MT104-Linear Algebra

Matrices and Matrix Operations

DEFINITION 1 A ***matrix*** is a rectangular array of numbers. The numbers in the array are called the ***entries*** in the matrix.

Examples of Matrices

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

A matrix with m rows and n columns is referred to as an $\boldsymbol{m \times n}$ **matrix** or as having **size** $\boldsymbol{m \times n}$.

A general $m \times n$ matrix is

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

When a compact notation is desired, the preceding matrix can be written as

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

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

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

and for the matrix

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

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

Row Vector and Column Vector

A matrix with only one row is called a ***row vector*** (or a ***row matrix***) , and a matrix with only one column is called a ***column vector*** (or a ***column matrix***).

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

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

Square Matrix and Main Diagonal

A matrix A with n rows and n columns is called a ***square matrix of order n***, and the shaded entries $a_{11}, a_{22}, \ldots, a_{nn}$ in (2) are said to be on the ***main diagonal*** of A.

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

Operations on Matrices

DEFINITION 2 Two matrices are defined to be ***equal*** if they have the same size and their corresponding entries are equal.

Problem

Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$, discuss the possibility that

$A = B$, $B = C$, $A = C$.

Solution ▶ $A = B$ is impossible because A and B are of different sizes: A is 2×2 whereas B is 2×3. Similarly, $B = C$ is impossible. But $A = C$ is possible provided that corresponding entries are equal: $\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}$ means $a = 1$, $b = 0$, $c = -1$, and $d = 2$.

DEFINITION 3 If A and B are matrices of the same size, then the ***sum*** $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A, and the ***difference*** $A - B$ is the matrix obtained by subtracting the entries of B from the corresponding entries of A. Matrices of different sizes cannot be added or subtracted.

Consider the matrices

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

Then

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

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

Scalar Multiples of a Matrix

DEFINITION 4 If A is any matrix and c is any scalar, then the ***product*** cA is the matrix obtained by multiplying each entry of the matrix A by c. The matrix cA is said to be a ***scalar multiple*** of A.

For the matrices

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

we have

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

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

Product of Two Matrices

DEFINITION 5 If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the ***product*** AB is the $m \times n$ matrix whose entries are determined as follows: To find the entry in row i and column j of AB, single out row i from the matrix A and column j from the matrix B. Multiply the corresponding entries from the row and column together, and then add up the resulting products.

Consider the matrices

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

Since A is a 2×3 matrix and B is a 3×4 matrix, the product AB is a 2×4 matrix.

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

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

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

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

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

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

The computations for the remaining entries are

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

$$AB = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

In general, if $A = [a_{ij}]$ is an $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then, as illustrated by the shading in the following display,

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ir} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn} \end{bmatrix} \tag{4}$$

the entry $(AB)_{ij}$ in row i and column j of AB is given by

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj} \tag{5}$$

Formula (5) is called the ***row-column rule*** for matrix multiplication.

Partitioned Matrices

A matrix can be subdivided or ***partitioned*** into smaller matrices by inserting horizontal and vertical rules between selected rows and columns. **e.g**

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

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

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

- Partitioning has many uses, one of which is for finding particular rows or columns of a matrix product AB without computing the entire product.

Matrix Multiplication by Columns and by Rows

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

(*AB* computed column by column)

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

(*AB* computed row by row)

In words, these formulas state that

$$j\text{th column vector of } AB = A[\,j\text{th column vector of } B]$$

$$i\text{th row vector of } AB = [i\text{th row vector of } A]B$$

Problem

Compute AB if $A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 8 & 9 \\ 7 & 2 \\ 6 & 1 \end{bmatrix}$

Solution ▶ The columns of B are $\mathbf{b}_1 = \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 9 \\ 2 \\ 1 \end{bmatrix}$

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 67 \\ 78 \\ 55 \end{bmatrix} \text{ and } A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 9 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 29 \\ 24 \\ 10 \end{bmatrix}.$$

$$AB = [A\mathbf{b}_1 \;\; A\mathbf{b}_2] = \begin{bmatrix} 67 & 29 \\ 78 & 24 \\ 55 & 10 \end{bmatrix}.$$

Powers of a Matrix

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

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

For instance , in order to find A^2 we will multiply A with A.

Transpose of a Matrix

DEFINITION 7 If A is any $m \times n$ matrix, then the ***transpose of A***, denoted by A^T, is defined to be the $n \times m$ matrix that results by interchanging the rows and columns of A; that is, the first column of A^T is the first row of A, the second column of A^T is the second row of A, and so forth.

The following are some examples of matrices and their transposes.

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

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

Properties of Transpose

THEOREM 1.4.8 *If the sizes of the matrices are such that the stated operations can be performed, then:*

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

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

(c) $(A - B)^T = A^T - B^T$

(d) $(kA)^T = kA^T$

(e) $(AB)^T = B^T A^T$

Properties of Matrix Addition and Scalar Multiplication

THEOREM 1.4.1 Properties of Matrix Arithmetic

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

(a) $A + B = B + A$ **[Commutative law for matrix addition]**

(b) $A + (B + C) = (A + B) + C$ **[Associative law for matrix addition]**

(c) $A(BC) = (AB)C$ **[Associative law for matrix multiplication]**

(d) $A(B + C) = AB + AC$ **[Left distributive law]**

(e) $(B + C)A = BA + CA$ **[Right distributive law]**

(f) $A(B - C) = AB - AC$

(g) $(B - C)A = BA - CA$

(h) $a(B + C) = aB + aC$

(i) $a(B - C) = aB - aC$

(j) $(a + b)C = aC + bC$

(k) $(a - b)C = aC - bC$

(l) $a(bC) = (ab)C$

(m) $a(BC) = (aB)C = B(aC)$

WARNINGS:

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

Example on Point 1

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

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

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

$\Rightarrow$ $AB \neq BA$

Example on Point 2

Consider the matrices

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

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

But $$B \neq C$$

Example on Point 3

Here are two matrices for which $AB = 0$, but $A \neq 0$ and $B \neq 0$:

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

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

Trace of a Matrix

DEFINITION 8 If A is a square matrix, then the ***trace of A***, denoted by $\operatorname{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A. The trace of A is undefined if A is not a square matrix.

The following are examples of matrices and their traces.

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

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

Properties of Trace

1. $tr(A + B) = tr(A) + tr(B)$
2. $tr(AB) = tr(BA)$
3. $tr(cA) = c\,tr(A)$
4. $tr(A^t) = tr(A)$

Zero Matrices

A matrix whose entries are all zero is called a ***zero matrix***. Some examples are

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

Identity Matrices

A square matrix with 1's on the main diagonal and zeros elsewhere is called an ***identity matrix***. Some examples are

$$[1], \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

An identity matrix is denoted by the letter I. If it is important to emphasize the size, we will write I_n for the $n \times n$ identity matrix.

Triangular Matrices

A square matrix in which all the entries above the main diagonal are zero is called ***lower triangular***, and a square matrix in which all the entries below the main diagonal are zero is called ***upper triangular***. A matrix that is either upper triangular or lower triangular is called ***triangular***.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

A general 4×4 upper triangular matrix

$$\begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

A general 4×4 lower triangular matrix

$$, \quad I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Symmetric Matrices

A square matrix A is said to be ***symmetric*** if $A = A^T$.

The following matrices are symmetric, since each is equal to its own transpose

$$\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix}$$

THEOREM 1.7.2 *If A and B are symmetric matrices with the same size, and if k is any scalar, then:*

(a) A^T *is symmetric.*

(b) $A + B$ *and* $A - B$ *are symmetric.*

(c) kA *is symmetric.*

THEOREM 1.7.3 *The product of two symmetric matrices is symmetric if and only if the matrices commute.*

Equivalent Matrices

DEFINITION 1 Matrices A and B are said to be ***row equivalent*** if either (hence each) can be obtained from the other by a sequence of elementary row operations.

e.g I_2 **and the matrix** $A = \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ **are equivalent because we can obtain A from identity matrix by multiplying row 2 by -3**

Elementary Matrix

DEFINITION 2 A matrix E is called an ***elementary matrix*** if it can be obtained from an identity matrix by performing a *single* elementary row operation.

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

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

Multiply the second row of I_2 by -3.

Interchange the second and fourth rows of I_4.

Add 3 times the third row of I_3 to the first row.

Multiply the first row of I_3 by 1.

Determinants

The determinant of a 1×1 matrix $[a_{11}]$ is

$$\det[a_{11}] = a_{11}$$

For 2×2 Matrices the formula is

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

Let's talk about minor and cofactor of an element with the help of which we will calculate determinants of higher order matrices.

Minor & Cofactor of an element
(In a Square Matrix)

DEFINITION 1 If A is a square matrix, then the ***minor of entry*** $\boldsymbol{a_{ij}}$ is denoted by M_{ij} and is defined to be the determinant of the submatrix that remains after the ith row and jth column are deleted from A. The number $(-1)^{i+j}M_{ij}$ is denoted by C_{ij} and is called the ***cofactor of entry*** $\boldsymbol{a_{ij}}$.

Let

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

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 16$$

The cofactor of a_{11} is

$$C_{11} = (-1)^{1+1}M_{11} = M_{11} = 16$$

Similarly, the minor of entry a_{32} is

$$M_{32} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

The cofactor of a_{32} is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -26 \blacktriangleleft$$

Definition of a General Determinant

DEFINITION 2 If A is an $n \times n$ matrix, then the number obtained by multiplying the entries in any row or column of A by the corresponding cofactors and adding the resulting products is called the ***determinant of A***, and the sums themselves are called ***cofactor expansions of A***. That is,

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{7}$$

[cofactor expansion along the *j*th column]

and

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \tag{8}$$

[cofactor expansion along the *i*th row]

Problem

Find the determinant of the matrix

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

Solution **By cofactor expansion along the first Row**

$$\det(A) = \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} = 3\begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} - 1\begin{vmatrix} -2 & 3 \\ 5 & -2 \end{vmatrix} + 0\begin{vmatrix} -2 & -4 \\ 5 & 4 \end{vmatrix}$$

$$= 3(-4) - (1)(-11) + 0 = -1$$

By cofactor expansion along the first coulmn

Solution

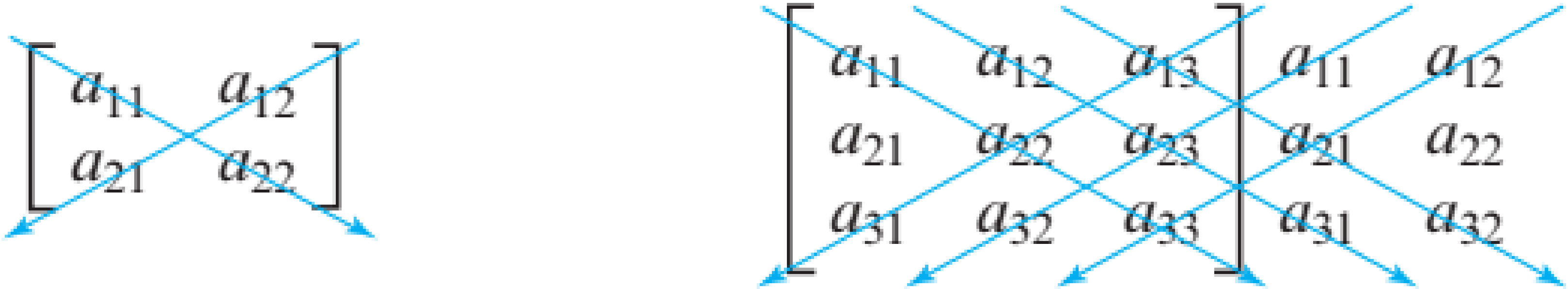
$$\det(A) = \begin{vmatrix} 3 & 1 & 0 \\ -2 & -4 & 3 \\ 5 & 4 & -2 \end{vmatrix} = 3\begin{vmatrix} -4 & 3 \\ 4 & -2 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 4 & -2 \end{vmatrix} + 5\begin{vmatrix} 1 & 0 \\ -4 & 3 \end{vmatrix}$$

$$= 3(-4) - (-2)(-2) + 5(3) = -1$$

Tthe best strategy for co-factor expansion is to expand along a row or column with the most zeros.

A Useful Technique for Evaluating 2×2 and 3×3 Determinants

Determinants of 2×2 and 3×3 matrices can be evaluated very efficiently using the pattern suggested in Figure 2.1.1.



▶ Figure 2.1.1

e. g

$$\begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = (3)(-2) - (1)(4) = -10$$

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = \left|\begin{matrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{matrix}\right| \begin{matrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{matrix}$$

$$= [45 + 84 + 96] - [105 - 48 - 72] = 240$$

Properties of Determinants

Theorem 8.5.1 Determinant of a Transpose

If $\mathbf{A}^T$ is the transpose of the $n \times n$ matrix $\mathbf{A}$, then $\det \mathbf{A}^T = \det \mathbf{A}$.

For example, for the matrix $\mathbf{A} = \begin{pmatrix} 5 & 7 \\ 3 & -4 \end{pmatrix}$, we have $\mathbf{A}^T = \begin{pmatrix} 5 & 3 \\ 7 & -4 \end{pmatrix}$. Observe that

$$\det \mathbf{A} = \begin{vmatrix} 5 & 7 \\ 3 & -4 \end{vmatrix} = -41 \quad \text{and} \quad \det \mathbf{A}^T = \begin{vmatrix} 5 & 3 \\ 7 & -4 \end{vmatrix} = -41.$$

Theorem 8.5.2 Two Identical Rows

If any two rows (columns) of an $n \times n$ matrix $\mathbf{A}$ are the same, then $\det \mathbf{A} = 0$.

EXAMPLE 1 Matrix with Two Identical Rows

Since the second and third columns in the matrix $\mathbf{A} = \begin{pmatrix} 6 & 2 & 2 \\ 4 & 2 & 2 \\ 9 & 2 & 2 \end{pmatrix}$ are the same, it follows from Theorem 8.5.2 that

$$\det \mathbf{A} = \begin{vmatrix} 6 & 2 & 2 \\ 4 & 2 & 2 \\ 9 & 2 & 2 \end{vmatrix} = 0.$$

Theorem 8.5.3 Zero Row or Column

If all the entries in a row (column) of an $n \times n$ matrix $\mathbf{A}$ are zero, then $\det \mathbf{A} = 0$.

$$\text{zero row} \rightarrow \begin{vmatrix} 0 & 0 \\ 7 & -6 \end{vmatrix} = 0 \quad \text{and} \quad \begin{array}{c} \text{zero column} \downarrow \\ \begin{vmatrix} 4 & 6 & 0 \\ 1 & 5 & 0 \\ 8 & -1 & 0 \end{vmatrix} \end{array} = 0.$$

Theorem 8.5.4 Interchanging Rows

If $\mathbf{B}$ is the matrix obtained by interchanging any two rows (columns) of an $n \times n$ matrix $\mathbf{A}$, then $\det \mathbf{B} = -\det \mathbf{A}$.

For example, if $\mathbf{B}$ is the matrix obtained by interchanging the first and third rows of

$$\mathbf{A} = \begin{pmatrix} 4 & -1 & 9 \\ 6 & 0 & 7 \\ 2 & 1 & 3 \end{pmatrix}$$

then $$\det \mathbf{B} = \begin{vmatrix} 2 & 1 & 3 \\ 6 & 0 & 7 \\ 4 & -1 & 9 \end{vmatrix} = -\begin{vmatrix} 4 & -1 & 9 \\ 6 & 0 & 7 \\ 2 & 1 & 3 \end{vmatrix} = -\det \mathbf{A}.$$

Theorem 8.5.5 Constant Multiple of a Row

If $\mathbf{B}$ is the matrix obtained from an $n \times n$ matrix $\mathbf{A}$ by multiplying a row (column) by a nonzero real number k, then $\det \mathbf{B} = k \det \mathbf{A}$.

$$\begin{vmatrix} 5 & 8 \\ 20 & 16 \end{vmatrix} = \overset{\substack{\text{from}\\ \text{first column}\\ \downarrow}}{5} \begin{vmatrix} 1 & 8 \\ 4 & 16 \end{vmatrix} = 5 \cdot \overset{\substack{\text{from}\\ \text{second column}\\ \downarrow}}{8} \begin{vmatrix} 1 & 1 \\ 4 & 2 \end{vmatrix} = 5 \cdot 8 \cdot \overset{\substack{\text{from}\\ \text{second row}\\ \downarrow}}{2} \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} = 80(1 - 2) = -80$$

$$\begin{vmatrix} 4 & 2 & -1 \\ 5 & -2 & 1 \\ 7 & 4 & -2 \end{vmatrix} = (-2) \begin{vmatrix} 4 & -1 & -1 \\ 5 & 1 & 1 \\ 7 & -2 & -2 \end{vmatrix} = (-2) \cdot 0 = 0$$

Theorem 8.5.6 Determinant of a Matrix Product

If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ matrices, then $\det \mathbf{AB} = \det \mathbf{A} \cdot \det \mathbf{B}$.

In other words, the determinant of a product of two $n \times n$ matrices is the same as the product of the determinants.

Suppose $\mathbf{A} = \begin{pmatrix} 2 & 6 \\ 1 & -1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 3 & -4 \\ -3 & 5 \end{pmatrix}$. Then $\mathbf{AB} = \begin{pmatrix} -12 & 22 \\ 6 & -9 \end{pmatrix}$. Now $\det \mathbf{AB} = -24$, $\det \mathbf{A} = -8$, $\det \mathbf{B} = 3$, and so we see that

$$\det \mathbf{A} \cdot \det \mathbf{B} = (-8)(3) = -24 = \det \mathbf{AB}. \quad \equiv$$

Theorem 8.5.7 Determinant Is Unchanged

Suppose **B** is the matrix obtained from an $n \times n$ matrix **A** by multiplying the entries in a row (column) by a nonzero real number k and adding the result to the corresponding entries in another row (column). Then $\det \mathbf{B} = \det \mathbf{A}$.

A Multiple of a Row Added to Another

Suppose $\mathbf{A} = \begin{pmatrix} 5 & 1 & 2 \\ 3 & 0 & 7 \\ 4 & -1 & 4 \end{pmatrix}$ and suppose the matrix **B** is defined as that matrix obtained from **A** by the elementary row operation

$$\mathbf{A} = \begin{pmatrix} 5 & 1 & 2 \\ 3 & 0 & 7 \\ 4 & -1 & 4 \end{pmatrix} \overset{-3R_1 + R_3}{\Rightarrow} \begin{pmatrix} 5 & 1 & 2 \\ 3 & 0 & 7 \\ -11 & -4 & -2 \end{pmatrix} = \mathbf{B}.$$

Expanding by cofactors along, say, the second column, we find $\det \mathbf{A} = 45$ and $\det \mathbf{B} = 45$.

Theorem 8.5.8 Determinant of a Triangular Matrix

Suppose $\mathbf{A}$ is an $n \times n$ triangular matrix (upper or lower). Then

$$\det \mathbf{A} = a_{11}a_{22} \cdots a_{nn},$$

where $a_{11}, a_{22}, \ldots, a_{nn}$ are the entries on the main diagonal of $\mathbf{A}$.

$$\begin{vmatrix} 3 & 0 & 0 & 0 \\ 2 & 6 & 0 & 0 \\ 5 & 9 & -4 & 0 \\ 7 & 2 & 4 & -2 \end{vmatrix} = 3 \cdot 6 \cdot (-4) \cdot (-2) = 144.$$

$$\begin{vmatrix} -3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 4 \end{vmatrix} = (-3) \cdot 6 \cdot 4 = -72.$$

THEOREM 2.3.1 *Let A, B, and C be $n \times n$ matrices that differ only in a single row, say the rth, and assume that the rth row of C can be obtained by adding corresponding entries in the rth rows of A and B. Then*

$$\det(C) = \det(A) + \det(B)$$

The same result holds for columns.

$$\det\begin{bmatrix} 1 & 7 & 5 \\ 2 & 0 & 3 \\ 1+0 & 4+1 & 7+(-1) \end{bmatrix} = \det\begin{bmatrix} 1 & 7 & 5 \\ 2 & 0 & 3 \\ 1 & 4 & 7 \end{bmatrix} + \det\begin{bmatrix} 1 & 7 & 5 \\ 2 & 0 & 3 \\ 0 & 1 & -1 \end{bmatrix}$$

$$\det(kA) = k^n \det(A)$$

In general

$$\det(A + B) \neq \det(A) + \det(B)$$

Consider

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

We have $\det(A) = 1$, $\det(B) = 8$, and $\det(A + B) = 23$; thus

$$\det(A + B) \neq \det(A) + \det(B) \quad \blacktriangleleft$$

Evaluating Determinants by Row Reduction

Evaluating the determinant of an $n \times n$ matrix by the method of cofactor expansion requires a Herculean effort when the order of the matrix is large. To expand the determinant of, say, a 5×5 matrix with nonzero entries requires evaluating five cofactors that are determinants of 4×4 submatrices; each of these in turn requires four additional cofactors that are determinants of 3×3 submatrices, and so on.

The idea of the method is to reduce the given matrix to upper triangular form by elementary row operations, then compute the determinant of the upper triangular matrix (an easy computation), and then relate that determinant to that of the original matrix.

Evaluate the determinant of $\mathbf{A} = \begin{pmatrix} 6 & 2 & 7 \\ -4 & -3 & 2 \\ 2 & 4 & 8 \end{pmatrix}$.

SOLUTION

$$\det \mathbf{A} = \begin{vmatrix} 6 & 2 & 7 \\ -4 & -3 & 2 \\ 2 & 4 & 8 \end{vmatrix}$$

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

2 is a common factor in third row

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

interchange first and third rows

$$= -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 18 \\ 6 & 2 & 7 \end{vmatrix}$$

4 times first row added to second row

$$= -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 18 \\ 0 & -10 & -17 \end{vmatrix}$$

-6 times first row added to third row

$$= -2\begin{vmatrix} 1 & 2 & 4 \\ 0 & 5 & 18 \\ 0 & 0 & 19 \end{vmatrix}$$

2 times second row added to third row

$$= (-2)(1)(5)(19) = -190$$

Compute the determinant of $$A = \begin{bmatrix} 1 & 0 & 0 & 3 \\ 2 & 7 & 0 & 6 \\ 0 & 6 & 3 & 0 \\ 7 & 3 & 1 & -5 \end{bmatrix}$$

Adjoint Matrix

DEFINITION 1 If A is any $n \times n$ matrix and C_{ij} is the cofactor of a_{ij}, then the matrix

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

is called the ***matrix of cofactors from A***. The transpose of this matrix is called the ***adjoint of A*** and is denoted by $\text{adj}(A)$.

Let

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

$$\begin{array}{lll} C_{11} = 12 & C_{12} = 6 & C_{13} = -16 \\ C_{21} = 4 & C_{22} = 2 & C_{23} = 16 \\ C_{31} = 12 & C_{32} = -10 & C_{33} = 16 \end{array}$$

so the matrix of cofactors is

$$\begin{bmatrix} 12 & 6 & -16 \\ 4 & 2 & 16 \\ 12 & -10 & 16 \end{bmatrix}$$

and the adjoint of A is

$$\operatorname{adj}(A) = \begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix} \blacktriangleleft$$

Inverse of a Matrix

Definition 8.6.1 Inverse of a Matrix

Let $\mathbf{A}$ be an $n \times n$ matrix. If there exists an $n \times n$ matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}, \tag{1}$$

where $\mathbf{I}$ is the $n \times n$ identity, then the matrix $\mathbf{A}$ is said to be **nonsingular** or **invertible**. The matrix $\mathbf{B}$ is said to be the **inverse** of $\mathbf{A}$.

For example, the matrix $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ is nonsingular or invertible since the matrix $\mathbf{B} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ is its inverse. To verify this, observe that

$$\mathbf{AB} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}$$

and

$$\mathbf{BA} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}.$$

Inverse of a 2 × 2 Matrix

THEOREM 1.4.5 *The matrix*

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

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

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

Determine whether the matrix is invertible. If so, find its inverse. $A = \begin{bmatrix} 6 & 1 \\ 5 & 2 \end{bmatrix}$

Solution The determinant of A is $\det(A) = (6)(2) - (1)(5) = 7$, which is nonzero.

Thus, A is invertible, and its inverse is

$$A^{-1} = \frac{1}{7}\begin{bmatrix} 2 & -1 \\ -5 & 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{5}{7} & \frac{6}{7} \end{bmatrix}$$

THEOREM 2.3.6 Inverse of a Matrix Using Its Adjoint

If A is an invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) \tag{6}$$

Problem

Find the inverse of the matrix $A = \begin{bmatrix} 3 & 2 & -1 \\ 1 & 6 & 3 \\ 2 & -4 & 0 \end{bmatrix}$

Sol

$$\begin{array}{lll} C_{11} = 12 & C_{12} = 6 & C_{13} = -16 \\ C_{21} = 4 & C_{22} = 2 & C_{23} = 16 \\ C_{31} = 12 & C_{32} = -10 & C_{33} = 16 \end{array} \qquad \operatorname{adj}(A) = \begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix} \qquad \det(A) = 64$$

$$A^{-1} = \frac{1}{\det(A)}\operatorname{adj}(A) = \frac{1}{64}\begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix} = \begin{bmatrix} \frac{12}{64} & \frac{4}{64} & \frac{12}{64} \\ \frac{6}{64} & \frac{2}{64} & -\frac{10}{64} \\ -\frac{16}{64} & \frac{16}{64} & \frac{16}{64} \end{bmatrix}$$

Another method for finding inverse of a matrix is illustrated in the coming slides.

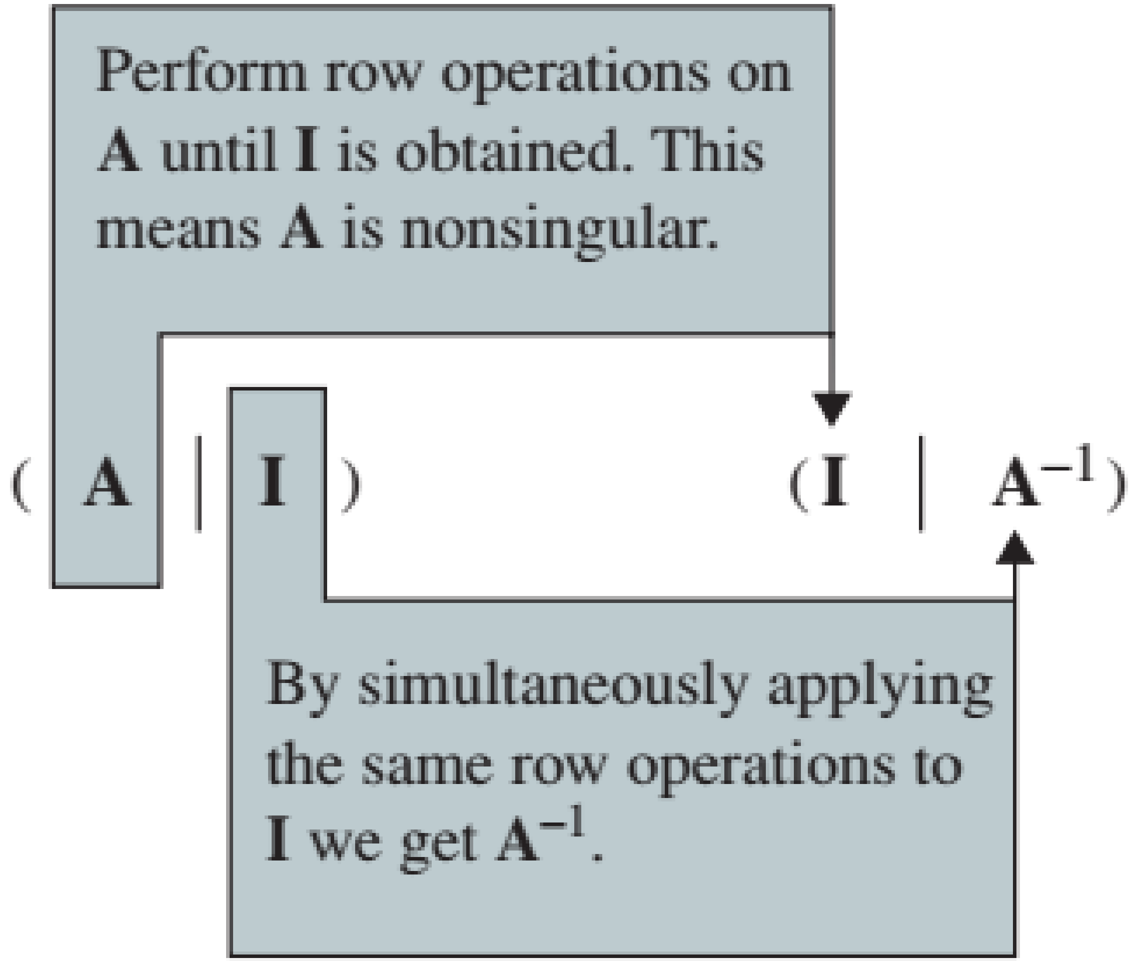
Theorem 8.6.4 Finding the Inverse

If an $n \times n$ matrix $\mathbf{A}$ can be transformed into the $n \times n$ identity $\mathbf{I}$ by a sequence of elementary row operations, then $\mathbf{A}$ is nonsingular. The same sequence of operations that transforms $\mathbf{A}$ into the identity $\mathbf{I}$ will also transform $\mathbf{I}$ into $\mathbf{A}^{-1}$.

It is convenient to carry out these row operations on $\mathbf{A}$ and $\mathbf{I}$ simultaneously by means of an $n \times 2n$ matrix obtained by augmenting $\mathbf{A}$ with the identity $\mathbf{I}$ as shown here:

$$(\mathbf{A}|\mathbf{I}) = \left(\begin{array}{cccc|cccc} a_{11} & a_{12} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array}\right).$$

The procedure for finding $\mathbf{A}^{-1}$ is outlined in the following diagram:



Find the inverse of

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

The computations are as follows:

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

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

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

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right] \longleftarrow \text{We added 2 times the second row to the third.}$$

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \longleftarrow \text{We multiplied the third row by } -1.$$

$$\left[\begin{array}{rrr|rrr} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \longleftarrow \text{We added 3 times the third row to the second and } -3 \text{ times the third row to the first.}$$

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

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

Thus,

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

Note

Often it will not be known in advance if a given $n \times n$ matrix A is invertible.

However, if it is not, then it will be impossible to reduce A to I_n by elementary row operations. This will be signaled by a row of zeros appearing on the *left side* of the partition at some stage of the inversion algorithm. If this occurs, then you can stop the computations and conclude that A is not invertible.

▶ EXAMPLE 5 Showing That a Matrix Is Not Invertible

Consider the matrix

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

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

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

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

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

We added the second row to the third.

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

Properties

Let **A** and **B** be nonsingular matrices. Then

(*i*) $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$

(*ii*) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

(*iii*) $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$

THEOREM 2.3.5 *If A is invertible, then*

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

THEOREM 1.4.9 *If A is an invertible matrix, then* A^T *is also invertible and*

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

THEOREM 1.4.7 *If A is invertible and n is a nonnegative integer, then:*

(*a*) A^{-1} *is invertible and* $(A^{-1})^{-1} = A$.

(*b*) A^n *is invertible and* $(A^n)^{-1} = A^{-n} = (A^{-1})^n$.

(*c*) kA *is invertible for any nonzero scalar* k, *and* $(kA)^{-1} = k^{-1}A^{-1}$.

Diagonal Matrices

A square matrix in which all the entries off the main diagonal are zero is called a ***diagonal matrix***. Here are some examples:

$$\begin{bmatrix} 2 & 0 \\ 0 & -5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Remark Observe that diagonal matrices are both upper triangular and lower triangular since they have zeros below and above the main diagonal. Observe also that a *square* matrix in row echelon form is upper triangular since it has zeros below the main diagonal.

General Form of a Diagonal Matrix

A general $n \times n$ diagonal matrix D can be written as

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

Inverse of a Diagonal Matrix

A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero; in this case the inverse of (1) is

$$D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix}$$

Power of a Diagonal Matrix

Powers of diagonal matrices are easy to compute; we leave it for you to verify that if D is the diagonal matrix (1) and k is a positive integer, then

$$D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix}$$

If

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

then

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad A^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -243 & 0 \\ 0 & 0 & 32 \end{bmatrix}, \quad A^{-5} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{243} & 0 \\ 0 & 0 & \frac{1}{32} \end{bmatrix}$$

- *nonzero* row or column in a matrix means a row or column that contains at least one nonzero entry;
- a **leading entry** of a row refers to the leftmost nonzero entry (in a nonzero row).

Echelon Forms

A matrix is said to be in Echelon form if it has the following properties

1. **Every nonzero row precedes every zero row**.
2. **In each successive nonzero row, the number of zeros before the leading entry of a row increases row by row.**

If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form**):

- The leading entry in each nonzero row is 1.
- Each leading 1 is the only nonzero entry in its column.

The matrices

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$ are in echelon form.

In fact, the second matrix is in reduced echelon form.

THEOREM

Uniqueness of the Reduced Echelon Form

Each matrix is row equivalent to one and only one reduced echelon matrix.

Rank of a Matrix

The number of nonzero rows in the echelon form of a matrix is called rank of the matrix.

THEOREM

An $n \times n$ square matrix $\mathbf{A}$ has rank n if and only if

$$\det \mathbf{A} \neq 0.$$

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Linear System

A general linear system of m equations in the n unknowns $x_1, x_2, \ldots, x_n$ can be written as

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

A ***solution*** of a linear system in n unknowns $x_1, x_2, \ldots, x_n$ is a sequence of n numbers $s_1, s_2, \ldots, s_n$ for which the substitution

$$x_1 = s_1, \quad x_2 = s_2, \ldots, \quad x_n = s_n$$

makes each equation a true statement.

For example $x = 1, \quad y = -2$ is a solution of $\begin{array}{l} 5x + y = 3 \\ 2x - y = 4 \end{array}$

In matrix form, we can write (★) as

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

where,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

A is called the **coefficient matrix**

$$\tilde{\mathbf{A}} = \left[\begin{array}{ccc:c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

$\tilde{\mathbf{A}}$ is called the **augmented matrix** of the system

Homogeneous Linear Systems

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

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

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

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

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

Important Points about Linear Systems

1. There are a lot of methods for solving linear systems, some methods work in special cases while some works for every type of linear system.

2. If in a linear system the number of equations is equal to the number of unknowns (i.e. if the coefficient matrix is a square matrix i.e. if $m = n$) then from determinant we can comment about the number of solution(s).

3. Any linear system in which $m = n$, and $|A| \neq 0$, has a unique solution. We can obtain this unique solution by any of the following methods:

(i) Matrix Inversion Method (ii) Cramer's Rule

(iii) Gauss Elimination (iv) Gauss-Jordan Elimination

Furthermore, if the system is homogeneous then this unique solution is the trivial solution (zero solution).

4. A homogeneous system has always a zero solution called trivial solution. Thus a homogeneous system is always consistent.

5. A homogeneous linear system in which $m = n$, and $|A| = 0$, has infinite solutions (i.e. nontrivial solutions also exist).

5. In case of nonhomogeneous linear system in which $m = n$ and $|A| = 0$, there is possibility that the system has

(i) No Solution (ii) Infinite Solutions

6. If in a linear system <u>the number of equations is not equal to the number of unknowns</u> (i.e. if the coefficient matrix is not a square matrix i.e. if $m \neq n$) then we can comment about the number of solution(s) by looking into the rank of coefficient matrix, rank of augmented matrix and number of unknowns in the system. In this case we can find the solution(s) (if exist) by either Gauss Elimination or by Gauss-Jordan Elimination.

THEOREM 1

Fundamental Theorem for Linear Systems

(a) Existence. *A linear system of* m *equations in* n *unknowns* $x_1, \cdots, x_n$

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

is **consistent**, *that is, has solutions, if and only if the coefficient matrix* $\mathbf{A}$ *and the augmented matrix* $\widetilde{\mathbf{A}}$ *have the same rank. Here,*

$$
\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \widetilde{\mathbf{A}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdots & \cdot & \cdot \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]
$$

(b) Uniqueness. *The system* (1) *has precisely one solution if and only if this common rank* r *of* $\mathbf{A}$ *and* $\widetilde{\mathbf{A}}$ *equals* n.

(c) Infinitely many solutions. *If this common rank r is less than n, the system (1) has infinitely many solutions. All of these solutions are obtained by determining r suitable unknowns in terms of the remaining n − r unknowns, to which arbitrary values can be assigned.*

(d) Gauss elimination *If solutions exist, they can all be obtained by the Gauss elimination.* (This method will automatically reveal whether or not solutions exist.

THEOREM 1.2.2 *A homogeneous linear system with more unknowns than equations has infinitely many solutions.*

(Because for homogeneous linear system Rank of the coefficient matrix is always equal to Rank of the augmented Matrix and if number of equations is less than unknowns this mean that common rank < n, so by third part of fundamental theorem, system will have infinite solutions.

Methods for solving a linear System

Gaussian Elimination (or Gauss Elimination)

Gaussian Elimination with Back-Substitution

1. Write the augmented matrix of the system of linear equations.
2. Use elementary row operations to rewrite the 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.

Problem

Solve the system.

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

SOLUTION The augmented matrix for this system is

$$
\left[\begin{array}{rrrr|r}
0 & 1 & 1 & -2 & -3 \\
1 & 2 & -1 & 0 & 2 \\
2 & 4 & 1 & -3 & -2 \\
1 & -4 & -7 & -1 & -19
\end{array}\right].
$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 2 & 4 & 1 & -3 & -2 \\ 1 & -4 & -7 & -1 & -19 \end{array}\right] \qquad R_1 \leftrightarrow R_2$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 1 & -4 & -7 & -1 & -19 \end{array}\right] \qquad R_3 + (-2)R_1$$

$$\left[\begin{array}{rrrr|r} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & -6 & -6 & -1 & -21 \end{array}\right] \quad R_4 + (-1)R_1$$

$$\left[\begin{array}{rrrr|r} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 3 & -3 & -6 \\ 0 & 0 & 0 & -13 & -39 \end{array}\right] \quad R_4 + (6)R_2$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & -13 & -39 \end{array}\right] \qquad \left(\tfrac{1}{3}\right)R_3$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 2 \\ 0 & 1 & 1 & -2 & -3 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \qquad \left(-\tfrac{1}{13}\right)R_4$$

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

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

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

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

Gauss-Jordan Elimination

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 Friedrich Gauss and Wilhelm Jordan (1842–1899), continues the reduction process until a *reduced* row-echelon form is obtained.

Use Gauss-Jordan elimination to solve the system.

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

Associated Augmented Matrix

$$\left[\begin{array}{rrr|r} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{array}\right]$$

Add the first row to the second row to produce a new second row.

$$\left[\begin{array}{rrr|r} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 2 & -5 & 5 & 17 \end{array}\right] \quad R_2 + R_1 \rightarrow R_2$$

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

$$\left[\begin{array}{rrr|r} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{array}\right] \quad R_3 + (-2)R_1 \rightarrow R_3$$

Add the second row to the third row to produce a new third row.

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{array}\right] \quad R_3 + R_2 \rightarrow R_3$$

Multiply the third row by $\frac{1}{2}$ to produce a new third row.

$$\left[\begin{array}{ccc|c} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \quad \left(\tfrac{1}{2}\right)R_3 \rightarrow R_3$$

Now, apply elementary row operations until you obtain zeros above each of the leading 1's, as shown below.

$$\left[\begin{array}{ccc|c} 1 & 0 & 9 & 19 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad R_1 + (2)R_2 \rightarrow R_1$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 9 & 19 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad R_2 + (-3)R_3 \rightarrow R_2$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{array}\right] \qquad R_1 + (-9)R_3 \rightarrow R_1$$

The matrix is now in reduced row-echelon form. Converting back to a system of linear equations, you have

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

Problem

Solve by Gauss–Jordan elimination.

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

Solution The augmented matrix for the system is

$$\left[\begin{array}{rrrrrr|r}
1 & 3 & -2 & 0 & 2 & 0 & 0 \\
2 & 6 & -5 & -2 & 4 & -3 & -1 \\
0 & 0 & 5 & 10 & 0 & 15 & 5 \\
2 & 6 & 0 & 8 & 4 & 18 & 6
\end{array}\right]$$

If we reduced the augmented matrix to reduce echelon form we have

$$\left[\begin{array}{cccccc|c} 1 & 3 & 0 & 4 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The corresponding system of equations is

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

Solving for the leading variables, we obtain

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

THEOREM 3.11 Cramer's Rule

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

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

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

Problem

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

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

SOLUTION

The determinant of the coefficient matrix is $|A| = \begin{vmatrix} -1 & 2 & -3 \\ 2 & 0 & 1 \\ 3 & -4 & 4 \end{vmatrix} = 10.$

The determinant is nonzero, so you know that the solution is unique. Apply Cramer's Rule to solve for x, as shown below.

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

$$y = \frac{\det(A_2)}{\det(A)} =$$

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

$$y = \frac{\begin{vmatrix} -1 & 1 & -3 \\ 2 & 0 & 1 \\ 3 & 2 & 4 \end{vmatrix}}{10} =$$

$A_2 =$

$$|A_2| = (-1)(-2) - 1(5) - (3)(4)$$

$$= 2 - 5 - 12$$

$$\boxed{|A_2| = -15}$$

$$y = \frac{-\cancel{15}^{\,3}}{\cancel{10}_{\,2}} = -3/2$$

$$A_3 = \begin{bmatrix} -1 & 2 & 1 \\ 2 & 0 & 0 \\ 3 & -4 & 2 \end{bmatrix}$$

$$z = \frac{\det(A_3)}{\det(A)} = \frac{\boxed{}}{10} = \bigcirc$$

Solving Linear Systems by Matrix Inversion

$n \times n$ $\quad$ $n \times 1$

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

$A^{-1}Ax = A^{-1}b$

$Ix = A^{-1}b$

$\Rightarrow \boxed{x = A^{-1}b} \rightarrow (1)$

$n \times 1$ $\quad$ $n \times$

e.g.

Consider the system of linear equations

$$
\left.\begin{aligned}
x_1 + 2x_2 + 3x_3 &= 5 \\
2x_1 + 5x_2 + 3x_3 &= 3 \\
x_1 + 8x_3 &= 17
\end{aligned}\right\} \text{(A)}
$$

In matrix form this system can be written as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix}$$

$$Ax = b$$

$$\Rightarrow x = A^{-1} b \longrightarrow (1)$$

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

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

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

$$\operatorname{adj}(C) = \begin{bmatrix} & & \end{bmatrix}^T$$

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

or $x_1 = 1$, $x_2 = -1$, $x_3 = 2$.

CHAPTER 4

General Vector Spaces

4.1 Real Vector Spaces

DEFINITION 1 Let V be an arbitrary nonempty set of objects on which two operations are defined: addition, and multiplication by numbers called ***scalars***. By ***addition*** we mean a rule for associating with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V an object $\mathbf{u}+\mathbf{v}$, called the ***sum*** of $\mathbf{u}$ and $\mathbf{v}$; by ***scalar multiplication*** we mean a rule for associating with each scalar k and each object $\mathbf{u}$ in V an object $k\mathbf{u}$, called the ***scalar multiple*** of $\mathbf{u}$ by k. If the following axioms are satisfied by all objects $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all scalars k and m, then we call V a ***vector space*** and we call the objects in V ***vectors***.

1. If $\mathbf{u}$ and $\mathbf{v}$ are objects in V, then $\mathbf{u}+\mathbf{v}$ is in V.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$
4. There is an object $\mathbf{0}$ in V, called a ***zero vector*** for V, such that $\mathbf{0}+\mathbf{u}=\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}$ in V.
5. For each $\mathbf{u}$ in V, there is an object $-\mathbf{u}$ in V, called a ***negative*** of $\mathbf{u}$, such that $\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}$.

6. If k is any scalar and $\mathbf{u}$ is any object in V, then $k\mathbf{u}$ is in V.

7. $k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$

8. $(k + m)\mathbf{u} = k\mathbf{u} + m\mathbf{u}$

9. $k(m\mathbf{u}) = (km)(\mathbf{u})$

10. $1\mathbf{u} = \mathbf{u}$

To Show That a Set with Two Operations Is a Vector Space

Step 1. Identify the set V of objects that will become vectors.

Step 2. Identify the addition and scalar multiplication operations on V.

Step 3. Verify Axioms 1 and 6; that is, adding two vectors in V produces a vector in V, and multiplying a vector in V by a scalar also produces a vector in V. Axiom 1 is called ***closure under addition***, and Axiom 6 is called ***closure under scalar multiplication***.

Step 4. Confirm that Axioms 2, 3, 4, 5, 7, 8, 9, and 10 hold.

R^n with the Standard Operations Is a Vector Space

Let $V = R^n$, and define the vector space operations on V to be the usual operations of addition and scalar multiplication of n-tuples; that is,

$$\mathbf{u} + \mathbf{v} = (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$
$$k\mathbf{u} = (ku_1, ku_2, \ldots, ku_n)$$

The set $V = R^n$ is closed under addition and scalar multiplication because the foregoing operations produce n-tuples as their end result, and these operations satisfy Axioms 2, 3, 4, 5, 7, 8, 9, and 10 .

$V = R^3$

$V = R^3$

$u = \{u_1, u_2, u_3\}, \quad v = \{v_1, v_2, v_3\}, \quad w = \{w_1, w_2, w_3\} \in V = R^3$

(i) $u + v = \{u_1 + v_1, u_2 + v_2, u_3 + v_3\} \in V = R^3$

(ii) $u + v = \{u_1 + v_1, u_2 + v_2, u_3 + v_3\}$

$= \{v_1 + u_1, v_2 + u_2, v_3 + u_3\}$

$u + v = v + u \Rightarrow$

(iii) $u + (v + w) = u + \{v_1 + w_1, v_2 + w_2, v_3 + w_3\}$

$= \{u_1, u_2, u_3\} + \{v_1 + w_1, v_2 + w_2, v_3 + w_3\}$

$= \{u_1 + v_1 + w_1, u_2 + v_2 + w_2, u_3 + v_3 + w_3\}$

$= \{u_1 + v_1, u_2 + v_2, u_3 + v_3\} + \{w_1, w_2, w_3\}$

$= (u + v) + w$

(iv) $0 = (0, 0, 0) \in R^3$

$u = \{u_1, u_2, u_3\}$

$0 + u = 0 + u = u$

(v) $u \in R^3$ $\quad -u \in R^3$

$u + (-u) = (-u) + u = 0$

(vi) let k is any scalar and $u \in R^3$ $\quad ku \in R^3$

$ku = k(u_1, u_2, u_3) = \{ku_1, ku_2\}$

$ku \in R^3$

(vii) $k(u+v) = k(u_1+v_1, u_2+v_2, u_3+v_3)$

$= \{ku_1 + kv_1, ku_2 + kv_2, ku_3 + kv_3)$

$= (ku_1, ku_2, ku_3\} + \{kv_1, kv_2, kv_3\}$

$= k\{u_1, u_2, u_3\} + k(v_1, v_2, v_3)$

$k(u+v) = ku + kv$ ✓

viii $(k+m)u = (k+m)(u_1, u_2, u_3) = \{(k+m)u_1, (k+m)u_2, (k+m)u_3\}$

$= \{ku_1 + mu_1, ku_2 + mu_2, ku_3 + mu_3\}$

$\boxed{ku + mu}$

$= \{ku_1, ku_2, ku_3\} + \{mu_1, mu_2, mu_3\}$

$= k\{u_1, u_2, u_3\} + m\{u_1, u_2, u_3\}$

$= ku + mu$ ✓

ix $k(mu) = k(mu_1, mu_2, mu_3) = \{kmu_1, kmu_2, kmu_3\}$

$= (km)\{u_1, u_2, u_3\} = (km)u$

$1 \in \mathbb{R}$, $u = \{u_1, u_2, u_3\}$

$1u = 1\{u_1, u_2, u_3\} = \{1 u_1, 1 \cdot u_2, 1 \cdot u_3\} = \{u_1, u_2, u_3\} = u$

$1u = u$.

$\Rightarrow V = R^3$ is a V. space.

$P_1(x) =$ $p(x) + Q(x)$

$Q =$

$\boxed{\int_{x_1}^{x_1} \; diff}$

$V = R^n$

$V = P(x)$

$V = P_2(x)$

The Vector Space of All Polynomials of Degree 2 or Less

Let P_2 be the set of all polynomials of the form $p(x) = a_0 + a_1x + a_2x^2$, where a_0, a_1, and a_2 are real numbers. The *sum* of two polynomials $p(x) = a_0 + a_1x + a_2x^2$ and $q(x) = b_0 + b_1x + b_2x^2$ is defined in the usual way,

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

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

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

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 example, the set of real numbers is closed under addition, so it follows that $a_0 + b_0$, $a_1 + b_1$, and $a_2 + b_2$ are real numbers, and

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

is in the set P_2 because it is a polynomial of degree 2 or less. So, P_2 is closed under addition. To verify the commutative property of addition, write

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

Can you see where the commutative property of addition of real numbers was used? The zero vector in this space is the zero polynomial $\mathbf{0}(x) = 0 + 0x + 0x^2$. Verify the other vector space axioms to show that P_2 is a vector space.

Important Vector Spaces

R = set of all real numbers

R^2 = set of all ordered pairs

R^3 = set of all ordered triples

R^n = set of all n-tuples

$C(-\infty, \infty)$ = set of all continuous functions defined on the real number line

$C[a, b]$ = set of all continuous functions defined on a closed interval $[a, b]$, where $a \neq b$ =

P = set of all polynomials

P_n = set of all polynomials of degree $\leq n$ (together with the zero polynomial)

M_{mn} = set of all $m \times n$ matrices

M_{nn} = set of all $n \times n$ square matrices

$F(-\infty, \infty)$ = set of real-valued functions that are defined at each x in the interval $(-\infty, \infty)$.

A Set That Is Not a Vector Space

Let $V = R^2$ and define addition and scalar multiplication operations as follows: If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, then define

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

$u+v = (u_1+v_1, 2v_2)$

$ku = (0, 2u_2)$

and if k is any real number, then define

$$k\mathbf{u} = (ku_1, 0)$$

For example, if $\mathbf{u} = (2, 4)$, $\mathbf{v} = (-3, 5)$, and $k = 7$, then

$$\mathbf{u} + \mathbf{v} = (2 + (-3), 4 + 5) = (-1, 9)$$

$$k\mathbf{u} = 7\mathbf{u} = (7 \cdot 2, 0) = (14, 0)$$

Axiom 10 fails to hold for certain vectors. For example, if $\mathbf{u} = (u_1, u_2)$ is such that $u_2 \neq 0$, then

$1u = u$ ✓

$$1\mathbf{u} = 1(u_1, u_2) = (1 \cdot u_1, 0) = (u_1, 0) \neq \mathbf{u}$$

Thus, V is not a vector space with the stated operations.

4.2 Subspaces

$W = \bigcirc \subseteq V$

V-space

DEFINITION 1 A subset W of a vector space V is called a ***subspace*** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

THEOREM 4.2.1 *If W is a set of one or more vectors in a vector space V, then W is a subspace of V if and only if the following conditions are satisfied.*

(a) *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* W, *then* $\mathbf{u} + \mathbf{v}$ *is in* W.

(b) *If* k *is a scalar and* $\mathbf{u}$ *is a vector in* W, *then* $k\mathbf{u}$ *is in* W.

$u, v \in W$

$u + v \in W$

$ku \in W$

Subspaces of M_{nn}

$W =$ $D_{3\times 3}$ matr

$V = M_{nn}$

We know that the sum of two symmetric $n \times n$ matrices is symmetric and that a scalar multiple of a symmetric $n \times n$ matrix is symmetric. Thus, the set of symmetric $n \times n$ matrices is closed under addition and scalar multiplication and hence is a subspace of M_{nn}. Similarly, the sets of upper triangular matrices, lower triangular matrices, and diagonal matrices are subspaces of M_{nn}.

$$A = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix} \in D_{3\times 3}$$

(i) $$A + B = \begin{bmatrix} a_{11}+b_{11} & 0 & 0 \\ 0 & a_{22}+b_{22} & 0 \\ 0 & 0 & a_{33}+b_{33} \end{bmatrix} \in D_{3\times 3}$$

(ii) $$kA = \begin{bmatrix} ka_{11} & 0 & 0 \\ 0 & ka_{22} & 0 \\ 0 & 0 & ka_{33} \end{bmatrix} \in D_{3\times 3}$$

So $D_{3\times 3} \subseteq M_{n\times n}$

A Subset of M_{nn} That Is Not a Subspace

The set W of invertible $n \times n$ matrices is not a subspace of M_{nn}, failing on two counts—it is not closed under addition and not closed under scalar multiplication. We will illustrate this with an example in M_{22} that you can readily adapt to M_{nn}. Consider the matrices

$$U = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -1 & 2 \\ -2 & 5 \end{bmatrix}$$

$u+v = \begin{bmatrix} 0 & 4 \\ 0 & 10 \end{bmatrix} = 0$

The matrix $0U$ is the 2×2 zero matrix and hence is not invertible, and the matrix $U + V$ has a column of zeros so it also is not invertible.

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

The Subspace $C(-\infty, \infty)$

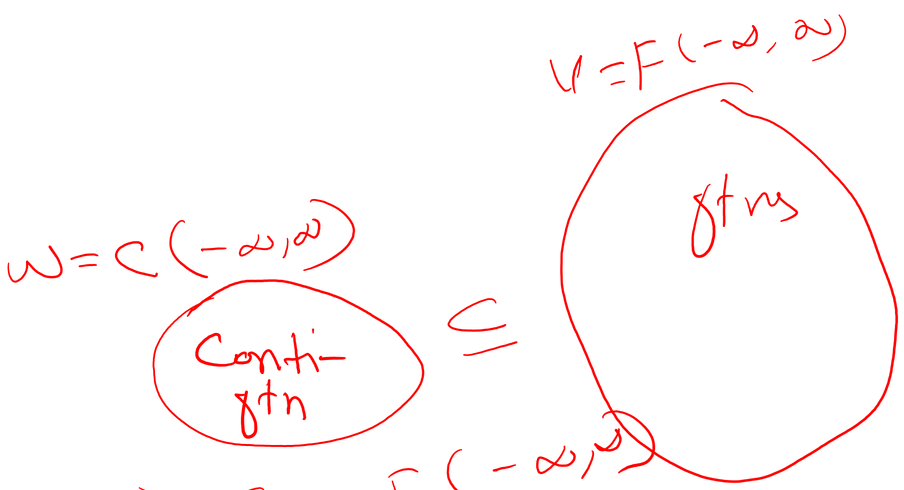
There is a theorem in calculus which states that a sum of continuous functions is continuous and that a constant times a continuous function is continuous. Rephrased in vector language, the set of continuous functions on $(-\infty, \infty)$ is a subspace of $F(-\infty, \infty)$. We will denote this subspace by $C(-\infty, \infty)$.

① Let $f_1(x)$, and $f_2(x) \in W = C(-\infty, \infty)$

$f_1(x) + f_2(x) \in W$

$\Rightarrow$ continuous

② $kf_1(x) \Rightarrow$ continuous



$C(-\infty, \infty) \subseteq F(-\infty, \infty)$

THEOREM 4.2.4 *The solution set of a homogeneous linear system* $\mathbf{Ax} = \mathbf{0}$ *of m equations in n unknowns is a subspace of* R^n.

$S = \{ \quad \}$

$\boxed{x = 2,\ y = -1}$

$s_1 + s_2 \in V$

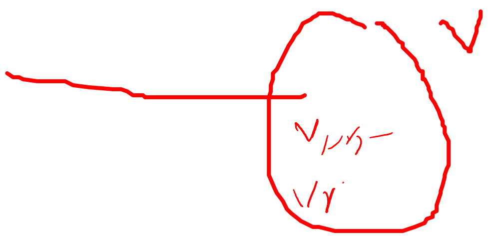
$x + 2y = 0$

s_1, s_2

$k(\ )$

Linear Combination

$w \in V$



DEFINITION 2 If $\mathbf{w}$ is a vector in a vector space V, then $\mathbf{w}$ is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in V if $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \tag{2}$$

where $k_1, k_2, \ldots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

THEOREM 4.2.3 *If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then:*

The set W of all possible linear combinations of the vectors in S is a subspace of V.

Examples of Linear Combinations

a. For the set of vectors in R^3

$$
\begin{array}{c}
\quad\;\; \mathbf{v}_1 \qquad\quad\;\; \mathbf{v}_2 \qquad\qquad \mathbf{v}_3 \\
S = \{(1, 3, 1), (0, 1, 2), (1, 0, -5)\}
\end{array}
$$

$\mathbf{v}_1$ is a linear combination of $\mathbf{v}_2$ and $\mathbf{v}_3$ because

$$\mathbf{v}_1 = 3\mathbf{v}_2 + \mathbf{v}_3 = 3(0, 1, 2) + (1, 0, -5) = (1, 3, 1). = v_1$$

v_1 is l-combinat of v_2 and v_3

$S = \{w_1, w_2, \ldots, w_r\}$ — V

W = all l-combinat

$w \in V$

$S = \{v_1, v_2, \ldots, v_r\}$

$w = k_1v_1 + k_2v_2 + \ldots + k_rv_r$

b. For the set of vectors in M_{22}

$$S = \left\{ \begin{array}{cccc} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ \begin{bmatrix} 0 & 8 \\ 2 & 1 \end{bmatrix}, & \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, & \begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix}, & \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix} \end{array} \right\}$$

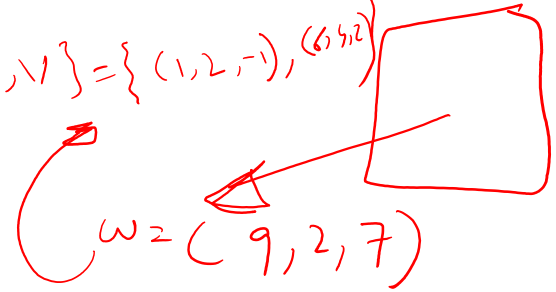
$\mathbf{v}_1$ is a linear combination of $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ because

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{v}_2 + 2\mathbf{v}_3 - \mathbf{v}_4 \\ &= \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix} + 2\begin{bmatrix} -1 & 3 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 8 \\ 2 & 1 \end{bmatrix}. \end{aligned}$$

Problem

Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in R^3. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and that $\mathbf{w}' = (4, -1, 8)$ is *not* a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

$S = \{u, v\} = \{(1,2,-1), (6,4,2)\}$ $\quad V = R^3$



$w = k_1 u + k_2 v$

$(9,2,7) = k_1(1,2,-1) + k_2(6,4,2)$

$= (k_1, 2k_1, -k_1) + (6k_2, 4k_2, 2k_2)$

$(9,2,7) = (k_1 + 6k_2,\ 2k_1 + 4k_2,\ -k_1 + 2k_2)$

$$\begin{aligned} k_1 + 6k_2 &= 9 \\ 2k_1 + 4k_2 &= 2 \\ -k_1 + 2k_2 &= 7 \end{aligned}$$

$$A = \left[\begin{array}{cc|c} 1 & 6 & 9 \\ 2 & 4 & 2 \\ -1 & 2 & 7 \end{array}\right] \underset{R}{\sim} \left[\begin{array}{cc|c} 1 & 6 & 9 \\ 0 & -8 & -16 \\ 0 & 8 & 16 \end{array}\right] \quad \begin{array}{l} R_2 - 2R_1 \\ R_3 + R_1 \end{array}$$

$$\underset{R}{\sim} \left[\begin{array}{cc|c} 1 & 6 & 9 \\ 0 & 1 & 2 \\ 0 & 1 & \end{array}\right]$$

$k_1 + 6k_2 = 9$

$k_1 + 12 = 9$

$k_1 = 9 - 12$

$k_1 = -3$

$$R \quad \begin{bmatrix} 1 & 6 & 9 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{Rank } A = 2 \quad , \quad \text{Rank } \tilde{A} = 2$$

Spanning Set

DEFINITION 3 If $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then the subspace W of V that consists of all possible linear combinations of the vectors in S is called the subspace of V ***generated*** by S, and we say that the vectors $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r$ ***span*** W. We denote this subspace as

$$W = \operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_r\} \quad \text{or} \quad W = \operatorname{span}(S)$$

$S = \{w_1, w_2, \dots w_r\}$

$W = \{\text{all lin.}\} \subset V$

Two Major Problems to Address

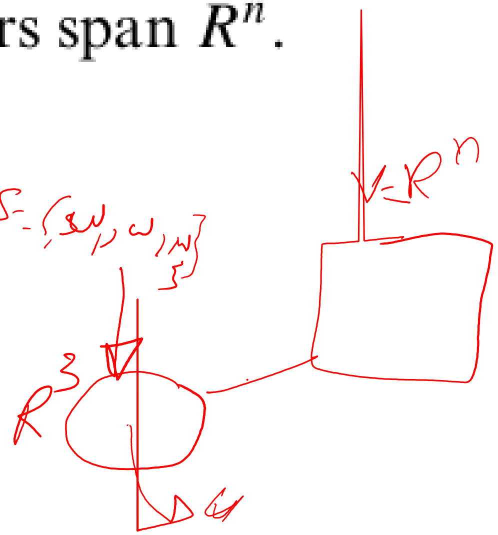
- Given a nonempty set S of vectors in R^n and a vector $\mathbf{v}$ in R^n, determine whether $\mathbf{v}$ is a linear combination of the vectors in S.
- Given a nonempty set S of vectors in R^n, determine whether the vectors span R^n.

Problem

$w_1 \quad w_2 \quad w_3$

Show that the set $S = \{(1, 2, 3), (0, 1, 2), (-2, 0, 1)\}$ spans R^3.

$S = \{w_1, w_2, w_3\}$



SOLUTION

Let $\mathbf{u} = (u_1, u_2, u_3)$ be *any* vector in R^3. Find scalars c_1, c_2, and c_3 such that

$$(u_1, u_2, u_3) = c_1(1, 2, 3) + c_2(0, 1, 2) + c_3(-2, 0, 1)$$
$$= (c_1 - 2c_3, 2c_1 + c_2, 3c_1 + 2c_2 + c_3).$$

$\underline{u} = c_1 w_1 + c_2 w_2 + c_3 w_3$

This vector equation produces the system

$$
\begin{aligned}
c_1 \qquad\qquad\quad - 2c_3 &= u_1 \\
2c_1 + \quad c_2 \qquad\qquad &= u_2 \\
3c_1 + 2c_2 + \quad c_3 &= u_3.
\end{aligned}
$$

The coefficient matrix of this system has a nonzero determinant (verify that it is equal to -1), the system has a unique solution. So, any vector in R^3 can be written as a linear combination of the vectors in S, and you can conclude that the set S spans R^3.

$R^3 = \text{span}(S)$

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

The Standard Unit Vectors Span R^n

$S = \{e_1, e_2, e_3, \dots e_n\}$ R^n

the standard unit vectors in R^n are

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \dots, \quad \mathbf{e}_n = (0, 0, 0, \dots, 1)$$

These vectors span R^n since every vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$ in R^n can be expressed as

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

Let $U = \{u_1, u_2, \dots, u_n\} \in R^n$

$S = \{e_1, \dots, e_n\} \rightarrow R^n$

$$u = c_1 e_1 + c_2 e_2 + \dots + c_n e_n$$

$$(u_1, u_2, u_3, \dots, u_n) = c_1(1, 0, \dots, 0) + c_2(0, 1, 0, \dots, 0) + \dots + c_n(0, 0, \dots, 1)$$

$$= \underbrace{(c_1, 0, \dots, 0) + (0, c_2, \dots, 0)}_{} + \dots + (0, 0, \dots, c_n)$$

$$u_1 = c_1, \quad u_2 = c_2, \quad \dots, \quad u_n = c_n$$

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

$$\Rightarrow R^n = \operatorname{span}(S) = \operatorname{span}(e_1, \dots, e_n).$$

4.3 Linear Independence

DEFINITION 1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a set of two or more vectors in a vector space V, then S is said to be a ***linearly independent set*** if no vector in S can be expressed as a linear combination of the others. A set that is not linearly independent is said to be ***linearly dependent***.

THEOREM 4.3.1 *A nonempty set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ in a vector space V is linearly independent if and only if the only coefficients satisfying the vector equation*

$$\boxed{k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}}$$

are $k_1 = 0, k_2 = 0, \ldots, k_r = 0$.

The most basic linearly independent set in R^n is the set of standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

$$k_1 e_1 + k_2 e_2 + \cdots + k_n e_n = (0, 0, \ldots, 0)$$

$$k_1 (1, 0, 0, \ldots, 0) + k_2 (0, 1, 0, \ldots, 0) + \cdots + k_n (0, 0, \ldots, 1) = (0, 0, \ldots, 0)$$

$$(k_1, 0, 0, \ldots, 0) + (0, k_2, 0, \ldots, 0) + \cdots + (0, 0, \ldots, k_n) = (0, 0, 0, \ldots, 0)$$

$$(k_1, k_2, k_3, \ldots, k_n) = (0, 0, \ldots, 0)$$

$$\boxed{k_1 = 0, \quad k_2 = 0, \quad k_3 = 0, \ldots, \quad k_n = 0}$$

Problem

Determine whether the vectors

$$\mathbf{v}_1 = (1, -2, 3), \quad \mathbf{v}_2 = (5, 6, -1), \quad \mathbf{v}_3 = (3, 2, 1)$$

$S = \{v_1, v_2, v_3\}$

are linearly independent or linearly dependent in R^3.

$k_1 v_1 + k_2 v_2 + k_3 v_3 = 0$

$k_1(1, -2, 3) + k_2(5, 6, -1) + k_3(3, 2, 1) = (0, 0, 0)$

$(k_1, -2k_1, 3k_1) + (5k_2, 6k_2, -k_2) + (3k_3, 2k_3, k_3) = (0, 0, 0)$

$$\left.\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0 \end{aligned}\right] \longrightarrow A$$

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

$k_1, k_2 = ,$

$$\sim \left[\begin{array}{ccc|c} 1 & 5 & 3 & 0 \\ 0 & 16 & 8 & 0 \\ 0 & -16 & -8 & 0 \end{array}\right] \quad \begin{array}{l} R_2 + 2R_1 \\ R_3 - 3R_1 \end{array}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 5 & 3 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & -2 & -1 & 0 \end{array}\right] \quad \begin{array}{l} \frac{1}{8}R_2 \\ \frac{1}{8}R_3 \end{array}$$

$$\sim \left[\begin{array}{ccc|c} 1 & 5 & 3 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad R_3 + R_2$$

$$\boxed{|A| = 0}$$

$k_1 + 5k_2 + 3k_3 = 0 \longrightarrow (1)$

$2k_2 + k_3 = 0 \longrightarrow (11)$

let $\boxed{k_3 = t}$

$\boxed{k_2 = -\frac{1}{2}t}$

$\boxed{k_1 = -\frac{1}{2}t}$

Problem Determine whether the vectors

$$\mathbf{v}_1 = (1, -2, 3), \quad \mathbf{v}_2 = (5, 6, -1), \quad \mathbf{v}_3 = (3, 2, 1)$$

are linearly independent or linearly dependent in R^3.

Solution The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0} \tag{3}$$

can be satisfied with coefficients that are not all zero. To see whether this is so, let us rewrite (3) in the component form

$$k_1(1, -2, 3) + k_2(5, 6, -1) + k_3(3, 2, 1) = (0, 0, 0)$$

Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0 \end{aligned} \tag{4}$$

Thus, our problem reduces to determining whether this system has nontrivial solutions. There are various ways to do this; one possibility is to simply solve the system, which yields

$$k_1 = -\tfrac{1}{2}t, \quad k_2 = -\tfrac{1}{2}t, \quad k_3 = t$$

(we omit the details). This shows that the system has nontrivial solutions and hence that the vectors are linearly dependent.

THEOREM 4.3.2

(a) *A finite set that contains* **0** *is linearly dependent.*

(b) *A set with exactly one vector is linearly independent if and only if that vector is not* **0**.

(c) *A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.*

V_1, V $\quad V_2 = kV_1$ or $V_1 = kV_2$ $\quad$ L.dep.

$V_1 = (0, 1, 2)$, $V_2 = (0, 2, 4) \longrightarrow$ L. dep.

$V_2 = \underline{\underline{2}} V_1$

THEOREM 4.3.3 *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a set of vectors in R^n. If $r > n$, then S is linearly dependent.*

$S = \{ \underset{v_1}{(0,1,0)}, \underset{v_2}{(1,2,3)}, \underset{v_3}{(3,-1,0)}, \underset{v_4}{(5,0,1)} \}$

$r = 4$

$n = 3$

$4 > 3$

$r > n$

▶ EXAMPLE 4 An Important Linearly Independent Set in P_n

Show that the polynomials

$$1, \quad x, \quad x^2, \ldots, \quad x^n$$

form a linearly independent set in P_n.

$a_0, a_1, \ldots, a$

$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = P$

soln:

$p_0 = 1, \; p_1 = x, \; p_2 = x^2, \; \ldots \; p_n = x^n$

$k_0 p_0 + k_1 p_1 + k_2 p_2 + k_3 p_3 + \cdots + k_n p_n = 0$

$\Rightarrow k_0 + k_1 x + k_2 x^2 + \cdots + k_n x^n = 0$

$0x^2 + 0x + 0 = 0$

$\Rightarrow k_0 = k_1 = \cdots k_n = 0$

L. indep.

EXAMPLE 5 Linear Independence of Polynomials

Determine whether the polynomials

$$\mathbf{p}_1 = 1 - x, \quad \mathbf{p}_2 = 5 + 3x - 2x^2, \quad \mathbf{p}_3 = 1 + 3x - x^2$$

are linearly dependent or linearly independent in P_2.

$$k_1 p_1 + k_2 p_2 + k_3 p_3 = 0$$

$$\Rightarrow k_1(1-x) + k_2(5+3x-2x^2) + k_3(1+3x-x^2) = 0$$

$$\Rightarrow k_1 - k_1 x + 5k_2 + 3k_2 x - 2k_2 x^2 + k_3 + 3k_3 x - k_3 x^2 = 0$$

$$\Rightarrow k_1 + 5k_2 + k_3 + (-k_1 + 3k_2 + 3k_3)x + (-2k_2 - k_3)x^2 = 0$$

$$\left.\begin{array}{r} k_1 + 5k_2 + k_3 = 0 \\ -k_1 + 3k_2 + 3k_3 = 0 \\ -2k_2 - k_3 = 0 \end{array}\right\} \Rightarrow A = \begin{bmatrix} 1 & 5 & 1 \\ -1 & 3 & 3 \\ 0 & -2 & -1 \end{bmatrix}$$

$$|A| = 3 - 5(1) + 1(2) = \underline{\underline{0}}$$

singular

L. dep.

Wronskian

DEFINITION 2 If $\mathbf{f}_1 = f_1(x), \mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are functions that are $n-1$ times differentiable on the interval $(-\infty, \infty)$, then the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the ***Wronskian*** of $f_1, f_2, \ldots, f_n$.

EX: $f_1 = 1$, $f_2 = e^x$, $f_3 = e^{2x}$ $\longrightarrow$ $n-1 = 3-1 = 2$

$$W = \begin{vmatrix} f_1 & f_2 & f_3 \\ f_1' & f_2' & f_3' \\ f_1'' & f_2'' & f_3'' \end{vmatrix} = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix}$$

$$= 1 \begin{vmatrix} e^x & 2e^{2x} \\ e^x & 4e^{2x} \end{vmatrix} = 4e^{3x} - 2e^{3x}$$

$$= 2e^{3x} \neq 0$$

$e^x \cdot e^{2x} = e^{x+2x} = e^{3x}$

e^x, $(-\infty, \infty)$

$e^x \neq 0$

$\Rightarrow W \neq 0$ ✓

L. indep

THEOREM 4.3.4 *If the functions* $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ *have* $n-1$ *continuous derivatives on the interval* $(-\infty, \infty)$, *and if the Wronskian of these functions is not identically zero on* $(-\infty, \infty)$, *then these functions form a linearly independent set of vectors in* $C^{(n-1)}(-\infty, \infty)$.

WARNING The converse of Theorem 4.3.4 is false. If the Wronskian of $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ is identically zero on $(-\infty, \infty)$, then no conclusion can be reached about the linear independence of $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$—this set of vectors may be linearly independent or linearly dependent.

Basis for a Vector Space

$S = \{v_1 = (0,1,2),\ v_2, v_3, v_4\}$ R^3

$(4) > n^3$

DEFINITION 1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors in a finite-dimensional vector space V, then S is called a ***basis*** for V if:

(a) S spans V.

(b) S is linearly independent.

$S = \{v_1 \ldots v_n\}$ V

$w = \{\text{All - lin}\}$

$w = k_1 v_1 + k_2 v_2 + \cdots + k_n v_n$ ✓

$0 = k_1 v_1 + k_2 v_2 \cdots + k_n v_n$ ✓

$k_1 = k_2 \cdots k_n = 0$

$v_1 \cdots v_n$

L-indep.

▶ EXAMPLE 3 Another Basis for R^3

Show that the vectors $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 9, 0)$, and $\mathbf{v}_3 = (3, 3, 4)$ form a basis for R^3.

$S = \{v_1, v_2, v_3\}$

$V = R^3$

$S = \{v_1, v_2, v_3\}$

$\omega \leftarrow V$

$\omega = k_1 v_1 + k_2 v_2 + k_3 v_3$

① S span $V = R^3$

let $\omega = (a, b, c) \in \mathbb{R}^3$.

$$\omega = k_1 v_1 + k_2 v_2 + k_3 v_3$$

$$(a, b, c) = k_1(1, 2, 1) + k_2(2, 9, 0) + k_3(3, 3, 4)$$

$$(a, b, c) = (k_1, 2k_1, k_1) + (2k_2, 9k_2, 0) + (3k_3, 3k_3, 4k_3)$$

$$(a, b, c) = (k_1 + 2k_2 + 3k_3,\ 2k_1 + 9k_2 + 3k_3,\ k_1 + 4k_3)$$

$$k_1 + 2k_2 + 3k_3 = a$$
$$2k_1 + 9k_2 + 3k_3 = b$$
$$k_1 + 0k_2 + 4k_3 = c$$
$\longrightarrow (*)$

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 2 & 9 & 3 & b \\ 0 & 0 & 0 & c \end{array}\right]$$

a, b $c = 0$

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

$|A| = -1 \neq 0$

$\Rightarrow S = \{v_1, v_2, v_3\}$ span R^3

& generate R^3

(11) L. Indep.

$$k_1 v_1 + k_2 v_2 + k_3 v_3 = 0$$

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

$= 0$

$= 0$

$= 0$

$|A| \neq 0 \Rightarrow$ Trivial soln. $\Rightarrow$ L. Indep

$$\underline{k_1 = k_2 = k_3 = 0}$$

$S = \{v_1, v_2, v_3\}$ basis for R^3.

The Standard Basis for R^n

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

$V = P_n$

$W = k_1 v_1 + \cdots + k_n v_n$

The Standard Basis for P_n

$$S = \{1, x, x^2, \ldots, x^n\}$$

(1) S span P_n ??

let $p \in P_n$ $\quad p = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$

$\Rightarrow S$ span P_n.

$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$

$a_0, a_1 = \cdots = a_n = 0$

$a_0 = 0$

$x \neq 0$

$a_1 x = 0$

a_n

$\Rightarrow a_1 = 0$

$a_2 x^2 = 0$

$a_0 = a_1 = a_2 = \cdots = a_n = 0$

$\Rightarrow a_2$

L. indep

THEOREM 4.5.1 *All bases for a finite-dimensional vector space have the same number of vectors.*

EXAMPLE 4 The Standard Basis for M_{mn}

Show that the matrices

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis for the vector space M_{22} of 2×2 matrices.

$\boxed{V = M_{2\times 2}}$

Exercise Set 4.4

Dimension

DEFINITION 1 The ***dimension*** of a finite-dimensional vector space V is denoted by $\dim(V)$ and is defined to be the number of vectors in a basis for V. In addition, the zero vector space is defined to have dimension zero.

Dimensions of Some Familiar Vector Spaces

$\dim(R^n) = n$ **[The standard basis has n vectors.]**

$\dim(P_n) = n + 1$ **[The standard basis has $n + 1$ vectors.]**

$\dim(M_{mn}) = mn$ **[The standard basis has mn vectors.]**

THEOREM 4.5.2 *Let V be an n-dimensional vector space, and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis.*

(a) *If a set in V has more than n vectors, then it is linearly dependent.*

(b) *If a set in V has fewer than n vectors, then it does not span V.*

THEOREM 4.5.4 *Let V be an n-dimensional vector space, and let S be a set in V with exactly n vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.*

4.7 Row Space, Column Space, and Null Space

DEFINITION 1 For an $m \times n$ matrix

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

the vectors

$$\begin{aligned} \mathbf{r}_1 &= [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \\ \mathbf{r}_2 &= [a_{21} \quad a_{22} \quad \cdots \quad a_{2n}] \\ &\vdots \\ \mathbf{r}_m &= [a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}] \end{aligned}$$

in R^n that are formed from the rows of A are called the ***row vectors*** of A

vectors

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \quad \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

in R^m formed from the columns of A are called the ***column vectors*** of A.

Let

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

The row vectors of A are

$$\mathbf{r}_1 = [2 \quad 1 \quad 0] \quad \text{and} \quad \mathbf{r}_2 = [3 \quad -1 \quad 4]$$

and the column vectors of A are

$$\mathbf{c}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \text{and} \quad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 4 \end{bmatrix} \quad \blacktriangleleft$$

DEFINITION 2 If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is called the ***row space*** of A, and the subspace of R^m spanned by the column vectors of A is called the ***column space*** of A. The solution space of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the ***null space*** of A.

We will sometimes denote the row space of A, the column space of A, and the null space of A by $\text{row}(A)$, $\text{col}(A)$, and $\text{null}(A)$, respectively.

Question 1. What relationships exist among the solutions of a linear system $A\mathbf{x} = \mathbf{b}$ and the row space, column space, and null space of the coefficient matrix A?

Starting with the first question, suppose that

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

$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n \tag{1}$$

Thus, a linear system, $A\mathbf{x} = \mathbf{b}$, of m equations in n unknowns can be written as

$$x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n = \mathbf{b} \tag{2}$$

from which we conclude that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is expressible as a linear combination of the column vectors of A.

THEOREM 4.7.1 *A system of linear equations* $A\mathbf{x} = \mathbf{b}$ *is consistent if and only if* $\mathbf{b}$ *is in the column space of* A.

Problem

Let $A\mathbf{x} = \mathbf{b}$ be the linear system

$$\begin{bmatrix} -1 & 3 & 2 \\ 1 & 2 & -3 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -9 \\ -3 \end{bmatrix}$$

Show that $\mathbf{b}$ is in the column space of A by expressing it as a linear combination of the column vectors of A.

Solution Solving the system by Gaussian elimination yields (verify)

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

It follows from this and Formula (2) that

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

THEOREM 4.7.5 *If a matrix R is in row echelon form, then the row vectors with the leading 1's (the nonzero row vectors) form a basis for the row space of R, and the column vectors with the leading 1's of the row vectors form a basis for the column space of R.*

Problem Find bases for the row and column spaces of the matrix

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

Solution Since the matrix R is in row echelon form, it follows from Theorem 4.7.5 that the vectors

$$\begin{aligned}
\mathbf{r}_1 &= [1 \quad -2 \quad 5 \quad 0 \quad 3] \\
\mathbf{r}_2 &= [0 \quad 1 \quad 3 \quad 0 \quad 0] \\
\mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 1 \quad 0]
\end{aligned}$$

form a basis for the row space of R, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

THEOREM 4.14

Basis for the Row Space of a Matrix

If an $m \times n$ matrix A is row-equivalent to an $m \times n$ matrix B, then the row space of A is equal to the row space of B.

THEOREM 4.15

Row and Column Spaces Have Equal Dimensions

If a matrix A is row-equivalent to a matrix B in row-echelon form, then the nonzero row vectors of B form a basis for the row space of A.

Problem Find a basis for the row space of the matrix

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

Solution

Reducing A to row echelon form, we obtain

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

$$\begin{aligned}
\mathbf{r}_1 &= [1 \quad -3 \quad 4 \quad -2 \quad 5 \quad 4] \\
\mathbf{r}_2 &= [0 \quad 0 \quad 1 \quad 3 \quad -2 \quad -6] \\
\mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 5]
\end{aligned}$$

form a basis for the row space of A.

THEOREM 4.7.6 *If A and B are row equivalent matrices, then:*

(*a*) *A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.*

(*b*) *A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B form a basis for the column space of B.*

Problem Find a basis for the column space of the matrix

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

Solution

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

is a row echelon form of A.

▶ EXAMPLE 3 Dimension of a Solution Space

Find a basis for and the dimension of the solution space of the homogeneous system

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

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

The corresponding system of equations is

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

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= 0 \end{aligned} \tag{7}$$

If we now assign the free variables x_2, x_4, and x_5 arbitrary values r, s, and t, respectively, then we can express the solution set parametrically as

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = 0$$

Solution In Example 6 of Section 1.2 we found the solution of this system to be

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = 0$$

which can be written in vector form as

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (-3r - 4s - 2t, r, -2s, s, t, 0)$$

or, alternatively, as

$$(x_1, x_2, x_3, x_4, x_5, x_6) = r(-3, 1, 0, 0, 0, 0) + s(-4, 0, -2, 1, 0, 0) + t(-2, 0, 0, 0, 1, 0)$$

This shows that the vectors

$$\mathbf{v}_1 = (-3, 1, 0, 0, 0, 0), \quad \mathbf{v}_2 = (-4, 0, -2, 1, 0, 0), \quad \mathbf{v}_3 = (-2, 0, 0, 0, 1, 0)$$

span the solution space. We leave it for you to check that these vectors are linearly independent by showing that none of them is a linear combination of the other two (but see the remark that follows). Thus, the solution space has dimension 3. ◀

Problem Find the nullspace of the matrix.

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

SOLUTION

The nullspace of A is the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$.

To solve this system, you need to write the augmented matrix $[A \;\vdots\; \mathbf{0}]$ in reduced row-echelon form. Because the system of equations is homogeneous, the right-hand column of the augmented matrix consists entirely of zeros and will not change as you do row operations. It is sufficient to find the reduced row-echelon form of A.

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

The system of equations corresponding to the reduced row-echelon form is

$$\begin{aligned} x_1 + 2x_2 \phantom{{}+x_3} + 3x_4 &= 0 \\ x_3 + x_4 &= 0. \end{aligned}$$

Choose x_2 and x_4 as free variables to represent the solutions in this parametric form.

$$x_1 = -2s - 3t, \quad x_2 = s, \quad x_3 = -t, \quad x_4 = t$$

This means that the solution space of $A\mathbf{x} = \mathbf{0}$ consists of all solution vectors $\mathbf{x}$ of the form shown below.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2s - 3t \\ s \\ -t \\ t \end{bmatrix} = s\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

A basis for the nullspace of A consists of the vectors

$$\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

THEOREM 4.15

Row and Column Spaces Have Equal Dimensions

If A is an $m \times n$ matrix, then the row space and column space of A have the same dimension.

DEFINITION 1 The common dimension of the row space and column space of a matrix A is called the ***rank*** of A and is denoted by $\text{rank}(A)$; the dimension of the null space of A is called the ***nullity*** of A and is denoted by $\text{nullity}(A)$.

THEOREM 4.8.2 Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$rank(A) + nullity(A) = n \tag{4}$$

Problem Find the rank and nullity of the matrix

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

Solution The reduced row echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

dim (A) = 2
dim col = 2

Ran(A) = 2

Quiz -2

Find the rank and nullity of the matrix

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

THEOREM 4.8.2 Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$rank(A) + nullity(A) = n \tag{4}$$

$$x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 = 0$$

$$x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 = 0$$

$$\Rightarrow x_1 = 4x_3 + 28x_4 + 37x_5 - 13x_6$$

$$x_2 = \underline{2x_3} + \underline{12x_4} + \underline{16x_5} - \underline{5x_6}$$

Let $x_6 = s$, $\quad x_3 = u$

$x_5 = t$

$x_4 = r$

$$x_1 = 4u + 28r + 37t - 13s$$

$$x_2 = 2u + 12r + 16t - 5s$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 4u + 28r + 37t - 13s \\ 2u + 12r + 16t - 5s \\ u \\ r \\ t \\ s \end{bmatrix} = u \overset{v_1}{\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + r \overset{v_2}{\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}} + t \overset{v_3}{\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}$$

$$+ s \begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \rightarrow v_4$$

nullity $A = 4$

$\text{Ran}\,A + \text{nullty}(A) = n$

$2 + 4 = 6$

Assignments -2

Ex-4.1 to Ex.4.8

Due date of submission
9/11/2021

CHAPTER 6

Inner Product Spaces

DEFINITION 4 If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then the ***dot product*** (also called the ***Euclidean inner product***) of $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n \tag{17}$$

THEOREM 3.2.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *and if* k *is a scalar, then:*

(*a*) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ **[Symmetry property]**

(*b*) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ **[Distributive property]**

(*c*) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ **[Homogeneity property]**

(*d*) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$ **[Positivity property]**

6.1 Inner Products

DEFINITION 1 An ***inner product*** on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **[Symmetry axiom]**
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ **[Additivity axiom]**
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ **[Homogeneity axiom]**
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ **[Positivity axiom]**

A real vector space with an inner product is called a ***real inner product space***.

A Different Inner Product for R^2

Show that the following function defines an inner product on R^2, where $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$.

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_2 v_2$$

weighted Euclidean inner product

More precisely, if

$$w_1, w_2, \ldots, w_n$$

are *positive* real numbers, which we will call ***weights***, and if $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in R^n, then it can be shown that the formula

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1u_1v_1 + w_2u_2v_2 + \cdots + w_nu_nv_n \tag{2}$$

defines an inner product on R^n that we call the ***weighted Euclidean inner product with weights*** $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$.

▶ EXAMPLE 1 Weighted Euclidean Inner Product

Let $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ be vectors in R^2. Verify that the weighted Euclidean inner product

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2 \tag{3}$$

satisfies the four inner product axioms.

① symmetry Axiom:

$$\langle u, v \rangle = 3u_1v_1 + 2u_2v_2$$

$$= 3v_1u_1 + 2v_2u_2$$

$$= \langle v, u \rangle$$

$= \langle v, u \rangle$

$u_1v_1 = v_1u_1$

$= 3v_1u_1 + 2v_2u_2$

(ii) Additivity axiom: $\langle u+v, w \rangle = 3(u_1+v_1)w_1 + 2(u_2+v_2)w_2$

Let $w = (w_1, w_2) \in \mathbb{R}^2$

$$= 3u_1w_1 + 3v_1w_1 + 2u_2w_2 + 2v_2w_2$$

$$= 3u_1w_1 + 2u_2w_2 + 3v_1w_1 + 2v_2w_2$$

$$= \langle u, w \rangle + \langle v, w \rangle$$

$= \langle u, w \rangle + \langle v, w \rangle$

$u + v = (u_1 + v_1, u_2 + v_2)$

$u_1 \quad u_2$

(iii) $\langle ku, v\rangle = 3(ku_1)v_1 + 2(ku_2)v_2$ $\qquad k\langle u, v\rangle = ???$

$= 3ku_1v_1 + 2ku_2v_2$

$= k(3u_1v_1 + 2u_2v_2)$

$= k\langle u, v\rangle$

(iv) positivity axiom: $\langle v, v\rangle \geq 0$

$\langle v, v\rangle = 3v_1v_1 + 2v_2v_2$

$= 3v_1^2 + 2v_2^2 \geq 0$

$\langle v, v\rangle = 0$ iff $v = 0$

$3v_1^2 + 2v_2^2 = 0$

$\Rightarrow v_1 = v_2 = 0$

A Function That Is Not an Inner Product

Show that the following function is not an inner product on R^3, where $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$.

$$\langle \mathbf{u}, \mathbf{v} \rangle = u_1v_1 - 2u_2v_2 + u_3v_3$$

SOLUTION Observe that Axiom 4 is not satisfied. For example, let $\mathbf{v} = (1, 2, 1)$. Then $\langle \mathbf{v}, \mathbf{v} \rangle = (1)(1) - 2(2)(2) + (1)(1) = -6$, which is less than zero.

$\langle v, v \rangle = v_1 v_1 - 2 v_2 v_2 + v_3 v_3$

$= v_1^2 - 2v_2^2 + v_3^2 \not\geq 0$

$\langle v, v \rangle = 1 - 2(4) + 1$

$= 1 - 8 + 1$

$= 2 - 8 = -6$

$\langle v, v \rangle \geq 0$

An Inner Product on $M_{2,2}$

Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

be matrices in the vector space $M_{2,2}$. The function

$$\langle A, B \rangle = a_{11}b_{11} + a_{21}b_{21} + a_{12}b_{12} + a_{22}b_{22}$$

is an inner product on $M_{2,2}$. The verification of the four inner product axioms is left to you.

$\langle A, B \rangle = A^T B$

$B A^T$

$A^T B \neq B A^T$

$\langle A, B \rangle = a_{11} b_{11} + a_{21} b_{21} + a_{12} b_{12} + a_{22} b_{22}$

$= b_{11} a_{11} + b_{21} a_{21} + b_{12} a_{12} + b_{22} a_{22}$

$\neq \langle B, A \rangle$

$\langle B, A \rangle$

$A_3 = \langle kp, q \rangle = (ka_0)b_0 + (ka_1)b_1 + \text{---} + (ka_n)b_n.$

$= ka_0b_0 + ka_1b_1 + \text{---} + ka_nb_n.$

$= k(\underline{a_0b_0 + \text{---} + a_nb_n})$

$= k\langle p, q \rangle.$ ✓.

$\underline{\underline{A_4}}$ - $\langle p, p \rangle = a_0a_0 + a_1a_1 + \text{---} + a_na_n.$

$= a_0^2 + a_1^2 + \text{---} + a_n^2 \quad \underline{\underline{\geq 0}}.$

$\Rightarrow \langle p, p \rangle \geq \underline{0}$

and $\langle p, p \rangle = 0$ iff

$\Rightarrow a_0^2 + a_1^2 + \text{---} + \underline{\underline{a_n}} = 0$

$\Rightarrow a_0 = 0 = a_1 = \text{----} \underline{\underline{a_n}} = 0.$

$p+q = (a_0+b_0) + (a_1+b_1)x + (\cdots + (a_n+b_n)x^n$

The Standard Inner Product on P_n

If

$$\mathbf{p} = a_0 + a_1 x + \cdots + a_n x^n \quad \text{and} \quad \mathbf{q} = b_0 + b_1 x + \cdots + b_n x^n$$

are polynomials in P_n, then the following formula defines an inner product on P_n that we will call the ***standard inner product*** on this space:

$$\langle \mathbf{p}, \mathbf{q} \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$$

① $\langle p, q \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$

$= b_0 a_0 + b_1 a_1 + \cdots + b_n a_n$

$= \langle q, p \rangle$.

$\langle p, r \rangle + \langle q, r \rangle$

let $r = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n \in P_n$.

A_2: $\langle p+q, r \rangle = (a_0+b_0)c_0 + (a_1+b_1)c_1 + \cdots + (a_n+b_n)c_n$

$= a_0 c_0 + b_0 c_0 + a_1 c_1 + b_1 c_1 + \cdots + a_n c_n + b_n c_n$

$= a_0 c_0 + a_1 c_1 + \cdots + a_n c_n + b_0 c_0 + b_1 c_1 + \cdots + b_n c_n$

$= \langle p, r \rangle + \langle q, r \rangle =$

$A_3 =$

Practice Problems

▶ In Exercises 33–34, let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$. Show that the expression does *not* define an inner product on R^3.

33. $\langle \mathbf{u}, \mathbf{v} \rangle = u_1^2 v_1^2 + u_2^2 v_2^2 + u_3^2 v_3^2$

34. $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - u_2 v_2 + u_3 v_3$

Norm of a vector and Distance b/w two vectors in an Inner Product Space

DEFINITION 2 If V is a real inner product space, then the ***norm*** (or ***length***) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the ***distance*** between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a ***unit vector***.

▶ EXAMPLE 2 Calculating with a Weighted Euclidean Inner Product

It is important to keep in mind that norm and distance depend on the inner product being used. If the inner product is changed, then the norms and distances between vectors also change. For example, for the vectors $\mathbf{u} = (1, 0)$ and $\mathbf{v} = (0, 1)$ in R^2 with the Euclidean inner product we have

$$\|\mathbf{u}\| = \sqrt{1^2 + 0^2} = 1$$

and

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \|(1, -1)\| = \sqrt{1^2 + (-1)^2} = \sqrt{2}$$

but if we change to the weighted Euclidean inner product

$$\boxed{\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2}$$

we have

$$\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{1/2} = [3(1)(1) + 2(0)(0)]^{1/2} = \sqrt{3}$$

$$\|u - v\| = \|(1,0) - (0,1)\|$$
$$= \|(1,-1)\|$$

$$v = (v_1, v_2)$$

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

$$\langle u, u \rangle = 3u_1u_1 + 2u_2u_2$$
$$= 3(1)(1) + 2(0)(0)$$

and

$$
\begin{aligned}
d(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \langle (1, -1), (1, -1) \rangle^{1/2} \\
&= [3(1)(1) + 2(-1)(-1)]^{1/2} = \sqrt{5}
\end{aligned}
$$

6.2 Angle and Orthogonality in Inner Product Spaces

Angle Between Vectors

$$\theta = \cos^{-1}\left(\frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)$$

Let M_{22} have the standard inner product. Find the cosine of the angle between the vectors

$$\mathbf{u} = U = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = V = \begin{bmatrix} -1 & 0 \\ 3 & 2 \end{bmatrix}$$

Solution We showed in Example 6 of the previous section that

$$\langle \mathbf{u}, \mathbf{v} \rangle = 16, \quad \|\mathbf{u}\| = \sqrt{30}, \quad \|\mathbf{v}\| = \sqrt{14}$$

from which it follows that

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \frac{16}{\sqrt{30}\sqrt{14}} \approx 0.78 \quad \blacktriangleleft$$

Orthogonality

DEFINITION 1 Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space V called ***orthogonal*** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Orthogonality Depends on the Inner Product

The vectors $\mathbf{u} = (1, 1)$ and $\mathbf{v} = (1, -1)$ are orthogonal with respect to the Euclidean inner product on R^2 since

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

However, they are not orthogonal with respect to the weighted Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = 3u_1v_1 + 2u_2v_2$ since

$$\langle \mathbf{u}, \mathbf{v} \rangle = 3(1)(1) + 2(1)(-1) = 1 \neq 0$$

Orthogonal and Orthonormal Sets

DEFINITION 1 A set of two or more vectors in a real inner product space is said to be ***orthogonal*** if all pairs of distinct vectors in the set are orthogonal. An orthogonal set in which each vector has norm 1 is said to be ***orthonormal***.

EXAMPLE 1 An Orthogonal Set in R^3

Let

$$\mathbf{v}_1 = (0, 1, 0), \quad \mathbf{v}_2 = (1, 0, 1), \quad \mathbf{v}_3 = (1, 0, -1)$$

and assume that R^3 has the Euclidean inner product. It follows that the set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthogonal since $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathbf{v}_3 \rangle = \langle \mathbf{v}_2, \mathbf{v}_3 \rangle = 0$. ◀

Technique for making an orthogonal set an orthonormal Set

It frequently happens that one has found a set of orthogonal vectors in an inner product space but what is actually needed is a set of *orthonormal* vectors. A simple way to convert an orthogonal set of nonzero vectors into an orthonormal set is to multiply each vector $\mathbf{v}$ in the orthogonal set by the reciprocal of its length to create a vector of norm 1 (called a ***unit vector***). To see why this works, suppose that $\mathbf{v}$ is a nonzero vector in an inner product space, and let

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} \tag{1}$$

Then it follows from Theorem 6.1.1(b) with $k = \|\mathbf{v}\|$ that

$$\|\mathbf{u}\| = \left\| \frac{1}{\|\mathbf{v}\|}\mathbf{v} \right\| = \left| \frac{1}{\|\mathbf{v}\|} \right| \|\mathbf{v}\| = \frac{1}{\|\mathbf{v}\|}\|\mathbf{v}\| = 1$$

This process of multiplying a vector $\mathbf{v}$ by the reciprocal of its length is called ***normalizing*** $\mathbf{v}$.

$$\|\mathbf{v}_1\| = 1, \quad \|\mathbf{v}_2\| = \sqrt{2}, \quad \|\mathbf{v}_3\| = \sqrt{2}$$

Consequently, normalizing $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ yields

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = (0, 1, 0), \quad \mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right),$$

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

Problem

Do there exist scalars k and l such that the vectors

$$\mathbf{p}_1 = 2 + kx + 6x^2, \quad \mathbf{p}_2 = l + 5x + 3x^2, \quad \mathbf{p}_3 = 1 + 2x + 3x^2$$

are mutually orthogonal with respect to the standard inner product on P_2?

THEOREM 6.3.1 *If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space, then S is linearly independent.*

Simply speaking, orthogonality implies linear independency

$$\mathbf{u}_1 = (0, 1, 0), \quad \mathbf{u}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \quad \text{and} \quad \mathbf{u}_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$

form an orthonormal set with respect to the Euclidean inner product on R^3.

these vectors form a linearly independent set, and since R^3 is three-dimensional, $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for R^3.

The Gram–Schmidt Process

To convert a basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$, perform the following computations:

Step 1. $\mathbf{v}_1 = \mathbf{u}_1$

Step 2. $\mathbf{v}_2 = \mathbf{u}_2 - \dfrac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$

Step 3. $\mathbf{v}_3 = \mathbf{u}_3 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$

Step 4. $\mathbf{v}_4 = \mathbf{u}_4 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 - \dfrac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2}\mathbf{v}_3$

$$\vdots$$

(continue for r steps)

Optional Step. To convert the orthogonal basis into an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_r\}$, normalize the orthogonal basis vectors.

Problem

Assume that the vector space R^3 has the Euclidean inner product. Apply the Gram–Schmidt process to transform the basis vectors

$$\mathbf{u}_1 = (1, 1, 1), \quad \mathbf{u}_2 = (0, 1, 1), \quad \mathbf{u}_3 = (0, 0, 1)$$

into an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$.

Solution

Step 1. $\mathbf{v}_1 = \mathbf{u}_1 = (1, 1, 1)$

Step 2. $$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = (0, 1, 1) - \frac{2}{3}(1, 1, 1) = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

Step 3. $$\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$$

$$= (0, 0, 1) - \frac{1}{3}(1, 1, 1) - \frac{1/3}{2/3}\left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right)$$

$$= \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

Thus,

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = \left(-\frac{2}{3}, \frac{1}{3}, \frac{1}{3}\right), \quad \mathbf{v}_3 = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

form an orthogonal basis for R^3. The norms of these vectors are

$$\|\mathbf{v}_1\| = \sqrt{3}, \quad \|\mathbf{v}_2\| = \frac{\sqrt{6}}{3}, \quad \|\mathbf{v}_3\| = \frac{1}{\sqrt{2}}$$

so an orthonormal basis for R^3 is

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \quad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \left(-\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right),$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Practice Problem

Let $\mathbb{R}^3$ has the inner product defined by $< \boldsymbol{u}, \boldsymbol{v} >= u_1 v_1 + 2u_2 v_2 + 3u_3 v_3$. Use the Gram-Schmidt process to transform the basis vectors $\boldsymbol{u_1} = (1, 1, 1)$, $\boldsymbol{u_2} = (1, 1, 0)$, $\boldsymbol{u_3} = (1, 0,\ 0)$ into an orthogonal basis $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}\}$, and then normalize the orthogonal basis vectors to obtain an orthonormal basis $\{\boldsymbol{w_1}, \boldsymbol{w_2}, \boldsymbol{w_3}\}$.

▶ In Exercises 29–30, let R^3 have the Euclidean inner product and use the Gram–Schmidt process to transform the basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ into an orthonormal basis. ◀

29. $\mathbf{u}_1 = (1, 1, 1)$, $\mathbf{u}_2 = (-1, 1, 0)$, $\mathbf{u}_3 = (1, 2, 1)$

30. $\mathbf{u}_1 = (1, 0, 0)$, $\mathbf{u}_2 = (3, 7, -2)$, $\mathbf{u}_3 = (0, 4, 1)$

Orthogonal Complements

DEFINITION 2 If W is a subspace of a real inner product space V, then the set of all vectors in V that are orthogonal to every vector in W is called the ***orthogonal complement*** of W and is denoted by the symbol $W^{\perp}$.

THEOREM 6.2.4 *If W is a subspace of a real inner product space V, then:*

(*a*) $W^{\perp}$ *is a subspace of* V.

(*b*) $W \cap W^{\perp} = \{\mathbf{0}\}$.

▶ EXAMPLE 6 Basis for an Orthogonal Complement

Let W be the subspace of R^6 spanned by the vectors

$$\mathbf{w}_1 = (1, 3, -2, 0, 2, 0), \qquad \mathbf{w}_2 = (2, 6, -5, -2, 4, -3),$$
$$\mathbf{w}_3 = (0, 0, 5, 10, 0, 15), \qquad \mathbf{w}_4 = (2, 6, 0, 8, 4, 18)$$

Find a basis for the orthogonal complement of W.

Solution The subspace W is the same as the row space of the matrix

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

Since the row space and null space of A are orthogonal complements, our problem reduces to finding a basis for the null space of this matrix. In Example 4 of Section 4.7 we showed that

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for this null space. Expressing these vectors in comma-delimited form (to match that of $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, and $\mathbf{w}_4$), we obtain the basis vectors

$$\mathbf{v}_1 = (-3, 1, 0, 0, 0, 0), \quad \mathbf{v}_2 = (-4, 0, -2, 1, 0, 0), \quad \mathbf{v}_3 = (-2, 0, 0, 0, 1, 0)$$

You may want to check that these vectors are orthogonal to $\mathbf{w}_1$, $\mathbf{w}_2$, $\mathbf{w}_3$, and $\mathbf{w}_4$ by computing the necessary dot products. ◀

See slide on page 176

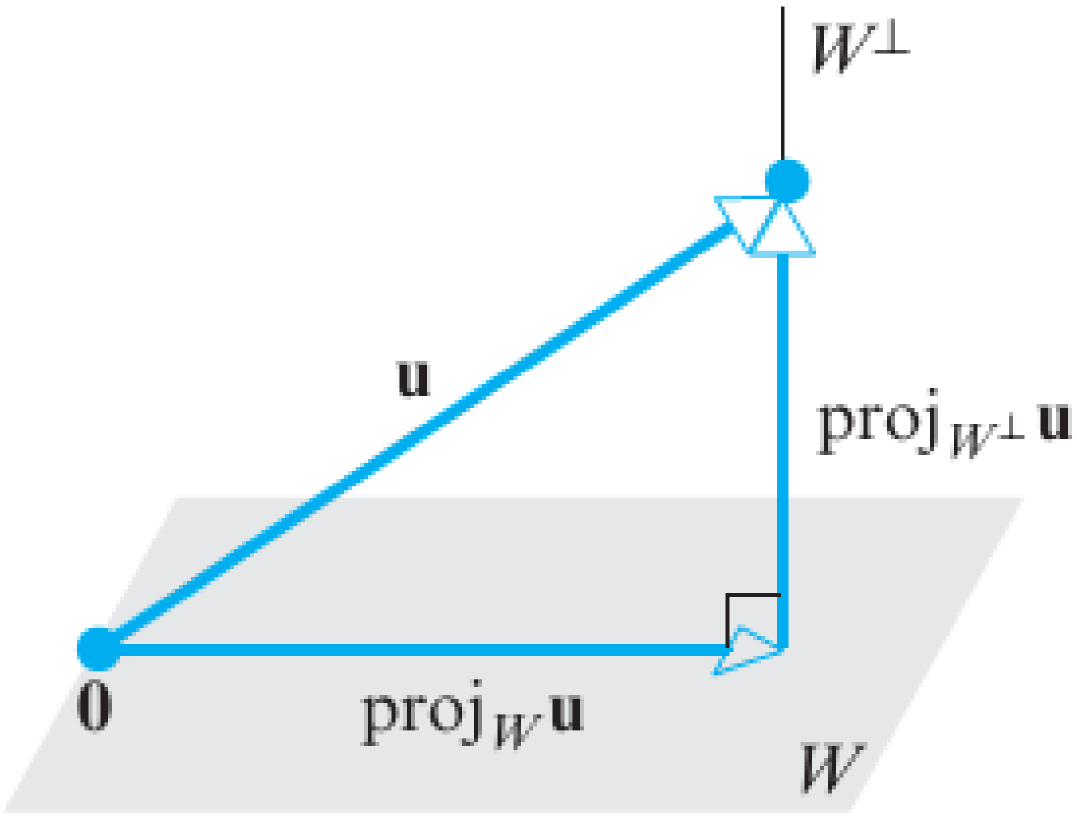
Orthogonal Projections

THEOREM 6.3.3 Projection Theorem

If W is a finite-dimensional subspace of an inner product space V, then every vector $\mathbf{u}$ in V can be expressed in exactly one way as

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 \tag{8}$$

where $\mathbf{w}_1$ is in W and $\mathbf{w}_2$ is in $W^\perp$.



$W^{\perp}$
$\mathbf{u}$
$\text{proj}_{W^{\perp}}\mathbf{u}$
$\mathbf{0}$
$\text{proj}_{W}\mathbf{u}$
W

▲ Figure 6.3.1

The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in Formula (8) are commonly denoted by

$$\mathbf{w}_1 = \text{proj}_W \mathbf{u} \quad \text{and} \quad \mathbf{w}_2 = \text{proj}_{W^\perp} \mathbf{u} \tag{9}$$

These are called the ***orthogonal projection of* u *on W*** and the ***orthogonal projection of* u *on* $W^\perp$**, respectively. The vector $\mathbf{w}_2$ is also called the ***component of* u *orthogonal to W***. Using the notation in (9), Formula (8) can be expressed as

$$\mathbf{u} = \text{proj}_W \mathbf{u} + \text{proj}_{W^\perp} \mathbf{u} \tag{10}$$

6.4 Best Approximation; Least Squares

There are many applications in which some linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns should be consistent on physical grounds but fails to be so because of measurement errors in the entries of A or $\mathbf{b}$. In such cases one looks for vectors that come as close as possible to being solutions in the sense that they minimize $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. In this section we will discuss methods for finding such minimizing vectors.

Least Squares Solutions of Linear Systems

Least Squares Problem Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, find a vector $\mathbf{x}$ in R^n that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ with respect to the Euclidean inner product on R^m. We call such a vector, if it exists, a ***least squares solution*** of $A\mathbf{x} = \mathbf{b}$, we call $\mathbf{b} - A\mathbf{x}$ the ***least squares error vector***, and we call $\|\mathbf{b} - A\mathbf{x}\|$ the ***least squares error***.

To explain the terminology in this problem, suppose that the column form of $\mathbf{b} - A\mathbf{x}$ is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$

The term "least squares solution" results from the fact that minimizing $\|\mathbf{b} - A\mathbf{x}\|$ also has the effect of minimizing $\|\mathbf{b} - A\mathbf{x}\|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$.

What is important to keep in mind about the least squares problem is that for every vector $\mathbf{x}$ in R^n, the product $A\mathbf{x}$ is in the column space of A because it is a linear combination of the column vectors of A. That being the case, to find a least squares solution of $A\mathbf{x} = \mathbf{b}$ is equivalent to finding a vector $A\hat{\mathbf{x}}$ in the column space of A that is closest to $\mathbf{b}$ in the sense that it minimizes the length of the vector $\mathbf{b} - A\mathbf{x}$.

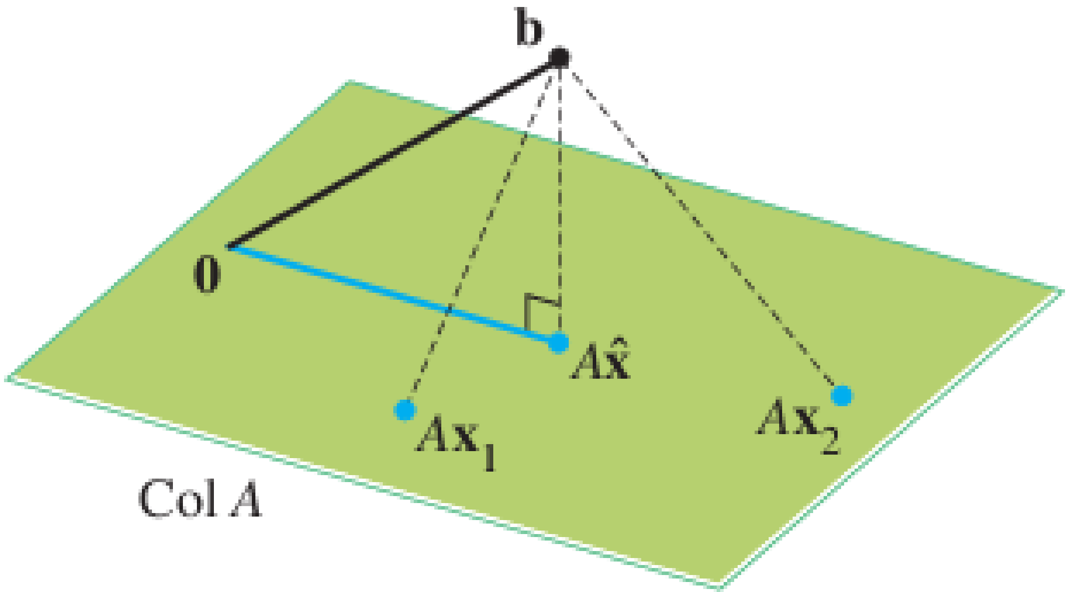


FIGURE 1 The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

THEOREM 6.4.1 Best Approximation Theorem

*If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V, then $\mathrm{proj}_W\,\mathbf{b}$ is the **best approximation** to $\mathbf{b}$ from W in the sense that*

$$\|\mathbf{b} - \mathrm{proj}_W\,\mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ in W that is different from $\mathrm{proj}_W\,\mathbf{b}$.

Finding Least Squares Solutions

THEOREM 6.4.2 *For every linear system $A\mathbf{x} = \mathbf{b}$, the associated normal system*

$$A^T A\mathbf{x} = A^T\mathbf{b} \tag{5}$$

is consistent, and all solutions of (5) are least squares solutions of $A\mathbf{x} = \mathbf{b}$. Moreover, if W is the column space of A, and $\mathbf{x}$ is any least squares solution of $A\mathbf{x} = \mathbf{b}$, then the orthogonal projection of $\mathbf{b}$ on W is

$$\text{proj}_W \mathbf{b} = A\mathbf{x} \tag{6}$$

Problem

Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Also find least squares error vector and least square error.

Sol:

$$A^T A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TA\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Row operations can be used to solve this system, but since A^TA is invertible and 2×2, it is probably faster to compute

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^TA\mathbf{x} = A^T\mathbf{b}$ as

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{and} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

Hence

$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

and

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

Problem

Find the least squares solution, the least squares error vector, and the least squares error of the linear system

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

Solution It will be convenient to express the system in the matrix form $A\mathbf{x} = \mathbf{b}$, where

$$
A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \tag{7}
$$

It follows that

$$A^T A = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$$

$$A^T \mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

so the normal system $A^T A\mathbf{x} = A^T \mathbf{b}$ is

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Solving this system yields a unique least squares solution, namely,

$$x_1 = \frac{17}{95}, \quad x_2 = \frac{143}{285}$$

The least squares error vector is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{95}{57} \end{bmatrix} = \begin{bmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{bmatrix}$$

and the least squares error is

$$\|\mathbf{b} - A\mathbf{x}\| \approx 4.556 \quad ◀$$

Note

If a linear system is consistent, then its exact solutions are the same as its least squares solutions, in which case the least squares error is zero.

Practice Problem

Find the least squares solutions, the least squares error vector, and the least squares error of the linear system

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 2 \\ x_1 - 4x_2 + 3x_3 &= -2 \\ x_1 + 10x_2 - 7x_3 &= 1 \end{aligned}$$

Orthogonal Matrices

DEFINITION 1 A square matrix A is said to be ***orthogonal*** if its transpose is the same as its inverse, that is, if

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

or, equivalently, if

$$AA^T = A^TA = I \tag{1}$$

The matrix

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

is orthogonal since

$$A^T A = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

THEOREM 7.1.2

(a) *The transpose of an orthogonal matrix is orthogonal.*

(b) *The inverse of an orthogonal matrix is orthogonal.*

(c) *A product of orthogonal matrices is orthogonal.*

(d) *If A is orthogonal, then* $\det(A) = 1$ *or* $\det(A) = -1$.

Coordinates Relative to a Basis

THEOREM 4.4.1 Uniqueness of Basis Representation

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, then every vector $\mathbf{v}$ in V can be expressed in the form $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ in exactly one way.

DEFINITION 2 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, and

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \ldots, c_n$ are called the ***coordinates*** of $\mathbf{v}$ relative to the basis S. The vector $(c_1, c_2, \ldots, c_n)$ in R^n constructed from these coordinates is called the ***coordinate vector of*** $\mathbf{v}$ ***relative to*** $\boldsymbol{S}$; it is denoted by

$$(\mathbf{v})_S = (c_1, c_2, \ldots, c_n) \tag{6}$$

Coordinates Relative to the Standard Basis for R^n

In the special case where $V = R^n$ and S is the *standard basis*, the coordinate vector $(\mathbf{v})_S$ and the vector $\mathbf{v}$ are the same; that is,

$$\mathbf{v} = (\mathbf{v})_S$$

For example, in R^3 the representation of a vector $\mathbf{v} = (a, b, c)$ as a linear combination of the vectors in the standard basis $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is

$$\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

so the coordinate vector relative to this basis is $(\mathbf{v})_S = (a, b, c)$, which is the same as the vector $\mathbf{v}$.

Coordinate Vectors Relative to Standard Bases

Problem

(a) Find the coordinate vector for the polynomial

$$\mathbf{p}(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$$

relative to the standard basis for the vector space P_n.

(b) Find the coordinate vector of

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

relative to the standard basis for M_{22}.

Solution (a) The given formula for $\mathbf{p}(x)$ expresses this polynomial as a linear combination of the standard basis vectors $S = \{1, x, x^2, \ldots, x^n\}$. Thus, the coordinate vector for $\mathbf{p}$ relative to S is

$$(\mathbf{p})_S = (c_0, c_1, c_2, \ldots, c_n)$$

Solution (b)

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

so the coordinate vector of B relative to S is

$$(B)_S = (a, b, c, d)$$

Problem

(a) the vectors

$$\mathbf{v}_1 = (1, 2, 1), \quad \mathbf{v}_2 = (2, 9, 0), \quad \mathbf{v}_3 = (3, 3, 4)$$

form a basis for R^3. Find the coordinate vector of $\mathbf{v} = (5, -1, 9)$ relative to the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(b) Find the vector $\mathbf{v}$ in R^3 whose coordinate vector relative to S is $(\mathbf{v})_S = (-1, 3, 2)$.

Solution (a) To find $(\mathbf{v})_S$ we must first express $\mathbf{v}$ as a linear combination of the vectors in S; that is, we must find values of c_1, c_2, and c_3 such that

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

or, in terms of components,

$$(5, -1, 9) = c_1(1, 2, 1) + c_2(2, 9, 0) + c_3(3, 3, 4)$$

Equating corresponding components gives

$$\begin{aligned} c_1 + 2c_2 + 3c_3 &= 5 \\ 2c_1 + 9c_2 + 3c_3 &= -1 \\ c_1 \qquad\quad + 4c_3 &= 9 \end{aligned}$$

Solving this system we obtain $c_1 = 1, c_2 = -1, c_3 = 2$

Therefore,

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

Solution (b) Using the definition of $(\mathbf{v})_S$, we obtain

$$\begin{aligned} \mathbf{v} &= (-1)\mathbf{v}_1 + 3\mathbf{v}_2 + 2\mathbf{v}_3 \\ &= (-1)(1, 2, 1) + 3(2, 9, 0) + 2(3, 3, 4) = (11, 31, 7) \end{aligned}$$

4.6 Change of Basis

A basis that is suitable for one problem may not be suitable for another, so it is a common process in the study of vector spaces to change from one basis to another. Because a basis is the vector space generalization of a coordinate system, changing bases is akin to changing coordinate axes in R^2 and R^3. In this section we will study problems related to changing bases.

In this section we will find it convenient to express coordinate vectors in the matrix form

$$[\mathbf{v}]_S = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

where the square brackets emphasize the matrix notation

There are many applications in which it is necessary to work with more than one coordinate system. In such cases it becomes important to know how the coordinates of a fixed vector relative to each coordinate system are related. This leads to the following problem.

The Change-of-Basis Problem If $\mathbf{v}$ is a vector in a finite-dimensional vector space V, and if we change the basis for V from a basis B to a basis B', how are the coordinate vectors $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ related?

Remark To solve this problem, it will be convenient to refer to B as the "old basis" and B' as the "new basis." Thus, our objective is to find a relationship between the old and new coordinates of a fixed vector $\mathbf{v}$ in V.

For simplicity, we will solve this problem for two-dimensional spaces. The solution for n-dimensional spaces is similar. Let

$$B = \{\mathbf{u}_1, \mathbf{u}_2\} \quad \text{and} \quad B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$$

be the old and new bases, respectively. We will need the coordinate vectors for the new basis vectors relative to the old basis. Suppose they are

$$[\mathbf{u}'_1]_B = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad [\mathbf{u}'_2]_B = \begin{bmatrix} c \\ d \end{bmatrix} \tag{3}$$

That is,

$$\begin{aligned} \mathbf{u}'_1 &= a\mathbf{u}_1 + b\mathbf{u}_2 \\ \mathbf{u}'_2 &= c\mathbf{u}_1 + d\mathbf{u}_2 \end{aligned} \tag{4}$$

Now let $\mathbf{v}$ be any vector in V, and let

$$[\mathbf{v}]_{B'} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \tag{5}$$

be the new coordinate vector, so that

$$\mathbf{v} = k_1\mathbf{u}'_1 + k_2\mathbf{u}'_2 \tag{6}$$

In order to find the old coordinates of $\mathbf{v}$, we must express $\mathbf{v}$ in terms of the old basis B. To do this, we substitute (4) into (6). This yields

$$\mathbf{v} = k_1(a\mathbf{u}_1 + b\mathbf{u}_2) + k_2(c\mathbf{u}_1 + d\mathbf{u}_2)$$

or

$$\mathbf{v} = (k_1a + k_2c)\mathbf{u}_1 + (k_1b + k_2d)\mathbf{u}_2$$

Thus, the old coordinate vector for $\mathbf{v}$ is

$$[\mathbf{v}]_B = \begin{bmatrix} k_1a + k_2c \\ k_1b + k_2d \end{bmatrix}$$

which, by using (5), can be written as

$$[\mathbf{v}]_B = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} [\mathbf{v}]_{B'}$$

This equation states that the old coordinate vector $[\mathbf{v}]_B$ results when we multiply the new coordinate vector $[\mathbf{v}]_{B'}$ on the left by the matrix

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

Since the columns of this matrix are the coordinates of the new basis vectors relative to the old basis [see (3)], we have the following solution of the change-of-basis problem.

Solution of the Change-of-Basis Problem If we change the basis for a vector space V from an old basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ to a new basis $B' = \{\mathbf{u}'_1, \mathbf{u}'_2, \ldots, \mathbf{u}'_n\}$, then for each vector $\mathbf{v}$ in V, the old coordinate vector $[\mathbf{v}]_B$ is related to the new coordinate vector $[\mathbf{v}]_{B'}$ by the equation

$$[\mathbf{v}]_B = P[\mathbf{v}]_{B'} \tag{7}$$

where the columns of P are the coordinate vectors of the new basis vectors relative to the old basis; that is, the column vectors of P are

$$[\mathbf{u}'_1]_B, \quad [\mathbf{u}'_2]_B, \ldots, \quad [\mathbf{u}'_n]_B \tag{8}$$

Transition Matrices

The matrix P in Equation (7) is called the ***transition matrix*** from B' to B. For emphasis, we will often denote it by $P_{B' \to B}$. It follows from (8) that this matrix can be expressed in terms of its column vectors as

$$P_{B' \to B} = \left[[\mathbf{u}'_1]_B \mid [\mathbf{u}'_2]_B \mid \cdots \mid [\mathbf{u}'_n]_B \right] \tag{9}$$

Similarly, the transition matrix from B to B' can be expressed in terms of its column vectors as

$$P_{B \to B'} = \left[[\mathbf{u}_1]_{B'} \mid [\mathbf{u}_2]_{B'} \mid \cdots \mid [\mathbf{u}_n]_{B'} \right] \tag{10}$$

Remark There is a simple way to remember both of these formulas using the terms "old basis" and "new basis" defined earlier in this section: In Formula (9) the old basis is B' and the new basis is B, whereas in Formula (10) the old basis is B and the new basis is B'. Thus, both formulas can be restated as follows:

In short

The columns of the transition matrix from an old basis to a new basis are the coordinate vectors of the old basis relative to the new basis.

Problem

Consider the bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$ for R^2, where

$$\mathbf{u}_1 = (1, 0), \quad \mathbf{u}_2 = (0, 1), \quad \mathbf{u}'_1 = (1, 1), \quad \mathbf{u}'_2 = (2, 1)$$

(a) Find the transition matrix $P_{B' \to B}$ from B' to B.

(b) Find the transition matrix $P_{B \to B'}$ from B to B'.

Solution (a) Here the old basis vectors are $\mathbf{u}_1'$ and $\mathbf{u}_2'$ and the new basis vectors are $\mathbf{u}_1$ and $\mathbf{u}_2$. We want to find the coordinate matrices of the old basis vectors $\mathbf{u}_1'$ and $\mathbf{u}_2'$ relative to the new basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. To do this, observe that

$$\begin{aligned} \mathbf{u}_1' &= \mathbf{u}_1 + \mathbf{u}_2 \\ \mathbf{u}_2' &= 2\mathbf{u}_1 + \mathbf{u}_2 \end{aligned}$$

from which it follows that

$$[\mathbf{u}_1']_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{u}_2']_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and hence that

$$P_{B' \to B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Solution (*b*) Here the old basis vectors are $\mathbf{u}_1$ and $\mathbf{u}_2$ and the new basis vectors are $\mathbf{u}'_1$ and $\mathbf{u}'_2$.

observe that

$$\begin{aligned} \mathbf{u}_1 &= -\mathbf{u}'_1 + \mathbf{u}'_2 \\ \mathbf{u}_2 &= 2\mathbf{u}'_1 - \mathbf{u}'_2 \end{aligned}$$

from which it follows that

$$[\mathbf{u}_1]_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{u}_2]_{B'} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

and hence that

$$P_{B \to B'} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$$

Suppose now that B and B' are bases for a finite-dimensional vector space V. Since multiplication by $P_{B'\to B}$ maps coordinate vectors relative to the basis B' into coordinate vectors relative to a basis B, and $P_{B\to B'}$ maps coordinate vectors relative to B into coordinate vectors relative to B', it follows that for every vector $\mathbf{v}$ in V we have

$$[\mathbf{v}]_B = P_{B'\to B}[\mathbf{v}]_{B'} \tag{11}$$

$$[\mathbf{v}]_{B'} = P_{B\to B'}[\mathbf{v}]_B \tag{12}$$

Practice Problem

Consider the bases $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3\}$ for R^3, where

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{u}'_1 = \begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix}, \quad \mathbf{u}'_2 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \quad \mathbf{u}'_3 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

(a) Find the transition matrix B to B'.

(b) Compute the coordinate vector $[\mathbf{w}]_B$, where

$$\mathbf{w} = \begin{bmatrix} -5 \\ 8 \\ -5 \end{bmatrix}$$

CHAPTER 5

Eigenvalues and Eigenvectors

DEFINITION 1 If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an ***eigenvector*** of A (or of the matrix operator T_A) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is,

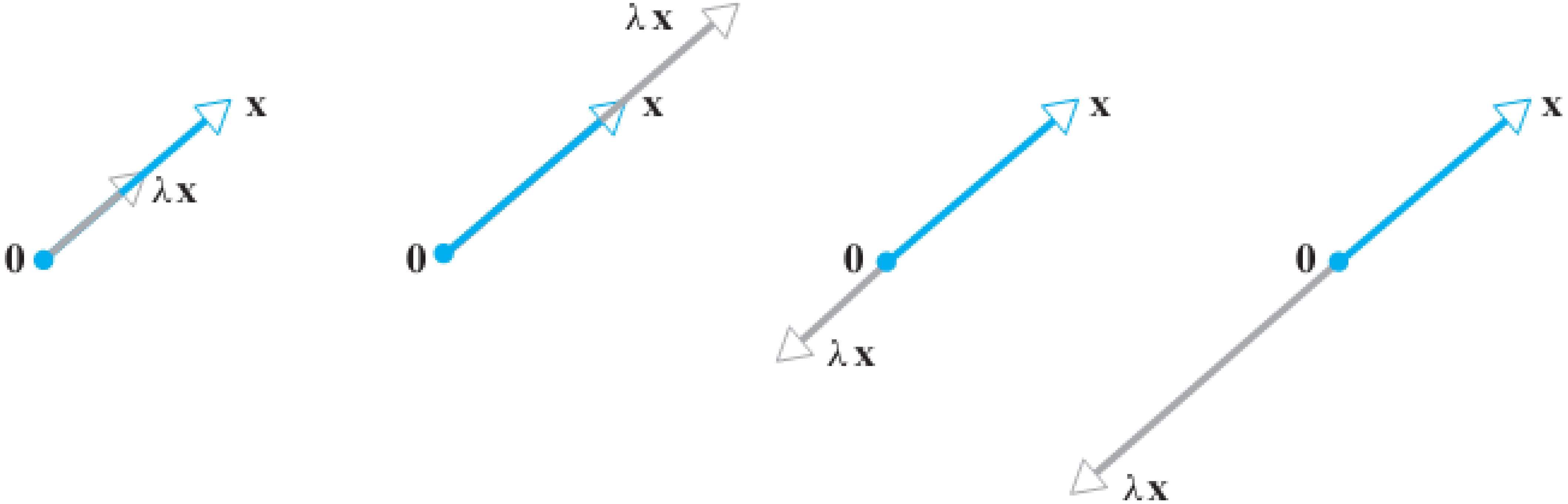
$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar λ. The scalar λ is called an ***eigenvalue*** of A (or of T_A), and $\mathbf{x}$ is said to be an ***eigenvector corresponding to*** $\boldsymbol{\lambda}$.

The requirement that an eigenvector be nonzero is imposed to avoid the unimportant case $A\mathbf{0} = \lambda\mathbf{0}$, which holds for every A and λ.

An eigen*value* of $\lambda = 0$, however, is possible.

In general, the image of a vector $\mathbf{x}$ under multiplication by a square matrix A differs from $\mathbf{x}$ in both magnitude and direction. However, in the special case where $\mathbf{x}$ is an eigenvector of A, multiplication by A leaves the direction unchanged. For example, in R^2 or R^3 multiplication by A maps each eigenvector $\mathbf{x}$ of A (if any) along the same line through the origin as $\mathbf{x}$. Depending on the sign and magnitude of the eigenvalue λ corresponding to $\mathbf{x}$, the operation $A\mathbf{x} = \lambda\mathbf{x}$ compresses or stretches $\mathbf{x}$ by a factor of λ, with a reversal of direction in the case where λ is negative (Figure 5.1.1).



$(a)\ 0 \le \lambda \le 1$ $(b)\ \lambda \ge 1$ $(c)\ -1 \le \lambda \le 0$ $(d)\ \lambda \le -1$

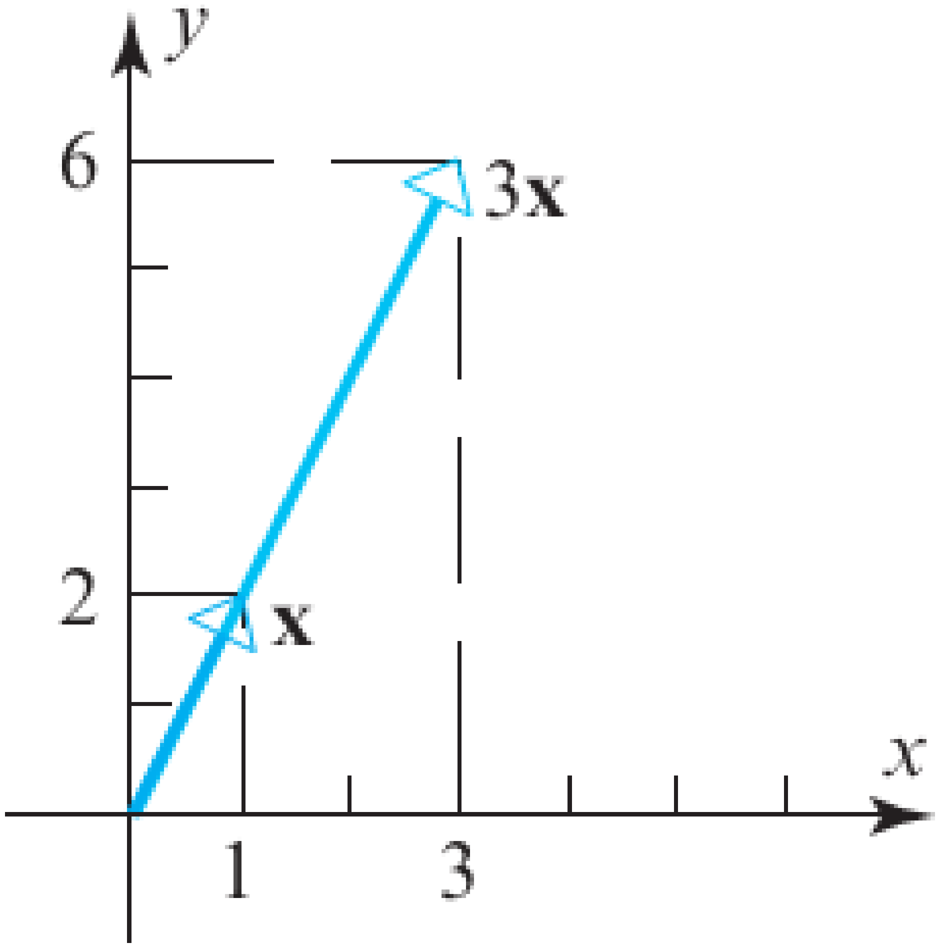
The vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of

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

corresponding to the eigenvalue $\lambda = 3$, since

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$

Geometrically, multiplication by A has stretched the vector $\mathbf{x}$ by a factor of 3


y
6
3x
2
x
x
1
3

Problem

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 $\mathbf{x}_1$ on the left by A produces

$$A\mathbf{x}_1 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \underbrace{2}_{\boxed{\text{Eigenvalue}}}\underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\boxed{\text{Eigenvector}}}.$$

So, $\mathbf{x}_1 = (1, 0)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_1 = 2$.

Similarly, multiplying $\mathbf{x}_2$ on the left by A produces

$$A\mathbf{x}_2 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -1 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So, $\mathbf{x}_2 = (0, 1)$ is an eigenvector of A corresponding to the eigenvalue $\lambda_2 = -1$.

The question is : How to find the eigenvalues and eigenvectors of a given matrix.

To find the eigenvalues and eigenvectors of an $n \times n$ matrix A, let I be the $n \times n$ identity matrix. Rewriting $A\mathbf{x} = \lambda\mathbf{x}$ as $\lambda I\mathbf{x} = A\mathbf{x}$ and rearranging gives $(\lambda I - A)\mathbf{x} = \mathbf{0}$. This homogeneous system of equations has nonzero solutions if and only if the coefficient matrix $(\lambda I - A)$ is *not* invertible—that is, if and only if its determinant is zero. The next theorem formally states this.

THEOREM

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 $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

Note that a square matrix A has *many* eigenvectors associated with any given eigenvalue λ. In fact *every* nonzero solution $\mathbf{x}$ of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ is an eigenvector.

The equation $\det(\lambda I - A) = 0$ is the **characteristic equation** of A. Moreover, when expanded to polynomial form, the polynomial

$$|\lambda I - A| = \lambda^n + c_{n-1}\lambda^{n-1} + \cdot \cdot \cdot + c_2\lambda^2 + c_1\lambda + c_0$$

is the **characteristic polynomial** of A. So, the eigenvalues of an $n \times n$ matrix A correspond to the roots of the characteristic polynomial of A.

Problem

Find the eigenvalues and corresponding eigenvectors of $A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$.

SOLUTION

The characteristic polynomial of A is

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix} = \lambda^2 + 3\lambda - 10 + 12 = (\lambda + 1)(\lambda + 2).$$

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, 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

$$(-2)I - A = \begin{bmatrix} -2-2 & 12 \\ -1 & -2+5 \end{bmatrix} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}.$$

Letting $x_2 = t$, you can conclude that every eigenvector of λ_2 is of the form

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3t \\ t \end{bmatrix} = t\begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad t \neq 0.$$

Problem

Find the eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

Solution The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4$$

The eigenvalues of A must therefore satisfy the cubic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0 \tag{5}$$

To solve this equation, we will begin by searching for integer solutions. This task can be simplified by exploiting the fact that all integer solutions (if there are any) of a polynomial equation with *integer coefficients*

$$\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$

must be divisors of the constant term, c_n.

the divisors of -4, that is, $\pm 1, \pm 2, \pm 4$. Successively substituting these values in (5) shows that $\lambda = 4$ is an integer solution and hence that $\lambda - 4$ is a factor of the left side of (5). Dividing $\lambda - 4$ into $\lambda^3 - 8\lambda^2 + 17\lambda - 4$ shows that (5) can be rewritten as

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0$$

Thus, the remaining solutions of (5) satisfy the quadratic equation

$$\lambda^2 - 4\lambda + 1 = 0$$

which can be solved by the quadratic formula. Thus, the eigenvalues of A are

$$\lambda = 4, \quad \lambda = 2 + \sqrt{3}, \quad \text{and} \quad \lambda = 2 - \sqrt{3}$$

THEOREM 5.1.2 *If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then the eigenvalues of A are the entries on the main diagonal of A.*

the eigenvalues of the lower triangular matrix

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

are $\lambda = \frac{1}{2}$, $\lambda = \frac{2}{3}$, and $\lambda = -\frac{1}{4}$.

Finding Eigenvectors and Bases for Eigenspaces

By definition, the eigenvectors of A corresponding to an eigenvalue λ are the nonzero vectors that satisfy

$$(\lambda I - A)\mathbf{x} = \mathbf{0}$$

Thus, we can find the eigenvectors of A corresponding to λ by finding the nonzero vectors in the solution space of this linear system. This solution space, which is called the ***eigenspace*** of A corresponding to λ, can also be viewed as:

Sated anther way, the eigenspace of A corresponding to the eigenvalues is the solution space of the homogeneous system $(\lambda I - A)\mathbf{x} = \mathbf{0}$

Problem

Find bases for the eigenspaces of the matrix

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

Solution The characteristic equation of A is

$$\begin{vmatrix} \lambda + 1 & -3 \\ -2 & \lambda \end{vmatrix} = \lambda(\lambda + 1) - 6 = (\lambda - 2)(\lambda + 3) = 0$$

so the eigenvalues of A are $\lambda = 2$ and $\lambda = -3$. Thus, there are two eigenspaces of A, one for each eigenvalue.

By definition,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an eigenvector of A corresponding to an eigenvalue λ if and only if $(\lambda I - A)\mathbf{x} = \mathbf{0}$, that is,

$$\begin{bmatrix} \lambda + 1 & -3 \\ -2 & \lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In the case where $\lambda = 2$ this equation becomes

$$\begin{bmatrix} 3 & -3 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

whose general solution is

$$x_1 = t, \quad x_2 = t$$

Since this can be written in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

it follows that

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

is a basis for the eigenspace corresponding to $\lambda = 2$.

Similarly

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

is a basis for the eigenspace corresponding to $\lambda = -3$.

▶ EXAMPLE 7 Eigenvectors and Bases for Eigenspaces

Find bases for the eigenspaces of

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

THEOREM 5.1.4 *A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.*

5.2 Diagonalization

Similar Matrices

DEFINITION 1 If A and B are square matrices, then we say that ***B is similar to A*** if there is an invertible matrix P such that $B = P^{-1}AP$.

Note that if B is similar to A, then it is also true that A is similar to B since we can express A as $A = Q^{-1}BQ$ by taking $Q = P^{-1}$. This being the case, we will usually say that A and B are ***similar matrices*** if either is similar to the other.

$$A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$$ and $$B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ are similar matrices

because there is $$P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$ such that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = B$$

Diagonalizable Matrix

DEFINITION 2 A square matrix A is said to be ***diagonalizable*** if it is similar to some diagonal matrix; that is, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. In this case the matrix P is said to ***diagonalize*** A.

e.g

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

is diagonalizable

A is similar to the diagonal matrix

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

(Detail in previous Slide)

THEOREM 5.2.1 *If A is an $n \times n$ matrix, the following statements are equivalent.*

(*a*) *A is diagonalizable.*

(*b*) *A has n linearly independent eigenvectors.*

A Procedure for Diagonalizing an $n \times n$ Matrix

Step 1. Determine first whether the matrix is actually diagonalizable by searching for n linearly independent eigenvectors. One way to do this is to find a basis for each eigenspace and count the total number of vectors obtained. If there is a total of n vectors, then the matrix is diagonalizable, and if the total is less than n, then it is not.

Step 2. If you ascertained that the matrix is diagonalizable, then form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ whose column vectors are the n basis vectors you obtained in Step 1.

Step 3. $P^{-1}AP$ will be a diagonal matrix whose successive diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ that correspond to the successive columns of P.

Problem

Find a matrix P that diagonalizes

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

Solution The characteristic equation of A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$

Thus, the distinct eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so there are two eigenspaces of A.

By definition,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is an eigenvector of A corresponding to λ if and only if $\mathbf{x}$ is a nontrivial solution of $(\lambda I - A)\mathbf{x} = \mathbf{0}$, or in matrix form,

$$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{6}$$

In the case where $\lambda = 2$, Formula (6) becomes

$$\begin{bmatrix} 2 & 0 & 2 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system using Gaussian elimination yields

$$x_1 = -s, \quad x_2 = t, \quad x_3 = s$$

Thus, the eigenvectors of A corresponding to $\lambda = 2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

If $\lambda = 1$, then (6) becomes

$$\begin{bmatrix} 1 & 0 & 2 \\ -1 & -1 & -1 \\ -1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system yields

$$x_1 = -2s, \quad x_2 = s, \quad x_3 = s$$

Thus, the eigenvectors corresponding to $\lambda = 1$ are the nonzero vectors of the form

$$\begin{bmatrix} -2s \\ s \\ s \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \quad \text{so that} \quad \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 1$.

$$\lambda = 2: \quad \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 1: \quad \mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

There are three basis vectors in total, so the matrix

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

diagonalizes A.

As a check

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Problem Show that the following matrix is not diagonalizable:

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

Solution The characteristic polynomial of A is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & 0 \\ 3 & -5 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$$

so the characteristic equation is

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

and the distinct eigenvalues of A are $\lambda = 1$ and $\lambda = 2$.

bases for the eigenspaces are

$$\lambda = 1: \quad \mathbf{p}_1 = \begin{bmatrix} \frac{1}{8} \\ -\frac{1}{8} \\ 1 \end{bmatrix}; \qquad \lambda = 2: \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Since A is a 3×3 matrix and there are only two basis vectors in total, A is not diagonalizable.

THEOREM 5.2.2

(*a*) *If* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *are distinct eigenvalues of a matrix* A, *and if* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ *are corresponding eigenvectors, then* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *is a linearly independent set.*

(*b*) *An* $n \times n$ *matrix with* n *distinct eigenvalues is diagonalizable.*

Computing Powers of a Matrix

The problem of computing powers of a matrix is greatly simplified when the matrix is diagonalizable. To see why this is so, suppose that A is a diagonalizable $n \times n$ matrix, that P diagonalizes A, and that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = D$$

Squaring both sides of this equation yields

$$(P^{-1}AP)^2 = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} = D^2$$

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P$$

from which we obtain the relationship $P^{-1}A^2P = D^2$.

More generally, if k is a positive integer then a similar computation will show that

$$P^{-1}A^kP = D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$$

which we can rewrite as

$$A^k = PD^kP^{-1} = P\begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}P^{-1}$$

Theorem

If $A = PDP^{-1}$ *then* $A^k = PD^kP^{-1}$ *for each* $k = 1, 2, \ldots.$

Problem

find A^{13}, where $A = \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{bmatrix}$

Solution

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

(Previous Slides)

$$D = P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A^{13} = PD^{13}P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix}$$

7.2 Orthogonal Diagonalization

DEFINITION 1 If A and B are square matrices, then we say that B is ***orthogonally similar*** to A if there is an orthogonal matrix P such that $B = P^TAP$.

Note that if B is orthogonally similar to A, then it is also true that A is orthogonally similar to B since we can express A as $A = Q^TBQ$ by taking $Q = P^T$ (verify). This being the case we will say that A and B are ***orthogonally similar matrices*** if either is orthogonally similar to the other.

If A is orthogonally similar to some diagonal matrix, say

$$P^TAP = D$$

then we say that A is ***orthogonally diagonalizable*** and that P ***orthogonally diagonalizes*** A.

Conditions for Orthogonal Diagonalizability

THEOREM 7.2.1 *If A is an $n \times n$ matrix with real entries, then the following are equivalent.*

(a) *A is orthogonally diagonalizable.*

(b) *A has an orthonormal set of n eigenvectors.*

(c) *A is symmetric.*

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix P whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^TAP$ will be in the same order as their corresponding eigenvectors in P.

Problem Find an orthogonal matrix P that diagonalizes

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

Solution

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 4 & -2 & -2 \\ -2 & \lambda - 4 & -2 \\ -2 & -2 & \lambda - 4 \end{bmatrix} = (\lambda - 2)^2(\lambda - 8) = 0$$

Thus, the distinct eigenvalues of A are $\lambda = 2$ and $\lambda = 8$.

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for the eigenspace corresponding to $\lambda = 2$.

Applying the Gram–Schmidt process to $\{\mathbf{u}_1, \mathbf{u}_2\}$ yields the following orthonormal eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

The eigenspace corresponding to $\lambda = 8$ has

$$\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

as a basis. Applying the Gram–Schmidt process to $\{\mathbf{u}_3\}$ (i.e., normalizing $\mathbf{u}_3$) yields

$$\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Finally, using $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ as column vectors, we obtain

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

which orthogonally diagonalizes A. to confirm that

$$P^TAP = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

CHAPTER 8

General Linear Transformations

Definition of a Linear Transformation

Let V and W be vector spaces. The function

$$T: V \to W$$

is a **linear transformation** of V into W when the two properties below are true for all $\mathbf{u}$ and $\mathbf{v}$ in V and for any scalar c.

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ **[Additivity property]**
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ **[Homogeneity property]**

In the special case where $V = W$, the linear transformation T is called a ***linear operator*** on the vector space V.

Problem

Show that the function

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

is a linear transformation from R^2 into R^2.

SOLUTION

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

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

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

Problem

For any vector $\mathbf{v} = (v_1, v_2)$ in R^2, define $T: R^2 \to R^2$ 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 image of $\mathbf{v} = (0, 0)$.

c. 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 $\mathbf{v} = (0, 0)$, then

$$T(0, 0) = (0 - 0, 0 + 2(0)) = (0, 0).$$

c. 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 $(3, 4)$.

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 example,

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

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 example, $(1 + 2)^2 \neq 1^2 + 2^2$.

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

EXAMPLE 2 The Zero Transformation

Let V and W be any two vector spaces. The mapping $T: V \to W$ such that $T(\mathbf{v}) = \mathbf{0}$ for every $\mathbf{v}$ in V is a linear transformation called the ***zero transformation***. To see that T is linear, observe that

$$T(\mathbf{u} + \mathbf{v}) = \mathbf{0}, \quad T(\mathbf{u}) = \mathbf{0}, \quad T(\mathbf{v}) = \mathbf{0}, \quad \text{and} \quad T(k\mathbf{u}) = \mathbf{0}$$

Therefore,

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{and} \quad T(k\mathbf{u}) = kT(\mathbf{u})$$

EXAMPLE 3 The Identity Operator

Let V be any vector space. The mapping $I: V \to V$ defined by $I(\mathbf{v}) = \mathbf{v}$ is called the ***identity operator*** on V. We will leave it for you to verify that I is linear.

EXAMPLE 5 A Linear Transformation from P_n to P_{n+1}

Let $\mathbf{p} = p(x) = c_0 + c_1x + \cdots + c_nx^n$ be a polynomial in P_n, and define the transformation $T: P_n \to P_{n+1}$ by

$$T(\mathbf{p}) = T(p(x)) = xp(x) = c_0x + c_1x^2 + \cdots + c_nx^{n+1}$$

This transformation is linear because for any scalar k and any polynomials $\mathbf{p}_1$ and $\mathbf{p}_2$ in P_n we have

$$T(k\mathbf{p}) = T(kp(x)) = x(kp(x)) = k(xp(x)) = kT(\mathbf{p})$$

and

$$\begin{aligned} T(\mathbf{p}_1 + \mathbf{p}_2) &= T(p_1(x) + p_2(x)) = x(p_1(x) + p_2(x)) \\ &= xp_1(x) + xp_2(x) = T(\mathbf{p}_1) + T(\mathbf{p}_2) \end{aligned}$$

▶ EXAMPLE 6 A Linear Transformation Using the Dot Product

Let $\mathbf{v}_0$ be any fixed vector in R^n, and let $T: R^n \to R$ be the transformation

$$T(\mathbf{x}) = \langle \mathbf{x} \cdot \mathbf{v}_0 \rangle$$

that maps a vector $\mathbf{x}$ to its dot product with $\mathbf{v}_0$. This transformation is linear, for if k is any scalar, and if $\mathbf{u}$ and $\mathbf{v}$ are any vectors in R^n, then it follows from properties of the dot product in Theorem 3.2.2 that

$$T(k\mathbf{u}) = (k\mathbf{u}) \cdot \mathbf{v}_0 = k(\mathbf{u} \cdot \mathbf{v}_0) = kT(\mathbf{u})$$

$$T(\mathbf{u} + \mathbf{v}) = (\mathbf{u} + \mathbf{v}) \cdot \mathbf{v}_0 = (\mathbf{u} \cdot \mathbf{v}_0) + (\mathbf{v} \cdot \mathbf{v}_0) = T(\mathbf{u}) + T(\mathbf{v})$$

EXAMPLE 10 Computing with Images of Basis Vectors

Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for R^3, where

$$\mathbf{v}_1 = (1, 1, 1), \quad \mathbf{v}_2 = (1, 1, 0), \quad \mathbf{v}_3 = (1, 0, 0)$$

Let $T : R^3 \to R^2$ be the linear transformation for which

$$T(\mathbf{v}_1) = (1, 0), \quad T(\mathbf{v}_2) = (2, -1), \quad T(\mathbf{v}_3) = (4, 3)$$

Find a formula for $T(x_1, x_2, x_3)$, and then use that formula to compute $T(2, -3, 5)$.

Solution We first need to express $\mathbf{x} = (x_1, x_2, x_3)$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. If we write

$$(x_1, x_2, x_3) = c_1(1, 1, 1) + c_2(1, 1, 0) + c_3(1, 0, 0)$$

then on equating corresponding components, we obtain

$$\begin{aligned} c_1 + c_2 + c_3 &= x_1 \\ c_1 + c_2 &= x_2 \\ c_1 &= x_3 \end{aligned}$$

which yields $c_1 = x_3$, $c_2 = x_2 - x_3$, $c_3 = x_1 - x_2$, so

$$\begin{aligned} (x_1, x_2, x_3) &= x_3(1, 1, 1) + (x_2 - x_3)(1, 1, 0) + (x_1 - x_2)(1, 0, 0) \\ &= x_3\mathbf{v}_1 + (x_2 - x_3)\mathbf{v}_2 + (x_1 - x_2)\mathbf{v}_3 \end{aligned}$$

Thus

$$\begin{aligned} T(x_1, x_2, x_3) &= x_3T(\mathbf{v}_1) + (x_2 - x_3)T(\mathbf{v}_2) + (x_1 - x_2)T(\mathbf{v}_3) \\ &= x_3(1, 0) + (x_2 - x_3)(2, -1) + (x_1 - x_2)(4, 3) \\ &= (4x_1 - 2x_2 - x_3, 3x_1 - 4x_2 + x_3) \end{aligned}$$

From this formula we obtain

$$T(2, -3, 5) = (9, 23)$$

Problem

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

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

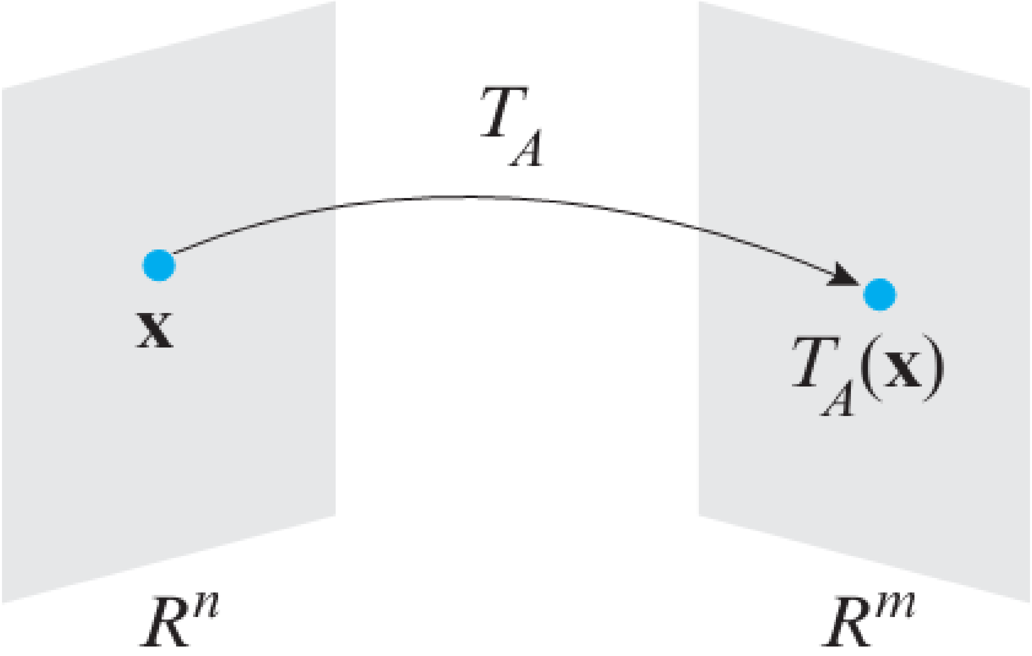
$$\begin{aligned} T(2, 3, -2) &= 2T(1, 0, 0) + 3T(0, 1, 0) - 2T(0, 0, 1) \\ &= 2(2, -1, 4) + 3(1, 5, -2) - 2(0, 3, 1) \\ &= (7, 7, 0). \end{aligned}$$

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.



$$T_A : R^n \rightarrow R^m$$

We denote ***matrix transformation*** by $T_A : R^n \rightarrow R^m$

We call this a ***matrix operator*** in the special case where $m = n$

Problem

Define the function $T: R^2 \rightarrow R^3$ as

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

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

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

SOLUTION

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

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

which means that $T(2, -1) = (6, 3, 0)$.

$$T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T(\mathbf{u}) + T(\mathbf{v}).$$

$$T(c\mathbf{u}) = A(c\mathbf{u}) = c(A\mathbf{u}) = cT(\mathbf{u}).$$

Problem

Let M_{nn} be the vector space of $n \times n$ matrices. In each part determine whether the transformation is linear.

(a) $T_1(A) = A^T$ (b) $T_2(A) = \det(A)$

Solution (a)

$$T_1(kA) = (kA)^T = kA^T = kT_1(A)$$
$$T_1(A + B) = (A + B)^T = A^T + B^T = T_1(A) + T_1(B)$$

so T_1 is linear.

Solution (b)

$$T_2(kA) = \det(kA) = k^n \det(A) = k^n T_2(A)$$

Thus, T_2 is not homogeneous and hence not linear if $n > 1$.

Definition of Kernel of a Linear Transformation

Let $T: V \rightarrow W$ be a linear transformation. Then the set of all vectors $\mathbf{v}$ in V that satisfy $T(\mathbf{v}) = \mathbf{0}$ is the **kernel** of T and is denoted by $\ker(T)$.

Problem

Find the kernel of the linear transformation $T: R^2 \to R^3$ represented by

$$T(x_1, x_2) = (x_1 - 2x_2, 0, -x_1).$$

SOLUTION

To find $\ker(T)$, you need to find all $\mathbf{x} = (x_1, x_2)$ in R^2 such that

$$T(x_1, x_2) = (x_1 - 2x_2, 0, -x_1) = (0, 0, 0).$$

This leads to the homogeneous system

$$\begin{aligned} x_1 - 2x_2 &= 0 \\ 0 &= 0 \\ -x_1 \qquad &= 0 \end{aligned}$$

which has only the trivial solution $(x_1, x_2) = (0, 0)$. So, you have

$$\ker(T) = \{(0, 0)\} = \{\mathbf{0}\}.$$

Problem

Find the kernel of the linear transformation $T: R^3 \to R^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$, where

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

SOLUTION

The kernel of T is the set of all $\mathbf{x} = (x_1, x_2, x_3)$ in R^3 such that $T(x_1, x_2, x_3) = (0, 0)$. From this equation, you can write the homogeneous system

$$\begin{bmatrix} 1 & -1 & -2 \\ -1 & 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \begin{aligned} x_1 - x_2 - 2x_3 &= 0 \\ -x_1 + 2x_2 + 3x_3 &= 0. \end{aligned}$$

Writing the augmented matrix of this system in reduced row-echelon form produces

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{aligned} x_1 &= x_3 \\ x_2 &= -x_3. \end{aligned}$$

Using the parameter $t = x_3$ produces the family of solutions

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ -t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

So, the kernel of T is

$$\ker(T) = \{t(1, -1, 1): t \text{ is a real number}\} = \operatorname{span}\{(1, -1, 1)\}.$$

Practice Problem

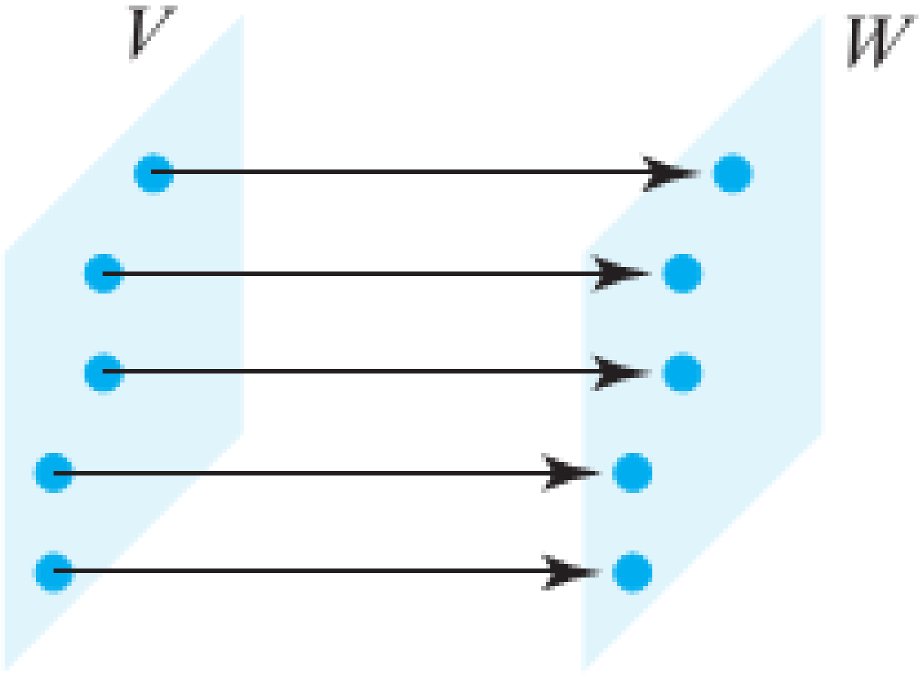
Determine whether the function is a linear transformation

$$T: P_2 \longrightarrow P_2, \qquad T(a_0 + a_1 x + a_2 x^2) = (a_0 + a_1 + a_2) + (a_1 + a_2)x + a_2 x^2.$$

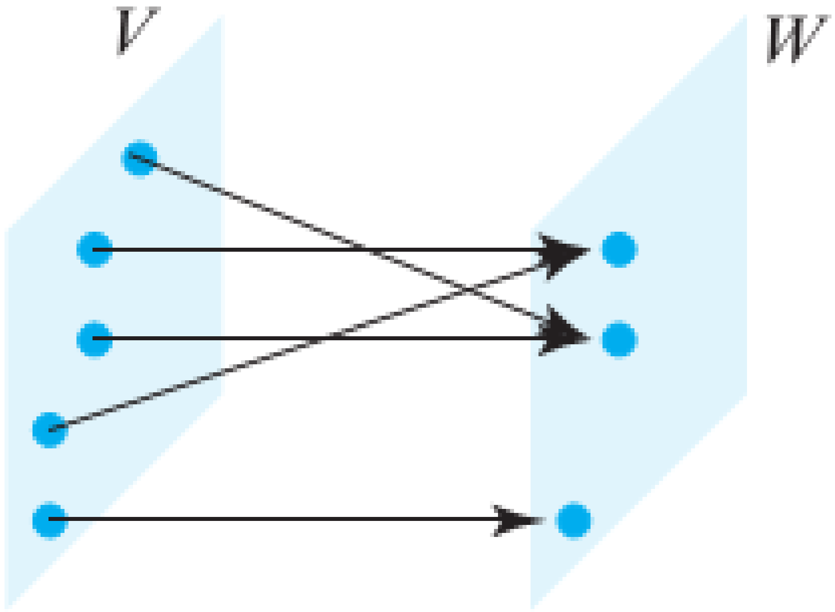
Inverse Transformations

DEFINITION 1 If $T: V \to W$ is a linear transformation from a vector space V to a vector space W, then T is said to be ***one-to-one*** if T maps distinct vectors in V into distinct vectors in W.

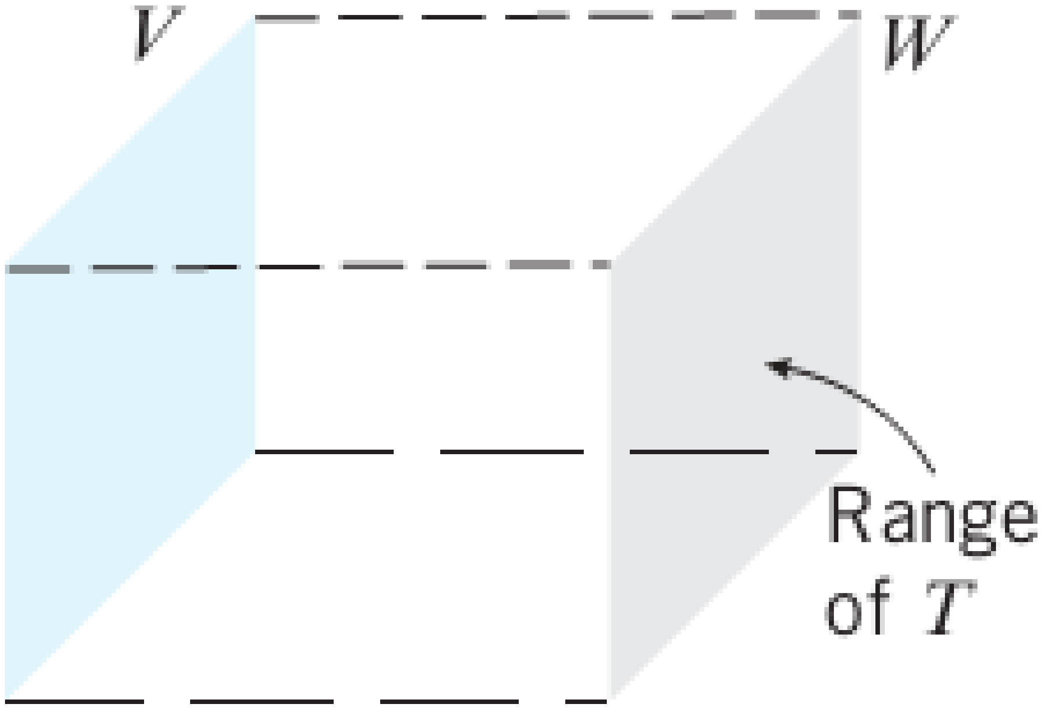
DEFINITION 2 If $T: V \to W$ is a linear transformation from a vector space V to a vector space W, then T is said to be ***onto*** (or ***onto*** W) if every vector in W is the image of at least one vector in V.



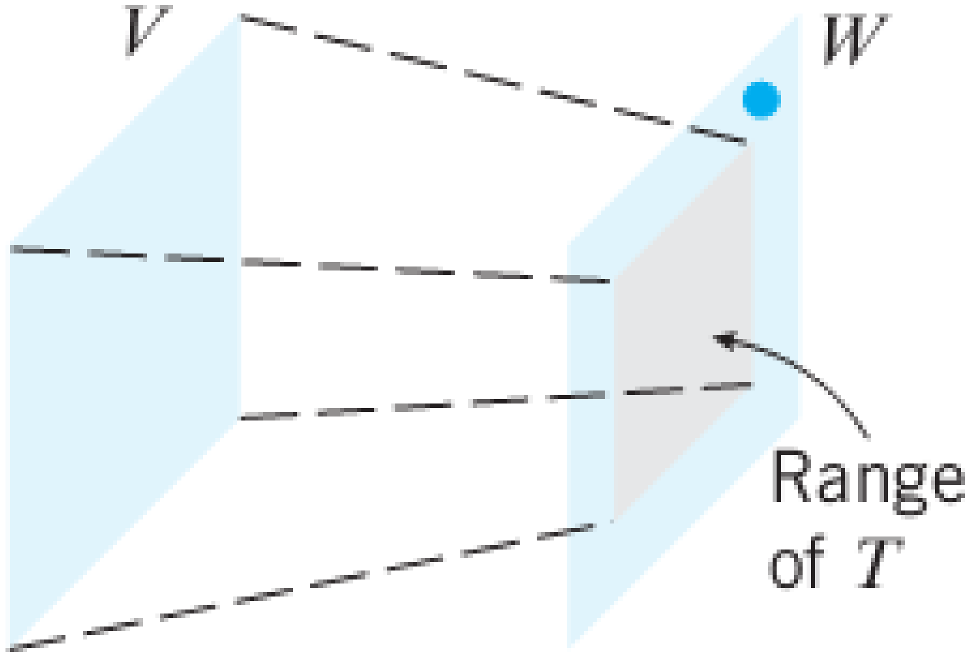
One-to-one. Distinct vectors in V have distinct images in W.



Not one-to-one. There exist distinct vectors in V with the same image.



Onto W. Every vector in W is the image of some vector in V.



Not onto W. Not every vector in W is the image of some vector in V.

THEOREM 8.2.2 *If V and W are finite-dimensional vector spaces with the same dimension, and if $T: V \to W$ is a linear transformation, then the following statements are equivalent.*

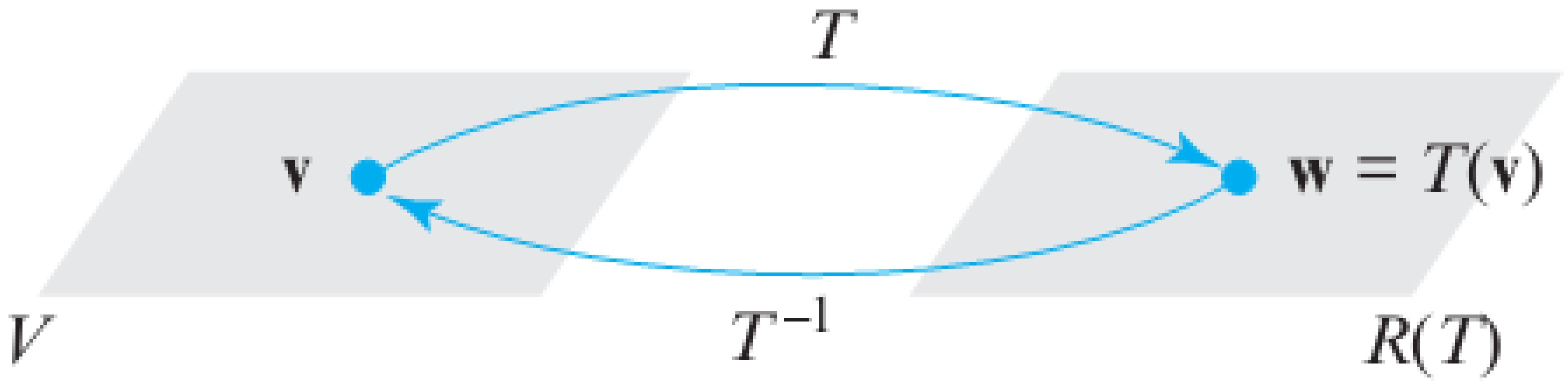
(a) T *is one-to-one.*

(b) $\ker(T) = \{\mathbf{0}\}$.

(c) T *is onto* [*i.e.*, $R(T) = W$].

Inverse Linear Transformations

If $T : V \to W$ is a one-to-one linear transformation with range $R(T)$, and if $\mathbf{w}$ is any vector in $R(T)$, then the fact that T is one-to-one means that there is *exactly one* vector $\mathbf{v}$ in V for which $T(\mathbf{v}) = \mathbf{w}$. This fact allows us to define a new function, called the ***inverse of T*** (and denoted by T^{-1}), that is defined on the range of T and that maps $\mathbf{w}$ back into $\mathbf{v}$ (Figure 8.2.4).



▶ Figure 8.2.4 The inverse of T maps $T(\mathbf{v})$ back into $\mathbf{v}$.

THEOREM 1.8.3 *Every linear transformation from R^n to R^m is a matrix transformation, and conversely, every matrix transformation from R^n to R^m is a linear transformation.*

Now since every linear transformation from R^n to R^m is a matrix transformation and for a matrix transformation there should be a matrix which when multiplied with any domain vector give image. We will call such matrix the **standard matrix** for the linear transformation.

Method for finding standard Matrix for a Linear Transformation from R^n to R^m

Step 1. Find the images of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n.

Step 2. Construct the matrix that has the images obtained in Step 1 as its successive columns. This matrix is the standard matrix for the transformation.

Problem

Find the standard matrix A for the linear transformation $T: R^2 \to R^2$ defined by the formula

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 3x_2 \\ -x_1 + x_2 \end{bmatrix}$$

Solution

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$$

the standard matrix is

$$A = [T(\mathbf{e}_1) \mid T(\mathbf{e}_2)] = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ -11 \\ 3 \end{bmatrix}$$

Practice Problem

Find the standard matrix for the operator T defined by the formula.

(a) $T(x_1, x_2) = (2x_1 - x_2, x_1 + x_2)$

(b) $T(x_1, x_2) = (x_1, x_2)$

THEOREM 4.10.1 *If A is an $n \times n$ matrix and $T_A: R^n \to R^n$ is the corresponding matrix operator, then the following statements are equivalent.*

(a) *A is invertible.*

(b) *The kernel of T_A is $\{\mathbf{0}\}$.*

(c) *The range of T_A is R^n.*

(d) *T_A is one-to-one.*

$$[T^{-1}] = [T]^{-1}$$

Problem

Let $T : R^3 \to R^3$ be the linear operator defined by the formula

$$T(x_1, x_2, x_3) = (3x_1 + x_2, -2x_1 - 4x_2 + 3x_3, 5x_1 + 4x_2 - 2x_3)$$

Determine whether T is one-to-one; if so, find $T^{-1}(x_1, x_2, x_3)$.

Solution the standard matrix for T is

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

This matrix is invertible, **(Because determinant of the standard matrix for T is non zero)**

the standard matrix for T^{-1} is

$$[T^{-1}] = [T]^{-1} = \begin{bmatrix} 4 & -2 & -3 \\ -11 & 6 & 9 \\ -12 & 7 & 10 \end{bmatrix}$$

$$T^{-1}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = [T^{-1}]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 & -2 & -3 \\ -11 & 6 & 9 \\ -12 & 7 & 10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4x_1 - 2x_2 - 3x_3 \\ -11x_1 + 6x_2 + 9x_3 \\ -12x_1 + 7x_2 + 10x_3 \end{bmatrix}$$

Expressing this result in horizontal notation yields

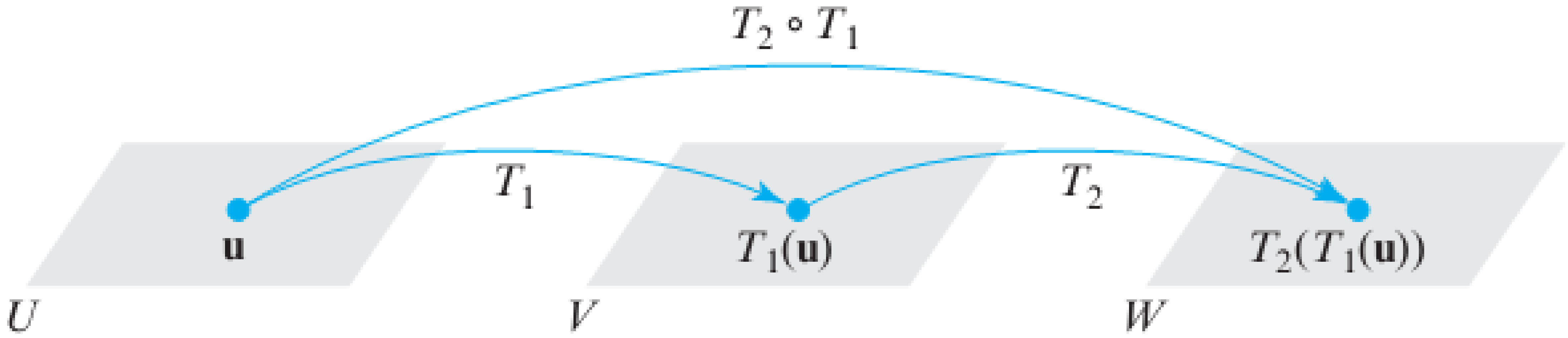
$$T^{-1}(x_1, x_2, x_3) = (4x_1 - 2x_2 - 3x_3, -11x_1 + 6x_2 + 9x_3, -12x_1 + 7x_2 + 10x_3)$$

DEFINITION 3 If $T_1 : U \to V$ and $T_2 : V \to W$ are linear transformations, then the ***composition of T_2 with T_1***, denoted by $T_2 \circ T_1$ (which is read "T_2 circle T_1"), is the function defined by the formula

$$(T_2 \circ T_1)(\mathbf{u}) = T_2(T_1(\mathbf{u})) \tag{1}$$

where $\mathbf{u}$ is a vector in U.

Remark Observe that this definition requires that the domain of T_2 (which is V) contain the range of T_1. This is essential for the formula $T_2(T_1(\mathbf{u}))$ to make sense (Figure 8.2.2).



▲ Figure 8.2.2 The composition of T_2 with T_1.

THEOREM 8.2.3 *If* $T_1: U \to V$ *and* $T_2: V \to W$ *are linear transformations, then* $(T_2 \circ T_1): U \to W$ *is also a linear transformation.*

Let $T_1: P_1 \to P_2$ and $T_2: P_2 \to P_2$ be the linear transformations given by the formulas

$$T_1(p(x)) = xp(x) \quad \text{and} \quad T_2(p(x)) = p(2x+4)$$

Then the composition $(T_2 \circ T_1): P_1 \to P_2$ is given by the formula

$$(T_2 \circ T_1)(p(x)) = T_2(T_1(p(x))) = T_2(xp(x)) = (2x+4)p(2x+4)$$

In particular, if $p(x) = c_0 + c_1 x$, then

$$\begin{aligned}(T_2 \circ T_1)(p(x)) = (T_2 \circ T_1)(c_0 + c_1 x) &= (2x+4)(c_0 + c_1(2x+4)) \\ &= c_0(2x+4) + c_1(2x+4)^2\end{aligned}$$

Composition of One-to-One Linear Transformations

THEOREM 8.2.4 *If $T_1: U \to V$ and $T_2: V \to W$ are one-to-one linear transformations, then:*

(a) $T_2 \circ T_1$ *is one-to-one.*

(b) $(T_2 \circ T_1)^{-1} = T_1^{-1} \circ T_2^{-1}$.

Practice Problem 1

Let $T_1: M_{22} \to R$ and $T_2: M_{22} \to M_{22}$ be the linear transformations given by $T_1(A) = \mathrm{tr}(A)$ and $T_2(A) = A^T$.

(a) Find $(T_1 \circ T_2)(A)$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

(b) Can you find $(T_2 \circ T_1)(A)$? Explain.

Answer:

(a) $a + d$

(b) $(T_2 \circ T_1)(A)$ does not exist since $T_1(A)$ is not a 2×2 matrix.

Practice Problem 2

Let $T: P_1 \to R^2$ be the function defined by the formula

$$T(p(x)) = (p(0),\ p(1))$$

(a) Find $T(1-2x)$.

(b) Show that T is a linear transformation.

(c) Show that T is one-to-one.

(d) Find $T^{-1}(2, 3)$, and sketch its graph.

Answers

(a) $(1,\ -1)$

(d) $T^{-1}(2, 3) = 2 + x$

8.5 Similarity

Effect of Changing Bases on Matrices of Linear Operators

Problem If B and B' are two bases for a finite-dimensional vector space V, and if $T: V \to V$ is a linear operator, what relationship, if any, exists between the matrices $[T]_B$ and $[T]_{B'}$?

THEOREM 8.5.2 *Let $T: V \rightarrow V$ be a linear operator on a finite-dimensional vector space V, and let B and B' be bases for V. Then*

$$[T]_{B'} = P^{-1}[T]_B P \tag{11}$$

where $P = P_{B' \rightarrow B}$ and $P^{-1} = P_{B \rightarrow B'}$.

$$[T]_{B'} = P_{B \rightarrow B'} [T]_B P_{B' \rightarrow B}$$

Exterior subscripts

matrices representing the same linear operator relative to different bases must be similar.

The End