\mathbf{v} = \mathbf{0}$, as follows.

$$\mathbf{v} = 1\mathbf{v} = \left(\frac{1}{c}\right)(c)\mathbf{v} = \frac{1}{c}(c\mathbf{v}) = \frac{1}{c}(\mathbf{0}) = \mathbf{0}$$

Note that the last step uses Property 2 (the one you just proved). The proofs of the first and fourth properties are left as exercises.

The Set of Integers Is Not a Vector Space

The set of all integers (with the standard operations) does not form a vector space because it is not closed under scalar multiplication. For example,

$$\underbrace{\tfrac{1}{2}}_{\text{Scalar}}\underbrace{(1)}_{\text{Integer}} = \underbrace{\tfrac{1}{2}}_{\text{Noninteger}}.$$

The Set of Second-Degree Polynomials Is Not a Vector Space

The set of all second-degree polynomials is not a vector space because it is not closed under addition. To see this, consider the second-degree polynomials

$$p(x) = x^2 \qquad \text{and} \qquad q(x) = -x^2 + x + 1,$$

whose sum is the first-degree polynomial

$$p(x) + q(x) = x + 1.$$

A Set That Is Not a Vector Space

Let $V = R^2$, the set of all ordered pairs of real numbers, with the standard operation of addition and the *nonstandard* definition of scalar multiplication listed below.

$$c(x_1, x_2) = (cx_1, 0)$$

Show that V is not a vector space.

SOLUTION In this example, the operation of scalar multiplication is not the standard one. For instance, the product of the scalar 2 and the ordered pair $(3, 4)$ does not equal $(6, 8)$. Instead, the second component of the product is 0,

$$2(3, 4) = (2 \cdot 3, 0) = (6, 0).$$

This example is interesting because it actually satisfies the first nine axioms of the definition of a vector space (try showing this). The tenth axiom is where you get into trouble. In attempting to verify that axiom, the nonstandard definition of scalar multiplication gives you

$$1(1, 1) = (1, 0) \neq (1, 1).$$

The tenth axiom is not verified and the set (together with the two operations) is not a vector space.

Subspaces of Vector Spaces

- In most important applications in linear algebra, vector spaces occur as **subspaces** of larger spaces.
- For instance, you will see that the solution set of a homogeneous system of linear equations in n variables is a subspace of R^n.

A nonempty subset W of a vector space V is called a **subspace** of V if W is a vector space under the operations of addition and scalar multiplication defined in V.

REMARK: Note that if W is a subspace of V, it must be closed under the operations inherited from V.

A Subspace of R^3

Show that the set $W = \{(x_1, 0, x_3): x_1 \text{ and } x_3 \text{ are real numbers}\}$ is a subspace of R^3 with the standard operations.

SOLUTION The set W is nonempty because it contains the zero vector $(0, 0, 0)$.

Graphically, the set W can be interpreted as simply the xz-plane, as shown in Figure 4.9. The set W is closed under addition because the sum of any two vectors in the xz-plane must also lie in the xz-plane. That is, if $(x_1, 0, x_3)$ and $(y_1, 0, y_3)$ are in W, then their sum $(x_1 + y_1, 0, x_3 + y_3)$ is also in W (because the second component is zero). Similarly, to see that W is closed under scalar multiplication, let $(x_1, 0, x_3)$ be in W and let c be a scalar. Then $c(x_1, 0, x_3) = (cx_1, 0, cx_3)$ has zero as its second component and must be in W. The other eight vector space axioms can be verified as well, and these verifications are left to you.

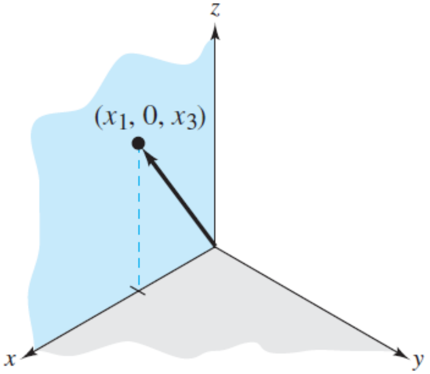


Figure 4.9

Test for a Subspace

If W is a nonempty subset of a vector space V, then W is a subspace of V if and only if the following closure conditions hold.

1. If $\mathbf{u}$ and $\mathbf{v}$ are in W, then $\mathbf{u} + \mathbf{v}$ is in W.
2. If $\mathbf{u}$ is in W and c is any scalar, then $c\mathbf{u}$ is in W.

PROOF The proof of this theorem in one direction is straightforward. That is, if W is a subspace of V, then W is a vector space and must be closed under addition and scalar multiplication.

To prove the theorem in the other direction, assume that W is closed under addition and scalar multiplication. Note that if $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are in W, then they are also in V. Consequently, vector space axioms 2, 3, 7, 8, 9, and 10 are satisfied automatically. Because W is closed under addition and scalar multiplication, it follows that for any $\mathbf{v}$ in W and scalar $c = 0$,

$$c\mathbf{v} = \mathbf{0}$$

and

$$(-1)\mathbf{v} = -\mathbf{v}$$

both lie in W, which satisfies axioms 4 and 5.

The Subspace of $M_{2,2}$

Let W be the set of all 2×2 symmetric matrices. Show that W is a subspace of the vector space $M_{2,2}$, with the standard operations of matrix addition and scalar multiplication.

SOLUTION Recall that a matrix is called *symmetric* if it is equal to its own transpose. Because $M_{2,2}$ is a vector space, you only need to show that W (a subset of $M_{2,2}$) satisfies the conditions of Theorem 4.5. Begin by observing that W is *nonempty*. W is closed under addition because $A_1 = A_1^T$ and $A_2 = A_2^T$, which implies that

$$(A_1 + A_2)^T = A_1^T + A_2^T = A_1 + A_2.$$

So, if A_1 and A_2 are symmetric matrices of order 2, then so is $A_1 + A_2$. Similarly, W is closed under scalar multiplication because $A = A^T$ implies that $(cA)^T = cA^T = cA$. If A is a symmetric matrix of order 2, then so is cA.

The Set of Singular Matrices Is Not a Subspace of $M_{n,n}$

Let W be the set of singular matrices of order 2. Show that W is not a subspace of $M_{2,2}$ with the standard operations.

SOLUTION By Theorem 4.5, you can show that a subset W is not a subspace by showing that W is empty, W is not closed under addition, or W is not closed under scalar multiplication. For this particular set, W is nonempty and closed under scalar multiplication, but it is not closed under addition. To see this, let A and B be

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

Then A and B are both singular (noninvertible), but their sum

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is nonsingular (invertible). So W is not closed under addition, and by Theorem 4.5 you can conclude that it is not a subspace of $M_{2,2}$.

The Set of First Quadrant Vectors Is Not a Subspace of R^2

Show that $W = \{(x_1, x_2): x_1 \geq 0 \text{ and } x_2 \geq 0\}$, with the standard operations, is not a subspace of R^2.

SOLUTION This set is nonempty and closed under addition. It is not, however, closed under scalar multiplication. To see this, note that $(1, 1)$ is in W, but the scalar multiple

$$(-1)(1, 1) = (-1, -1)$$

is not in W. So W is not a subspace of R^2.

Linear Dependence and Linear Independence

A set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V is called **linearly independent** if the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

has only the trivial solution, $c_1 = 0, c_2 = 0, \ldots, c_k = 0$. If there are also nontrivial solutions, then S is called **linearly dependent.**

Examples of Linearly Dependent Sets

(a) The set $S = \{(1, 2), (2, 4)\}$ in R^2 is linearly dependent because

$$-2(1, 2) + (2, 4) = (0, 0).$$

(b) The set $S = \{(1, 0), (0, 1), (-2, 5)\}$ in R^2 is linearly dependent because

$$2(1, 0) - 5(0, 1) + (-2, 5) = (0, 0).$$

(c) The set $S = \{(0, 0), (1, 2)\}$ in R^2 is linearly dependent because

$$1(0, 0) + 0(1, 2) = (0, 0).$$

Testing for Linear Independence

Determine whether the set of vectors in R^3 is linearly independent or linearly dependent.

$$S = \{\overset{\mathbf{v}_1}{(1, 2, 3)}, \overset{\mathbf{v}_2}{(0, 1, 2)}, \overset{\mathbf{v}_3}{(-2, 0, 1)}\}$$

SOLUTION To test for linear independence or linear dependence, form the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}.$$

If the only solution of this equation is

$$c_1 = c_2 = c_3 = 0,$$

then the set S is linearly independent. Otherwise, S is linearly dependent. Expanding this equation, you have

$$\begin{aligned} c_1(1, 2, 3) + c_2(0, 1, 2) + c_3(-2, 0, 1) &= (0, 0, 0) \\ (c_1 - 2c_3, 2c_1 + c_2, 3c_1 + 2c_2 + c_3) &= (0, 0, 0), \end{aligned}$$

which yields the homogeneous system of linear equations in c_1, c_2, and c_3 shown below.

$$\begin{aligned} c_1 \qquad - 2c_3 &= 0 \\ 2c_1 + c_2 \qquad &= 0 \\ 3c_1 + 2c_2 + c_3 &= 0 \end{aligned}$$

The augmented matrix of this system reduces by Gauss-Jordan elimination as follows.

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

This implies that the only solution is the trivial solution

$$c_1 = c_2 = c_3 = 0.$$

So, S is linearly independent.

Testing for Linear Independence and Linear Dependence

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, . . . , \mathbf{v}_k\}$ be a set of vectors in a vector space V. To determine whether S is linearly independent or linearly dependent, perform the following steps.

1. From the vector equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdot \cdot \cdot + c_k\mathbf{v}_k = \mathbf{0}$, write a homogeneous system of linear equations in the variables $c_1, c_2, . . . ,$ and c_k.
2. Use Gaussian elimination to determine whether the system has a unique solution.
3. If the system has only the trivial solution, $c_1 = 0, c_2 = 0, . . . , c_k = 0$, then the set S is linearly independent. If the system also has nontrivial solutions, then S is linearly dependent.

Testing for Linear Independence

Determine whether the set of vectors in P_2 is linearly independent or linearly dependent.

$$S = \{\underbrace{1 + x - 2x^2}_{\mathbf{v}_1}, \underbrace{2 + 5x - x^2}_{\mathbf{v}_2}, \underbrace{x + x^2}_{\mathbf{v}_3}\}$$

SOLUTION Expanding the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ produces

$$c_1(1 + x - 2x^2) + c_2(2 + 5x - x^2) + c_3(x + x^2) = 0 + 0x + 0x^2$$
$$(c_1 + 2c_2) + (c_1 + 5c_2 + c_3)x + (-2c_1 - c_2 + c_3)x^2 = 0 + 0x + 0x^2.$$

Equating corresponding coefficients of equal powers of x produces the homogeneous system of linear equations in c_1, c_2, and c_3 shown below.

$$\begin{aligned} c_1 + 2c_2 \phantom{{}+c_3} &= 0 \\ c_1 + 5c_2 + c_3 &= 0 \\ -2c_1 - c_2 + c_3 &= 0 \end{aligned}$$

The augmented matrix of this system reduces by Gaussian elimination as follows.

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

This implies that the system has an infinite number of solutions. So, the system must have nontrivial solutions, and you can conclude that the set S is linearly dependent.

One nontrivial solution is

$$c_1 = 2, \quad c_2 = -1, \quad \text{and} \quad c_3 = 3,$$

which yields the nontrivial linear combination

$$(2)(1 + x - 2x^2) + (-1)(2 + 5x - x^2) + (3)(x + x^2) = 0.$$

Testing for Linear Independence

Determine whether the set of vectors in $M_{2,2}$ is linearly independent or linearly dependent.

$$S = \left\{ \overset{\mathbf{v}_1}{\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}}, \overset{\mathbf{v}_2}{\begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}}, \overset{\mathbf{v}_3}{\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}} \right\}$$

SOLUTION From the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0},$$

you have

$$c_1\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} + c_2\begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which produces the system of linear equations in c_1, c_2, and c_3 shown below.

$$\begin{aligned} 2c_1 + 3c_2 + c_3 &= 0 \\ c_1 &= 0 \\ 2c_2 + 2c_3 &= 0 \\ c_1 + c_2 &= 0 \end{aligned}$$

Using Gaussian elimination, the augmented matrix of this system reduces as follows.

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

The system has only the trivial solution and you can conclude that the set S is linearly independent.

A Property of Linearly Dependent Sets

A set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, $k \geq 2$, is linearly dependent if and only if at least one of the vectors $\mathbf{v}_j$ can be written as a linear combination of the other vectors in S.

PROOF To prove the theorem in one direction, assume S is a linearly dependent set. Then there exist scalars $c_1, c_2, c_3, \ldots, c_k$ (not all zero) such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_k\mathbf{v}_k = \mathbf{0}.$$

Because one of the coefficients must be nonzero, no generality is lost by assuming $c_1 \neq 0$. Then solving for $\mathbf{v}_1$ as a linear combination of the other vectors produces

$$c_1\mathbf{v}_1 = -c_2\mathbf{v}_2 - c_3\mathbf{v}_3 - \cdots - c_k\mathbf{v}_k$$

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 - \frac{c_3}{c_1}\mathbf{v}_3 - \cdots - \frac{c_k}{c_1}\mathbf{v}_k.$$

Conversely, suppose the vector $\mathbf{v}_1$ in S is a linear combination of the other vectors. That is,

$$\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_k\mathbf{v}_k.$$

Then the equation $-\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ has at least one coefficient, -1, that is nonzero, and you can conclude that S is linearly dependent.

Writing a Vector as a Linear Combination of Other Vectors

In Example 9, you determined that the set

$$S = \{\overset{\mathbf{v}_1}{1 + x - 2x^2}, \overset{\mathbf{v}_2}{2 + 5x - x^2}, \overset{\mathbf{v}_3}{x + x^2}\}$$

is linearly dependent. Show that one of the vectors in this set can be written as a linear combination of the other two.

SOLUTION In Example 9, the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0}$ produced the system

$$\begin{aligned} c_1 + 2c_2 \qquad &= 0 \\ c_1 + 5c_2 + c_3 &= 0 \\ -2c_1 - \quad c_2 + c_3 &= 0. \end{aligned}$$

This system has an infinite number of solutions represented by $c_3 = 3t$, $c_2 = -t$, and $c_1 = 2t$. Letting $t = 1$ results in the equation $2\mathbf{v}_1 - \mathbf{v}_2 + 3\mathbf{v}_3 = \mathbf{0}$. So, $\mathbf{v}_2$ can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_3$ as follows.

$$\mathbf{v}_2 = 2\mathbf{v}_1 + 3\mathbf{v}_3$$

A check yields

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

Corollary

Two vectors **u** and **v** in a vector space V are linearly dependent if and only if one is a scalar multiple of the other.

REMARK: The zero vector is always a scalar multiple of another vector in a vector space.

Testing for Linear Dependence of Two Vectors

(a) The set

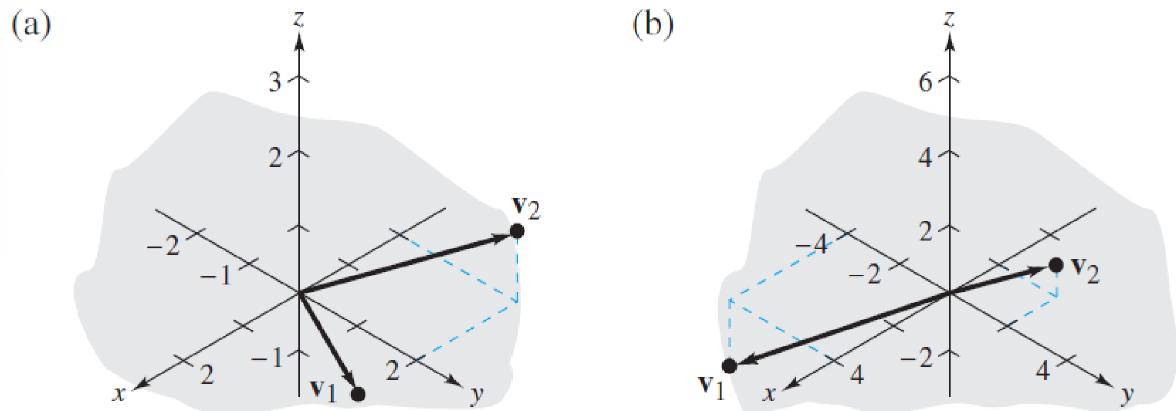
$$S = \{\overset{\mathbf{v}_1}{(1, 2, 0)}, \overset{\mathbf{v}_2}{(-2, 2, 1)}\}$$

is linearly independent because $\mathbf{v}_1$ and $\mathbf{v}_2$ are not scalar multiples of each other, as shown in Figure 4.17(a).

(b) The set

$$S = \{\overset{\mathbf{v}_1}{(4, -4, -2)}, \overset{\mathbf{v}_2}{(-2, 2, 1)}\}$$

is linearly dependent because $\mathbf{v}_1 = -2\mathbf{v}_2$, as shown in Figure 4.17(b).

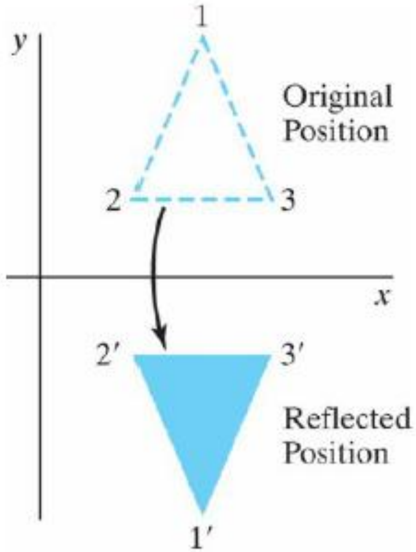


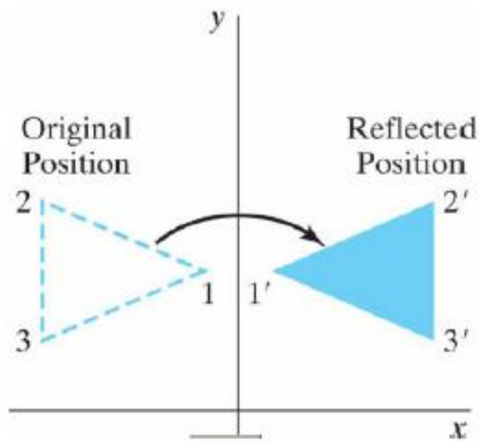
$S = \{(1, 2, 0), (-2, 2, 1)\}$
The set S is linearly independent.

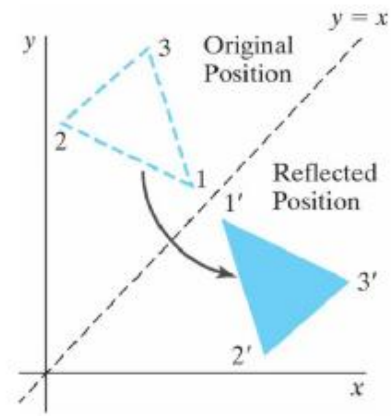
$S = \{(4, -4, -2), (-2, 2, 1)\}$
The set S is linearly dependent because $\mathbf{v}_1 = -2\mathbf{v}_2$.

Figure 4.17

การประยุกต์การแปลงเชิงเส้น : Computer Graphics

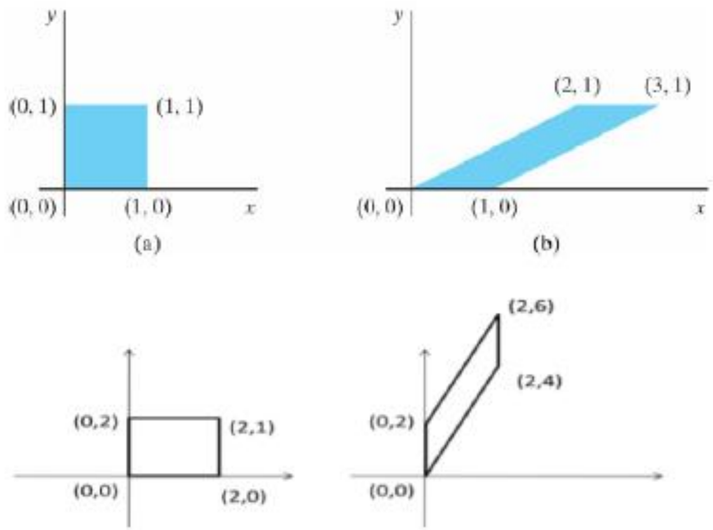




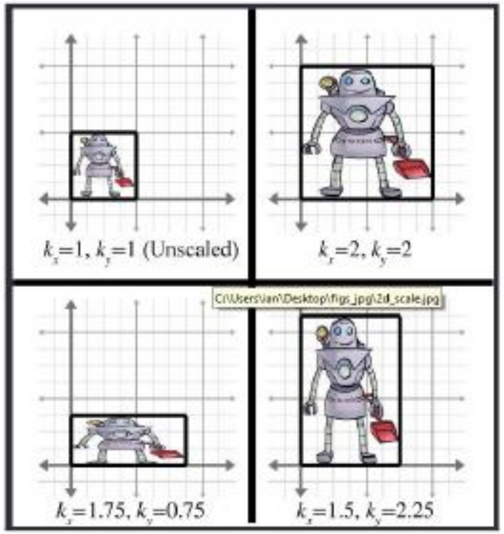


การสะท้อน (Reflection)

การประยุกต์การแปลงเชิงเส้น : Computer Graphics

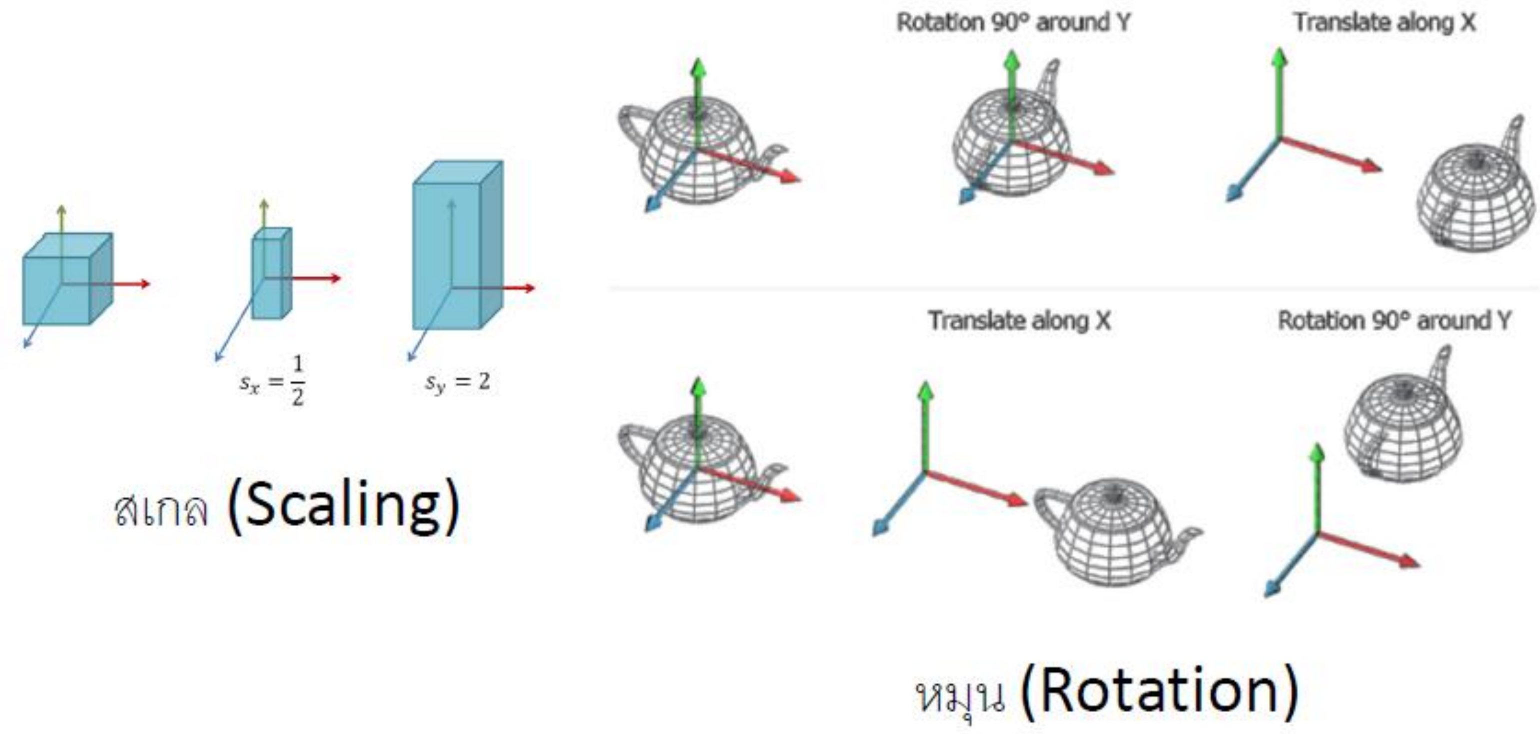


ตัดเฉือน (Shearing)



สเกล (Scaling)

การประยุกต์การแปลงเชิงเส้น : Computer Graphics



Introduction to Linear Transformations

A function that maps a vector space V into a vector space W is denoted by

$$T: V \to W.$$

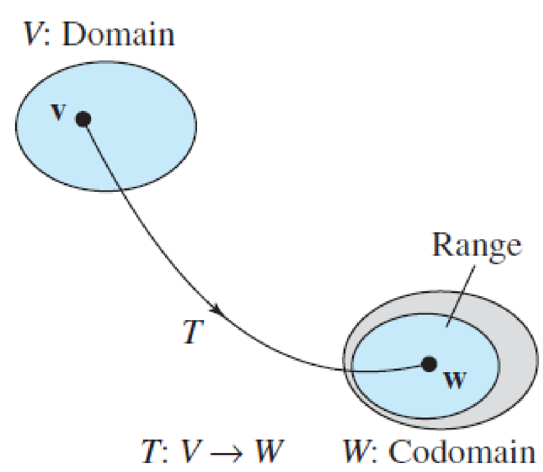
The standard function terminology is used for such functions. For instance, V is called the domain of T, and W is called the codomain of T. If $\mathbf{v}$ is in V and $\mathbf{w}$ is in W such that

$$T(\mathbf{v}) = \mathbf{w},$$

then $\mathbf{w}$ is called the image of $\mathbf{v}$ under T. The set of all images of vectors in V is called the range of T, and the set of all $\mathbf{v}$ in V such that $T(\mathbf{v}) = \mathbf{w}$ is called the preimage of $\mathbf{w}$.

REMARK: For a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in R^n, it would be technically correct to use double parentheses to denote $T(\mathbf{v})$ as $T(\mathbf{v}) = T((v_1, v_2, \ldots, v_n))$. For convenience, however, one set of parentheses is dropped, producing

$$T(\mathbf{v}) = T(v_1, v_2, \ldots, v_n).$$



A Function from R^2 into R^2

For any vector $\mathbf{v} = (v_1, v_2)$ in R^2, let $T: R^2 \to R^2$ be defined by

$$T(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2).$$

(a) Find the image of $\mathbf{v} = (-1, 2)$.

(b) Find the preimage of $\mathbf{w} = (-1, 11)$.

SOLUTION (a) For $\mathbf{v} = (-1, 2)$ you have

$$T(-1, 2) = (-1 - 2, -1 + 2(2)) = (-3, 3).$$

(b) If $T(\mathbf{v}) = (v_1 - v_2, v_1 + 2v_2) = (-1, 11)$, then

$$\begin{aligned} v_1 - v_2 &= -1 \\ v_1 + 2v_2 &= 11. \end{aligned}$$

This system of equations has the unique solution $v_1 = 3$ and $v_2 = 4$. So, the preimage of $(-1, 11)$ is the set in R^2 consisting of the single vector $(3, 4)$.

Definition of a Linear Transformation

Let V and W be vector spaces. The function $T: V \rightarrow W$ is called a **linear transformation** of V into W if the following two properties are true for all $\mathbf{u}$ and $\mathbf{v}$ in V and for any scalar c.

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2. $T(c\mathbf{u}) = cT(\mathbf{u})$

A linear transformation is said to be *operation preserving*, because the same result occurs whether the operations of addition and scalar multiplication are performed before or after the linear transformation is applied. Although the same symbols are used to denote the vector operations in both V and W, you should note that the operations may be different, as indicated in the diagram below.

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \qquad T(c\mathbf{u}) = cT(\mathbf{u})$$

Verifying a Linear Transformation from R^2 into R^2

Show that the function given in Example 1 is a linear transformation from R^2 into R^2.

$$T(v_1, v_2) = (v_1 - v_2, v_1 + 2v_2)$$

SOLUTION To show that the function T is a linear transformation, you must show that it preserves vector addition and scalar multiplication. To do this, let $\mathbf{v} = (v_1, v_2)$ and $\mathbf{u} = (u_1, u_2)$ be vectors in R^2 and let c be any real number. Then, using the properties of vector addition and scalar multiplication, you have the two statements below.

1. Because $\mathbf{u} + \mathbf{v} = (u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$, you have

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T(u_1 + v_1, u_2 + v_2) \\ &= ((u_1 + v_1) - (u_2 + v_2), (u_1 + v_1) + 2(u_2 + v_2)) \\ &= ((u_1 - u_2) + (v_1 - v_2), (u_1 + 2u_2) + (v_1 + 2v_2)) \\ &= (u_1 - u_2, u_1 + 2u_2) + (v_1 - v_2, v_1 + 2v_2) \\ &= T(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

REMARK: A linear transformation $T: V \to V$ from a vector space into itself (as in Example 2) is called a **linear operator.**

2. Because $c\mathbf{u} = c(u_1, u_2) = (cu_1, cu_2)$, you have

$$\begin{aligned} T(c\mathbf{u}) &= T(cu_1, cu_2) \\ &= (cu_1 - cu_2, cu_1 + 2cu_2) \\ &= c(u_1 - u_2, u_1 + 2u_2) \\ &= cT(\mathbf{u}). \end{aligned}$$

So, T is a linear transformation.

Some Functions That Are Not Linear Transformations

(a) $f(x) = \sin x$ is not a linear transformation from R into R because, in general,

$$\sin(x_1 + x_2) \neq \sin x_1 + \sin x_2.$$

For instance, $\sin[(\pi/2) + (\pi/3)] \neq \sin(\pi/2) + \sin(\pi/3)$.

(b) $f(x) = x^2$ is not a linear transformation from R into R because, in general,

$$(x_1 + x_2)^2 \neq x_1^2 + x_2^2.$$

For instance, $(1 + 2)^2 \neq 1^2 + 2^2$.

(c) $f(x) = x + 1$ is not a linear transformation from R into R because

$$f(x_1 + x_2) = x_1 + x_2 + 1$$

whereas

$$f(x_1) + f(x_2) = (x_1 + 1) + (x_2 + 1) = x_1 + x_2 + 2.$$

So $f(x_1 + x_2) \neq f(x_1) + f(x_2)$.

Properties of Linear Transformations

Two simple linear transformations are the **zero transformation** and the **identity transformation,** which are defined as follows.

1. $T(\mathbf{v}) = \mathbf{0}$, for all $\mathbf{v}$ — Zero transformation $(T: V \to W)$
2. $T(\mathbf{v}) = \mathbf{v}$, for all $\mathbf{v}$ — Identity transformation $(T: V \to V)$

Let T be a linear transformation from V into W, where $\mathbf{u}$ and $\mathbf{v}$ are in V. Then the following properties are true.

1. $T(\mathbf{0}) = \mathbf{0}$
2. $T(-\mathbf{v}) = -T(\mathbf{v})$
3. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$
4. If $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdot\cdot\cdot + c_n\mathbf{v}_n$,

   then

$$\begin{aligned} T(\mathbf{v}) &= T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdot\cdot\cdot + c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdot\cdot\cdot + c_nT(\mathbf{v}_n). \end{aligned}$$

Linear Transformations and Bases

Let $T: R^3 \to R^3$ be a linear transformation such that

$$T(1, 0, 0) = (2, -1, 4)$$
$$T(0, 1, 0) = (1, 5, -2)$$
$$T(0, 0, 1) = (0, 3, 1).$$

Find $T(2, 3, -2)$.

SOLUTION Because $(2, 3, -2)$ can be written as

$$(2, 3, -2) = 2(1, 0, 0) + 3(0, 1, 0) - 2(0, 0, 1),$$

you can use Property 4 of Theorem 6.1 to write

$$\begin{aligned} T(2, 3, -2) &= 2T(1, 0, 0) + 3T(0, 1, 0) - 2T(0, 0, 1) \\ &= 2(2, -1, 4) + 3(1, 5, -2) - 2(0, 3, 1) \\ &= (7, 7, 0). \end{aligned}$$

A Linear Transformation Defined by a Matrix

The function $T\colon R^2 \to R^3$ is defined as follows.

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

(a) Find $T(\mathbf{v})$, where $\mathbf{v} = (2, -1)$.

(b) Show that T is a linear transformation from R^2 into R^3.

SOLUTION (a) Because $\mathbf{v} = (2, -1)$, you have

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \\ -1 & -2 \end{bmatrix} \underset{\boxed{\text{Vector in } R^2}}{\begin{bmatrix} 2 \\ -1 \end{bmatrix}} = \underset{\boxed{\text{Vector in } R^3}}{\begin{bmatrix} 6 \\ 3 \\ 0 \end{bmatrix}}.$$

So, you have $T(2, -1) = (6, 3, 0)$.

(b) Begin by observing that T does map a vector in R^2 to a vector in R^3. To show that T is a linear transformation, use the properties of matrix multiplication, as shown in Theorem 2.3. For any vectors $\mathbf{u}$ and $\mathbf{v}$ in R^2, the distributive property of matrix multiplication over addition produces

$$T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v}).$$

Similarly, for any vector $\mathbf{u}$ in R^2 and any scalar c, the commutative property of scalar multiplication with matrix multiplication produces

$$T(c\mathbf{u}) = A(c\mathbf{u}) = c(A\mathbf{u}) = cT(\mathbf{u}).$$

The Linear Transformation Given by a Matrix

Let A be an $m \times n$ matrix. The function T defined by

$$T(\mathbf{v}) = A\mathbf{v}$$

is a linear transformation from R^n into R^m. In order to conform to matrix multiplication with an $m \times n$ matrix, the vectors in R^n are represented by $n \times 1$ matrices and the vectors in R^m are represented by $m \times 1$ matrices.

REMARK: The $m \times n$ zero matrix corresponds to the zero transformation from R^n into R^m, and the $n \times n$ identity matrix I_n corresponds to the identity transformation from R^n into R^n.

Be sure you see that an $m \times n$ matrix A defines a linear transformation from R^n into R^m.

$$A\mathbf{v} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}}_{\text{Vector in } R^n} = \underbrace{\begin{bmatrix} a_{11}v_1 + a_{12}v_2 + \cdots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \cdots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \cdots + a_{mn}v_n \end{bmatrix}}_{\text{Vector in } R^m}$$

Linear Transformation Given by Matrices

The linear transformation $T: R^n \rightarrow R^m$ is defined by $T(\mathbf{v}) = A\mathbf{v}$. Find the dimensions of R^n and R^m for the linear transformation represented by each matrix.

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

(b) $A = \begin{bmatrix} 2 & -3 \\ -5 & 0 \\ 0 & -2 \end{bmatrix}$

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

SOLUTION (a) Because the size of this matrix is 3×3, it defines a linear transformation from R^3 into R^3.

$$A\mathbf{v} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 3 & 0 \\ 4 & 2 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}}_{\text{Vector in } R^3} = \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}}_{\text{Vector in } R^3}$$

(b) Because the size of this matrix is 3×2, it defines a linear transformation from R^2 into R^3.

(c) Because the size of this matrix is 2×4, it defines a linear transformation from R^4 into R^2.

Rotation in the Plane

Show that the linear transformation $T: R^2 \to R^2$ represented by the matrix

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

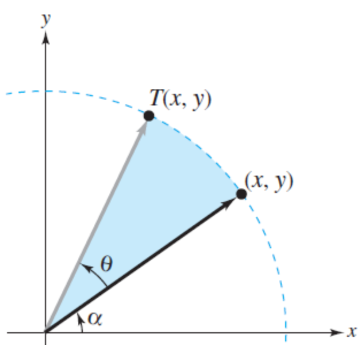
has the property that it rotates every vector in R^2 counterclockwise about the origin through the angle θ.

SOLUTION From Theorem 6.2, you know that T is a linear transformation. To show that it rotates every vector in R^2 counterclockwise through the angle θ, let $\mathbf{v} = (x, y)$ be a vector in R^2. Using polar coordinates, you can write $\mathbf{v}$ as

$$\mathbf{v} = (x, y) = (r\cos\alpha, r\sin\alpha),$$

where r is the length of $\mathbf{v}$ and α is the angle from the positive x-axis counterclockwise to the vector $\mathbf{v}$. Now, applying the linear transformation T to $\mathbf{v}$ produces

$$\begin{aligned} T(\mathbf{v}) = A\mathbf{v} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix} \\ &= \begin{bmatrix} r\cos\theta\cos\alpha - r\sin\theta\sin\alpha \\ r\sin\theta\cos\alpha + r\cos\theta\sin\alpha \end{bmatrix} \\ &= \begin{bmatrix} r\cos(\theta+\alpha) \\ r\sin(\theta+\alpha) \end{bmatrix}. \end{aligned}$$



Rotation in the Plane

Figure 6.2

So, the vector $T(\mathbf{v})$ has the same length as $\mathbf{v}$. Furthermore, because the angle from the positive x-axis to $T(\mathbf{v})$ is $\theta + \alpha$, $T(\mathbf{v})$ is the vector that results from rotating the vector $\mathbf{v}$ counterclockwise through the angle θ, as shown in Figure 6.2 on the previous page.

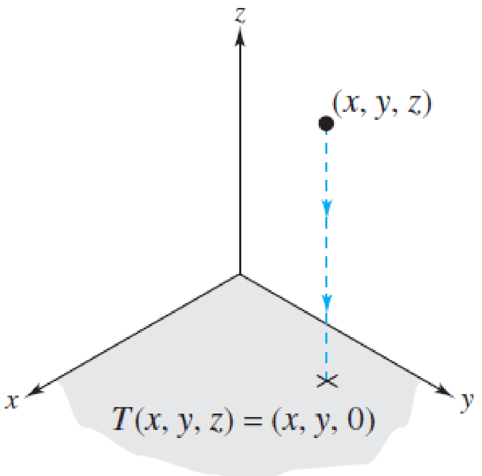
REMARK: The linear transformation in Example 7 is called a **rotation** in R^2. Rotations in R^2 preserve both vector length and the angle between two vectors. That is, the angle between $\mathbf{u}$ and $\mathbf{v}$ is equal to the angle between $T(\mathbf{u})$ and $T(\mathbf{v})$.

A Projection in R^3

The linear transformation $T: R^3 \rightarrow R^3$ represented by

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

is called a **projection** in R^3. If $\mathbf{v} = (x, y, z)$ is a vector into R^3, then $T(\mathbf{v}) = (x, y, 0)$. In other words, T maps every vector in R^3 to its orthogonal projection in the xy-plane, as shown in Figure 6.3.



Projection onto xy-plane

Figure 6.3

A Linear Transformation from $M_{m,n}$ into $M_{n,m}$

Let $T: M_{m,n} \to M_{n,m}$ be the function that maps an $m \times n$ matrix A to its transpose. That is,

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

Show that T is a linear transformation.

SOLUTION Let A and B be $m \times n$ matrices. From Theorem 2.6 you have

$$\begin{aligned} T(A + B) &= (A + B)^T \\ &= A^T + B^T \\ &= T(A) + T(B) \end{aligned}$$

and

$$\begin{aligned} T(cA) &= (cA)^T \\ &= c(A^T) \\ &= cT(A). \end{aligned}$$

So, T is a linear transformation from $M_{m,n}$ into $M_{n,m}$.

Matrices for Linear Transformations

Which representation of $T: R^3 \to R^3$ is better,

$$T(x_1, x_2, x_3) = (2x_1 + x_2 - x_3, -x_1 + 3x_2 - 2x_3, 3x_2 + 4x_3)$$

or

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 3 & -2 \\ 0 & 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}?$$

The second representation is better than the first for at least three reasons: it is simpler to write, simpler to read, and more easily adapted for computer use. Later you will see that matrix representation of linear transformations also has some theoretical advantages. In this section you will see that for linear transformations involving finite-dimensional vector spaces, matrix representation is always possible.

The key to representing a linear transformation $T: V \to W$ by a matrix is to determine how it acts on a basis of V. Once you know the image of every vector in the basis, you can use the properties of linear transformations to determine $T(\mathbf{v})$ for any $\mathbf{v}$ in V.

For convenience, the first three theorems in this section are stated in terms of linear transformations from R^n into R^m, relative to the standard bases in R^n and R^m. At the end of the section these results are generalized to include nonstandard bases and general vector spaces.

Recall that the standard basis for R^n, written in column vector notation, is represented by

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

Standard Matrix for a Linear Transformation

Let $T: R^n \to R^m$ be a linear transformation such that

$$T(\mathbf{e}_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, T(\mathbf{e}_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, T(\mathbf{e}_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

Then the $m \times n$ matrix whose n columns correspond to $T(\mathbf{e}_i)$,

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

is such that $T(\mathbf{v}) = A\mathbf{v}$ for every $\mathbf{v}$ in R^n. A is called the **standard matrix** for T.

Finding the Standard Matrix for a Linear Transformation

Find the standard matrix for the linear transformation $T: R^3 \to R^2$ defined by

$$T(x, y, z) = (x - 2y, 2x + y).$$

SOLUTION Begin by finding the images of $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$.

Vector Notation	*Matrix Notation*
$T(\mathbf{e}_1) = T(1, 0, 0) = (1, 2)$	$T(\mathbf{e}_1) = T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\end{bmatrix}$
$T(\mathbf{e}_2) = T(0, 1, 0) = (-2, 1)$	$T(\mathbf{e}_2) = T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}-2\\1\end{bmatrix}$
$T(\mathbf{e}_3) = T(0, 0, 1) = (0, 0)$	$T(\mathbf{e}_3) = T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\0\end{bmatrix}$

By Theorem 6.10, the columns of A consist of $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$, and you have

$$A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots T(\mathbf{e}_3)] = \begin{bmatrix}1 & -2 & 0\\2 & 1 & 0\end{bmatrix}.$$

As a check, note that

$$A\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}1 & -2 & 0\\2 & 1 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}x - 2y\\2x + y\end{bmatrix},$$

which is equivalent to $T(x, y, z) = (x - 2y, 2x + y)$.

Finding the Standard Matrix for a Linear Transformation

The linear transformation $T: R^2 \to R^2$ is given by projecting each point in R^2 onto the x-axis, as shown in Figure 6.8. Find the standard matrix for T.

SOLUTION This linear transformation is represented by

$$T(x, y) = (x, 0).$$

So, the standard matrix for T is

$$A = [T(1, 0) \vdots T(0, 1)]$$

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

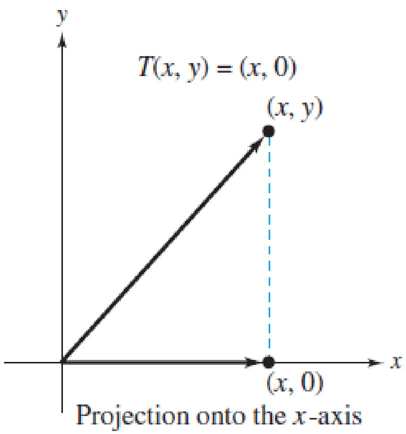
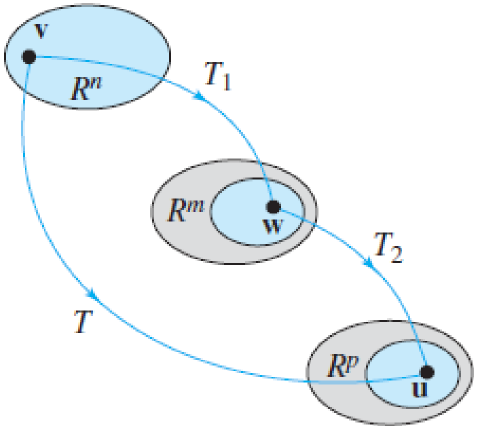


Figure 6.8

Composition of Linear Transformations

Let $T_1: R^n \to R^m$ and $T_2: R^m \to R^p$ be linear transformations with standard matrices A_1 and A_2. The **composition** $T: R^n \to R^p$, defined by $T(\mathbf{v}) = T_2(T_1(\mathbf{v}))$, is a linear transformation. Moreover, the standard matrix A for T is given by the matrix product

$$A = A_2 A_1.$$



Composition of Transformations

The Standard Matrix for a Composition

Let T_1 and T_2 be linear transformations from R^3 into R^3 such that

$$T_1(x, y, z) = (2x + y, 0, x + z) \quad \text{and} \quad T_2(x, y, z) = (x - y, z, y).$$

Find the standard matrices for the compositions $T = T_2 \circ T_1$ and $T' = T_1 \circ T_2$.

SOLUTION The standard matrices for T_1 and T_2 are

$$A_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

By Theorem 6.11, the standard matrix for T is

$$A = A_2 A_1 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

and the standard matrix for T' is

$$A' = A_1 A_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Definition of Inverse Linear Transformation

If $T_1: R^n \rightarrow R^n$ and $T_2: R^n \rightarrow R^n$ are linear transformations such that for every $\mathbf{v}$ in R^n

$$T_2(T_1(\mathbf{v})) = \mathbf{v} \qquad \text{and} \qquad T_1(T_2(\mathbf{v})) = \mathbf{v},$$

then T_2 is called the **inverse** of T_1, and T_1 is said to be **invertible.**

Not every linear transformation has an inverse. If the transformation T_1 is invertible, however, then the inverse is unique and is denoted by T_1^{-1}.

Just as the inverse of a function of a real variable can be thought of as undoing what the function did, the inverse of a linear transformation T can be thought of as undoing the mapping done by T. For instance, if T is a linear transformation from R^3 onto R^3 such that

$$T(1, 4, -5) = (2, 3, 1)$$

and if T^{-1} exists, then T^{-1} maps $(2, 3, 1)$ back to its preimage under T. That is,

$$T^{-1}(2, 3, 1) = (1, 4, -5).$$

Existence of an Inverse Transformation

Let $T: R^n \rightarrow R^n$ be a linear transformation with standard matrix A. Then the following conditions are equivalent.

1. T is invertible.
2. T is an isomorphism.
3. A is invertible.

And, if T is invertible with standard matrix A, then the standard matrix for T^{-1} is A^{-1}.

Finding the Inverse of a Linear Transformation

The linear transformation $T: R^3 \to R^3$ is defined by

$$T(x_1, x_2, x_3) = (2x_1 + 3x_2 + x_3, 3x_1 + 3x_2 + x_3, 2x_1 + 4x_2 + x_3).$$

Show that T is invertible, and find its inverse.

SOLUTION The standard matrix for T is

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

Using the techniques for matrix inversion (see Section 2.3), you can find that A is invertible and its inverse is

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

So, T is invertible and its standard matrix is A^{-1}.

Using the standard matrix for the inverse, you can find the rule for T^{-1} by computing the image of an arbitrary vector $\mathbf{v} = (x_1, x_2, x_3)$.

$$A^{-1}\mathbf{v} = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 6 & -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_1 + x_2 \\ -x_1 + x_3 \\ 6x_1 - 2x_2 - 3x_3 \end{bmatrix}$$

In other words,

$$T^{-1}(x_1, x_2, x_3) = (-x_1 + x_2, -x_1 + x_3, 6x_1 - 2x_2 - 3x_3).$$

Elementary Matrices for Linear Transformations in the Plane

Reflection in y-Axis

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

Reflection in x-Axis

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

Reflection in Line $y = x$

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

Horizontal Expansion $(k > 1)$
or Contraction $(0 < k < 1)$

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

Vertical Expansion $(k > 1)$
or Contraction $(0 < k < 1)$

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

Horizontal Shear

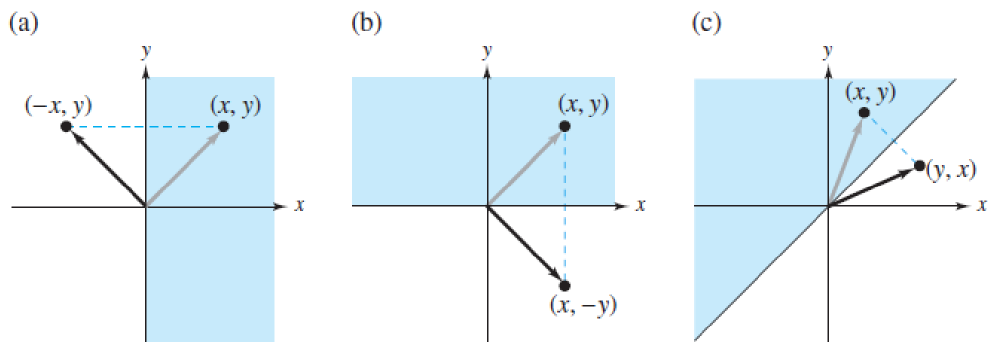
$$A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$$

Vertical Shear

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

Reflections in the Plane

The transformations defined by the matrices listed below are called **reflections.** Reflections have the effect of mapping a point in the xy-plane to its "mirror image" with respect to one of the coordinate axes or the line $y = x$, as shown in Figure 6.11.



Reflections in the Plane

Figure 6.11

(a) Reflection in the y-axis:

$$T(x, y) = (-x, y)$$

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$$

(b) Reflection in the x-axis:

$$T(x, y) = (x, -y)$$

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

(c) Reflection in the line $y = x$:

$$T(x, y) = (y, x)$$

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

Expansions and Contractions in the Plane

The transformations defined by the matrices below are called **expansions** or **contractions,** depending on the value of the positive scalar k.

(a) Horizontal contractions and expansions:

$$T(x, y) = (kx, y)$$

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

(b) Vertical contractions and expansions:

$$T(x, y) = (x, ky)$$

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

Note that in Figures 6.12 and 6.13, the distance the point (x, y) is moved by a contraction or an expansion is proportional to its x- or y-coordinate. For instance, under the transformation represented by $T(x, y) = (2x, y)$, the point $(1, 3)$ would be moved one unit to the right, but the point $(4, 3)$ would be moved four units to the right.

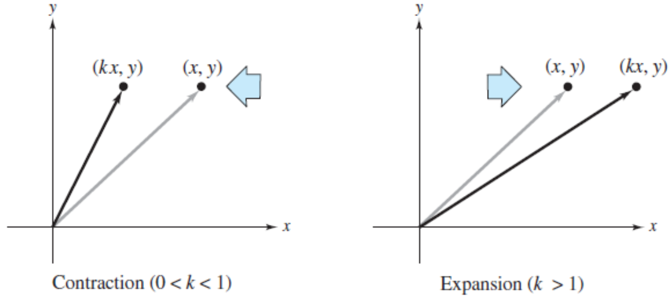


Figure 6.12

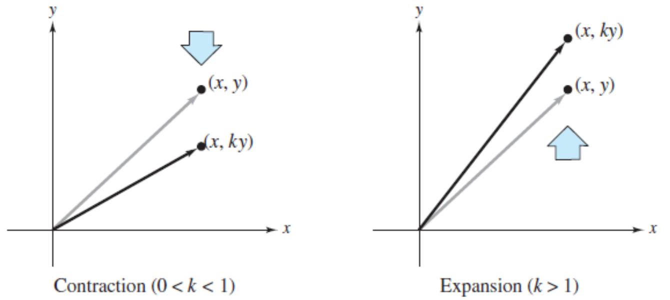


Figure 6.13

Shears in the Plane

The transformations defined by the following matrices are shears.

$$T(x, y) = (x + ky, y) \qquad\qquad T(x, y) = (x, y + kx)$$

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + ky \\ y \end{bmatrix} \qquad\qquad \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ kx + y \end{bmatrix}$$

(a) The horizontal shear represented by $T(x, y) = (x + 2y, y)$ is shown in Figure 6.14. Under this transformation, points in the upper half-plane are "sheared" to the right by amounts proportional to their y-coordinates. Points in the lower half-plane are "sheared" to the left by amounts proportional to the absolute values of their y-coordinates. Points on the x-axis are unmoved by this transformation.

(b) The vertical shear represented by $T(x, y) = (x, y + 2x)$ is shown in Figure 6.15. Here, points in the right half-plane are "sheared" upward by amounts proportional to their x-coordinates. Points in the left half-plane are "sheared" downward by amounts proportional to the absolute values of their x-coordinates. Points on the y-axis are unmoved.

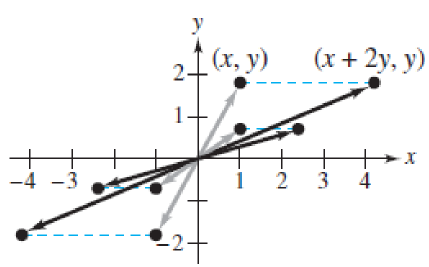


Figure 6.14

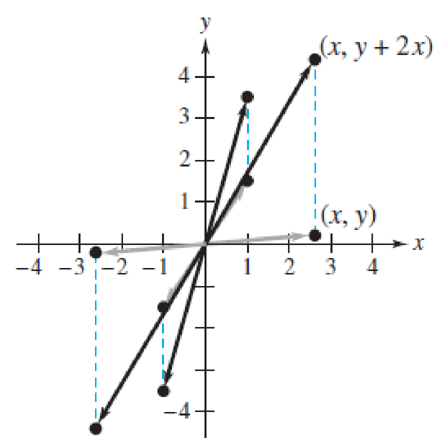


Figure 6.15

Applications of Linear Transformations

- Applying data transformations like rotation and scaling
- Reducing dimensionality using Principal Component Analysis (PCA)
- Preprocessing features in machine learning workflows

Computer Graphics

Linear transformations are useful in computer graphics. In Example 7 in Section 6.1, you saw how a linear transformation could be used to rotate figures in the plane. Here you will see how linear transformations can be used to rotate figures in three-dimensional space.

Suppose you want to rotate the point (x, y, z) counterclockwise about the z-axis through an angle θ, as shown in Figure 6.16. Letting the coordinates of the rotated point be (x', y', z'), you have

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \\ z \end{bmatrix}.$$

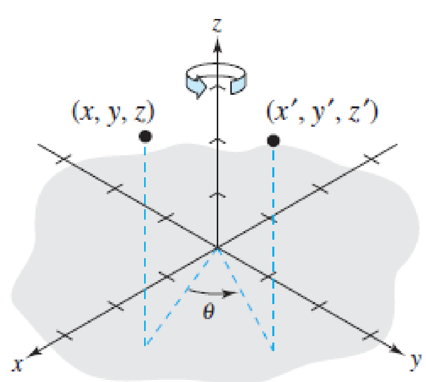


Figure 6.16

Rotation About the z-Axis

The eight vertices of a rectangular box having sides of lengths 1, 2, and 3 are as follows.

$$V_1 = (0, 0, 0), \quad V_2 = (1, 0, 0), \quad V_3 = (1, 2, 0), \quad V_4 = (0, 2, 0),$$
$$V_5 = (0, 0, 3), \quad V_6 = (1, 0, 3), \quad V_7 = (1, 2, 3), \quad V_8 = (0, 2, 3)$$

Find the coordinates of the box when it is rotated counterclockwise about the z-axis through each angle.

(a) $\theta = 60°$ (b) $\theta = 90°$ (c) $\theta = 120°$

SOLUTION

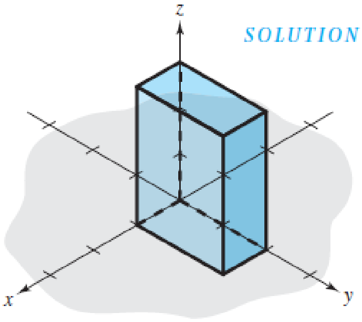


Figure 6.17

The original box is shown in Figure 6.17.

(a) The matrix that yields a rotation of 60° is

$$A = \begin{bmatrix} \cos 60° & -\sin 60° & 0 \\ \sin 60° & \cos 60° & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Rotation About the z-Axis (cont.)

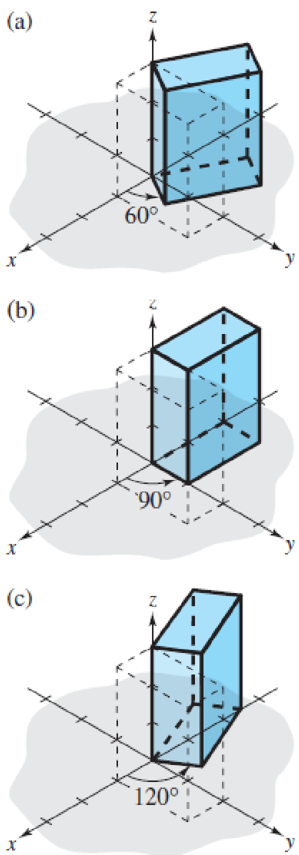


Figure 6.18

Multiplying this matrix by the eight vertices produces the rotated vertices listed below

Original Vertex	*Rotated Vertex*
$V_1 = (0, 0, 0)$	$(0, 0, 0)$
$V_2 = (1, 0, 0)$	$(0.5, 0.87, 0)$
$V_3 = (1, 2, 0)$	$(-1.23, 1.87, 0)$
$V_4 = (0, 2, 0)$	$(-1.73, 1, 0)$
$V_5 = (0, 0, 3)$	$(0, 0, 3)$
$V_6 = (1, 0, 3)$	$(0.5, 0.87, 3)$
$V_7 = (1, 2, 3)$	$(-1.23, 1.87, 3)$
$V_8 = (0, 2, 3)$	$(-1.73, 1, 3)$

A computer-generated graph of the rotated box is shown in Figure 6.18(a). Note that in this graph, line segments representing the sides of the box are drawn between images of pairs of vertices connected in the original box. For instance, because V_1 and V_2 are connected in the original box, the computer is told to connect the images of V_1 and V_2 in the rotated box.

(b) The matrix that yields a rotation of 90° is

$$A = \begin{bmatrix} \cos 90^\circ & -\sin 90^\circ & 0 \\ \sin 90^\circ & \cos 90^\circ & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and the graph of the rotated box is shown in Figure 6.18(b).

(c) The matrix that yields a rotation of 120° is

$$A = \begin{bmatrix} \cos 120^\circ & -\sin 120^\circ & 0 \\ \sin 120^\circ & \cos 120^\circ & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and the graph of the rotated box is shown in Figure 6.18(c).

Rotation About the x- or y-Axis

In Example 4, matrices were used to perform rotations about the z-axis. Similarly, you can use matrices to rotate figures about the x- or y-axis. All three types of rotations are summarized as follows.

Rotation About the x-Axis	*Rotation About the y-Axis*	*Rotation About the z-Axis*
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$	$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$

In each case the rotation is oriented counterclockwise relative to a person facing the negative direction of the indicated axis, as shown in Figure 6.19.

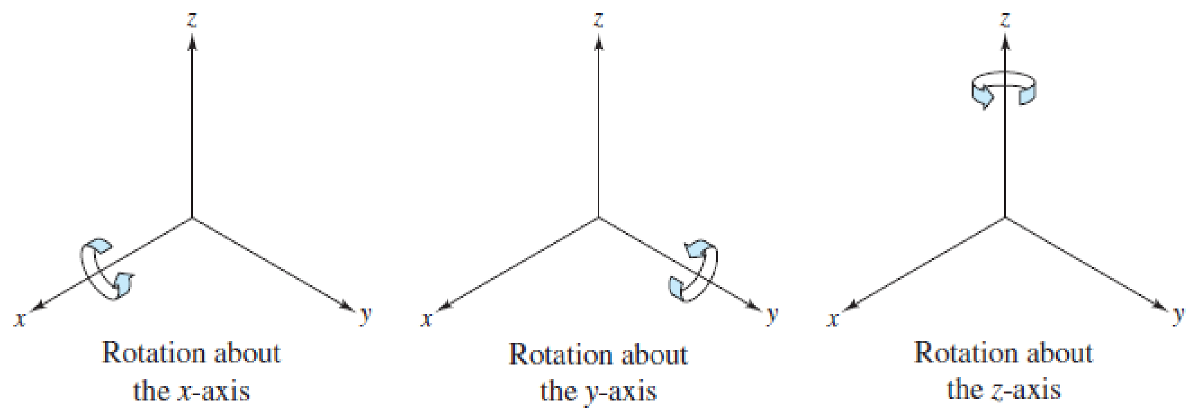


Figure 6.19

Rotation About the x-Axis and y-Axis

(a) The matrix that yields a rotation of 90° about the x-axis is

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix},$$

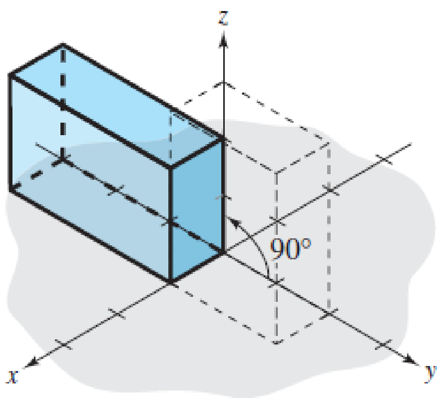
and the graph of the rotated box from Example 4 is shown in Figure 6.20(a) below.

(b) The matrix that yields a rotation of 90° about the y-axis is

$$A = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

and the graph of the rotated box from Example 4 is shown in Figure 6.20(b) below.

(a)



(b)

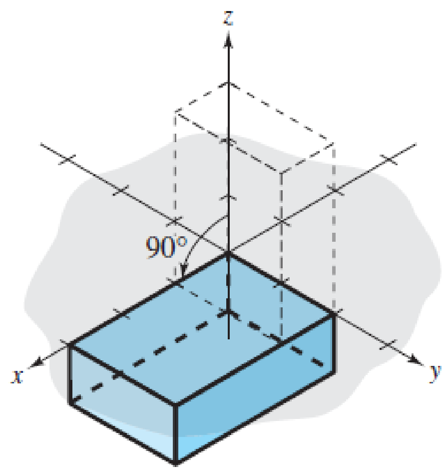


Figure 6.20

Rotations

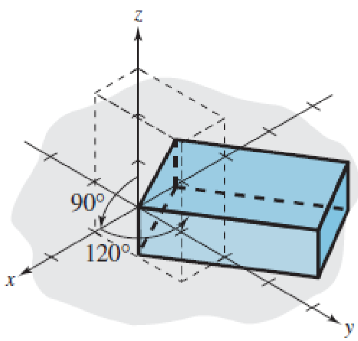


Figure 6.21

Rotations about the coordinate axes can be combined to produce any desired view of a figure. For instance, Figure 6.21 shows the rotation produced by first rotating the box (from Example 4) 90° about the *y*-axis, then further rotating the box 120° about the *z*-axis.

The use of computer graphics has become common among designers in many fields. By simply entering the coordinates that form the outline of an object into a computer, a designer can see the object before it is created. As a simple example, the images of the toy boat shown in Figure 6.22 were created using only 27 points in space. Once the points have been stored in the computer, the boat can be viewed from any perspective.

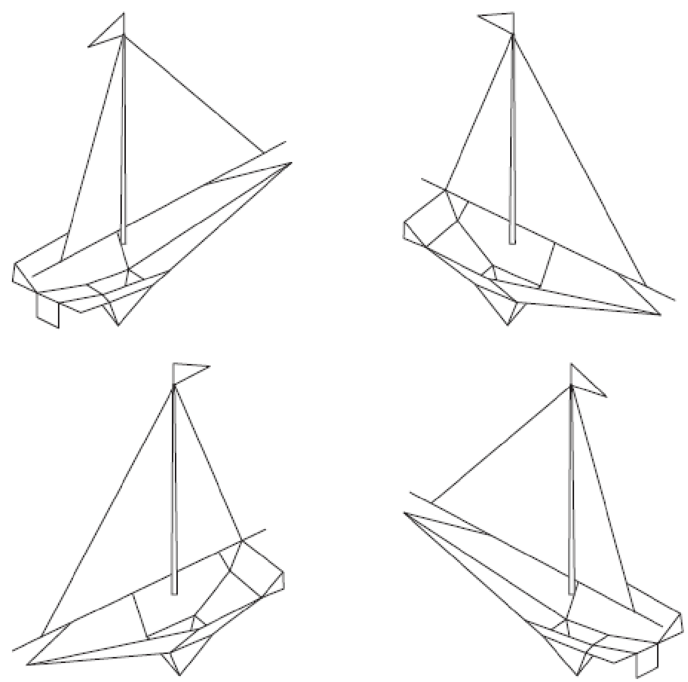


Figure 6.22

Eigenvalues and Eigenvectors

- If A is an $n \times n$ matrix, do nonzero vectors $\mathbf{x}$ in R^n exist such that $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$?
- The scalar, denoted by the Greek letter lambda (λ), is called an **eigenvalue** of the matrix A and the nonzero vector $\mathbf{x}$ is called an **eigenvector** of A corresponding to λ.
- The terms *eigenvalue* and *eigenvector* are derived from the German word *Eigenwert*, meaning "proper value." So, you have

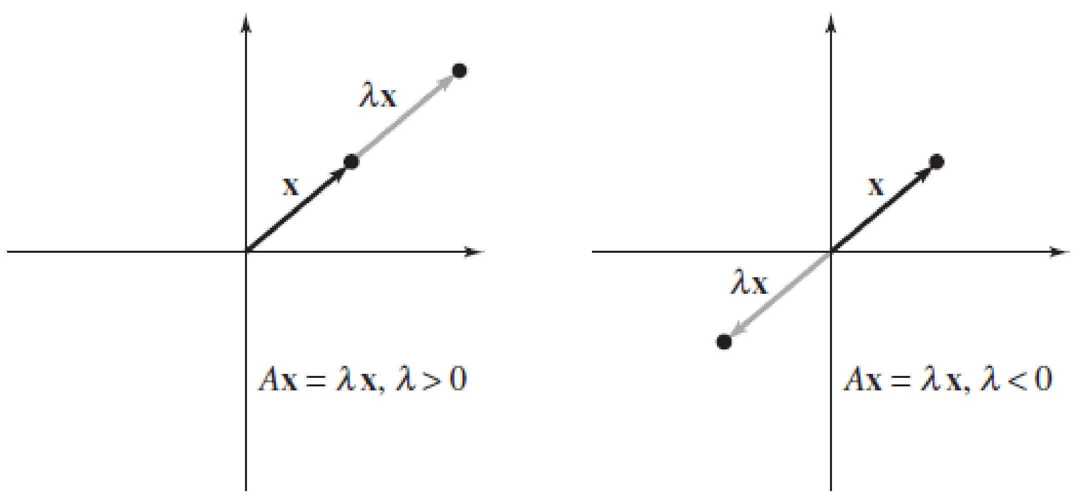
$$\boxed{\textit{Eigenvalue}}$$

$$A\mathbf{x} = \lambda\mathbf{x}.$$

$$\boxed{\textit{Eigenvector}}$$

Eigenvalues and Eigenvectors (cont.)

- Eigenvalues and eigenvectors have many important applications.
- For now you will consider a geometric interpretation of the problem in R^2.
- If λ is an eigenvalue of a matrix A and $\mathbf{x}$ is an eigenvector of A corresponding to λ, then multiplication of $\mathbf{x}$ by the matrix A produces a vector $\lambda\mathbf{x}$ that is parallel to $\mathbf{x}$.



Definitions of Eigenvalue and Eigenvector

Let A be an $n \times n$ matrix. The scalar λ is called an **eigenvalue** of A if there is a *nonzero* vector $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The vector $\mathbf{x}$ is called an **eigenvector** of A corresponding to λ.

REMARK: Note that an eigen*vector* cannot be zero. Allowing $\mathbf{x}$ to be the zero vector would render the definition meaningless, because $A\mathbf{0} = \lambda\mathbf{0}$ is true for all real values of λ. An eigen*value* of $\lambda = 0$, however, is possible.

Verifying Eigenvalues and Eigenvectors

For the matrix

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

verify that $\mathbf{x}_1 = (1, 0)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_1 = 2$, and that $\mathbf{x}_2 = (0, 1)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_2 = -1$.

SOLUTION Multiplying $\mathbf{x}_1$ by A produces

$$\begin{aligned} A\mathbf{x}_1 &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \end{bmatrix} \\ &= \underset{\text{Eigenvalue}}{2}\underset{\text{Eigenvector}}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}. \end{aligned}$$

So, $\mathbf{x}_1 = (1, 0)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_1 = 2$. Similarly, multiplying $\mathbf{x}_2$ by A produces

$$\begin{aligned} A\mathbf{x}_2 &= \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\ &= -1\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \end{aligned}$$

So, $\mathbf{x}_2 = (0, 1)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_2 = -1$.

Verifying Eigenvalues and Eigenvectors

For the matrix

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

verify that

$$\mathbf{x}_1 = (-3, -1, 1) \quad \text{and} \quad \mathbf{x}_2 = (1, 0, 0)$$

are eigenvectors of A and find their corresponding eigenvalues.

SOLUTION Multiplying $\mathbf{x}_1$ by A produces

$$A\mathbf{x}_1 = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}.$$

So, $\mathbf{x}_1 = (-3, -1, 1)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_1 = 0$. Similarly, multiplying $\mathbf{x}_2$ by A produces

$$A\mathbf{x}_2 = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

So, $\mathbf{x}_2 = (1, 0, 0)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_2 = 1$.

Eigenspaces

Although Examples 1 and 2 list only one eigenvector for each eigenvalue, each of the four eigenvalues in Examples 1 and 2 has an infinite number of eigenvectors. For instance, in Example 1 the vectors $(2, 0)$ and $(-3, 0)$ are eigenvectors of A corresponding to the eigenvalue 2. In fact, if A is an $n \times n$ matrix with an eigenvalue λ and a corresponding eigenvector $\mathbf{x}$, then every nonzero scalar multiple of $\mathbf{x}$ is also an eigenvector of A. This may be seen by letting c be a nonzero scalar, which then produces

$$A(c\mathbf{x}) = c(A\mathbf{x}) = c(\lambda\mathbf{x}) = \lambda(c\mathbf{x}).$$

It is also true that if $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors corresponding to the *same* eigenvalue λ, then their sum is also an eigenvector corresponding to λ, because

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \lambda\mathbf{x}_1 + \lambda\mathbf{x}_2 = \lambda(\mathbf{x}_1 + \mathbf{x}_2).$$

In other words, the set of all eigenvectors of a given eigenvalue λ, together with the zero vector, is a subspace of R^n. This special subspace of R^n is called the **eigenspace** of λ.

Eigenvectors of λ Form a Subspace

If A is an $n \times n$ matrix with an eigenvalue λ, then the set of all eigenvectors of λ, together with the zero vector

$$\{\mathbf{0}\} \cup \{\mathbf{x}: \mathbf{x} \text{ is an eigenvector of } \lambda\},$$

is a subspace of R^n. This subspace is called the **eigenspace** of λ.

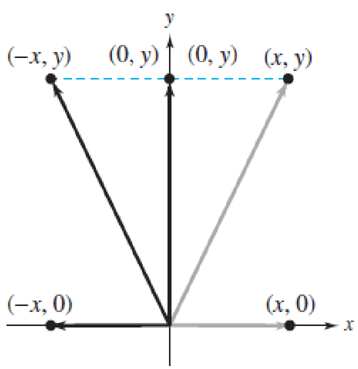
An Example of Eigenspaces in the Plane

Find the eigenvalues and corresponding eigenspaces of

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

SOLUTION Geometrically, multiplying a vector (x, y) in R^2 by the matrix A corresponds to a reflection in the y-axis. That is, if $\mathbf{v} = (x, y)$, then

$$A\mathbf{v} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}.$$



A reflects vectors in the *y*-axis.

Figure 7.2

Figure 7.2 illustrates that the only vectors reflected onto scalar multiples of themselves are those lying on either the x-axis or the y-axis.

For a vector on the x-axis

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

$\boxed{\text{Eigenvalue is } \lambda_1 = -1.}$

For a vector on the y-axis

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

$\boxed{\text{Eigenvalue is } \lambda_2 = 1.}$

So, the eigenvectors corresponding to $\lambda_1 = -1$ are the nonzero vectors on the x-axis, and the eigenvectors corresponding to $\lambda_2 = 1$ are the nonzero vectors on the y-axis. This implies that the eigenspace corresponding to $\lambda_1 = -1$ is the x-axis, and that the eigenspace corresponding to $\lambda_2 = 1$ is the y-axis.

Eigenvalues and Eigenvectors of a Matrix

Let A be an $n \times n$ matrix.

1. An eigenvalue of A is a scalar λ such that

$$\det(\lambda I - A) = 0.$$

2. The eigenvectors of A corresponding to λ are the nonzero solutions of

$$(\lambda I - A)\mathbf{x} = \mathbf{0}.$$

The equation $\det(\lambda I - A) = 0$ is called the **characteristic equation** of A. Moreover, when expanded to polynomial form, the polynomial

$$|\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

is called the **characteristic polynomial** of A. This definition tells you that the eigenvalues of an $n \times n$ matrix A correspond to the roots of the characteristic polynomial of A. Because the characteristic polynomial of A is of degree n, A can have at most n distinct eigenvalues.

REMARK: The Fundamental Theorem of Algebra states that an nth-degree polynomial has precisely n roots. These n roots, however, include both repeated and complex roots. In this chapter you will be concerned only with the real roots of characteristic polynomials—that is, real eigenvalues.

Finding Eigenvalues and Eigenvectors

Find the eigenvalues and corresponding eigenvectors of

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}.$$

SOLUTION The characteristic polynomial of A is

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix} \\ &= (\lambda - 2)(\lambda + 5) - (-12) \\ &= \lambda^2 + 3\lambda - 10 + 12 \\ &= \lambda^2 + 3\lambda + 2 \\ &= (\lambda + 1)(\lambda + 2). \end{aligned}$$

So, the characteristic equation is $(\lambda + 1)(\lambda + 2) = 0$, which gives $\lambda_1 = -1$ and $\lambda_2 = -2$ as the eigenvalues of A. To find the corresponding eigenvectors, use Gauss-Jordan elimination to solve the homogeneous linear system represented by $(\lambda I - A)\mathbf{x} = \mathbf{0}$ twice: first for $\lambda = \lambda_1 = -1$, and then for $\lambda = \lambda_2 = -2$. For $\lambda_1 = -1$, the coefficient matrix is

$$(-1)I - A = \begin{bmatrix} -1 - 2 & 12 \\ -1 & -1 + 5 \end{bmatrix} = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix},$$

which row reduces to

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

showing that $x_1 - 4x_2 = 0$. Letting $x_2 = t$, you can conclude that every eigenvector of λ_1 is of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad t \neq 0.$$

For $\lambda_2 = -2$, you have

$$(-2)I - A = \begin{bmatrix} -2 - 2 & 12 \\ -1 & -2 + 5 \end{bmatrix} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}.$$

Letting $x_2 = t$, you can conclude that every eigenvector of λ_2 is of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad t \neq 0.$$

Try checking $A\mathbf{x} = \lambda_i\mathbf{x}$ for the eigenvalues and eigenvectors in this example.

Finding Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix.

1. Form the characteristic equation $|\lambda I - A| = 0$. It will be a polynomial equation of degree n in the variable λ.
2. Find the real roots of the characteristic equation. These are the eigenvalues of A.
3. For each eigenvalue λ_i, find the eigenvectors corresponding to λ_i by solving the homogeneous system $(\lambda_i I - A)\mathbf{x} = \mathbf{0}$. This requires row reducing of an $n \times n$ matrix. The resulting reduced row-echelon form must have at least one row of zeros.

Finding the eigenvalues of an $n \times n$ matrix can be difficult because it involves the factorization of an nth-degree polynomial. Once an eigenvalue has been found, however, finding the corresponding eigenvectors is a straightforward application of Gauss-Jordan reduction.

Finding Eigenvalues and Eigenvectors

Find the eigenvalues and corresponding eigenvectors of

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

What is the dimension of the eigenspace of each eigenvalue?

SOLUTION The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & -1 & 0 \\ 0 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 2)^3.$$

So, the characteristic equation is $(\lambda - 2)^3 = 0$.

So, the only eigenvalue is $\lambda = 2$. To find the eigenvectors of $\lambda = 2$, solve the homogeneous linear system represented by $(2I - A)\mathbf{x} = \mathbf{0}$.

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

This implies that $x_2 = 0$. Using the parameters $s = x_1$ and $t = x_3$, you can find that the eigenvectors of $\lambda = 2$ are of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad s \text{ and } t \text{ not both zero.}$$

Because $\lambda = 2$ has two linearly independent eigenvectors, the dimension of its eigenspace is 2.

Finding Eigenvalues and Eigenvectors

Find the eigenvalues of

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

and find a basis for each of the corresponding eigenspaces.

SOLUTION The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 0 & 0 & 0 \\ 0 & \lambda - 1 & -5 & 10 \\ -1 & 0 & \lambda - 2 & 0 \\ -1 & 0 & 0 & \lambda - 3 \end{vmatrix}$$

$$= (\lambda - 1)^2(\lambda - 2)(\lambda - 3).$$

So, the characteristic equation is $(\lambda - 1)^2(\lambda - 2)(\lambda - 3) = 0$ and the eigenvalues are $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$. (Note that $\lambda_1 = 1$ has a multiplicity of 2.)

You can find a basis for the eigenspace of $\lambda_1 = 1$ as follows.

$$(1)I - A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 10 \\ -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Letting $s = x_2$ and $t = x_4$ produces

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0s - 2t \\ s + 0t \\ 0s + 2t \\ 0s + t \end{bmatrix} = s\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix}.$$

A basis for the eigenspace corresponding to $\lambda_1 = 1$ is

$B_1 = \{(0, 1, 0, 0), (-2, 0, 2, 1)\}$. **Basis for $\lambda_1 = 1$**

For $\lambda_2 = 2$ and $\lambda_3 = 3$, follow the same pattern to obtain the eigenspace bases

$B_2 = \{(0, 5, 1, 0)\}$ **Basis for $\lambda_2 = 2$**

$B_3 = \{(0, -5, 0, 1)\}$. **Basis for $\lambda_3 = 3$**

Eigenvalues of Triangular Matrices

There are a few types of matrices for which eigenvalues are easy to find. The next theorem states that the eigenvalues of an $n \times n$ triangular matrix are the entries on the main diagonal. Its proof follows from the fact that the determinant of a triangular matrix is the product of its diagonal elements.

If A is an $n \times n$ triangular matrix, then its eigenvalues are the entries on its main diagonal.

Finding Eigenvalues of Diagonal and Triangular Matrices

Find the eigenvalues of each matrix.

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

SOLUTION (a) Without using Theorem 7.3, you can find that

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 0 & 0 \\ 1 & \lambda - 1 & 0 \\ -5 & -3 & \lambda + 3 \end{vmatrix}$$

$$= (\lambda - 2)(\lambda - 1)(\lambda + 3).$$

So, the eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 1$, and $\lambda_3 = -3$, which are simply the main diagonal entries of A.

(b) In this case, use Theorem 7.3 to conclude that the eigenvalues are the main diagonal entries $\lambda_1 = -1$, $\lambda_2 = 2$, $\lambda_3 = 0$, $\lambda_4 = -4$, and $\lambda_5 = 3$.

Eigenvalues and Eigenvectors of Linear Transformations

This section began with definitions of eigenvalues and eigenvectors in terms of matrices. They can also be defined in terms of linear transformations. A number λ is called an **eigenvalue** of a linear transformation $T: V \to V$ if there is a nonzero vector $\mathbf{x}$ such that $T(\mathbf{x}) = \lambda\mathbf{x}$. The vector $\mathbf{x}$ is called an **eigenvector** of T corresponding to λ, and the set of all eigenvectors of λ (with the zero vector) is called the **eigenspace** of λ.

Consider the linear transformation $T: R^3 \to R^3$, whose matrix relative to the standard basis is

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

Standard basis:
$B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

In Example 5 of Section 6.4, you found that the matrix of T relative to the basis B' is the diagonal matrix

$$A' = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Nonstandard basis:
$B' = \{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$

The question now is: "For a given transformation T, can you find a basis B' whose corresponding matrix is diagonal?" The next example gives an indication of the answer.

Finding Eigenvalues and Eigenspaces

Find the eigenvalues and corresponding eigenspaces of

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

SOLUTION Because

$$\begin{aligned} |\lambda I - A| &= \begin{vmatrix} \lambda - 1 & -3 & 0 \\ -3 & \lambda - 1 & 0 \\ 0 & 0 & \lambda + 2 \end{vmatrix} \\ &= (\lambda + 2)[(\lambda - 1)^2 - 9] \\ &= (\lambda + 2)(\lambda^2 - 2\lambda - 8) = (\lambda + 2)^2(\lambda - 4), \end{aligned}$$

the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = -2$. The eigenspaces for these two eigenvalues are as follows.

$B_1 = \{(1, 1, 0)\}$ **Basis for $\lambda_1 = 4$**

$B_2 = \{(1, -1, 0), (0, 0, 1)\}$ **Basis for $\lambda_2 = -2$**

Applications of Eigenvalues and Eigenvectors

- Extracting key components in PCA for large datasets
- Analyzing network structure via eigenvector centrality
- Studying dynamic system stability in forecasting models

Diagonalization

- In this section, you will look at another classic problem in linear algebra called the **diagonalization problem**.
- Expressed in terms of matrices, the problem is this: "For a square matrix A, does there exist an invertible matrix P such that $P^{-1}AP$ is diagonal?"
- Two square matrices A and B are called **similar** if there exists an invertible matrix P such that $B = P^{-1}AP$.
- Matrices that are similar to diagonal matrices are called **diagonalizable.**

Definition of a Diagonalizable Matrix

An $n \times n$ matrix A is **diagonalizable** if A is similar to a diagonal matrix. That is, A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix.

Provided with this definition, the diagonalization problem can be stated as follows: "Which square matrices are diagonalizable?" Clearly, every diagonal matrix D is diagonalizable, because the identity matrix I can play the role of P to yield $D = I^{-1}DI$. Example 1 shows another example of a diagonalizable matrix.

A Diagonalizable Matrix

The matrix from Example 5 in Section 6.4,

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

is diagonalizable because

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

has the property

$$P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Similar Matrices Have the Same Eigenvalues

If A and B are similar $n \times n$ matrices, then they have the same eigenvalues.

PROOF Because A and B are similar, there exists an invertible matrix P such that $B = P^{-1}AP$. By the properties of determinants, it follows that

$$
\begin{aligned}
|\lambda I - B| = |\lambda I - P^{-1}AP| &= |P^{-1}\lambda IP - P^{-1}AP| \\
&= |P^{-1}(\lambda I - A)P| \\
&= |P^{-1}||\lambda I - A||P| \\
&= |P^{-1}||P||\lambda I - A| \\
&= |P^{-1}P||\lambda I - A| \\
&= |\lambda I - A|.
\end{aligned}
$$

But this means that A and B have the same characteristic polynomial. So, they must have the same eigenvalues.

Finding Eigenvalues of Similar Matrices

The matrices A and D are similar.

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

Use Theorem 7.4 to find the eigenvalues of A and D.

SOLUTION Because D is a diagonal matrix, its eigenvalues are simply the entries on its main diagonal—that is,

$$\lambda_1 = 1,$$
$$\lambda_2 = 2, \text{ and}$$
$$\lambda_3 = 3.$$

Moreover, because A is said to be similar to D, you know from Theorem 7.4 that A has the same eigenvalues. Check this by showing that the characteristic polynomial of A is

$$|\lambda I - A| = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

REMARK: Example 2 simply states that matrices A and D are similar. Try checking $D = P^{-1}AP$ using the matrices

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

In fact, the columns of P are precisely the eigenvectors of A corresponding to the eigenvalues 1, 2, and 3.

Condition for Diagonalization

An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors.

Diagonalizable Matrices

(a) The matrix in Example 1 has the eigenvalues and corresponding eigenvectors listed below.

$$\lambda_1 = 4, \mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}; \quad \lambda_2 = -2, \mathbf{p}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \quad \lambda_3 = -2, \mathbf{p}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The matrix P whose columns correspond to these eigenvectors is

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Moreover, because P is row-equivalent to the identity matrix, the eigenvectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are linearly independent.

(b) The matrix in Example 2 has the eigenvalues and corresponding eigenvectors listed below.

$$\lambda_1 = 1, \mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}; \quad \lambda_2 = 2, \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}; \quad \lambda_3 = 3, \mathbf{p}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

The matrix P whose columns correspond to these eigenvectors is

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

Again, because P is row-equivalent to the identity matrix, the eigenvectors $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ are linearly independent.

Steps for Diagonalizing an $n \times n$ Square Matrix

Let A be an $n \times n$ matrix.

1. Find n linearly independent eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ for A with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. If n linearly independent eigenvectors do not exist, then A is not diagonalizable.
2. If A has n linearly independent eigenvectors, let P be the $n \times n$ matrix whose columns consist of these eigenvectors. That is,

$$P = [\mathbf{p}_1 \vdots \mathbf{p}_2 \vdots \cdots \vdots \mathbf{p}_n].$$

3. The diagonal matrix $D = P^{-1}AP$ will have the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ on its main diagonal (and zeros elsewhere). Note that the order of the eigenvectors used to form P will determine the order in which the eigenvalues appear on the main diagonal of D.

A Matrix That Is Not Diagonalizable

Show that the matrix A is not diagonalizable.

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

SOLUTION Because A is triangular, the eigenvalues are simply the entries on the main diagonal. So, the only eigenvalue is $\lambda = 1$. The matrix $(I - A)$ has the reduced row-echelon form shown below.

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

This implies that $x_2 = 0$, and letting $x_1 = t$, you can find that every eigenvector of A has the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} t \\ 0 \end{bmatrix} = t\begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

So, A does not have two linearly independent eigenvectors, and you can conclude that A is not diagonalizable.

Diagonalizing a Matrix

Show that the matrix A is diagonalizable.

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

Then find a matrix P such that $P^{-1}AP$ is diagonal.

SOLUTION The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 1 & 1 \\ -1 & \lambda - 3 & -1 \\ 3 & -1 & \lambda + 1 \end{vmatrix} = (\lambda - 2)(\lambda + 2)(\lambda - 3).$$

So, the eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = -2$, and $\lambda_3 = 3$. From these eigenvalues you obtain the reduced row-echelon forms and corresponding eigenvectors shown below.

$$\begin{array}{lccc} & & & \textit{Eigenvector} \\ 2I - A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 3 & -1 & 3 \end{bmatrix} & \longrightarrow & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \\ -2I - A = \begin{bmatrix} -3 & 1 & 1 \\ -1 & -5 & -1 \\ 3 & -1 & -1 \end{bmatrix} & \longrightarrow & \begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix} \\ 3I - A = \begin{bmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 3 & -1 & 4 \end{bmatrix} & \longrightarrow & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \end{array}$$

Diagonalizing a Matrix

Form the matrix P whose columns are the eigenvectors just obtained.

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

This matrix is nonsingular, which implies that the eigenvectors are linearly independent and A is diagonalizable. The inverse of P is

$$P^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ \frac{1}{5} & 0 & \frac{1}{5} \\ \frac{1}{5} & 1 & \frac{1}{5} \end{bmatrix},$$

and it follows that

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Diagonalizing a Matrix

Show that the matrix A is diagonalizable.

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

Then find a matrix P such that $P^{-1}AP$ is diagonal.

SOLUTION In Example 6 in Section 7.1, you found that the three eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$, and $\lambda_3 = 3$ have the eigenvectors shown below.

$$\lambda_1: \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \qquad \lambda_2: \begin{bmatrix} 0 \\ 5 \\ 1 \\ 0 \end{bmatrix} \qquad \lambda_3: \begin{bmatrix} 0 \\ -5 \\ 0 \\ 1 \end{bmatrix}$$

The matrix whose columns consist of these eigenvectors is

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

Because P is invertible (check this), its column vectors form a linearly independent set.

$$P^{-1} = \begin{bmatrix} -\frac{5}{2} & 1 & -5 & 5 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix}$$

So, A is diagonalizable, and you have

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.$$

Sufficient Condition for Diagonalization

If an $n \times n$ matrix A has n *distinct* eigenvalues, then the corresponding eigenvectors are linearly independent and A is diagonalizable.

Determining Whether a Matrix Is Diagonalizable

Determine whether the matrix A is diagonalizable.

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

SOLUTION Because A is a triangular matrix, its eigenvalues are the main diagonal entries

$$\lambda_1 = 1, \qquad \lambda_2 = 0, \qquad \lambda_3 = -3.$$

Moreover, because these three values are distinct, you can conclude from Theorem 7.6 that A is diagonalizable.

Diagonalization and Linear Transformations

So far in this section, the diagonalization problem has been considered in terms of matrices. In terms of linear transformations, the diagonalization problem can be stated as follows. For a linear transformation

$$T: V \to V,$$

does there exist a basis B for V such that the matrix for T relative to B is diagonal? The answer is "yes," provided the standard matrix for T is diagonalizable.

Finding a Diagonal Matrix for a Linear Transformation

Let $T: R^3 \rightarrow R^3$ be the linear transformation represented by

$$T(x_1, x_2, x_3) = (x_1 - x_2 - x_3, x_1 + 3x_2 + x_3, -3x_1 + x_2 - x_3).$$

If possible, find a basis B for R^3 such that the matrix for T relative to B is diagonal.

SOLUTION The standard matrix for T is represented by

$$A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}.$$

From Example 5, you know that A is diagonalizable. So, the three linearly independent eigenvectors found in Example 5 can be used to form the basis B. That is,

$$B = \{(-1, 0, 1), (1, -1, 4), (-1, 1, 1)\}.$$

The matrix for T relative to this basis is

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

Applications of Diagonalization

- Simplifying matrix exponentiation in Markov Chains
- Modeling population growth in business simulations
- Reducing complexity in linear system analysis

Symmetric Matrices and Orthogonal Diagonalization

- For most matrices you must go through much of the diagonalization process before you can finally determine whether diagonalization is possible.
- One exception is a triangular matrix with distinct entries on the main diagonal.
- Such a matrix can be recognized as diagonalizable by simple inspection.
- In this section you will study another type of matrix that is guaranteed to be diagonalizable: a **symmetric** matrix.

A square matrix A is **symmetric** if it is equal to its transpose:

$$A = A^T.$$

Symmetric Matrices and Nonsymmetric Matrices

The matrices A and B are symmetric, but the matrix C is not.

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

$$B = \begin{bmatrix} 4 & 3 \\ 3 & 1 \end{bmatrix} \qquad \text{Symmetric}$$

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

Eigenvalues of Symmetric Matrices

If A is an $n \times n$ symmetric matrix, then the following properties are true.

1. A is diagonalizable.
2. All eigenvalues of A are real.
3. If λ is an eigenvalue of A with multiplicity k, then λ has k linearly independent eigenvectors. That is, the eigenspace of λ has dimension k.

REMARK: Theorem 7.7 is called the **Real Spectral Theorem,** and the set of eigenvalues of A is called the **spectrum** of A.

The Eigenvalues and Eigenvectors of a 2×2 Symmetric Matrix

Prove that a symmetric matrix

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

is diagonalizable.

SOLUTION The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - a & -c \\ -c & \lambda - b \end{vmatrix} = \lambda^2 - (a + b)\lambda + ab - c^2.$$

As a quadratic in λ, this polynomial has a discriminant of

$$\begin{aligned} (a + b)^2 - 4(ab - c^2) &= a^2 + 2ab + b^2 - 4ab + 4c^2 \\ &= a^2 - 2ab + b^2 + 4c^2 \\ &= (a - b)^2 + 4c^2. \end{aligned}$$

Because this discriminant is the sum of two squares, it must be either zero or positive. If $(a - b)^2 + 4c^2 = 0$, then $a = b$ and $c = 0$, which implies that A is already diagonal. That is,

$$A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}.$$

On the other hand, if $(a - b)^2 + 4c^2 > 0$, then by the Quadratic Formula the characteristic polynomial of A has two distinct real roots, which implies that A has two distinct real eigenvalues. So, A is diagonalizable in this case also.

Dimensions of the Eigenspaces of a Symmetric Matrix

Find the eigenvalues of the symmetric matrix

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

and determine the dimensions of the corresponding eigenspaces.

SOLUTION The characteristic polynomial of A is represented by

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & 2 & 0 & 0 \\ 2 & \lambda - 1 & 0 & 0 \\ 0 & 0 & \lambda - 1 & 2 \\ 0 & 0 & 2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2(\lambda - 3)^2.$$

So, the eigenvalues of A are $\lambda_1 = -1$ and $\lambda_2 = 3$. Because each of these eigenvalues has a multiplicity of 2, you know from Theorem 7.7 that the corresponding eigenspaces also have dimension 2. Specifically, the eigenspace of $\lambda_1 = -1$ has a basis of $B_1 = \{(1, 1, 0, 0), (0, 0, 1, 1)\}$ and the eigenspace of $\lambda_2 = 3$ has a basis of $B_2 = \{(1, -1, 0, 0), (0, 0, 1, -1)\}$.

Orthogonal Matrices

- To diagonalize a square matrix A, you need to find an *invertible* matrix P such that $P^{-1}AP$ is diagonal.
- For symmetric matrices, you will see that the matrix P can be chosen to have the special property that $P^{-1} = P^T$.
- This unusual matrix property is defined as follows.

A square matrix P is called **orthogonal** if it is invertible and if

$$P^{-1} = P^T.$$

Orthogonal Matrices

(a) The matrix

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

is orthogonal because

$$P^{-1} = P^T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

(b) The matrix

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

is orthogonal because

$$P^{-1} = P^T = \begin{bmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ -\frac{4}{5} & 0 & \frac{3}{5} \end{bmatrix}.$$

Property of Orthogonal Matrices

An $n \times n$ matrix P is orthogonal if and only if its column vectors form an orthonormal set.

PROOF Suppose the column vectors of P form an orthonormal set:

$$P = [\mathbf{p}_1 \vdots \mathbf{p}_2 \vdots \cdots \vdots \mathbf{p}_n]$$
$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}.$$

Then the product P^TP has the form

$$P^TP = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{n1} \\ p_{12} & p_{22} & \cdots & p_{n2} \\ \vdots & \vdots & & \vdots \\ p_{1n} & p_{2n} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

$$P^TP = \begin{bmatrix} \mathbf{p}_1 \cdot \mathbf{p}_1 & \mathbf{p}_1 \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_1 \cdot \mathbf{p}_n \\ \mathbf{p}_2 \cdot \mathbf{p}_1 & \mathbf{p}_2 \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_2 \cdot \mathbf{p}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{p}_n \cdot \mathbf{p}_1 & \mathbf{p}_n \cdot \mathbf{p}_2 & \cdots & \mathbf{p}_n \cdot \mathbf{p}_n \end{bmatrix}.$$

Because the set $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ is orthonormal, you have

$$\mathbf{p}_i \cdot \mathbf{p}_j = 0,\ i \neq j \quad \text{and} \quad \mathbf{p}_i \cdot \mathbf{p}_i = \|\mathbf{p}_i\|^2 = 1.$$

So, the matrix composed of dot products has the form

$$P^TP = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_n.$$

This implies that $P^T = P^{-1}$, and you can conclude that P is orthogonal.

Conversely, if P is orthogonal, you can reverse the steps above to verify that the column vectors of P form an orthonormal set.

An Orthogonal Matrix

Show that

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

is orthogonal by showing that $PP^T = I$. Then show that the column vectors of P form an orthonormal set.

SOLUTION Because

$$PP^T = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ -\frac{2}{3\sqrt{5}} & -\frac{4}{3\sqrt{5}} & \frac{5}{3\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & -\frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3,$$

it follows that $P^T = P^{-1}$, and you can conclude that P is orthogonal. Moreover, letting

$$\mathbf{p}_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{\sqrt{5}} \\ -\frac{2}{3\sqrt{5}} \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{\sqrt{5}} \\ -\frac{4}{3\sqrt{5}} \end{bmatrix}, \text{ and } \mathbf{p}_3 = \begin{bmatrix} \frac{2}{3} \\ 0 \\ \frac{5}{3\sqrt{5}} \end{bmatrix}$$

produces

$$\mathbf{p}_1 \cdot \mathbf{p}_2 = \mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}_2 \cdot \mathbf{p}_3 = 0$$

and

$$\|\mathbf{p}_1\| = \|\mathbf{p}_2\| = \|\mathbf{p}_3\| = 1.$$

So, $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is an orthonormal set, as guaranteed by Theorem 7.8.

Property of Symmetric Matrices

Let A be an $n \times n$ symmetric matrix. If λ_1 and λ_2 are distinct eigenvalues of A, then their corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

PROOF Let λ_1 and λ_2 be distinct eigenvalues of A with corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$. So,

$$A\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \quad \text{and} \quad A\mathbf{x}_2 = \lambda_2\mathbf{x}_2.$$

To prove the theorem, use the matrix form of the dot product shown below.

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = [x_{11} \;\; x_{12} \;\; \cdots \;\; x_{1n}]\begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2n} \end{bmatrix} = \mathbf{x}_1^T\mathbf{x}_2$$

Now you can write

$$\begin{aligned}
\lambda_1(\mathbf{x}_1 \cdot \mathbf{x}_2) &= (\lambda_1\mathbf{x}_1) \cdot \mathbf{x}_2 \\
&= (A\mathbf{x}_1) \cdot \mathbf{x}_2 \\
&= (A\mathbf{x}_1)^T\mathbf{x}_2 \\
&= (\mathbf{x}_1^T A^T)\mathbf{x}_2 \\
&= (\mathbf{x}_1^T A)\mathbf{x}_2 && \text{Because } A \text{ is symmetric, } A = A^T. \\
&= \mathbf{x}_1^T(A\mathbf{x}_2) \\
&= \mathbf{x}_1^T(\lambda_2\mathbf{x}_2) \\
&= \mathbf{x}_1 \cdot (\lambda_2\mathbf{x}_2) \\
&= \lambda_2(\mathbf{x}_1 \cdot \mathbf{x}_2).
\end{aligned}$$

This implies that $(\lambda_1 - \lambda_2)(\mathbf{x}_1 \cdot \mathbf{x}_2) = 0$, and because $\lambda_1 \neq \lambda_2$ it follows that $\mathbf{x}_1 \cdot \mathbf{x}_2 = 0$. So, $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

Eigenvectors of a Symmetric Matrix

Show that any two eigenvectors of

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

corresponding to distinct eigenvalues are orthogonal.

SOLUTION The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 3 & -1 \\ -1 & \lambda - 3 \end{vmatrix} = (\lambda - 2)(\lambda - 4),$$

which implies that the eigenvalues of A are $\lambda_1 = 2$ and $\lambda_2 = 4$. Every eigenvector corresponding to $\lambda_1 = 2$ is of the form

$$\mathbf{x}_1 = \begin{bmatrix} s \\ -s \end{bmatrix}, \quad s \neq 0$$

and every eigenvector corresponding to $\lambda_2 = 4$ is of the form

$$\mathbf{x}_2 = \begin{bmatrix} t \\ t \end{bmatrix}, \quad t \neq 0.$$

So,

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \begin{bmatrix} s \\ -s \end{bmatrix} \cdot \begin{bmatrix} t \\ t \end{bmatrix} = st - st = 0,$$

and you can conclude that $\mathbf{x}_1$ and $\mathbf{x}_2$ are orthogonal.

Orthogonal Diagonalization

- A matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix P such that $P^{-1}AP = D$ is diagonal.
- The following important theorem states that the set of orthogonally diagonalizable matrices is precisely the set of symmetric matrices.

Let A be an $n \times n$ matrix. Then A is orthogonally diagonalizable and has real eigenvalues if and only if A is symmetric.

Determining Whether a Matrix Is Orthogonally Diagonalizable

Which matrices are orthogonally diagonalizable?

$$A_1 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 5 & 2 & 1 \\ 2 & 1 & 8 \\ -1 & 8 & 0 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix} \qquad A_4 = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$$

SOLUTION By Theorem 7.10, the orthogonally diagonalizable matrices are the symmetric ones: A_1 and A_4.

Orthogonal Diagonalization of a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix.

1. Find all eigenvalues of A and determine the multiplicity of each.
2. For *each* eigenvalue of multiplicity 1, choose a unit eigenvector. (Choose any eigenvector and then normalize it.)
3. For each eigenvalue of multiplicity $k \geq 2$, find a set of k linearly independent eigenvectors. (You know from Theorem 7.7 that this is possible.) If this set is not orthonormal, apply the Gram-Schmidt orthonormalization process.
4. The composite of steps 2 and 3 produces an orthonormal set of n eigenvectors. Use these eigenvectors to form the columns of P. The matrix $P^{-1}AP = P^TAP = D$ will be diagonal. (The main diagonal entries of D are the eigenvalues of A.)

Orthogonal Diagonalization

Find an orthogonal matrix P that orthogonally diagonalizes

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

SOLUTION 1. The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda + 2 & -2 \\ -2 & \lambda - 1 \end{vmatrix} = (\lambda + 3)(\lambda - 2).$$

So the eigenvalues are $\lambda_1 = -3$ and $\lambda_2 = 2$.

2. For each eigenvalue, find an eigenvector by converting the matrix $\lambda I - A$ to reduced row-echelon form.

$$\begin{array}{lccc} & & & \textit{Eigenvector} \\ -3I - A = \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix} & \rightarrow \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} & \rightarrow & \begin{bmatrix} -2 \\ 1 \end{bmatrix} \\ 2I - A = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} & \rightarrow \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix} & \rightarrow & \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{array}$$

The eigenvectors $(-2, 1)$ and $(1, 2)$ form an *orthogonal* basis for R^2. Each of these eigenvectors is normalized to produce an *orthonormal* basis.

$$\mathbf{p}_1 = \frac{(-2, 1)}{\|(-2, 1)\|} = \frac{1}{\sqrt{5}}(-2, 1) = \left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$$

$$\mathbf{p}_2 = \frac{(1, 2)}{\|(1, 2)\|} = \frac{1}{\sqrt{5}}(1, 2) = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

3. Because each eigenvalue has a multiplicity of 1, go directly to step 4.
4. Using $\mathbf{p}_1$ and $\mathbf{p}_2$ as column vectors, construct the matrix P.

$$P = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

Verify that P is correct by computing $P^{-1}AP = P^TAP$.

$$P^TAP = \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$$

Orthogonal Diagonalization

Find an orthogonal matrix P that diagonalizes

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

SOLUTION

1. The characteristic polynomial of A, $|\lambda I - A| = (\lambda - 3)^2(\lambda + 6)$, yields the eigenvalues $\lambda_1 = -6$ and $\lambda_2 = 3$. λ_1 has a multiplicity of 1 and λ_2 has a multiplicity of 2.
2. An eigenvector for λ_1 is $\mathbf{v}_1 = (1, -2, 2)$, which normalizes to

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{3}, -\frac{2}{3}, \frac{2}{3}\right).$$

3. Two eigenvectors for λ_2 are $\mathbf{v}_2 = (2, 1, 0)$ and $\mathbf{v}_3 = (-2, 0, 1)$. Note that $\mathbf{v}_1$ is orthogonal to $\mathbf{v}_2$ and $\mathbf{v}_3$, as guaranteed by Theorem 7.9. The eigenvectors $\mathbf{v}_2$ and $\mathbf{v}_3$, however, are not orthogonal to each other. To find two orthonormal eigenvectors for λ_2, use the Gram-Schmidt process as follows.

$$\mathbf{w}_2 = \mathbf{v}_2 = (2, 1, 0)$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \left(\frac{\mathbf{v}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\right)\mathbf{w}_2 = \left(-\frac{2}{5}, \frac{4}{5}, 1\right)$$

These vectors normalize to

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}, 0\right)$$

$$\mathbf{u}_3 = \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \left(-\frac{2}{3\sqrt{5}}, \frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}\right).$$

4. The matrix P has $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ as its column vectors.

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

A check shows that

$$P^{-1}AP = P^TAP = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

1. Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for an inner product space V.
2. Let $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$, where $\mathbf{w}_i$ is given by

$$\mathbf{w}_1 = \mathbf{v}_1$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_3, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2$$

$$\vdots$$

$$\mathbf{w}_n = \mathbf{v}_n - \frac{\langle \mathbf{v}_n, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 - \frac{\langle \mathbf{v}_n, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 - \cdots - \frac{\langle \mathbf{v}_n, \mathbf{w}_{n-1} \rangle}{\langle \mathbf{w}_{n-1}, \mathbf{w}_{n-1} \rangle}\mathbf{w}_{n-1}.$$

   Then B' is an *orthogonal* basis for V.

3. Let $\mathbf{u}_i = \frac{\mathbf{w}_i}{\|\mathbf{w}_i\|}$. Then the set $B'' = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an *orthonormal* basis for V. Moreover, span $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ for $k = 1, 2, \ldots, n$.

Population Growth

Matrices can be used to form models for population growth. The first step in this process is to group the population into age classes of equal duration. For instance, if the maximum life span of a member is L years, then the age classes are represented by the n intervals shown below.

$$\underset{\substack{\text{First age}\\\text{class}}}{\left[0, \frac{L}{n}\right)}, \quad \underset{\substack{\text{Second age}\\\text{class}}}{\left[\frac{L}{n}, \frac{2L}{n}\right)}, \ldots, \underset{\substack{n\text{th age}\\\text{class}}}{\left[\frac{(n-1)L}{n}, L\right]}$$

The number of population members in each age class is then represented by the **age distribution vector**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \qquad \begin{matrix} \text{Number in first age class} \\ \text{Number in second age class} \\ \vdots \\ \text{Number in } n\text{th age class} \end{matrix}$$

Over a period of L/n years, the *probability* that a member of the ith age class will survive to become a member of the $(i + 1)$th age class is given by p_i, where

$$0 \le p_i \le 1, \quad i = 1, 2, \ldots, n - 1.$$

The *average number* of offspring produced by a member of the ith age class is given by b_i, where

$$0 \le b_i, \quad i = 1, 2, \ldots, n.$$

These numbers can be written in matrix form, as shown below.

$$A = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{n-1} & b_n \\ p_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_{n-1} & 0 \end{bmatrix}$$

Multiplying this **age transition matrix** by the age distribution vector for a specific time period produces the age distribution vector for the next time period. That is,

$$A\mathbf{x}_i = \mathbf{x}_{i+1}.$$

A Population Growth Model

A population of rabbits raised in a research laboratory has the characteristics listed below.

(a) Half of the rabbits survive their first year. Of those, half survive their second year. The maximum life span is 3 years.

(b) During the first year, the rabbits produce no offspring. The average number of offspring is 6 during the second year and 8 during the third year.

The laboratory population now consists of 24 rabbits in the first age class, 24 in the second, and 20 in the third. How many rabbits will be in each age class in 1 year?

SOLUTION The current age distribution vector is

$$\mathbf{x}_1 = \begin{bmatrix} 24 \\ 24 \\ 20 \end{bmatrix} \qquad \begin{array}{l} 0 \le \text{age} < 1 \\ 1 \le \text{age} < 2 \\ 2 \le \text{age} \le 3 \end{array}$$

and the age transition matrix is

$$A = \begin{bmatrix} 0 & 6 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}.$$

After 1 year the age distribution vector will be

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 0 & 6 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 24 \\ 24 \\ 20 \end{bmatrix} = \begin{bmatrix} 304 \\ 12 \\ 12 \end{bmatrix}. \qquad \begin{array}{l} 0 \le \text{age} < 1 \\ 1 \le \text{age} < 2 \\ 2 \le \text{age} \le 3 \end{array}$$

Finding a Stable Age Distribution Vector

Find a stable age distribution vector for the population in Example 1.

SOLUTION To solve this problem, find an eigenvalue λ and a corresponding eigenvector $\mathbf{x}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda & -6 & -8 \\ -0.5 & \lambda & 0 \\ 0 & -0.5 & \lambda \end{vmatrix} = \lambda^3 - 3\lambda - 2 = (\lambda + 1)^2(\lambda - 2),$$

which implies that the eigenvalues are -1 and 2. Choosing the positive value, let $\lambda = 2$. To find a corresponding eigenvector, row reduce the matrix $2I - A$ to obtain

$$\begin{bmatrix} 2 & -6 & -8 \\ -0.5 & 2 & 0 \\ 0 & -0.5 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -16 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, the eigenvectors of $\lambda = 2$ are of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 16t \\ 4t \\ t \end{bmatrix} = t\begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}.$$

For instance, if $t = 2$, then the initial age distribution vector would be

$$\mathbf{x}_1 = \begin{bmatrix} 32 \\ 8 \\ 2 \end{bmatrix} \quad \begin{array}{l} 0 \le \text{age} < 1 \\ 1 \le \text{age} < 2 \\ 2 \le \text{age} \le 3 \end{array}$$

and the age distribution vector for the next year would be

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 0 & 6 & 8 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}\begin{bmatrix} 32 \\ 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 64 \\ 16 \\ 4 \end{bmatrix}. \quad \begin{array}{l} 0 \le \text{age} < 1 \\ 1 \le \text{age} < 2 \\ 2 \le \text{age} \le 3 \end{array}$$

Notice that the ratio of the three age classes is still 16 : 4 : 1, and so the percent of the population in each age class remains the same.

Applications of Orthogonal Matrices

- Performing QR decomposition in regression problems
- Whitening features before training machine learning models
- Rotating and reflecting data in computer vision tasks

Q & A