(a) We begin by constructing a matrix that has $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_5$ as its column vectors:

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \\ \begin{array}{ccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 \end{array} \end{array} \tag{5}$$

The first part of our problem can be solved by finding a basis for the column space of this matrix. Reducing the matrix to *reduced* row echelon form and denoting the column vectors of the resulting matrix by $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$, and $\mathbf{w}_5$ yields

$$\begin{array}{c} \begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \begin{array}{ccccc} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 & \mathbf{w}_5 \end{array} \end{array} \tag{6}$$

The leading 1′s occur in columns 1, 2, and 4, so by Theorem 4.7.5,

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4\}$$

is a basis for the column space of 6, and consequently,

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$$

is a basis for the column space of 5.

(b) We will start by expressing $\mathbf{w}_3$ and $\mathbf{w}_5$ as linear combinations of the basis vectors $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4$. The simplest way of doing this is to express $\mathbf{w}_3$ and $\mathbf{w}_5$ in terms of basis vectors with smaller subscripts. Accordingly, we will express $\mathbf{w}_3$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$, and we will express $\mathbf{w}_5$ as a linear combination of $\mathbf{w}_1, \mathbf{w}_2$, and $\mathbf{w}_4$. By inspection of 6, these linear combinations are

$$\begin{aligned} \mathbf{w}_3 &= 2\mathbf{w}_1 - \mathbf{w}_2 \\ \mathbf{w}_5 &= \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_4 \end{aligned}$$

We call these the ***dependency equations***. The corresponding relationships in 5 are

$$\begin{aligned} \mathbf{v}_3 &= 2\mathbf{v}_1 - \mathbf{v}_2 \\ \mathbf{v}_5 &= \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_4 \end{aligned}$$

The following is a summary of the steps that we followed in our last example to solve the problem posed above.

Basis for Span(S)

Step 1. Form the matrix A having vectors in $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ as column vectors.

Step 2. Reduce the matrix A to reduced row echelon form R.

Step 3. Denote the column vectors of R by $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$.

Step 4. Identify the columns of R that contain the leading 1′s. The corresponding column vectors of A form a basis for span(S).
This completes the first part of the problem.

Step 5. Obtain a set of dependency equations by expressing each column vector of R that does not contain a leading 1 as a linear combination of preceding column vectors that do contain leading 1′s.

Step 6. Replace the column vectors of R that appear in the dependency equations by the corresponding column vectors of A.
This completes the second part of the problem.

Concept Review

- Row vectors
- Column vectors
- Row space
- Column space
- Null space
- General solution
- Particular solution
- Relationships among linear systems and row spaces, column spaces, and null spaces
- Relationships among the row space, column space, and null space of a matrix
- Dependency equations

Skills

- Determine whether a given vector is in the column space of a matrix; if it is, express it as a linear combination of the column vectors of the matrix.
- Find a basis for the null space of a matrix.
- Find a basis for the row space of a matrix.
- Find a basis for the column space of a matrix.
- Find a basis for the span of a set of vectors in R^n.

Exercise Set 4.7

1. List the row vectors and column vectors of the matrix

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

Answer:

$\mathbf{r}_1 = (2, -1, 0, 1)$, $\mathbf{r}_2 = (3, 5, 7, -1)$, $\mathbf{r}_3 = (1, 4, 2, 7)$;

$$\mathbf{c}_1 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -1 \\ 5 \\ 4 \end{bmatrix}, \mathbf{c}_3 = \begin{bmatrix} 0 \\ 7 \\ 2 \end{bmatrix}, \mathbf{c}_4 = \begin{bmatrix} 1 \\ -1 \\ 7 \end{bmatrix}$$

2. Express the product $A\mathbf{x}$ as a linear combination of the column vectors of A.

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

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

(c) $\begin{bmatrix} -3 & 6 & 2 \\ 5 & -4 & 0 \\ 2 & 3 & -1 \\ 1 & 8 & 3 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$

(d) $\begin{bmatrix} 2 & 1 & 5 \\ 6 & 3 & -8 \end{bmatrix}\begin{bmatrix} 3 \\ 0 \\ -5 \end{bmatrix}$

3. Determine whether $\mathbf{b}$ is in the column space of A, and if so, express $\mathbf{b}$ as a linear combination of the column vectors of A.

(a) $A = \begin{bmatrix} 1 & 3 \\ 4 & -6 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} -2 \\ 10 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 3 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & -1 & 1 \\ 9 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix}$

(d) $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$

(e) $A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 2 & 1 & 3 \\ 0 & 1 & 2 & 2 \end{bmatrix}; \quad \mathbf{b} = \begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix}$

Answer:

(a) $\begin{bmatrix} -2 \\ 10 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ -6 \end{bmatrix}$

(b) $\mathbf{b}$ is not in the column space of A.

(c) $\begin{bmatrix} 1 \\ 9 \\ 1 \end{bmatrix} - 3\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -1 \end{bmatrix}$

(d) $\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + (t-1)\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} + t\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

(e) $\begin{bmatrix} 4 \\ 3 \\ 5 \\ 7 \end{bmatrix} = -26\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + 13\begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix} - 7\begin{bmatrix} 0 \\ 2 \\ 1 \\ 2 \end{bmatrix} + 4\begin{bmatrix} 1 \\ 1 \\ 3 \\ 2 \end{bmatrix}$

4. Suppose that $x_1 = -1, x_2 = 2, x_3 = 4, x_4 = -3$ is a solution of a nonhomogeneous linear system $A\mathbf{x} = \mathbf{b}$ and that the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ is given by the formulas

$$x_1 = -3r + 4s, \quad x_2 = r - s, \quad x_3 = r, \quad x_4 = s$$

(a) Find a vector form of the general solution of $A\mathbf{x} = \mathbf{0}$.

(b) Find a vector form of the general solution of $A\mathbf{x} = \mathbf{b}$.

5. In parts (a)–(d), find the vector form of the general solution of the given linear system $A\mathbf{x} = \mathbf{b}$; then use that result to find the vector form of the general solution of $A\mathbf{x} = 0$.

(a) $$\begin{aligned} x_1 - 3x_2 &= 1 \\ 2x_1 - 6x_2 &= 2 \end{aligned}$$

(b) $$\begin{aligned} x_1 + x_2 + 2x_3 &= 5 \\ x_1 \qquad + x_3 &= -2 \\ 2x_1 + x_2 + 3x_3 &= 3 \end{aligned}$$

(c) $$\begin{aligned} x_1 - 2x_2 + x_3 + 2x_4 &= -1 \\ 2x_1 - 4x_2 + 2x_3 + 4x_4 &= -2 \\ -x_1 + 2x_2 - x_3 - 2x_4 &= 1 \\ 3x_1 - 6x_2 + 3x_3 + 6x_4 &= -3 \end{aligned}$$

(d) $$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 4 \\ -2x_1 + x_2 + 2x_3 + x_4 &= -1 \\ -x_1 + 3x_2 - x_3 + 2x_4 &= 3 \\ 4x_1 - 7x_2 \qquad - 5x_4 &= -5 \end{aligned}$$

Answer:

(a) $\begin{bmatrix} 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 1 \end{bmatrix}; \; t\begin{bmatrix} 3 \\ 1 \end{bmatrix}$

(b) $\begin{bmatrix} -2 \\ 7 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}; \; t\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}; \; r\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

(d) $\begin{bmatrix} \frac{6}{5} \\ \frac{7}{5} \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{1}{5} \\ -\frac{3}{5} \\ 0 \\ 1 \end{bmatrix}; \; s\begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} \frac{1}{5} \\ -\frac{3}{5} \\ 0 \\ 1 \end{bmatrix}$

6. Find a basis for the null space of A.

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

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

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

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

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

7. In each part, a matrix in row echelon form is given. By inspection, find bases for the row and column spaces of A.

(a) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

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

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

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

Answer:

(a)
$$\mathbf{r}_1 = [1\ 0\ 2],\ \mathbf{r}_2 = [0\ 0\ 1],\ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

(b)
$$\mathbf{r}_1 = [1\ -3\ 0\ 0],\ \mathbf{r}_2 = [0\ 1\ 0\ 0],\ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

(c) $\mathbf{r}_1 = [1\ 2\ 4\ 5],\ \mathbf{r}_2 = [0\ 1\ -3\ 0],\ \mathbf{r}_3 = [0\ 0\ 1\ -3],\ \mathbf{r}_4 = [0\ 0\ 0\ 1],$

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_3 = \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_4 = \begin{bmatrix} 5 \\ 0 \\ -3 \\ 1 \\ 0 \end{bmatrix}$$

(d) $\mathbf{r}_1 = [1\ 2\ -1\ 5]$, $\mathbf{r}_2 = [0\ 1\ 4\ 3]$, $\mathbf{r}_3 = [0\ 0\ 1\ -7]$, $\mathbf{r}_4 = [0\ 0\ 0\ 1]$

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix},\ \mathbf{c}_3 = \begin{bmatrix} -1 \\ 4 \\ 1 \\ 0 \end{bmatrix},\ \mathbf{c}_4 = \begin{bmatrix} 5 \\ 3 \\ -7 \\ 1 \end{bmatrix}$$

8. For the matrices in Exercise 6, find a basis for the row space of A by reducing the matrix to row echelon form.

9. By inspection, find a basis for the row space and a basis for the column space of each matrix.

(a) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

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

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

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

Answer:

(a)
$$\mathbf{r}_1 = [1\ \ 0\ \ 2];\ \mathbf{r}_2 = [0\ \ 0\ \ 1];\ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix};\ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

(b)
$$\mathbf{r}_1 = [1\ \ -3\ \ 0\ \ 0];\ \mathbf{r}_2 = [0\ \ 1\ \ 0\ \ 0];\ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix};\ \mathbf{c}_2 = \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

(c) $\mathbf{r}_1 = [1\ \ 2\ \ 4\ \ 5];\ \mathbf{r}_2 = [0\ \ 1\ \ -3\ \ 0];\ \mathbf{r}_3 = [0\ \ 0\ \ 1\ \ -3];$

$$\mathbf{r}_4 = [0\ \ 0\ \ 0\ \ 1];\ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix};\ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix};\ \mathbf{c}_3 = \begin{bmatrix} 4 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix};\ \mathbf{c}_4 = \begin{bmatrix} 5 \\ 0 \\ -3 \\ 1 \\ 0 \end{bmatrix}$$

(d) $\mathbf{r}_1 = [1\ \ 2\ \ -1\ \ 5];\ \mathbf{r}_2 = [0\ \ 1\ \ 4\ \ 3];\ \mathbf{r}_3 = [0\ \ 0\ \ 1\ \ -7];$

$$\mathbf{r}_4 = [0 \quad 0 \quad 0 \quad 1]; \ \mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \ \mathbf{c}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}; \ \mathbf{c}_3 = \begin{bmatrix} -1 \\ 4 \\ 1 \\ 0 \end{bmatrix}; \ \mathbf{c}_4 = \begin{bmatrix} 5 \\ 3 \\ -7 \\ 1 \end{bmatrix}$$

10. For the matrices in Exercise 6, find a basis for the row space of A consisting entirely of row vectors of A.

11. Find a basis for the subspace of R^4 spanned by the given vectors.

(a) $(1, 1, -4, -3), (2, 0, 2, -2), (2, -1, 3, 2)$

(b) $(-1, 1, -2, 0), (3, 3, 6, 0), (9, 0, 0, 3)$

(c) $(1, 1, 0, 0), \ (0, 0, 1, 1), \ (-2, 0, 2, 2), \ (0, -3, 0, 3)$

Answer:

(a) $(1, 1, \ -4 \ -3), \ (0, 1, \ -5, \ -2), \ \left(0, 0, 1, \ -\frac{1}{2}\right)$

(b) $(1, \ -1, 2, 0), \ (0, 1, 0, 0), \ \left(0, 0, 1, \ -\frac{1}{6}\right)$

(c) $(1, 1, 0, 0), (0, 1, 1, 1), (0, 0, 1, 1), (0, 0, 0, 1)$

12. Find a subset of the vectors that forms a basis for the space spanned by the vectors; then express each vector that is not in the basis as a linear combination of the basis vectors.

(a) $\mathbf{v}_1 = (1, 0, 1, 1), \ \mathbf{v}_2 = (-3, 3, 7, 1), \ \mathbf{v}_3 = (-1, 3, 9, 3), \ \mathbf{v}_4 = (-5, 3, 5, -1)$

(b) $\mathbf{v}_1 = (1, -2, 0, 3), \ \mathbf{v}_2 = (2, -4, 0, 6), \ \mathbf{v}_3 = (-1, 1, 2, 0), \ \mathbf{v}_4 = (0, -1, 2, 3)$

(c) $\mathbf{v}_1 = (1, -1, 5, 2), \ \mathbf{v}_2 = (-2, 3, 1, 0), \ \mathbf{v}_3 = (4, -5, 9, 4), \ \mathbf{v}_4 = (0, 4, 2, -3), \mathbf{v}_5 = (-7, 18, 2, -8)$

13. Prove that the row vectors of an $n \times n$ invertible matrix A form a basis for R^n.

14. Construct a matrix whose null space consists of all linear combinations of the vectors

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

15. (a) Let

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

Show that relative to an xyz-coordinate system in 3-space the null space of A consists of all points on the z-axis and that the column space consists of all points in the xy-plane (see the accompanying figure).

(b) Find a 3×3 matrix whose null space is the x-axis and whose column space is the yz-plane.

Figure Ex-15

Answer:

(b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

16. Find a 3×3 matrix whose null space is

(a) a point.

(b) a line.

(c) a plane.

17. (a) Find all 2×2 matrices whose null space is the line $3x - 5y = 0$.

(b) Sketch the null spaces of the following matrices:

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

Answer:

(a) $\begin{bmatrix} 3a & -5a \\ 3b & -5b \end{bmatrix}$ for all real numbers a, b not both 0.

(b) Since A and B are invertible, their null spaces are the origin. The null space of C is the line $3x + y = 0$. The null space of D is the entire xy-plane.

18. The equation $x_1 + x_2 + x_3 = 1$ can be viewed as a linear system of one equation in three unknowns. Express its general solution as a particular solution plus the general solution of the corresponding homogeneous system. [*Suggestion:* Write the vectors in column form.]

19. Suppose that A and B are $n \times n$ matrices and A is invertible.Invent and prove a theorem that describes how the row spaces of AB and B are related.

True-False Exercises

In parts (a)–(j) determine whether the statement is true or false, and justify your answer.

(a) The span of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is the column space of the matrix whose column vectors are $\mathbf{v}_1, \ldots, \mathbf{v}_n$.

Answer:

True

(b) The column space of a matrix A is the set of solutions of $A\mathbf{x} = \mathbf{b}$.

Answer:

False

(c) If R is the reduced row echelon form of A, then those column vectors of R that contain the leading 1′s form a basis for the column space of A.

Answer:

False

(d) The set of nonzero row vectors of a matrix A is a basis for the row space of A.

Answer:

False

(e) If A and B are $n \times n$ matrices that have the same row space, then A and B have the same column space.

Answer:

False

(f) If E is an $m \times m$ elementary matrix and A is an $m \times n$ matrix, then the null space of $E\,A$ is the same as the null space of A.

Answer:

True

(g) If E is an $m \times m$ elementary matrix and A is an $m \times n$ matrix, then the row space of $E\,A$ is the same as the row space of A.

Answer:

True

(h) If E is an $m \times m$ elementary matrix and A is an $m \times n$ matrix, then the column space of $E\,A$ is the same as the column space of A.

Answer:

False

(i) The system $A\mathbf{x} = \mathbf{b}$ is inconsistent if and only if $\mathbf{b}$ is not in the column space of A.

Answer:

True

(j) There is an invertible matrix A and a singular matrix B such that the row spaces of A and B are the same.

Answer:

False

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

4.8 Rank, Nullity, and the Fundamental Matrix Spaces

In the last section we investigated relationships between a system of linear equations and the row space, column space, and null space of its coefficient matrix. In this section we will be concerned with the dimensions of those spaces. The results weobtain will provide a deeper insight into the relationship between a linear system and its coefficient matrix.

Row and Column Spaces Have Equal Dimensions

In Examples 6 and 7 of Section 4.7 we found that the row and column spaces of the matrix

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

both have three basis vectors and hence are both three-dimensional. The fact that these spaces have the same dimension is not accidental, but rather a consequence of the following theorem.

THEOREM 4.8.1

The row space and column space of a matrix A have the same dimension.

Proof Let R be any row echelon form of A. It follows from Theorem 4.7.4 and Theorem 4.7.6 b that

$$\dim(\text{row space of } A) = \dim(\text{row space of } R)$$
$$\dim(\text{column space of } A) = \dim(\text{column space of } R)$$

so it suffices to show that the row and column spaces of R have the same dimension. But the dimension of the row space of R is the number of nonzero rows, and by Theorem 4.7.5 the dimension of the column space of R is the number of leading 1′s. Since these two numbers are the same, the row and column space have the same dimension.

Rank and Nullity

The dimensions of the row space, column space, and null space of a matrix are such important numbers that there is some notation and terminology associated with them.

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 rank(A); the dimension of the null space of A is called the ***nullity*** of A and is denoted by nullity(A).

The proof of Theorem 4.8.1 shows that the rank of A can be interpreted as the number of leading 1′s in any row echelon form of A.

EXAMPLE 1 Rank and Nullity of a 4 × 6 Matrix

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

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

(verify). Since this matrix has two leading 1′s, its row and column spaces are two-dimensional and rank $(A) = 2$. To find the nullity of A, we must find the dimension of the solution space of the linear system $A\mathbf{x} = 0$. This system can be solved by reducing its augmented matrix to reduced row echelon form. The resulting matrix will be identical to 1, except that it will have an additional last column of zeros, and hence the corresponding system of equations will be

$$\begin{aligned} x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 &= 0 \\ x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 &= 0 \end{aligned}$$

Solving these equations for the leading variables yields

$$\begin{aligned} x_1 &= 4x_3 + 28x_4 + 37x_5 - 13x_6 \\ x_2 &= 2x_3 + 12x_4 + 16x_5 - 5x_6 \end{aligned} \tag{2}$$

from which we obtain the general solution

$$\begin{aligned} x_1 &= 4r + 28s + 37t - 13u \\ x_2 &= 2r + 12s + 16t - 5u \\ x_3 &= r \\ x_4 &= s \\ x_5 &= t \\ x_6 &= u \end{aligned}$$

or in column vector form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

Because the four vectors on the right side of 3 form a basis for the solution space, nullity(A) = 4.

EXAMPLE 2 Maximum Value for Rank

What is the maximum possible rank of an $m \times n$ matrix A that is not square?

Solution Since the row vectors of A lie in R^n and the column vectors in R^m, the row space of A is at most n-dimensional and the column space is at most m-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. We denote this by writing

$$\text{rank}(A) \leq \min(m, n)$$

in which min (m, n) is the minimum of m and n.

The following theorem establishes an important relationship between the rank and nullity of a matrix.

THEOREM 4.8.2 Dimension Theorem for Matrices

If A is a matrix with n columns, then

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

Proof Since A has n columns, the homogeneous linear system $A\mathbf{x} = 0$ has n unknowns (variables). These fall into two distinct categories: the leading variables and the free variables. Thus,

$$\begin{bmatrix} \text{number of leading} \\ \text{variables} \end{bmatrix} + \begin{bmatrix} \text{number of free} \\ \text{variables} \end{bmatrix} = n$$

But the number of leading variables is the same as the number of leading 1′s in the reduced row echelon form of A, which is the rank of A; and the number of free variables is the same as the number of parameters in the general solution of $A\mathbf{x} = 0$, which is the nullity of A. This yields Formula 4.

EXAMPLE 3 The Sum of Rank and Nullity

The matrix

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

has 6 columns, so

$$\text{rank}(A) + \text{nullity}(A) = 6$$

This is consistent with Example 1, where we showed that

$$\text{rank}(A) = 2 \quad \text{and} \quad \text{nullity}(A) = 4$$

The following theorem, which summarizes results already obtained, interprets rank and nullity in the context of a homogeneous linear system.

THEOREM 4.8.3

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

(a) $\text{rank}(A) =$ the number of leading variables in the general solution of $A\mathbf{x} = 0$.

(b) $\text{nullity}(A) =$ the number of parameters in the general solution of $A\mathbf{x} = 0$

EXAMPLE 4 Number of Parameters in a General Solution

Find the number of parameters in the general solution of $A\mathbf{x} = 0$ if A is a 5×7 matrix of rank 3.

Solution From 4,

$$\text{nullity}(A) = n - \text{rank}(A) = 7 - 3 = 4$$

Thus there are four parameters.

Equivalence Theorem

In Theorem 2.3.8 we listed seven results that are equivalent to the invertibility of a square matrix A. We are now in a position to add eight more results to that list to produce a single theorem that summarizes most of the topics we have covered thus far.

THEOREM 4.8.4 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = 0$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

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

(h) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.

(j) The column vectors of A span R^n.

(k) The row vectors of A span R^n.

(l) The column vectors of A form a basis for R^n.

(m) The row vectors of A form a basis for R^n.

(n) A has rank n.

(o) A has nullity 0.

Proof The equivalence of (h) through (m) follows from Theorem 4.5.4 (we omit the details). To complete the proof we will show that (b), (n), and (o) are equivalent by proving the chain of implications
$(b) \Rightarrow (o) \Rightarrow (n) \Rightarrow (b)$.

$(b) \Rightarrow (o)$ If $A\mathbf{x} = 0$ has only the trivial solution, then there are no parameters in that solution, so nullity $(A) = 0$ by Theorem 4.8.3 *b*.

$(o) \Rightarrow (n)$ Theorem 4.8.2.

$(n) \Rightarrow (b)$ If A has rank n, then Theorem 4.8.3*a* implies that there are n leading variables (hence no free variables) in the general solution of $A\mathbf{x} = 0$. This leaves the trivial solution as the only possibility.

Overdetermined and Underdetermined Systems

In many applications the equations in a linear system correspond to physical constraints or conditions that must be satisfied. In general, the most desirable systems are those that have the same number of constraints as unknowns, since such systems often have a unique solution. Unfortunately, it is not always possible to match the number of constraints and unknowns, so researchers are often faced with linear systems that have more constraints than unknowns, called ***overdetermined systems***, or with fewer constraints than unknowns, called ***underdetermined systems***. The following two theorems will help us to analyze both overdetermined and underdetermined systems.

In engineering and other applications, the occurrence of an overdetermined or underdetermined linear system often signals that one or more variables were omitted in formulating the problem or that extraneous variables were included. This often leads to some kind of undesirable physical result.

THEOREM 4.8.5

If $A\mathbf{x} = \mathbf{b}$ is a consistent linear system of m equations in n unknowns, and if A has rank r, then the general solution of the system contains $n - r$ parameters.

Proof It follows from Theorem 4.7.2 that the number of parameters is equal to the nullity of A, which, by Theorem 4.8.2, is $n - r$.

THEOREM 4.8.6

Let A be an $m \times n$ matrix.

(a) **(Overdetermined Case)** If $m > n$, then the linear system $A\mathbf{x} = \mathbf{b}$ is inconsistent for at least one vector $\mathbf{b}$ in R^n.

(b) **(Underdetermined Case)** If $m < n$, then for each vector $\mathbf{b}$ in R^m the linear system $A\mathbf{x} = \mathbf{b}$ is either inconsistent or has infinitely many solutions.

Proof (a) Assume that $m > n$, in which case the column vectors of A cannot span R^m (fewer vectors than the dimension of R^m). Thus, there is at least one vector $\mathbf{b}$ in R^m that is not in the column space of A, and for that $\mathbf{b}$ the system $A\mathbf{x} = \mathbf{b}$ is inconsistent by Theorem 4.7.1.

Proof (b) Assume that $m < n$. For each vector $\mathbf{b}$ in R^n there are two possibilities: either the system $A\mathbf{x} = \mathbf{b}$ is consistent or it is inconsistent. If it is inconsistent, then the proof is complete. If it is consistent, then Theorem 4.8.5 implies that the general solution has $n - r$ parameters, where $r = \text{rank}(A)$. But rank (A) is the smaller of m and n, so

$$n - r = n - m > 0$$

This means that the general solution has at least one parameter and hence there are infinitely many solutions.

EXAMPLE 5 Overdetermined and Underdetermined Systems

(a) What can you say about the solutions of an overdetermined system $A\mathbf{x} = \mathbf{b}$ of 7 equations in 5 unknowns in which A has rank $r = 4$?

(b) What can you say about the solutions of an underdetermined system $A\mathbf{x} = \mathbf{b}$ of 5 equations in 7 unknowns in which A has rank $r = 4$?

Solution

(a) The system is consistent for some vector $\mathbf{b}$ in R^7, and for any such $\mathbf{b}$ the number of parameters in the general solution is $n - r = 5 - 4 = 1$.

(b) The system may be consistent or inconsistent, but if it is consistent for the vector $\mathbf{b}$ in R^5, then the general solution has $n - r = 7 - 4 = 3$ parameters.

EXAMPLE 6 An Overdetermined System

The linear system

$$\begin{array}{rcrcl} x_1 & - & 2x_2 & = & b_1 \\ x_1 & - & x_2 & = & b_2 \\ x_1 & + & x_2 & = & b_3 \\ x_1 & + & 2x_2 & = & b_4 \\ x_1 & + & 3x_2 & = & b_5 \end{array}$$

is overdetermined, so it cannot be consistent for all possible values of b_1, b_2, b_3, b_4, and b_5. Exact conditions under which the system is consistent can be obtained by solving the linear system by Gauss–Jordan elimination. We leave it for you to show that the augmented matrix is row equivalent to

$$\left[\begin{array}{ccccccc} 1 & 0 & & & 2b_2 & - & b_1 \\ 0 & 1 & & & b_2 & - & b_1 \\ 0 & 0 & b_3 & - & 3b_2 & + & 2b_1 \\ 0 & 0 & b_4 & - & 4b_2 & + & 3b_1 \\ 0 & 0 & b_5 & - & 5b_2 & + & 4b_1 \end{array}\right] \tag{5}$$

Thus, the system is consistent if and only if b_1, b_2, b_3, b_4, and b_5 satisfy the conditions

$$\begin{array}{rcrcrcrcrcl} 2b_1 & - & 3b_2 & + & b_3 & & & & & = & 0 \\ 3b_1 & - & 4b_2 & & & + & b_4 & & & = & 0 \\ 4b_1 & - & 5b_2 & & & & & + & b_5 & = & 0 \end{array}$$

Solving this homogeneous linear system yields

$$b_1 = 5r - 4s, \quad b_2 = 4r - 3s, \quad b_3 = 2r - s, \quad b_4 = r, \quad b_5 = s$$

where r and s are arbitrary.

Remark The coefficient matrix for the linear system in the last example has $n = 2$ columns, and it has rank $r = 2$ because there are two nonzero rows in its reduced row echelon form. This implies that when the system is consistent its general solution will contain $n - r = 0$ parameters; that is, the solution will be unique. With a moment's thought, you should be able to see that this is so from 5.

The Fundamental Spaces of a Matrix

There are six important vector spaces associated with a matrix A and its transpose A^T:

row space of A	row space of A^T
column space of A	column space of A^T
null space of A	null space of A^T

However, transposing a matrix converts row vectors into column vectors and conversely, so except for a difference in notation, the row space of A^T is the same as the column space of A, and the column space of A^T is the same as the row space of A. Thus, of the six spaces listed above, only the following four are distinct:

row space of A	column space of A
null space of A	null space of A^T

If A is an $m \times n$ matrix, then the row space and null space of A are subspaces of R^n, and the column space of A and the null space of A^T are subspaces of R^m.

These are called the ***fundamental spaces*** of a matrix A. We will conclude this section by discussing how these four subspaces are related.

Let us focus for a moment on the matrix A^T. Since the row space and column space of a matrix have the same dimension, and since transposing a matrix converts its columns to rows and its rows to columns, the following result should not be surprising.

THEOREM 4.8.7

If A is any matrix, then $\operatorname{rank}(A) = \operatorname{rank}(A^T)$.

Proof

$$\operatorname{rank}(A) = \dim(\text{row space of } A) = \dim(\text{column space of } A^T) = \operatorname{rank}(A^T).$$

This result has some important implications. For example, if A is an $m \times n$ matrix, then applying Formula 4 to the matrix A^T and using the fact that this matrix has m columns yields

$$\operatorname{rank}(A^T) + \operatorname{nullity}(A^T) = m$$

which, by virtue of Theorem 4.8.7, can be rewritten as

$$\operatorname{rank}(A) + \operatorname{nullity}(A^T) = m \tag{6}$$

This alternative form of Formula 4 in Theorem 4.8.2 makes it possible to express the dimensions of all four fundamental spaces in terms of the size and rank of A. Specifically, if $\operatorname{rank}(A) = r$, then

$$\begin{aligned} &\dim[\operatorname{row}(A)] = r && \dim[\operatorname{col}(A)] = r \\ &\dim[\operatorname{null}(A)] = n - r && \dim\left[\operatorname{null}(A^T)\right] = m - r \end{aligned} \tag{7}$$

The four formulas in 7 provide an *algebraic* relationship between the size of a matrix and the dimensions of its fundamental spaces. Our next objective is to find a *geometric* relationship between the fundamental spaces themselves. For this purpose recall from Theorem 3.4.3 that if A is an $m \times n$ matrix, then the null space of A consists of those vectors that are orthogonal to each of the row vectors of A. To develop that idea in more detail, we make the following definition.

DEFINITION 2

If W is a subspace of R^n, then the set of all vectors in R^n that are orthogonal to every vector in W is called the ***orthogonal complement*** of W and is denoted by the symbol $W^\perp$.

The following theorem lists three basic properties of orthogonal complements. We will omit the formal proof because a more general version of this theorem will be given later in the text.

THEOREM 4.8.8

If W is a subspace of R^n, then:

(a) $W^\perp$ is a subspace of R^n.

(b) The only vector common to W and $W^\perp$ is $\mathbf{0}$.

(c) The orthogonal complement of $W^\perp$ is W.

EXAMPLE 7 Orthogonal Complements

In R^2 the orthogonal complement of a line W through the origin is the line through the origin that is perpendicular to W (Figure 4.8.1*a*); and in R^3 the orthogonal complement of a plane W through the origin is the line through the origin that is perpendicular to that plane (Figure 4.8.1*b*).

Figure 4.8.1

Explain why $\{\mathbf{0}\}$ and R^n are orthogonal complements.

A Geometric Link Between the Fundamental Spaces

The following theorem provides a geometric link between the fundamental spaces of a matrix. Part (*a*) is essentially a restatement of Theorem 3.4.3 in the language of orthogonal complements, and part (*b*), whose proof is left as an exercise, follows from part (*a*). The essential idea of the theorem is illustrated in Figure 4.8.2.

THEOREM 4.8.9

If A is an $m \times n$ matrix, then:

(a) The null space of A and the row space of A are orthogonal complements in R^n.

(b) The null space of A^T and the column space of A are orthogonal complements in R^m.

Figure 4.8.2

More on the Equivalence Theorem

As our final result in this section, we will add two more statements to Theorem 4.8.4. We leave the proof that those statements are equivalent to the rest as an exercise.

THEOREM 4.8.10 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = 0$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

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

(h) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.

(j) The column vectors of A span R^n.

(k) The row vectors of A span R^n.

(l) The column vectors of A form a basis for R^n.

(m) The row vectors of A form a basis for R^n.

(n) A has $\operatorname{rank} n$.

(o) A has $\operatorname{nullity} 0$.

(p) The orthogonal complement of the null space of A is R^n.

(q) The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.

Applications of Rank

The advent of the Internet has stimulated research on finding efficient methods for transmitting large amounts of digital data over communications lines with limited bandwidths. Digital data are commonly stored in matrix form, and many techniques for improving transmission speed use the rank of a matrix in some way. Rank plays a role because it measures the "redundancy" in a matrix in the sense that if A is an $m \times n$ matrix of rank k, then $n - k$ of the column vectors and $m - k$ of the row vectors can be expressed in terms of k linearly independent column or row vectors. The essential idea in many data compression schemes is to approximate the original data set by a data set with smaller rank that conveys nearly the same information, then eliminate redundant vectors in the approximating set to speed up the transmission time.

Concept Review

- Rank
- Nullity
- Dimension Theorem
- Overdetermined system
- Underdetermined system
- Fundamental spaces of a matrix
- Relationships among the fundamental spaces
- Orthogonal complement
- Equivalent characterizations of invertible matrices

Skills

- Find the rank and nullity of a matrix.
- Find the dimension of the row space of a matrix.

Exercise Set 4.8

1. Verify that $\operatorname{rank}(A) = \operatorname{rank}(A^T)$.

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

Answer:

$\operatorname{Rank}(A) = \operatorname{Rank}(A^T) = 2$

2. Find the rank and nullity of the matrix; then verify that the values obtained satisfy Formula 4 in the Dimension Theorem.

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

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

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

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

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

3. In each part of Exercise 2, use the results obtained to find the number of leading variables and the number of parameters in the solution of $A\mathbf{x} = \mathbf{0}$ without solving the system.

Answer:

(a) 2; 1
(b) 1; 2
(c) 2; 2
(d) 2; 3
(e) 3; 2

4. In each part, use the information in the table to find the dimension of the row space of A, column space of A, null space of A, and null space of A^T.

	(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**
Size of A	3×3	3×3	3×3	5×9	9×5	4×4	6×2
Rank(A)	3	2	1	2	2	0	2

5. In each part, find the largest possible value for the rank of A and the smallest possible value for the nullity of A.

(a) A is 4×4
(b) A is 3×5
(c) A is 5×3

Answer:

(a) $\text{Rank} = 4, \ \text{nullity} = 0$
(b) $\text{Rank} = 3, \ \text{nullity} = 2$
(c) $\text{Rank} = 3, \ \text{nullity} = 0$

6. If A is an $m \times n$ matrix, what is the largest possible value for its rank and the smallest possible value for its nullity?

7. In each part, use the information in the table to determine whether the linear system $A\mathbf{x} = \mathbf{b}$ is consistent. If so, state the number of parameters in its general solution.

	(a)	**(b)**	**(c)**	**(d)**	**(e)**	**(f)**	**(g)**
Size of A	3×3	3×3	3×3	5×9	5×9	4×4	6×2
Rank (A)	3	2	1	2	2	0	2
Rank [A |**b**]	3	3	1	2	3	0	2

Answer:

(a) Yes, 0
(b) No
(c) Yes, 2
(d) Yes, 7
(e) No
(f) Yes, 4
(g) Yes, 0

8. For each of the matrices in Exercise 7, find the nullity of A, and determine the number of parameters in the general solution of the homogeneous linear system $A\mathbf{x} = 0$.

9. What conditions must be satisfied by b_1, b_2, b_3, b_4, and b_5 for the overdetermined linear system

$$\begin{aligned} x_1 - 3x_2 &= b_1 \\ x_1 - 2x_2 &= b_2 \\ x_1 + x_2 &= b_3 \\ x_1 - 4x_2 &= b_4 \\ x_1 + 5x_2 &= b_5 \end{aligned}$$

to be consistent?

Answer:

$b_1 = r,\ b_2 = s,\ b_3 = 4s - 3r,\ b_4 = 2r - s,\ b_5 = 8s - 7r$

10. Let

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

Show that A has $\text{rank } 2$ if and only if one or more of the determinants

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}, \quad \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

is nonzero.

11. Suppose that A is a 3×3 matrix whose null space is a line through the origin in 3-space. Can the row or column space of A also be a line through the origin? Explain.

Answer:

No

12. Discuss how the rank of A varies with t.

(a) $A = \begin{bmatrix} 1 & 1 & t \\ 1 & t & 1 \\ t & 1 & 1 \end{bmatrix}$

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

13. Are there values of r and s for which

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

has rank 1? Has rank 2? If so, find those values.

Answer:

Rank is 2 if $r = 2$ and $s = 1$; the rank is never 1.

14. Use the result in Exercise 10 to show that the set of points (x, y, z) in R^3 for which the matrix

$$\begin{bmatrix} x & y & z \\ 1 & x & y \end{bmatrix}$$

has $\text{rank } 1$ is the curve with parametric equations $x = t, y = t^2, z = t^3$.

15. Prove: If $k \neq 0$, then A and kA have the same rank.

16. (a) Give an example of a 3×3 matrix whose column space is a plane through the origin in 3-space.

(b) What kind of geometric object is the null space of your matrix?

(c) What kind of geometric object is the row space of your matrix?

17. (a) If A is a 3×5 matrix, then the number of leading 1′s in the reduced row echelon form of A is at most _____ . Why?

(b) If A is a 3×5 matrix, then the number of parameters in the general solution of $A\mathbf{x} = 0$ is at most _____ . Why?

(c) If A is a 5×3 matrix, then the number of leading 1′s in the reduced row echelon form of A is at most _____ . Why?

(d) If A is a 5×3 matrix, then the number of parameters in the general solution of $A\mathbf{x} = 0$ is at most _____ . Why?

Answer:

(a) 3

(b) 5

(c) 3

(d) 3

18. (a) If A is a 3×5 matrix, then the rank of A is at most _____ . Why?

(b) If A is a 3×5 matrix, then the nullity of A is at most _____ . Why?

(c) If A is a 3×5 matrix, then the rank of A^T is at most _____ . Why?

(d) If A is a 3×5 matrix, then the nullity of A^T is at most _____ . Why?

19. Find matrices A and B for which $\text{rank}(A) = \text{rank}(B)$, but $\text{rank}\left(A^2\right) \neq \text{rank}\left(B^2\right)$.

Answer:

$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$; $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

20. Prove: If a matrix A is not square, then either the row vectors or the column vectors of A are linearly dependent.

True-False Exercises

In parts (a)–(j) determine whether the statement is true or false, and justify your answer.

(a) Either the row vectors or the column vectors of a square matrix are linearly independent.

Answer:

False

(b) A matrix with linearly independent row vectors and linearly independent column vectors is square.

Answer:

True

(c) The nullity of a nonzero $m \times n$ matrix is at most m.

Answer:

False

(d) Adding one additional column to a matrix increases its rank by one.

Answer:

False

(e) The nullity of a square matrix with linearly dependent rows is at least one.

Answer:

True

(f) If A is square and $A\mathbf{x} = \mathbf{b}$ is inconsistent for some vector $\mathbf{b}$, then the nullity of A is zero.

Answer:

False

(g) If a matrix A has more rows than columns, then the dimension of the row space is greater than the dimension of the column space.

Answer:

False

(h) If $\text{rank}\left(A^T\right) = \text{rank}\left(A\right)$, then A is square.

Answer:

False

(i) There is no 3×3 matrix whose row space and null space are both lines in 3-space.

Answer:

True

(j) If V is a subspace of R^n and W is a subspace of V, then $W^{\perp}$ is a subspace of $V^{\perp}$.

Answer:

False

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

4.9 Matrix Transformations from R^n to R^m

In this section we will study functions of the form $\mathbf{w} = F(\mathbf{x})$, where the independent variable $\mathbf{x}$ is a vector in R^n and the dependent variable $\mathbf{w}$ is a vector in R^m. We will concentrate on a special class of such functions called "matrix transformations." Such transformations are fundamental in the study of linear algebra and have important applications in physics, engineering, social sciences, and various branches of mathematics.

Functions and Transformations

Recall that a ***function*** is a rule that associates with each element of a set A one and only one element in a set B. If f associates the element b with the element a, then we write

$$b = f(a)$$

and we say that b is the ***image*** of a under f or that $f(a)$ is the ***value*** of f at a. The set A is called the ***domain*** of f and the set B the ***codomain*** of f (Figure 4.9.1). The subset of the codomain that consists of all images of points in the domain is called the ***range*** of f.

Figure 4.9.1

For many common functions the domain and codomain are sets of real numbers, but in this text we will be concerned with functions for which the domain and codomain are vector spaces.

DEFINITION 1

If V and W are vector spaces, and if f is a function with domain V and codomain W, then we say that f is a ***transformation*** from V to W or that f ***maps*** V to W, which we denote by writing

$$f : V \to W$$

In the special case where $V = W$, the transformation is also called an ***operator*** on V.

In this section we will be concerned exclusively with transformations from R^n to R^m; transformations of general vector spaces will be considered in a later section. To illustrate one way in which such transformations can arise, suppose that $f_1, f_2, \ldots, f_m$ are real-valued functions of n variables, say

$$\begin{aligned} w_1 &= f_1(x_1, x_2, \ldots, x_n) \\ w_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ w_m &= f_m(x_1, x_2, \ldots, x_n) \end{aligned} \tag{1}$$

These m equations assign a unique point $(w_1, w_2, \ldots, w_m)$ in R^m to each point $(x_1, x_2, \ldots, x_n)$ in R^n and thus define a transformation from R^n to R^m. If we denote this transformation by T, then $T: R^n \to R^m$ and

$$T(x_1, x_2, \ldots, x_n) = (w_1, w_2, \ldots, w_m)$$

Matrix Transformations

In the special case where the equations in 1 are linear, they can be expressed in the form

$$\begin{aligned} w_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ w_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ w_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{aligned} \tag{2}$$

which we can write in matrix notation as

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{3}$$

or more briefly as

$$\mathbf{w} = A\mathbf{x} \tag{4}$$

Although we could view this as a linear system, we will view it instead as a transformation that maps the column vector $\mathbf{x}$ in R^n into the column vector $\mathbf{w}$ in R^m by multiplying $\mathbf{x}$ on the left by A. We call this a ***matrix transformation*** (or ***matrix operator*** if $m = n$), and we denote it by $T_A: R^n \to R^m$. With this notation, Equation 4 can be expressed as

$$\mathbf{w} = T_A(\mathbf{x}) \tag{5}$$

The matrix transformation T_A is called ***multiplication by A***, and the matrix A is called the ***standard matrix*** for the transformation.

We will also find it convenient, on occasion, to express 5 in the schematic form

$$\mathbf{x} \xrightarrow{T_A} \mathbf{w} \tag{6}$$

which is read "T_A maps $\mathbf{x}$ into $\mathbf{w}$."

EXAMPLE 1 A Matrix Transformation from R^4 to R^3

The matrix transformation $T: R^4 \to R^3$ defined by the equations

$$\begin{aligned} w_1 &= 2x_1 - 3x_2 + x_3 - 5x_4 \\ w_2 &= 4x_1 + x_2 - 2x_3 + x_4 \\ w_3 &= 5x_1 - x_2 + 4x_3 \end{aligned} \tag{7}$$

can be expressed in matrix form as

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{8}$$

so the standard matrix for T is

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

The image of a point (x_1, x_2, x_3, x_4) can be computed directly from the defining equations 7 or from 8 by matrix multiplication. For example, if

$$(x_1, x_2, x_3, x_4) = (1, -3, 0, 2)$$

then substituting in 7 yields $w_1 = 1, w_2 = 3, w_3 = 8$ (verify), or alternatively from 8,

$$\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 2 & -3 & 1 & -5 \\ 4 & 1 & -2 & 1 \\ 5 & -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 8 \end{bmatrix}$$

Some Notational Matters

Sometimes we will want to denote a matrix transformation without giving a name to the matrix itself. In such cases we will denote the standard matrix for $T: R^n \to R^m$ by the symbol $[T]$. Thus, the equation

$$T(\mathbf{x}) = [T]\mathbf{x} \tag{9}$$

is simply the statement that T is a matrix transformation with standard matrix $[T]$, and the image of $\mathbf{x}$ under this transformation is the product of the matrix $[T]$ and the column vector $\mathbf{x}$.

Properties of Matrix Transformations

The following theorem lists four basic properties of matrix transformations that follow from properties of matrix multiplication.

THEOREM 4.9.1

For every matrix A the matrix transformation $T_A: R^n \to R^m$ has the following properties for all vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and for every scalar k:

(a) $T_A(0) = 0$

(b) $T_A(k\mathbf{u}) = kT_A(\mathbf{u})$ [Homogeneity property]

(c) $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$ [Additivity property]

(d) $T_A(\mathbf{u}-\mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$

Proof All four parts are restatements of familiar properties of matrix multiplication:

$$A\mathbf{0} = \mathbf{0},\ A(k\mathbf{u}) = k(A\mathbf{u}),\ A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v},\ A(\mathbf{u}-\mathbf{v}) = A\mathbf{u} - A\mathbf{v}$$

It follows from Theorem 4.9.1 that a matrix transformation maps linear combinations of vectors in R^n into the corresponding linear combinations in R^m in the sense that

$$T_A(k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_r\mathbf{u}_r) = k_1T_A(\mathbf{u}_1) + k_2T_A(\mathbf{u}_2) + \cdots + k_rT_A(\mathbf{u}_r) \tag{10}$$

Depending on whether n-tuples and m-tuples are regarded as vectors or points, the geometric effect of a matrix transformation $T_A: R^n \to R^m$ is to map each vector (point) in R^n into a vector (point) in R^m (Figure 4.9.2).

Figure 4.9.2

The following theorem states that if two matrix transformations from R^n to R^m have the same image at each point of R^n, then the matrices themselves must be the same.

THEOREM 4.9.2

If $T_A: R^n \to R^m$ and $T_B: R^n \to R^m$ are matrix transformations, and if $T_A(\mathbf{x}) = T_B(\mathbf{x})$ for every vector $\mathbf{x}$ in R^n, then $A = B$.

Proof To say that $T_A(\mathbf{x}) = T_B(\mathbf{x})$ for every vector in R^n is the same as saying that

$$A\mathbf{x} = B\mathbf{x}$$

for every vector $\mathbf{x}$ in R^n. This is true, in particular, if $\mathbf{x}$ is any of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n; that is,

$$A\mathbf{e}_j = B\mathbf{e}_j \quad (j = 1, 2, \ldots, n) \tag{11}$$

Since every entry of $\mathbf{e}_j$ is 0 except for the jth, which is 1, it follows from Theorem 1.3.1 that $A\mathbf{e}_j$ is the jth column of A and $B\mathbf{e}_j$ is the jth column of B. Thus, it follows from 11 that corresponding columns of A and B are the same, and hence that $A = B$.

EXAMPLE 2 Zero Transformations

If 0 is the $m \times n$ zero matrix, then

$$T_0(\mathbf{x}) = 0\mathbf{x} = 0$$

so multiplication by zero maps every vector in R^n into the zero vector in R^m. We call T_0 the ***zero transformation*** from R^n to R^m.

EXAMPLE 3 Identity Operators

If I is the $n \times n$ identity matrix, then

$$T_I(\mathbf{x}) = I\mathbf{x} = \mathbf{x}$$

so multiplication by I maps every vector in R^n into itself. We call T_I the ***identity operator*** on R^n.

A Procedure for Finding Standard Matrices

There is a way of finding the standard matrix for a matrix transformation from R^n to R^m by considering the effect of that transformation on the standard basis vectors for R^n. To explain the idea, suppose that A is unknown and that

$$\mathbf{e}_1,\ \mathbf{e}_2, \ldots,\ \mathbf{e}_n$$

are the standard basis vectors for R^n. Suppose also that the images of these vectors under the transformation T_A are

$$T_A(\mathbf{e}_1) = A\mathbf{e}_1,\ T_A(\mathbf{e}_2) = A\mathbf{e}_2, \ldots,\ T_A(\mathbf{e}_n) = A\mathbf{e}_n$$

It follows from Theorem 1.3.1 that $A\mathbf{e}_j$ is a linear combination of the columns of A in which the successive coefficients are the entries of $\mathbf{e}_j$. But all entries of $\mathbf{e}_j$ are zero except the jth, so the product $A\mathbf{e}_j$ is just the jth column of the matrix A. Thus,

$$A = [T_A(\mathbf{e}_1)|T_A(\mathbf{e}_2)|\cdots|T_A(\mathbf{e}_n)] \tag{12}$$

In summary, we have the following procedure for finding the standard matrix for a matrix transformation:

Finding the Standard Matrix for a Matrix Transformation

Step 1. Find the images of the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ for R^n in column form.

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.

Reflection Operators

Some of the most basic matrix operators on R^2 and R^3 are those that map each point into its symmetric image about a fixed line or a fixed plane; these are called ***reflection operators***. Table 1 shows the standard matrices for the reflections about the coordinate axes in R^2, and Table 2 shows the standard matrices for the reflections about the coordinate planes in R^3. In each case the standard matrix was obtained by finding the images of the standard basis vectors, converting those images to column vectors, and then using those column vectors as successive columns of the standard matrix.

Table 1

Operator	**Illustration**	**Images of e_1 and e_2**	**Standard Matrix**
Reflection about the y-acis $T(x, y) = (-x, y)$		$T(e_1) = T(1, 0) = (-1, 0)$ $T(e_2) = T(0, 1) = (0, 1)$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection about the x-axis $T(x, y) = (x, -y)$		$T(e_1) = T(1, 0) = (1, 0)$ $T(e_2) = T(0, 1) = (0, -1)$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection about the line $y = x$ $T(x, y) = (y, x)$		$T(e_1) = T(1, 0) = (0, 1)$ $T(e_2) = T(0, 1) = (1, 0)$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Table 2

Operator	**Illustration**	**e_1, e_2, e_3**	**Standard Matrix**
Reflection about the xy-plane $T(x, y, z) = (x, y, -z)$		$T(e_1) = T(1, 0, 0) = (1, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, 1, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, -1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection about the xz-plane $T(x, y, z) = (x, -y, z)$		$T(e_1) = T(1, 0, 0) = (1, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, -1, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, 1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Reflection about the yz-plane $T(x, y, z) = (-x, y, z)$		$T(e_1) = T(1, 0, 0) = (-1, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, 1, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, 1)$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Projection Operators

Matrix operators on R^2 and R^3 that map each point into its orthogonal projection on a fixed line or plane are called ***projection operators*** (or more precisely, ***orthogonal projection*** operators). Table 3 shows the standard matrices for the orthogonal projections on the coordinate axes in R^2, and Table 4 shows the standard matrices for the orthogonal projections on the coordinate planes in R^3.

Table 3

Operator	**Illustration**	**Images of e_1 and e_2**	**Standard Matrix**
Orthogonal projection on the x-axis $T(x, y) = (x, 0)$		$T(e_1) = T(1, 0) = (1, 0)$ $T(e_2) = T(0, 1) = (0, 0)$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Orthogonal projection on the y-axis $T(x, y) = (0, y)$		$T(e_1) = T(1, 0) = (0, 0)$ $T(e_2) = T(0, 1) = (0, 1)$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

Table 4

Operator	**Illustration**	**Images of e_1, e_2, e_3**	**Standard Matrix**
Orthogonal projection on the xy-plane $T(x, y, z) = (x, y, 0)$		$T(e_1) = T(1, 0, 0) = (1, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, 1, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, 0)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Orthogonal projection on the xz-plane $T(x, y, z) = (x, 0, z)$		$T(e_1) = T(1, 0, 0) = (1, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, 0, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, 1)$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Orthogonal projection on the yz-plane $T(x, y, z) = (0, y, z)$		$T(e_1) = T(1, 0, 0) = (0, 0, 0)$ $T(e_2) = T(0, 1, 0) = (0, 1, 0)$ $T(e_3) = T(0, 0, 1) = (0, 0, 1)$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Rotation Operators

Matrix operators on R^2 and R^3 that move points along circular arcs are called ***rotation operators***. Let us consider how to find the standard matrix for the rotation operator $T: R^2 \to R^2$ that moves points *counterclockwise* about the origin through an angle θ (Figure 4.9.3). As illustrated in Figure 4.9.3, the images of the standard basis vectors are

$$T(\mathbf{e}_1) = T(1, 0) = (\cos\theta, \sin\theta) \text{ and } T(\mathbf{e}_2) = T(0, 1) = (-\sin\theta, \cos\theta)$$

so the standard matrix for T is

$$\left[T(\mathbf{e}_1) \middle| T(\mathbf{e}_2)\right] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Figure 4.9.3

In keeping with common usage we will denote this operator by R_θ and call

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{13}$$

the ***rotation matrix*** for R^2. If $\mathbf{x} = (x, y)$ is a vector in R^2, and if $\mathbf{w} = (w_1, w_2)$ is its image under the rotation, then the relationship $\mathbf{w} = R_\theta \mathbf{x}$ can be written in component form as

$$\begin{aligned} w_1 &= x\cos\theta - y\sin\theta \\ w_2 &= x\sin\theta + y\cos\theta \end{aligned} \tag{14}$$

These are called the ***rotation equations*** for R^2. These ideas are summarized in Table 5.

Table 5

Operator	**Illustration**	**Rotation Equations**	**Standard Matrix**
Rotation through an angle θ		$w_1 = x\cos\theta - y\sin\theta$ $w_2 = x\sin\theta + y\cos\theta$	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

In the plane, counterclockwise angles are positive and clockwise angles are negative. The rotation matrix for a *clockwise* rotation of $-\theta$ radians can be obtained by replacing θ by $-\theta$ in 12. After simplification this yields

$$R_{-\theta} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

EXAMPLE 4 A Rotation Operator

Find the image of $\mathbf{x} = (1, 1)$ under a rotation of $\pi / 6$ radians $\left(= 30^\circ\right)$ about the origin.

Solution It follows from 13 with $\theta = \pi / 6$ that

$$R_{\pi/6}\mathbf{x} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}-1}{2} \\ \frac{1+\sqrt{3}}{2} \end{bmatrix} \approx \begin{bmatrix} 0.37 \\ 1.37 \end{bmatrix}$$

or in comma-delimited notation, $R_{\pi/6}(1, 1) = (0.37, 1.37)$.

Rotations in R^3

A rotation of vectors in R^3 is usually described in relation to a ray emanating from the origin, called the ***axis of rotation***. As a vector revolves around the axis of rotation, it sweeps out some portion of a cone (Figure 4.9.4*a*). The ***angle of rotation***, which is measured in the base of the cone, is described as "clockwise" or "counterclockwise" in relation to a viewpoint that is along the axis of rotation *looking toward the origin*. For example, in Figure 4.9.4*a* the vector $\mathbf{w}$ results from rotating the vector $\mathbf{x}$ counterclockwise around the axis l through an angle θ. As in R^2, angles are *positive* if they are generated by counterclockwise rotations and *negative* if they are generated by clockwise rotations.

Figure 4.9.4

The most common way of describing a general axis of rotation is to specify a nonzero vector $\mathbf{u}$ that runs along the axis of rotation and has its initial point at the origin. The counterclockwise direction for a rotation about the axis can then be determined by a "right-hand rule" (Figure 4.9.4*b*): If the thumb of the right hand points in the direction of $\mathbf{u}$, then the cupped fingers point in a counterclockwise direction.

A ***rotation operator*** on R^3 is a matrix operator that rotates each vector in R^3 about some rotation axis through a fixed angle θ. In Table 6 we have described the rotation operators on R^3 whose axes of rotation are the positive coordinate axes. For each of these rotations one of the components is unchanged, and the relationships between the other components can be derived by the same procedure used to derive 14. For example, in the rotation about the z-axis, the z-components of $\mathbf{x}$ and $\mathbf{w} = T(\mathbf{x})$ are the same, and the x- and y-components are related as in 14. This yields the rotation equation shown in the last row of Table 6.

Table 6

Operator	Illustration	Rotation Equations	Standard Matrix
Counterclockwise rotation about the positive x-axis through an angle θ		$w_1 = x$ $w_2 = y\cos\theta - z\sin\theta$ $w_3 = y\sin\theta + z\cos\theta$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$
Counterclockwise rotation about the positive y-axis through an angle θ		$w_1 = x\cos\theta + z\sin\theta$ $w_2 = y$ $w_3 = -x\sin\theta + z\cos\theta$	$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$
Counterclockwise rotation about the positive z-axis through an angle θ		$w_1 = x\cos\theta - y\sin\theta$ $w_2 = x\sin\theta + y\cos\theta$ $w_3 = z$	$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$

For completeness, we note that the standard matrix for a counterclockwise rotation through an angle θ about an axis in R^3, which is determined by an arbitrary *unit vector* $\mathbf{u} = (a, b, c)$ that has its initial point at the origin, is

$$\begin{bmatrix} a^2\left(1-\cos\theta\right)+\cos\theta & \mathrm{ab}(1-\cos\theta)-c\sin\theta & \mathrm{ac}(1-\cos\theta)+b\sin\theta \\ \mathrm{ab}(1-\cos\theta)+c\sin\theta & b^2\left(1-\cos\theta\right)+\cos\theta & \mathrm{bc}(1-\cos\theta)-a\sin\theta \\ \mathrm{ac}(1-\cos\theta)-b\sin\theta & \mathrm{bc}(1-\cos\theta)+a\sin\theta & c^2\left(1-\cos\theta\right)+\cos\theta \end{bmatrix} \tag{15}$$

The derivation can be found in the book *Principles of Interactive Computer Graphics*, by W. M. Newman and R. F. Sproull (New York: McGraw-Hill, 1979). You may find it instructive to derive the results in Table 6 as special cases of this more general result.

Dilations and Contractions

If k is a nonnegative scalar, then the operator $T(\mathbf{x}) = k\mathbf{x}$ on R^2 or R^3 has the effect of increasing or decreasing the length of each vector by a factor of k. If $0 \le k < 1$ the operator is called a ***contraction*** with factor k, and if $k > 1$ it is

called a ***dilation*** with factor k (Figure 4.9.5). If $k = 1$, then T is the identity operator and can be regarded either as a contraction or a dilation. Tables 7 and 8 illustrate these operators.

Figure 4.9.5

Table 7

Operator	**Illustration** $T(x, y) = (kx, ky)$	**Effect on the Standard Basis**	**Standard Matrix**
Contraction with factor k on R^2 $(0 \le k < 1)$			$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$
Dilation with factor k on R^2 $(k > 1)$			

Table 8

Operator	**Illustration** $T(x, y, z) = (kx, ky, kz)$	**Standard Matrix**
Contraction with factor k on R^3 $(0 \le k \le 1)$		$\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}$
Dilation with factor k on R^3 $(k \ge 1)$		

Yaw, Pitch, and Roll

In aeronautics and astronautics, the orientation of an aircraft or space shuttle relative to an xyz-coordinate system is often described in terms of angles called ***yaw***, ***pitch***, and ***roll***. If, for example, an aircraft is flying

along the y-axis and the xy-plane defines the horizontal, then the aircraft's angle of rotation about the z-axis is called the ***yaw***, its angle of rotation about the x-axis is called the ***pitch***, and its angle of rotation about the y-axis is called the ***roll***. A combination of yaw, pitch, and roll can be achieved by a single rotation about some axis through the origin. This is, in fact, how a space shuttle makes attitude adjustments—it doesn't perform each rotation separately; it calculates one axis, and rotates about that axis to get the correct orientation. Such rotation maneuvers are used to align an antenna, point the nose toward a celestial object, or position a payload bay for docking.

Expansion and Compressions

In a dilation or contraction of R^2 or R^3, all coordinates are multiplied by a factor k. If only one of the coordinates is multiplied by k, then the resulting operator is called an ***expansion*** or ***compression*** with factor k. This is illustrated in Table 9 for R^2. You should have no trouble extending these results to R^3.

Table 9

Operator	**Illustration** $T(x, y) = (kx, y)$	**Effect on the Standard Basis**	**Standard Matrix**
Compression of R^2 in the x-direction with factor k $(0 \le k < 1)$			$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
Expansion of R^2 in the x-direction with factor k $(k > 1)$			
Operator	**Illustration** $T(x, y) = (x, ky)$	**Effect on the Standard Basis**	**Standard Matrix**
Compression of R^2 in the y-direction with factor k $(0 \le k < 1)$			$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

Operator	Illustration $T(x, y) = (kx, y)$	Effect on the Standard Basis	Standard Matrix
Expansion of R^2 in the y-direction with factor k $(k > 1)$			

Shears

A matrix operator of the form $T(x, y) = (x + ky, y)$ translates a point (x, y) in the xy-plane parallel to the x-axis by an amount ky that is proportional to the y-coordinate of the point. This operator leaves the points on the x-axis fixed (since $y = 0$), but as we progress away from the x-axis, the translation distance increases. We call this operator the ***shear in the x-direction with factor k***. Similarly, a matrix operator of the form $T(x, y) = (x, y + kx)$ is called the ***shear in the y-direction with factor k***. Table 10 illustrates the basic information about shears in R^2.

Table 10

Operator	Effect on the Standard Basis	Standard Matrix
Shear of R^2 in the x-direction with factor k $T(x, y) = (x + ky, y)$		$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Shear of R^2 in the y-direction with factor k $T(x, y) = (x, y + kx)$		$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

EXAMPLE 5 Some Basic Matrix Operators on R^2

In each part describe the matrix operator corresponding to A, and show its effect on the unit square.

(a) $A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ (b) $A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ (c) $A_3 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

Solution By comparing the forms of these matrices to those in Tables 7, 9, and 10, we see that the matrix A_1 corresponds to a shear in the x-direction with factor 2, the matrix A_2 corresponds to a dilation with factor 2, and A_3 corresponds to an expansion in the x-direction with factor 2. The effects of these operators on the unit square are shown in Figure 4.9.6.

Figure 4.9.6

OPTIONAL

Orthogonal Projections on Lines Through the Origin

In Table 3 we listed the standard matrices for the orthogonal projections on the coordinate axes in R^2. These are special cases of the more general operator $T: R^2 \to R^2$ that maps each point into its orthogonal projection on a line L through the origin that makes an angle θ with the positive x-axis (Figure 4.9.7). In Example 4 of Section 3.3 we used Formula 10 of that section to find the orthogonal projections of the standard basis vectors for R^2 on that line. Expressed in matrix form, we found those projections to be

$$T(\mathbf{e}_1) = \begin{bmatrix} \cos^2\theta \\ \sin\theta\cos\theta \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} \sin\theta\cos\theta \\ \sin^2\theta \end{bmatrix}$$

Figure 4.9.7

Thus, the standard matrix for T is

$$[T] = \left[T(\mathbf{e}_1) \middle| T(\mathbf{e}_2) \right] = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \sin^2\theta \end{bmatrix}$$

In keeping with common usage, we will denote this operator by

$$P_\theta = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \sin^2\theta \end{bmatrix} \tag{16}$$

We have included two versions of Formula 16 because both are commonly used. Whereas the first version involves only the angle θ, the second involves both θ and 2θ.

EXAMPLE 6 Orthogonal Projection on a Line Through the Origin

Use Formula 16 to find the orthogonal projection of the vector $\mathbf{x} = (1, 5)$ on the line through the origin that makes an angle of $\pi / 6 \quad \left(= 30^\circ\right)$ with the x-axis.

Solution Since $\sin(\pi / 6) = 1 / 2$ and $\cos\left(\pi / 6\right) = \sqrt{3} / 2$, it follows from 16 that the standard matrix for this projection is

$$P_{\pi/6} = \begin{bmatrix} \cos^2\left(\pi / 6\right) & \sin(\pi / 6)\cos(\pi / 6) \\ \sin(\pi / 6)\cos(\pi / 6) & \sin^2\left(\pi / 6\right) \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{1}{4} \end{bmatrix}$$

Thus,

$$P_{\pi/6}\mathbf{x} = \begin{bmatrix} \frac{3}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{3 + 5\sqrt{3}}{4} \\ \frac{\sqrt{3} + 5}{4} \end{bmatrix} \approx \begin{bmatrix} 2.91 \\ 1.68 \end{bmatrix}$$

or in comma-delimited notation, $P_{\pi/6}(1, 5) \approx (2.91, 1.68)$

Reflections About Lines Through the Origin

In Table 1 we listed the reflections about the coordinate axes in R^2. These are special cases of the more general operator $H_\theta : R^2 \to R^2$ that maps each point into its reflection about a line L through the origin that makes an angle θ with the positive x-axis (Figure 4.9.8). We could find the standard matrix for H_θ by finding the images of the standard basis vectors, but instead we will take advantage of our work on orthogonal projections by using the Formula 16 for P_θ to find a formula for H_θ.

Figure 4.9.8

You should be able to see from Figure 4.9.9 that for every vector $\mathbf{x}$ in R^n

$$P_\theta\mathbf{x} - \mathbf{x} = \frac{1}{2}\left(H_\theta\mathbf{x} - \mathbf{x}\right) \quad \text{or equivalently} \quad H_\theta\mathbf{x} = \left(2P_\theta - I\right)\mathbf{x}$$

Figure 4.9.9

Thus, it follows from Theorem 4.9.2 that

$$H_\theta = 2P_\theta - I \tag{17}$$

and hence from 16 that

$$H_\theta = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \tag{18}$$

EXAMPLE 7 Reflection About a Line Through the Origin

Find the reflection of the vector **x** = (1, 5) on the line through the origin that makes an angle of π/6(= 30°) with the x-axis.

Solution Since $\sin(\pi/3) = \sqrt{3}/2$ and $\cos(\pi/3) = 1/2$, it follows from 18 that the standard matrix for this projection is

$$H_{\pi/6} = \begin{bmatrix} \cos(\pi/3) & \sin(\pi/3) \\ \sin(\pi/3) & -\cos(\pi/3) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$$

Thus,

$$H_{\pi/6}\mathbf{x} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{1+5\sqrt{3}}{2} \\ \frac{\sqrt{3}-5}{2} \end{bmatrix} \approx \begin{bmatrix} 4.83 \\ -1.63 \end{bmatrix}$$

or in comma-delimited notation, $H_{\pi/6}(1, 5) \approx (4.83, -1.63)$

Show that the standard matrices in Tables 1 and 3 are special cases of 18 and 16.

Concept Review

- Function
- Image

- Value
- Domain
- Codomain
- Transformation
- Relationships among the fundamental spaces
- Operator
- Matrix transformation
- Matrix operator
- Standard matrix
- Properties of matrix transformations
- Zero transformation
- Identity operator
- Reflection operator
- Projection operator
- Rotation operator
- Rotation matrix
- Rotation equations
- Axis of rotation in 3-space
- Angle of rotation in 3-space
- Expansion operator
- Compression operator
- Shear
- Dilation
- Contraction

Skills

- Find the domain and codomain of a transformation, and determine whether the transformation is linear.
- Find the standard matrix for a matrix transformation.
- Describe the effect of a matrix operator on the standard basis in R^n.

Exercise Set 4.9

In Exercises 1–2, find the domain and codomain of the transformation $T_A(\mathbf{x}) = A\mathbf{x}$.

1. (a) A has size 3×2.
(b) A has size 2×3.
(c) A has size 3×3.
(d) A has size 1×6.

Answer:

(a) Domain: R^2; codomain: R^3

(b) Domain: R^3; codomain: R^2

(c) Domain: R^3; codomain: R^3

(d) Domain: R^6; codomain: R^1

2. (a) A has size 4×5.

(b) A has size 5×4.

(c) A has size 4×4.

(d) A has size 3×1.

3. If $T(x_1, x_2) = (x_1 + x_2, -x_2, 3x_1)$, then the domain of T is _____ , the codomain of T is _____ , and the image of $\mathbf{x} = (1, -2)$ under T is _____ .

Answer:

$R^2,\ R^3,\ (-1, 2, 3)$

4. If $T(x_1, x_2, x_3) = (x_1 + 2x_2, x_1 - 2x_2)$, then the domain of T is _____ , the codomain of T is _____ , and the image of $\mathbf{x} = (0, -1, 4)$ under T is _____ .

5. In each part, find the domain and codomain of the transformation defined by the equations, and determine whether the transformation is linear.

(a) $w_1 = 3x_1 - 2x_2 + 4x_3$
$w_2 = 5x_1 - 8x_2 + x_3$

(b) $w_1 = 2x_1x_2 - x_2$
$w_2 = x_1 + 3x_1x_2$
$w_3 = x_1 + x_2$

(c) $w_1 = 5x_1 - x_2 + x_3$
$w_2 = -x_1 + x_2 + 7x_3$
$w_3 = 2x_1 - 4x_2 - x_3$

(d) $w_1 = x_1^2 - 3x_2 + x_3 - 2x_4$
$w_2 = 3x_1 - 4x_2 - x_3^2 + x_4$

Answer:

(a) Linear; $R^3 \to R^2$

(b) Nonlinear; $R^2 \to R^3$

(c) Linear; $R^3 \to R^3$

(d) Nonlinear; $R^4 \to R^2$

6. In each part, determine whether T is a matrix transformation.

(a) $T(x, y) = (2x, y)$

(b) $T(x, y) = (-y, x)$

(c) $T(x, y) = (2x + y, x - y)$

(d) $T\left(x, y\right) = \left(x^2, y\right)$

(e) $T(x, y) = (x, y + 1)$

7. In each part, determine whether T is a matrix transformation.

(a) $T(x, y, z) = (0, 0)$

(b) $T(x, y, z) = (1, 1)$

(c) $T(x, y, z) = (3x - 4y, 2x - 5z)$

(d) $T(x, y, z) = (y^2, z)$

(e) $T(x, y, z) = (y - 1, x)$

Answer:

(a) and (c) are matrix transformations; (b), (d), and (e) are not matrix transformations.

8. Find the standard matrix for the transformation defined by the equations.

(a) $$\begin{aligned} w_1 &= 2x_1 - 3x_2 + x_4 \\ w_2 &= 3x_1 + 5x_2 - x_4 \end{aligned}$$

(b) $$\begin{aligned} w_1 &= 7x_1 + 2x_2 - 8x_3 \\ w_2 &= -x_2 + 5x_3 \\ w_3 &= 4x_1 + 7x_2 - x_3 \end{aligned}$$

(c) $$\begin{aligned} w_1 &= -x_1 + x_2 \\ w_2 &= 3x_1 - 2x_2 \\ w_3 &= 5x_1 - 7x_2 \end{aligned}$$

(d) $$\begin{aligned} w_1 &= x_1 \\ w_2 &= x_1 + x_2 \\ w_3 &= x_1 + x_2 + x_3 \\ w_4 &= x_1 + x_2 + x_3 + x_4 \end{aligned}$$

9. Find the standard matrix for the operator $T: R^3 \to R^3$ defined by

$$\begin{aligned} w_1 &= 3x_1 + 5x_2 - x_3 \\ w_2 &= 4x_1 - x_2 + x_3 \\ w_3 &= 3x_1 + 2x_2 - x_3 \end{aligned}$$

and then calculate $T(-1, 2, 4)$ by directly substituting in the equations and also by matrix multiplication.

Answer:

$$\begin{bmatrix} 3 & 5 & -1 \\ 4 & -1 & 1 \\ 3 & 2 & -1 \end{bmatrix}; T(-1, 2, 4) = (3, -2, -3)$$

10. 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) $T(x_1, x_2) = (x_1, x_2)$

(c) $T(x_1, x_2, x_3) = (x_1 + 2x_2 + x_3, x_1 + 5x_2, x_3)$

(d) $T(x_1, x_2, x_3) = (4x_1, 7x_2, -8x_3)$

11. Find the standard matrix for the transformation T defined by the formula.

(a) $T(x_1, x_2) = (x_2, -x_1, x_1 + 3x_2, x_1 - x_2)$

(b) $T(x_1, x_2, x_3, x_4) = (7x_1 + 2x_2 - x_3 + x_4, x_2 + x_3, -x_1)$

(c) $T(x_1, x_2, x_3) = (0, 0, 0, 0, 0)$

(d) $T(x_1, x_2, x_3, x_4) = (x_4, x_1, x_3, x_2, x_1 - x_3)$

Answer:

(a) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 3 \\ 1 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 7 & 2 & -1 & 1 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix}$

12. In each part, find $T(x)$, and express the answer in matrix form.

(a) $\begin{bmatrix} T \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}; \ \mathbf{x} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

(b) $\begin{bmatrix} T \end{bmatrix} = \begin{bmatrix} -1 & 2 & 0 \\ 3 & 1 & 5 \end{bmatrix}; \mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}$

(c) $\begin{bmatrix} T \end{bmatrix} = \begin{bmatrix} -2 & 1 & 4 \\ 3 & 5 & 7 \\ 6 & 0 & -1 \end{bmatrix}; \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

(d) $\begin{bmatrix} T \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 2 & 4 \\ 7 & 8 \end{bmatrix}; \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

13. In each part, use the standard matrix for T to find $T(\mathbf{x})$; then check the result by calculating $T(\mathbf{x})$ directly.

(a) $T(x_1, x_2) = (-x_1 + x_2, x_2); \mathbf{x} = (-1, 4)$

(b) $T(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, x_2 + x_3, 0); \mathbf{x} = (2, 1, -3)$

Answer:

(a) $T(-1, 4) = (5, 4)$

(b) $T(2, 1, -3) = (0, -2, 0)$

14. Use matrix multiplication to find the reflection of $(-1, 2)$ about

(a) the x-axis.

(b) the y-axis.

(c) the line $y = x$.

15. Use matrix multiplication to find the reflection of $(2, -5, 3)$ about

(a) the xy-plane.

(b) the xz-plane.

(c) the yz-plane.

Answer:

(a) $(2, -5, -3)$

(b) (2, 5, 3)

(c) $(-2, -5, 3)$

16. Use matrix multiplication to find the orthogonal projection of $(2, -5)$ on

(a) the x-axis.

(b) the y-axis.

17. Use matrix multiplication to find the orthogonal projection of $(-2, 1, 3)$ on

(a) the xy-plane.

(b) the xz-plane.

(c) the yz-plane.

Answer:

(a) $(-2, 1, 0)$

(b) $(-2, 0, 3)$

(c) (0, 1, 3)

18. Use matrix multiplication to find the image of the vector $(3, -4)$ when it is rotated through an angle of

(a) $\theta = 30^\circ$.

(b) $\theta = -60^\circ$.

(c) $\theta = 45^\circ$.

(d) $\theta = 90^\circ$.

19. Use matrix multiplication to find the image of the vector $(-2, 1, 2)$ if it is rotated

(a) 30° about the x-axis.

(b) 45° about the y-axis.

(c) 90° about the z-axis.

Answer:

(a) $\left(-2, \dfrac{\sqrt{3}-2}{2}, \dfrac{1+2\sqrt{3}}{2}\right)$

(b) $\left(0, 1, 2\sqrt{2}\right)$

(c) $(-1, -2, 2)$

20. Find the standard matrix for the operator that rotates a vector in R^3 through an angle of -60° about

(a) the x-axis.

(b) the y-axis.

(c) the z-axis.

21. Use matrix multiplication to find the image of the vector $(-2, 1, 2)$ if it is rotated

(a) -30° about the x-axis.

(b) -45° about the y-axis.

(c) -90° about the z-axis.

Answer:

(a) $\left(-2, \dfrac{\sqrt{3}+2}{2}, \dfrac{-1+2\sqrt{3}}{2}\right)$

(b) $\left(-2\sqrt{2}, 1, 0\right)$

(c) $(1, 2, 2)$

22. In R^3 the ***orthogonal projections*** on the x-axis, y-axis, and z-axis are defined by

$$T_1(x, y, z) = (x, 0, 0), \quad T_2(x, y, z) = (0, y, 0),$$
$$T_3(x, y, z) = (0, 0, z)$$

respectively.

(a) Show that the orthogonal projections on the coordinate axes are matrix operators, and find their standard matrices.

(b) Show that if $T: R^3 \to R^3$ is an orthogonal projection on one of the coordinate axes, then for every vector $\mathbf{x}$ in R^3, the vectors $T(\mathbf{x})$ and $\mathbf{x} - T(\mathbf{x})$ are orthogonal.

(c) Make a sketch showing $\mathbf{x}$ and $\mathbf{x} - T(\mathbf{x})$ in the case where T is the orthogonal projection on the x-axis.

23. Use Formula 15 to derive the standard matrices for the rotations about the x-axis, y-axis, and z-axis in R^3.

24. Use Formula 15 to find the standard matrix for a rotation of $\pi / 2$ radians about the axis determined by the vector $\mathbf{v} = (1, 1, 1)$. [Note: Formula 15 requires that the vector defining the axis of rotation have length 1.]

25. Use Formula 15 to find the standard matrix for a rotation of 180° about the axis determined by the vector $\mathbf{v} = (2, 2, 1)$. [Note: Formula 15 requires that the vector defining the axis of rotation have length 1.]

Answer:

$$\begin{bmatrix} -\frac{1}{9} & \frac{8}{9} & \frac{4}{9} \\ \frac{8}{9} & -\frac{1}{9} & \frac{4}{9} \\ \frac{4}{9} & \frac{4}{9} & -\frac{7}{9} \end{bmatrix}$$

26. It can be proved that if A is a 2×2 matrix with orthonormal column vectors and for which $\det(A) = 1$, then multiplication by A is a rotation through some angle θ. Verify that

$$A = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

satisfies the stated conditions and find the angle of rotation.

27. The result stated in Exercise 26 can be extended to R^3; that is, it can be proved that if A is a 3×3 matrix with orthonormal column vectors and for which $\det(A) = 1$, then multiplication by A is a rotation about some axis through some angle θ. Use Formula 15 to show that the angle of rotation satisfies the equation

$$\cos\theta = \frac{\operatorname{tr}(A) - 1}{2}$$

28. Let A be a 3×3 matrix (other than the identity matrix) satisfying the conditions stated in Exercise 27. It can be shown that if $\mathbf{x}$ is any nonzero vector in R^3, then the vector $\mathbf{u} = A\mathbf{x} + A^T\mathbf{x} + \left[1 - \operatorname{tr}(A)\right]\mathbf{x}$ determines an axis of rotation when $\mathbf{u}$ is positioned with its initial point at the origin. [See "The Axis of Rotation: Analysis, Algebra, Geometry," by Dan Kalman, *Mathematics Magazine*, Vol. 62, No. 4, October 1989.]

(a) Show that multiplication by

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

is a rotation.

(b) Find a vector of length 1 that defines an axis for the rotation.

(c) Use the result in Exercise 27 to find the angle of rotation about the axis obtained in part (b).

29. In words, describe the geometric effect of multiplying a vector $\mathbf{x}$ by the matrix A.

(a) $A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

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

Answer:

(a) Twice the orthogonal projection on the x-axis.

(b) Twice the reflection about the x-axis.

30. In words, describe the geometric effect of multiplying a vector $\mathbf{x}$ by the matrix A.

(a) $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

(b)

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

31. In words, describe the geometric effect of multiplying a vector $\mathbf{x}$ by the matrix

$$A = \begin{bmatrix} \cos^2\theta - \sin^2\theta & -2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \cos^2\theta - \sin^2\theta \end{bmatrix}$$

Answer:

Rotation through the angle 2θ.

32. If multiplication by A rotates a vector $\mathbf{x}$ in the xy-plane through an angle θ, what is the effect of multiplying $\mathbf{x}$ by A^T? Explain your reasoning.

33. Let $\mathbf{x}_0$ be a nonzero column vector in R^2, and suppose that $T: R^2 \to R^2$ is the transformation defined by the formula $T(\mathbf{x}) = \mathbf{x}_0 + R_\theta \mathbf{x}$, where R_θ is the standard matrix of the rotation of R^2 about the origin through the angle θ. Give a geometric description of this transformation. Is it a matrix transformation? Explain.

Answer:

Rotation through the angle θ and translation by $\mathbf{x}_0$; not a matrix transformation since $\mathbf{x}_0$ is nonzero.

34. A function of the form $f(x) = mx + b$ is commonly called a "linear function" because the graph of $y = mx + b$ is a line. Is f a matrix transformation on R?

35. Let $\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$ be a line in R^n, and let $T: R^n \to R^n$ be a matrix operator on R^n. What kind of geometric object is the image of this line under the operator T? Explain your reasoning.

Answer:

A line in R^n.

True-False Exercises

In parts (a)–(i) determine whether the statement is true or false, and justify your answer.

(a) If A is a 2×3 matrix, then the domain of the transformation T_A is R^2.

Answer:

False

(b) If A is an $m \times n$ matrix, then the codomain of the transformation T_A is R^n.

Answer:

False

(c) If $T: R^n \to R^m$ and $T(0) = 0$, then T is a matrix transformation.

Answer:

False

(d) If $T: R^n \to R^m$ and $T(c_1\mathbf{x} + c_2\mathbf{y}) = c_1T(\mathbf{x}) + c_2T(\mathbf{y})$ for all scalars c_1 and c_2 and all vectors $\mathbf{x}$ and $\mathbf{y}$ in R^n, then T is a matrix transformation.

Answer:

True

(e) There is only one matrix transformation $T: R^n \to R^m$ such that $T(-\mathbf{x}) = -T(\mathbf{x})$ for every vector $\mathbf{x}$ in R^n.

Answer:

False

(f) There is only one matrix transformation $T: R^n \to R^m$ such that $T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}-\mathbf{y})$ for all vectors $\mathbf{x}$ and $\mathbf{y}$ in R^n.

Answer:

True

(g) If $\mathbf{b}$ is a nonzero vector in R^n, then $T(\mathbf{x}) = \mathbf{x}+\mathbf{b}$ is a matrix operator on R^n.

Answer:

False

(h) The matrix $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ is the standard matrix for a rotation.

Answer:

False

(i) The standard matrices of the reflections about the coordinate axes in 2-space have the form $\begin{bmatrix} a & 0 \\ 0 & -a \end{bmatrix}$, where $a = \pm 1$.

Answer:

True

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

4.10 Properties of Matrix Transformations

In this section we will discuss properties of matrix transformations. We will show, for example, that if several matrix transformations are performed in succession, then the same result can be obtained by a single matrix transformation that is chosen appropriately. We will also explore the relationship between the invertibility of a matrix and properties of the corresponding transformation.

Compositions of Matrix Transformations

Suppose that T_A is a matrix transformation from R^n to R^k and T_B is a matrix transformation from R^k to R^m. If x is a vector in R^n, then T_A maps this vector into a vector $T_A(\mathbf{x})$ in R^k, and T_B, in turn, maps that vector into the vector $T_B(T_A(\mathbf{x}))$ in R^m. This process creates a transformation from R^n to R^m that we call the ***composition of T_B with T_A*** and denote by the symbol

$$T_B \circ T_A$$

which is read "T_B circle T_A". As illustrated in Figure 4.10.1, the transformation T_A in the formula is performed first; that is,

$$(T_B \circ T_A)(\mathbf{x}) = T_B(T_A(\mathbf{x})) \tag{1}$$

This composition is itself a matrix transformation since

$$(T_B \circ T_A)(\mathbf{x}) = T_B(T_A(\mathbf{x})) = B(T_A(\mathbf{x})) = B(A\mathbf{x}) = (BA)\mathbf{x}$$

which shows that it is multiplication by BA. This is expressed by the formula

$$T_B \circ T_A = T_{BA} \tag{2}$$

WARNING

Just as it is *not* true, in general, that

$$AB = BA$$

so it is *not* true, in general, that

$$T_B \circ T_A = T_A \circ T_B$$

That is, *order matters when matrix transformations are composed*.

Figure 4.10.1

Compositions can be defined for any finite succession of matrix transformations whose domains and ranges have

the appropriate dimensions. For example, to extend Formula 2 to three factors, consider the matrix transformations

$$T_A: R^n \to R^k,\ T_B: R^k \to R^l,\ T_C: R^l \to R^m$$

We define the composition $(T_C \circ T_B \circ T_A) R^n \to R^m$ by

$$(T_C \circ T_B \circ T_A)(\mathbf{x}) = T_C(T_B(T_A(\mathbf{x})))$$

As above, it can be shown that this is a matrix transformation whose standard matrix is CBA and that

$$T_C \circ T_B \circ T_A = T_{CBA} \tag{3}$$

As in Formula 9 of Section 4.9 , we can use square brackets to denote a matrix transformation without referencing a specific matrix. Thus, for example, the formula

$$[T_2 \circ T_1] = [T_2][T_1] \tag{4}$$

is a restatement of Formula 2 which states that the standard matrix for a composition is the product of the standard matrices in the appropriate order. Similarly,

$$[T_3 \circ T_2 \circ T_1] = [T_3][T_2][T_1] \tag{5}$$

is a restatement of Formula 3.

EXAMPLE 1 Composition of Two Rotations

Let $T_1 R^2 \to R^2$ and $T_2 R^2 \to R^2$ be the matrix operators that rotate vectors through the angles θ_1 and θ_2, respectively. Thus the operation

$$(T_2 \circ T_1)(\mathbf{x}) = T_2(T_1(\mathbf{x}))$$

first rotates $\mathbf{x}$ through the angle θ_1, then rotates $T_1(\mathbf{x})$ through the angle θ_2. It follows that the net effect of $T_2 \circ T_1$ is to rotate each vector in R^2 through the angle $\theta_1 + \theta_2$ (Figure 4.10.2). Thus, the standard matrices for these matrix operators are

$$[T_1] = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}, \quad \begin{bmatrix} T_2 \end{bmatrix} = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix},$$

$$\begin{bmatrix} T_2 \circ T_1 \end{bmatrix} = \begin{bmatrix} \cos(\theta_1 + \theta_2) & -\sin(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) & \cos(\theta_1 + \theta_2) \end{bmatrix}$$

These matrices should satisfy 4. With the help of some basic trigonometric identities, we can confirm that this is so as follows:

$$
\begin{aligned}
[T_2][T_1] &= \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & -(\cos\theta_2\sin\theta_1 + \sin\theta_2\cos\theta_1) \\ \sin\theta_2\cos\theta_1 + \cos\theta_2\sin\theta_1 & -\sin\theta_2\sin\theta_1 + \cos\theta_2\cos\theta_1 \end{bmatrix} \\
&= \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix} \\
&= [T_2 \circ T_1]
\end{aligned}
$$

Figure 4.10.2

EXAMPLE 2 Composition Is Not Commutative

Let $T_1 R^2 \to R^2$ be the reflection about the line $y = x$, and let $T_2 : R^2 \to R^2$ be the orthogonal projection on the y-axis. Figure 4.10.3 illustrates graphically that $T_1 \circ T_2$ and $T_2 \circ T_1$ have different effects on a vector $\mathbf{x}$. This same conclusion can be reached by showing that the standard matrices for T_1 and T_2 do not commute:

$$
\begin{aligned}
[T_1 \circ T_2] &= \begin{bmatrix} T_1 \end{bmatrix}\begin{bmatrix} T_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\
[T_2 \circ T_1] &= \begin{bmatrix} T_2 \end{bmatrix}\begin{bmatrix} T_1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}
\end{aligned}
$$

so $[T_2 \circ T_1] \neq [T_1 \circ T_2]$.

Figure 4.10.3

EXAMPLE 3 Composition of Two Reflections

Let $T_1: R^2 \to R^2$ be the reflection about the y-axis, and let $T_2: R^2 \to R^2$ be the reflection about the x-axis. In this case $T_1 \circ T_2$ and $T_2 \circ T_1$ are the same; both map every vector $\mathbf{x} = (x, y)$ into its negative $-\mathbf{x} = (-x, -y)$ (Figure 4.10.4):

$$\begin{aligned}(T_1 \circ T_2)(x, y) &= T_1(x, -y) = (-x, -y)\\ (T_2 \circ T_1)(x, y) &= T_2(-x, y) = (-x, -y)\end{aligned}$$

The equality of $T_1 \circ T_2$ and $T_2 \circ T_1$ can also be deduced by showing that the standard matrices for T_1 and T_2 commute:

$$\begin{aligned}[T_1 \circ T_2] &= \begin{bmatrix} T_1 \end{bmatrix}\begin{bmatrix} T_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\\ [T_2 \circ T_1] &= \begin{bmatrix} T_2 \end{bmatrix}\begin{bmatrix} T_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\end{aligned}$$

The operator $T(\mathbf{x}) = -\mathbf{x}$ on R^2 or R^3 is called the ***reflection about the origin***. As the foregoing computations show, the standard matrix for this operator on R^2 is

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

Figure 4.10.4

EXAMPLE 4 Composition of Three Transformations

Find the standard matrix for the operator $T: R^3 \to R^3$ that first rotates a vector counterclockwise about the z-axis through an angle θ, then reflects the resulting vector about the yz-plane, and then projects that vector orthogonally onto the xy-plane.

Solution The operator T can be expressed as the composition

$$T = T_3 \circ T_2 \circ T_1$$

where T_1 is the rotation about the *z*-axis, T_2 is the reflection about the *yz*-plane, and T_3 is the orthogonal projection on the *xy*-plane. From Tables 6, 2, and 4 of Section 4.9 , the standard matrices for these operators are

$$\begin{bmatrix} T_1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} T_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} T_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, it follows from 5 that the standard matrix for *T* is

$$\begin{aligned} [T] &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -\cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

One-to-One Matrix Transformations

Our next objective is to establish a link between the invertibility of a matrix *A* and properties of the corresponding matrix transformation T_A.

DEFINITION 1

A matrix transformation $T_A: R^n \to R^m$ is said to be ***one-to-one*** if T_A maps distinct vectors (points) in R^n into distinct vectors (points) in R^m.

(See Figure 4.10.5). This idea can be expressed in various ways. For example, you should be able to see that the following are just restatements of Definition 1:

1. T_A is one-to-one if for each vector *b* in the range of *A* there is exactly one vector $\mathbf{x}$ in R^n such that $T_A\mathbf{x} = \mathbf{b}$.
2. T_A is one-to-one if the equality $T_A(\mathbf{u}) = T_A(\mathbf{v})$ implies that $\mathbf{u} = \mathbf{v}$.

Figure 4.10.5

Rotation operators on R^2 are one-to-one since distinct vectors that are rotated through the same angle have distinct images (Figure 4.10.6). In contrast, the orthogonal projection of R^3 on the *xy*-plane is not one-to-one because it maps distinct points on the same vertical line into the same point (Figure 4.10.7).

Figure 4.10.6 Distinct vectors $\mathbf{u}$ and $\mathbf{v}$ are rotated into distinct vectors $T(\mathbf{u})$ and $T(\mathbf{v})$

Figure 4.10.7 The distinct points *P* and *Q* are mapped into the same point *M*

The following theorem establishes a fundamental relationship between the invertibility of a matrix and properties of the corresponding matrix transformation.

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 range of T_A is R^n.

(c) T_A is one-to-one.

Proof We will establish the chain of implications $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$.

$(a) \Rightarrow (b)$ Assume that *A* is invertible. By parts (*a*) and (*e*) of Theorem 4.8.10, the system $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$ in R^n. This implies that T_A maps $\mathbf{x}$ into the arbitrary vector $\mathbf{b}$ in R^n, which in turn implies that the range of T_A is all of R^n.

$(b) \Rightarrow (c)$ Assume that the range of T_A is R^n. This implies that for every vector $\mathbf{b}$ in R^n there is some vector $\mathbf{x}$ in R^n for which $T_A(\mathbf{x}) = \mathbf{b}$ and hence that the linear system $A\mathbf{x} = \mathbf{b}$ is consistent for every vector $\mathbf{b}$ in R^n. But the equivalence of parts (*e*) and (*f*) of Theorem 4.8.10 implies that $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector

$\mathbf{b}$ in R^n and hence that for every vector $\mathbf{b}$ in the range of T_A there is exactly one vector $\mathbf{x}$ in R^n such that $T_A\mathbf{x} = \mathbf{b}$.

$(c) \Rightarrow (a)$ Assume that T_A is one-to-one. Thus, if $\mathbf{b}$ is a vector in the range of T_A, there is a unique vector $\mathbf{x}$ in R^n for which $T_A(\mathbf{x}) = \mathbf{b}$. We leave it for you to complete the proof using Exercise 30.

EXAMPLE 5 Properties of a Rotation Operator

As indicated in Figure 4.10.6, the operator $T: R^n \to R^n$ that rotates vectors in R^2 through an angle θ is one-to-one. Confirm that $[T]$ is invertible in accordance with Theorem 4.10.1.

Solution From Table 5 of Section 4.9 the standard matrix for T is

$$\left[T\right] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This matrix is invertible because

$$\det\left[T\right] = \begin{vmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{vmatrix} = \cos^2\theta + \sin^2\theta = 1 \neq 0$$

EXAMPLE 6 Properties of a Projection Operator

As indicated in Figure 4.10.7, the operator $T: R^n \to R^n$ that projects each vector in R^3 orthogonally on the xy-plane is not one-to-one. Confirm that $[T]$ is not invertible in accordance with Theorem 4.10.1.

Solution From Table 4 of Section 4.9 the standard matrix for T is

$$\left[T\right] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is not invertible since $\det[T] = 0$.

Inverse of a One-to-One Matrix Operator

If $T_A: R^n \to R^n$ is a one-to-one matrix operator, then it follows from Theorem 4.10.1 that A is invertible. The matrix operator

$$T_{A^{-1}}: R^n \to R^n$$

that corresponds to A^{-1} is called the ***inverse operator*** or (more simply) the ***inverse*** of T_A. This terminology is appropriate because T_A and $T_{A^{-1}}$ cancel the effect of each other in the sense that if $\mathbf{x}$ is any vector in R^n, then

$$\begin{aligned} T_A(T_{A^{-1}}(\mathbf{x})) &= AA^{-1}\mathbf{x} = I\mathbf{x} = \mathbf{x} \\ T_{A^{-1}}(T_A(\mathbf{x})) &= A^{-1}A\mathbf{x} = I\mathbf{x} = \mathbf{x} \end{aligned}$$

or, equivalently,

$$\begin{aligned} T_A \circ T_{A^{-1}} &= T_{AA^{-1}} = T_I \\ T_{A^{-1}} \circ T_A &= T_{A^{-1}A} = T_I \end{aligned}$$

From a more geometric viewpoint, if $\mathbf{w}$ is the image of $\mathbf{x}$ under T_A, then $T_{A^{-1}}$ maps $\mathbf{w}$ back into $\mathbf{x}$, since

$$T_{A^{-1}}(\mathbf{w}) = T_{A^{-1}}(T_A(\mathbf{x})) = \mathbf{x}$$

(Figure 4.10.8).

Figure 4.10.8

Before considering examples, it will be helpful to touch on some notational matters. If $T_A: R^n \to R^n$ is a one-to-one matrix operator, and if $T_{A^{-1}}: R^n \to R^n$ is its inverse, then the standard matrices for these operators are related by the equation

$$T_{A^{-1}} = T_A^{-1} \tag{6}$$

In cases where it is preferable not to assign a name to the matrix, we will write this equation as

$$\left[T^{-1}\right] = [T]^{-1} \tag{7}$$

EXAMPLE 7 Standard Matrix for T^{-1}

Let $T: R^2 \to R^2$ be the operator that rotates each vector in R^2 through the angle θ, so from Table 5 of Section 4.9 ,

$$\left[T\right] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{8}$$

It is evident geometrically that to undo the effect of T, one must rotate each vector in R^2 through the angle $-\theta$. But this is exactly what the operator T^{-1} does, since the standard matrix for T^{-1} is

$$\left[T^{-1}\right] = [T]^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

(verify), which is the standard matrix for a rotation through the angle $-\theta$.

EXAMPLE 8 Finding T^{-1}

Show that the operator $T: R^2 \to R^2$ defined by the equations

$$\begin{aligned} w_1 &= 2x_1 + x_2 \\ w_2 &= 3x_1 + 4x_2 \end{aligned}$$

is one-to-one, and find $T^{-1}(w_1, w_2)$.

Solution The matrix form of these equations is

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

so the standard matrix for T is

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

This matrix is invertible (so T is one-to-one) and the standard matrix for T^{-1} is

$$\begin{bmatrix} T^{-1} \end{bmatrix} = [T]^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix}$$

Thus

$$\begin{bmatrix} T^{-1} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5}w_1 - \frac{1}{5}w_2 \\ -\frac{3}{5}w_1 + \frac{2}{5}w_2 \end{bmatrix}$$

from which we conclude that

$$T^{-1}(w_1, w_2) = \left(\frac{4}{5}w_1 - \frac{1}{5}w_2,\ -\frac{3}{5}w_1 + \frac{2}{5}w_2\right)$$

Linearity Properties

Up to now we have focused exclusively on matrix transformations from R^n to R^m. However, these are not the only kinds of transformations from R^n to R^m. For example, if $f_1, f_2, \ldots, f_m$ are any functions of the n variables $x_1, x_2, \ldots, x_n$, then the equations

$$\begin{aligned} w_1 &= f_1(x_1, x_2, \ldots, x_n) \\ w_2 &= f_2(x_1, x_2, \ldots, x_n) \\ &\vdots \\ w_m &= f_m(x_1, x_2, \ldots, x_n) \end{aligned}$$

define a transformation $T: R^n \to R^m$ that maps the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ into the vector $(w_1, w_2, \ldots, w_m)$. But it is only in the case where these equations are *linear* that T is a matrix transformation. The question that we

will now consider is this:

Question

Are there algebraic properties of a transformation $T: R^n \to R^m$ that can be used to determine whether T is a matrix transformation?

The answer is provided by the following theorem.

THEOREM 4.10.2

$T: R^n \to R^m$ is a matrix transformation if and only if the following relationships hold for all vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and for every scalar k:

(i) $T(\mathbf{u}+\mathbf{v}) = T((\mathbf{u}) + T(\mathbf{v}))$ **[Additivity property]**

(ii) $T(k\mathbf{u}) = kT(\mathbf{u})$ **[Homogeneity property]**

Proof If T is a matrix transformation, then properties (i) and (ii) follow respectively from parts (*c*) and (*b*) of Theorem 4.9.1.

Conversely, assume that properties (i) and (ii) hold. We must show that there exists an $m \times n$ matrix A such that

$$T(\mathbf{x}) = A\mathbf{x}$$

for every vector $\mathbf{x}$ in R^n. As a first step, recall from Formula (10) of Section 4.9 that the additivity and homogeneity properties imply that

$$T(k_1\mathbf{u}_1 + k_2\mathbf{u}_2 + \cdots + k_r\mathbf{u}_r) = k_1T(\mathbf{u}_1) + k_2T(\mathbf{u}_2) + \cdots + k_rT(\mathbf{u}_r) \tag{9}$$

for all scalars $k_1, k_2, \ldots, k_r$ and all vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ in R^n. Let A be the matrix

$$A = [T(\mathbf{e}_1)|T(\mathbf{e}_2)|\cdots|T(\mathbf{e}_n)]$$

in which $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard basis vectors for R^n.

It follows from Theorem 1.3.1 that $A\mathbf{x}$ is a linear combination of the columns of A in which the successive coefficients are the entries $x_1, x_2, \ldots, x_n$ of $\mathbf{x}$. That is,

$$A\mathbf{x} = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n)$$

Using 9 we can rewrite this as

$$A\mathbf{x} = T(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = T(\mathbf{x})$$

which completes the proof.

The additivity and homogeneity properties in Theorem 4.10.2 are called ***linearity conditions***, and a transformation that satisfies these conditions is called a ***linear transformation***. Using this terminology Theorem

4.10.2 can be restated as follows.

THEOREM 4.10.3

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.

More on the Equivalence Theorem

As our final result in this section, we will add parts (*b*) and (*c*) of Theorem 4.10.1 to Theorem 4.8.10.

THEOREM 4.10.4 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

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

(h) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.

(j) The column vectors of A span R^n.

(k) The row vectors of A span R^n.

(l) The column vectors of A form a basis for R^n.

(m) The row vectors of A form a basis for R^n.

(n) A has rank n.

(o) A has nullity 0.

(p) The orthogonal complement of the null space of A is R^n.

(q) The orthogonal complement of the row space of A is $\{0\}$.

(r) The range of T_A is R^n.

(s) T_A is one-to-one.

Concept Review

- Composition of matrix transformations
- Reflection about the origin
- One-to-one transformation
- Inverse of a matrix operator
- Linearity conditions
- Linear transformation
- Equivalent characterizations of invertible matrices

Skills

- Find the standard matrix for a composition of matrix transformations.
- Determine whether a matrix operator is one-to-one; if it is, then find the inverse operator.
- Determine whether a transformation is a linear transformation.

Exercise Set 4.10

In Exercises 1–2, let T_A and T_B be the operators whose standard matrices are given. Find the standard matrices for and $T_A \circ T_B$.

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

Answer:

$$T_B \circ T_A = \begin{bmatrix} 5 & -1 & 21 \\ 10 & -8 & 4 \\ 45 & 3 & 25 \end{bmatrix}, \; T_A \circ T_B = \begin{bmatrix} -8 & -3 & 1 \\ -5 & -15 & -8 \\ 44 & -11 & 45 \end{bmatrix}$$

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

3. Let $T_1(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$ and $T_2(x_1, x_2) = (3x_1, 2x_1 + 4x_2)$.

(a) Find the standard matrices for T_1 and T_2.

(b) Find the standard matrices for $T_2 \circ T_1$ and $T_1 \circ T_2$.

(c) Use the matrices obtained in part (b) to find formulas for $T_1(T_2(x_1, x_2))$ and $T_2(T_1(x_1, x_2))$.

Answer:

(a) $T_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $T_2 = \begin{bmatrix} 3 & 0 \\ 2 & 4 \end{bmatrix}$

(b) $T_2 \circ T_1 = \begin{bmatrix} 3 & 3 \\ 6 & -2 \end{bmatrix}$, $T_1 \circ T_2 = \begin{bmatrix} 5 & 4 \\ 1 & -4 \end{bmatrix}$

(c) $T_2(T_1(x_1,\ x_2)) = (3x_1 + 3x_2,\ 6x_1 - 2x_2)$,

$T_1(T_2(x_1,\ x_2)) = (5x_1 + 4x_2,\ x_1 - 4x_2)$

4. Let $T_1(x_1, x_2, x_3) = (4x_1,\ -2x_1 + x_2,\ -x_1 - 3x_2)$ and $T_2(x_1, x_2, x_3) = (x_1 + 2x_2,\ -x_3,\ 4x_1 - x_3)$.

(a) Find the standard matrices for T_1 and T_2.

(b) Find the standard matrices for $T_2 \circ T_1$ and $T_1 \circ T_2$.

(c) Use the matrices obtained in part (b) to find formulas for $T_1(T_2(x_1, x_2, x_3))$ and $T_2(T_1(x_1, x_2, x_3))$.

5. Find the standard matrix for the stated composition in R^2.

(a) A rotation of 90°, followed by a reflection about the line $y = x$.

(b) An orthogonal projection on the *y*-axis, followed by a contraction with factor $k = \frac{1}{2}$.

(c) A reflection about the *x*-axis, followed by a dilation with factor $k = 3$.

Answer:

(a) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$

(c) $\begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$

6. Find the standard matrix for the stated composition in R^2.

(a) A rotation of 60°, followed by an orthogonal projection on the *x*-axis, followed by a reflection about the line $y = x$.

(b) A dilation with factor $k = 2$, followed by a rotation of 45°, followed by a reflection about the *y*-axis.

(c) A rotation of 15°, followed by a rotation of 105°, followed by a rotation of 60°.

7. Find the standard matrix for the stated composition in R^3.

(a) A reflection about the *yz*-plane, followed by an orthogonal projection on the *xz*-plane.

(b) A rotation of 45° about the *y*-axis, followed by a dilation with factor $k = \sqrt{2}$.

(c) An orthogonal projection on the *xy*-plane, followed by a reflection about the *yz*-plane.

Answer:

(a) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

8. Find the standard matrix for the stated composition in R^3.

(a) A rotation of 30° about the *x*-axis, followed by a rotation of 30° about the *z*-axis, followed by a contraction with factor $k = \frac{1}{4}$.

(b) A reflection about the *xy*-plane, followed by a reflection about the *xz*-plane, followed by an orthogonal projection on the *yz*-plane.

(c) A rotation of 270° about the *x*-axis, followed by a rotation of 90° about the *y*-axis, followed by a rotation of 180° about the *z*-axis.

9. Determine whether $T_1 \circ T_2 = T_2 \circ T_1$.

(a) $T_1 : R^2 \to R^2$ is the orthogonal projection on the *x*-axis, and $T_2 : R^2 \to R^2$ is the orthogonal projection on the *y*-axis.

(b) $T_1 : R^2 \to R^2$ is the rotation through an angle θ_1, and $T_2 : R^2 \to R^2$ is the rotation through an angle θ_2.

(c) $T_1 : R^2 \to R^2$ is the orthogonal projection on the *x*-axis, and $T_2 : R^2 \to R^2$ is the rotation through an angle θ.

Answer:

(a) $T_1 \circ T_2 = T_2 \circ T_1$

(b) $T_1 \circ T_2 = T_2 \circ T_1$

(c) $T_1 \circ T_2 \neq T_2 \circ T_1$

10. Determine whether $T_1 \circ T_2 = T_2 \circ T_1$.

(a) $T_1 : R^3 \to R^3$ is a dilation by a factor *k*, and $T_2 : R^3 \to R^3$ is the rotation about the *z*-axis through an angle θ.

(b) $T_1 : R^3 \to R^3$ is the rotation about the *x*-axis through an angle θ_1, and $T_2 : R^3 \to R^3$ is the rotation about the *z*-axis through an angle θ_2.

11. By inspection, determine whether the matrix operator is one-to-one.

(a) the orthogonal projection on the *x*-axis in R^2

(b) the reflection about the *y*-axis in R^2

(c) the reflection about the line $y = x$ in R^2

(d) a contraction with factor $k > 0$ in R^2

(e) a rotation about the *z*-axis in R^3

(f) a reflection about the *xy*-plane in R^3

(g) a dilation with factor $k > 0$ in R^3

Answer:

(a) Not one-to-one

(b) One-to-one

(c) One-to-one

(d) One-to-one

(e) One-to-one

(f) One-to-one

(g) One-to-one

12. Find the standard matrix for the matrix operator defined by the equations, and use Theorem 4.10.4 to determine whether the operator is one-to-one.

(a) $\begin{aligned} w_1 &= 8x_1 + 4x_2 \\ w_2 &= 2x_1 + x_2 \end{aligned}$

(b) $\begin{aligned} w_1 &= 2x_1 - 3x_2 \\ w_2 &= 5x_1 + x_2 \end{aligned}$

(c) $\begin{aligned} w_1 &= -x_1 + 3x_2 + 2x_3 \\ w_2 &= 2x_1 \qquad + 4x_3 \\ w_3 &= x_1 + 3x_2 + 6x_3 \end{aligned}$

(d) $\begin{aligned} w_1 &= x_1 + 2x_2 + 3x_3 \\ w_2 &= 2x_1 + 5x_2 + 3x_3 \\ w_3 &= x_1 \qquad + 8x_3 \end{aligned}$

13. Determine whether the matrix operator $T: R^2 \to R^2$ defined by the equations is one-to-one; if so, find the standard matrix for the inverse operator, and find $T^{-1}(w_1, w_2)$.

(a) $\begin{aligned} w_1 &= x_1 + 2x_2 \\ w_2 &= -x_1 + x_2 \end{aligned}$

(b) $\begin{aligned} w_1 &= 4x_1 - 6x_2 \\ w_2 &= -2x_1 + 3x_2 \end{aligned}$

(c) $\begin{aligned} w_1 &= -x_2 \\ w_2 &= -x_1 \end{aligned}$

(d) $\begin{aligned} w_1 &= 3x_1 \\ w_2 &= -5x_1 \end{aligned}$

Answer:

(a) One-to-one; $\begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix}$; $T^{-1}(w_1, w_2) = \left(\frac{1}{3}w_1 - \frac{2}{3}w_2, \ \frac{1}{3}w_1 + \frac{1}{3}w_2\right)$

(b) Not one-to-one

(c) One-to-one; $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$; $T^{-1}(w_1, w_2) = (-w_2, \ -w_1)$

(d) Not one-to-one

14. Determine whether the matrix operator $T: R^3 \to R^3$ defined by the equations is one-to-one; if so, find the standard matrix for the inverse operator, and find $T^{-1}(w_1, w_2, w_3)$.

(a) $$\begin{aligned} w_1 &= x_1 - 2x_2 + 2x_3 \\ w_2 &= 2x_1 + x_2 + x_3 \\ w_3 &= x_1 + x_2 \end{aligned}$$

(b) $$\begin{aligned} w_1 &= x_1 - 3x_2 + 4x_3 \\ w_2 &= -x_1 + x_2 + x_3 \\ w_3 &= -2x_2 + 5x_3 \end{aligned}$$

(c) $$\begin{aligned} w_1 &= x_1 + 4x_2 - x_3 \\ w_2 &= 2x_1 + 7x_2 + x_3 \\ w_3 &= x_1 + 3x_2 \end{aligned}$$

(d) $$\begin{aligned} w_1 &= x_1 + 2x_2 + x_3 \\ w_2 &= -2x_1 + x_2 + 4x_3 \\ w_3 &= 7x_1 + 4x_2 - 5x_3 \end{aligned}$$

15. By inspection, find the inverse of the given one-to-one matrix operator.

(a) The reflection about the *x*-axis in R^2.

(b) The rotation through an angle of $\pi / 4$ in R^2.

(c) The dilation by a factor of 3 in R^2.

(d) The reflection about the *yz*-plane in R^3.

(e) The contraction by a factor of $\frac{1}{5}$ in R^3.

Answer:

(a) Reflection about the *x*-axis

(b) Rotation through the angle $-\frac{\pi}{4}$

(c) Contraction by a factor of $\frac{1}{3}$

(d) Reflection about the *yz*-plane

(e) Dilation by a factor of 5

In Exercises 16—17, use Theorem 4.10.2 to determine whether $T: R^2 \to R^2$ is a matrix operator.

16. (a) $T(x, y) = (2x, y)$

(b) $T(x, y) = (x^2, y)$

(c) $T(x, y) = (-y, x)$

(d) $T(x, y) = (x, 0)$

17. (a) $T(x, y) = (2x + y, x - y)$

(b) $T(x, y) = (x + 1, y)$

(c) $T(x, y) = (y, y)$

(d) $T\left(x, y\right) = \left(\sqrt[3]{x}, \sqrt[3]{y}\right)$

Answer:

(a) Matrix operator

(b) Not a matrix operator

(c) Matrix operator

(d) Not a matrix operator

In Exercises 18–19, use Theorem 4.10.2 to determine whether $T: R^3 \to R^2$ is a matrix transformation.

18. (a) $T(x, y, z) = (x, x + y + z)$

(b) $T(x, y, z) = (1, 1)$

19. (a) $T(x, y, z) = (0, 0)$

(b) $T(x, y, z) = (3x - 4y, 2x - 5z)$

Answer:

(a) Matrix transformation

(b) Matrix transformation

20. In each part, use Theorem 4.10.3 to find the standard matrix for the matrix operator from the images of the standard basis vectors.

(a) The reflection operators on R^2 in Table 1 of Section 4.9 .

(b) The reflection operators on R^3 in Table 2 of Section 4.9 .

(c) The projection operators on R^2 in Table 3 of Section 4.9 .

(d) The projection operators on R^3 in Table 4 of Section 4.9 .

(e) The rotation operators on R^2 in Table 5 of Section 4.9 .

(f) The dilation and contraction operators on R^3 in Table 8 of Section 4.9 .

21. Find the standard matrix for the given matrix operator.

(a) $T: R^2 \to R^2$ projects a vector orthogonally onto the *x*-axis and then reflects that vector about the *y*-axis.

(b) $T: R^2 \to R^2$ reflects a vector about the line $y = x$ and then reflects that vector about the *x*-axis.

(c) $T: R^2 \to R^2$ dilates a vector by a factor of 3, then reflects that vector about the line $y = x$, and then projects that vector orthogonally onto the *y*-axis.

Answer:

(a) $\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix}$

22. Find the standard matrix for the given matrix operator.

(a) $T: R^3 \to R^3$ reflects a vector about the *xz*-plane and then contracts that vector by a factor of $\frac{1}{5}$.

(b) $T: R^3 \to R^3$ projects a vector orthogonally onto the *xz*-plane and then projects that vector orthogonally onto the *xy*-plane.

(c) $T: R^3 \to R^3$ reflects a vector about the *xy*-plane, then reflects that vector about the *xz*-plane, and then reflects that vector about the *yz*-plane.

23. Let $T_A: R^3 \to R^3$ be multiplication by

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

and let $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ be the standard basis vectors for R^3. Find the following vectors by inspection.

(a) $T_A(\mathbf{e}_1)$, $T_A(\mathbf{e}_2)$, and $T_A(e_3)$

(b) $T_A(e_1 + e_2 + e_3)$

(c) $T_A(7e_3)$

Answer:

(a) $T_A(\mathbf{e}_1) = (-1, 2, 4),\ T_A(\mathbf{e}_2) = (3, 1, 5),\ T_A(\mathbf{e}_3) = (0, 2, -3)$

(b) $T_A(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) = (2, 5, 6)$

(c) $T_A(7\mathbf{e}_3) = (0, 14, -21)$

24. Determine whether multiplication by A is a one-to-one matrix transformation.

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

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

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

25. (a) Is a composition of one-to-one matrix transformations one-to-one? Justify your conclusion.

(b) Can the composition of a one-to-one matrix transformation and a matrix transformation that is not one-to-one be one-to-one? Account for both possible orders of composition and justify your conclusion.

Answer:

(a) Yes

(b) Yes

26. Show that $T(x, y) = (0, 0)$ defines a matrix operator on R^2 but $T(x, y) = (1, 1)$ does not.

27. (a) Prove: If $T: R^n \to R^m$ is a matrix transformation, then $T(\mathbf{0}) = \mathbf{0}$; that is, T maps the zero vector in R^n into the zero vector in R^m.

(b) The converse of this is not true. Find an example of a function that satisfies $T(0) = 0$ but is not a matrix transformation.

Answer:

(b) $T(x_1,\ x_2) = \left(x_1^2 + x_2^2,\ x_1 x_2\right)$

28. Prove: An $n \times n$ matrix A is invertible if and only if the linear system $A\mathbf{x} = \mathbf{w}$ has exactly one solution for every vector $\mathbf{w}$ in R^n for which the system is consistent.

29. Let A be an $n \times n$ matrix such that $\det(A) = \mathbf{0}$, and let $T: R^n \to R^n$ be multiplication by A.

(a) What can you say about the range of the matrix T? Give an example that illustrates your conclusion.

(b) What can you say about the number of vectors that T maps into $\mathbf{0}$?

Answer:

(a) The range of T is a proper subset of R^n.

(b) T must map infinitely many vectors to 0.

30. Prove: If the matrix transformation $T_A: R^n \to R^n$ is one-to-one, then A is invertible.

True-False Exercises

In parts (a)–(f) determine whether the statement is true or false, and justify your answer.

(a) If $T: R^n \to R^m$ and $T(\mathbf{0}) = \mathbf{0}$, then T is a matrix transformation.

Answer:

False

(b) If $T: R^n \to R^m$ and $T(c_1\mathbf{x} + c_2\mathbf{y}) = c_1 T(\mathbf{x}) + c_2 T(\mathbf{y})$ for all scalars c_1 and c_2 and all vectors $\mathbf{x}$ and $\mathbf{y}$ in R^n, then T is a matrix transformation.

Answer:

True

(c) If $T: R^n \to R^m$ is a one-to-one matrix transformation, then there are no distinct vectors $\mathbf{x}$ and $\mathbf{y}$ for which $T(\mathbf{x} - \mathbf{y}) = 0$.

Answer:

True

(d) If $T: R^n \to R^m$ is a matrix transformation and $m > n$, then T is one-to-one.

Answer:

False

(e) If $T: R^n \to R^m$ is a matrix transformation and $m = n$, then T is one-to-one.

Answer:

False

(f) If $T: R^n \to R^m$ is a matrix transformation and $m < n$, then T is one-to-one.

Answer:

False

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

4.11 Geometry of Matrix Operators on R^2

In this optional section we will discuss matrix operators on R^2 in a little more depth. The ideas that we will develop here have important applications to computer graphics.

Transformations of Regions

In Section 4.9 we focused on the effect that a matrix operator has on individual vectors in R^2 and R^3. However, it is also important to understand how such operators affect the shapes of regions. For example, Figure 4.11.1 shows a famous picture of Albert Einstein and three computer-generated modifications of that image that result from matrix operators on R^2. The original picture was scanned and then digitized to decompose it into a rectangular array of pixels. The pixels were then transformed as follows:

- The program MATLAB was used to assign coordinates and a gray level to each pixel.
- The coordinates of the pixels were transformed by matrix multiplication.
- The pixels were then assigned their original gray levels to produce the transformed picture.

Figure 4.11.1

The overall effect of a matrix operator on R^2 can often be ascertained by graphing the images of the vertices $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ of the unit square (Figure 4.11.2). Table 1 shows the effect that some of the matrix operators studied in Section 4.9 have on the unit square. For clarity, we have shaded a portion of the original square and its corresponding image.

Figure 4.11.2

Table 1

Operator	Standard Matrix	Effect on the Unit Square
Reflection about the y-axis	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	
Reflection about the x-axis	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	
Reflection about the line $y = x$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	
Counterclockwise rotation through an angle θ	$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$	
Compression in the x-direction by a factor of k $(0 < k < 1)$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$	
Expansion in the x-direction by a factor of k $(k > 1)$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$	
Shear in the x-direction with factor $k > 0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$	
Shear in the	$\begin{bmatrix} 1 & k \end{bmatrix}$	

EXAMPLE 1 Transforming with Diagonal Matrices

Suppose that the xy-plane first is compressed or expanded by a factor of k_1 in the x-direction and then is compressed or expanded by a factor of k_2 in the y-direction. Find a single matrix operator that performs both operations.

Solution The standard matrices for the two operations are

$$\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix}$$

x-compression (expansion) **y-compression (expansion)**

Thus, the standard matrix for the composition of the x-operation followed by the y-operation is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix}\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} \tag{1}$$

This shows that multiplication by a diagonal 2×2 matrix compresses or expands the plane in the x-direction and also in the y-direction. In the special case where k_1 and k_2 are the same, say $k_1 = k_2 = k$, Formula 1 simplifies to

$$A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$$

which is a contraction or a dilation (Table 7 of Section 4.9).

EXAMPLE 2 Finding Matrix Operators

(a) Find the standard matrix for the operator on R^2 that first shears by a factor of 2 in the x-direction and then reflects the result about the line $y = x$. Sketch the image of the unit square under this operator.

(b) Find the standard matrix for the operator on R^2 that first reflects about $y = x$ and then shears by a factor of 2 in the x-direction. Sketch the image of the unit square under this operator.

(c) Confirm that the shear and the reflection in parts (a) and (b) do not commute.

Solution

(a) The standard matrix for the shear is

$$A_1 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

and for the reflection is

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Thus, the standard matrix for the shear followed by the reflection is

$$A_2A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

(b) The standard matrix for the reflection followed by the shear is

$$A_1A_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$$

(c) The computations in Solutions (*a*) and (*b*) show that $A_1A_2 \neq A_2A_1$, so the standard matrices, and hence the operators, do not commute. The same conclusion follows from Figures 4.11.3 and 4.11.4, since the two operators produce different images of the unit square.

Figure 4.11.3

Figure 4.11.4

Geometry of One-to-One Matrix Operators

We will now turn our attention to one-to-one matrix operators on R^2, which are important because they map distinct points into distinct points. Recall from Theorem 4.10.4 (the Equivalence Theorem) that a matrix transformation T_A is one-to-one if and only if A can be expressed as a product of elementary matrices. Thus, we can analyze the effect of any one-to-one transformation T_A by first factoring the matrix A into a product of elementary matrices, say

$$A = E_1E_2 \ldots E_r$$

and then expressing T_A as the composition

$$T_A = T_{E_1E_2 \ldots E_r} = T_{E_1} \circ T_{E_2} \circ \ldots \circ T_{E_r} \tag{2}$$

The following theorem explains the geometric effect of matrix operators corresponding to elementary matrices.

THEOREM 4.11.1

If E is an elementary matrtix, then $T_E: R^2 \to R^2$ is one of the following:

(a) A shear along a coordinate axis.

(b) A reflection about $y = x$.

(c) A compression along a coordinate axis.

(d) An expansion along a coordinate axis.

(e) A reflection about a coordinate axis.

(f) A compression or expansion along a coordinate axis followed by a reflection about a coordinate axis.

Proof Because a 2×2 elementary matrix results from performing a single elementary row operation on the 2×2 identity matrix, such a matrix must have one of the following forms (verify):

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

The first two matrices represent shears along coordinate axes, and the third represents a reflection about $y = x$. If $k > 0$, the last two matrices represent compressions or expansions along coordinate axes, depending on whether $0 \le k < 1$ or $k > 1$. If $k < 0$, and if we express k in the form $k = -k_1$, where $k_1 > 0$, then the last two matrices can be written as

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -k_1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -k_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & k_1 \end{bmatrix} \tag{4}$$

Since $k_1 > 0$, the product in 3 represents a compression or expansion along the x-axis followed by a reflection about the y-axis, and 4 represents a compression or expansion along the y-axis followed by a reflection about the x-axis. In the case where $k = -1$, transformations 3 and 4 are simply reflections about the y-axis and x-axis, respectively.

Since every invertible matrix is a product of elementary matrices, the following result follows from Theorem 4.11.1 and Formula 2.

THEOREM 4.11.2

If $T_A: R^2 \to R^2$ is multiplication by an invertible matrix A, then the geometric effect of T_A is the same as an appropriate succession of shears, compressions, expansions, and reflections.

EXAMPLE 3 Analyzing the Geometric Effect of a Matrix Operator

Assuming that k_1 and k_2 are positive, express the diagonal matrix

$$A = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$$

as a product of elementary matrices, and describe the geometric effect of multiplication by A in terms of compressions and expansions.

Solution From Example 1 we have

$$A = \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & k_2 \end{bmatrix}\begin{bmatrix} k_1 & 0 \\ 0 & 1 \end{bmatrix}$$

which shows that multiplication by A has the geometric effect of compressing or expanding by a factor of k_1 in the x-direction and then compressing or expanding by a factor of k_2 in the y-direction.

EXAMPLE 4 Analyzing the Geometric Effect of a Matrix Operator

Express

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

as a product of elementary matrices, and then describe the geometric effect of multiplication by A in terms of shears, compressions, expansions, and reflections.

Solution A can be reduced to I as follows:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow[\text{Add } -3 \text{ times the first row to the second.}]{} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \xrightarrow[\text{Multiply the second row by } -\frac{1}{2}.]{} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \xrightarrow[\text{Add } -2 \text{ times the second row to the first.}]{} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The three successive row operations can be performed by multiplying A on the left successively by

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix},\ E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix},\ E_3 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

Inverting these matrices and using Formula 4 of Section 1.5 yields

$$A = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Reading from right to left and noting that

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

it follows that the effect of multiplying by A is equivalent to

1. shearing by a factor of 2 in the x-direction,
2. then expanding by a factor of 2 in the y-direction,
3. then reflecting about the x-axis,
4. then shearing by a factor of 3 in the y-direction.

Images of Lines Under Matrix Operators

Many images in computer graphics are constructed by connecting points with line segments. The following theorem, some of whose parts are proved in the exercises, is helpful for understanding how matrix operators transform such figures.

THEOREM 4.11.3

If $T: R^2 \to R^2$ is multiplication by an invertible matrix, then:

(a) The image of a straight line is a straight line.

(b) The image of a straight line through the origin is a straight line through the origin.

(c) The images of parallel straight lines are parallel straight lines.

(d) The image of the line segment joining points P and Q is the line segment joining the images of P and Q.

(e) The images of three points lie on a line if and only if the points themselves lie on a line.

Note that it follows from Theorem 4.11.3 that if A is an invertible 2×2 matrix, then multiplication by A maps triangles into triangles and parallelograms into parallelograms.

EXAMPLE 5 Image of a Square

Sketch the image of the square with vertices $(0, 0)$, $(1, 1)$, and $(0, 1)$ under multiplication by

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

Solution Since

$$\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

the image of the square is a parallelogram with vertices $(0, 0)$, $(-1, 2)$, $(2, -1)$, and $(1, 1)$ (Figure 4.11.5).

Figure 4.11.5

EXAMPLE 6 Image of a Line

According to Theorem 4.11.3, the invertible matrix

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

maps the line $y = 2x + 1$ into another line. Find its equation.

Solution Let (x, y) be a point on the line $y = 2x + 1$, and let (x', y') be its image under multiplication by A. Then

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \text{ and } \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}^{-1}\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix}$$

so

$$\begin{aligned} x &= x' - y' \\ y &= -2x' + 3y' \end{aligned}$$

Substituting in $y = 2x + 1$ yields

$$-2x' + 3y' = 2(x' - y') + 1 \text{ or equivalently } y' = \frac{4}{5}x' + \frac{1}{5}$$

Thus (x', y') satisfies

$$y = \frac{4}{5}x + \frac{1}{5}$$

which is the equation we want.

Concept Review

- Effect of a matrix operator on the unit square
- Geometry of one-to-one matrix operators
- Images of lines under matrix operators

Skills

- Find standard matrices for geometric transformations of R^2.
- Describe the geometric effect of an invertible matrix operator.
- Find the image of the unit square under a matrix operator.
- Find the image of a line under a matrix operator.

Exercise Set 4.11

1. Find the standard matrix for the operator $T: R^2 \to R^2$ that maps a point (x, y) into

(a) its reflection about the line $y = -x$.

(b) its reflection through the origin.

(c) its orthogonal projection on the x-axis.

(d) its orthogonal projection on the y-axis.

Answer:

(a) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

2. For each part of Exercise 1, use the matrix you have obtained to compute $T(2, 1)$. Check your answers geometrically by plotting the points $(2, 1)$ and $T(2, 1)$.

3. Find the standard matrix for the operator $T: R^3 \to R^3$ that maps a point (x, y, z) into

(a) its reflection through the xy-plane.

(b) its reflection through the xz-plane.

(c) its reflection through the yz-plane.

Answer:

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

4. For each part of Exercise 3, use the matrix you have obtained to compute $T(1, 1, 1)$. Check your answers geometrically by plotting the points $(1, 1, 1)$ and $T(1, 1, 1)$.

5. Find the standard matrix for the operator $T: R^3 \to R^3$ that

(a) rotates each vector 90° counterclockwise about the *z*-axis (looking along the positive *z*-axis toward the origin).

(b) rotates each vector 90° counterclockwise about the *x*-axis (looking along the positive *x*-axis toward the origin).

(c) rotates each vector 90° counterclockwise about the *y*-axis (looking along the positive *y*-axis toward the origin).

Answer:

(a) $\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$

6. Sketch the image of the rectangle with vertices $(0, 0)$, $(1, 0)$, $(1, 2)$, and $(0, 2)$ under

(a) a reflection about the *x*-axis.

(b) a reflection about the *y*-axis.

(c) a compression of factor $k = \frac{1}{4}$ in the *y*-direction.

(d) an expansion of factor $k = 2$ in the *x*-direction.

(e) a shear of factor $k = 3$ in the *x*-direction.

(f) a shear of factor $k = 2$ in the *y*-direction.

7. Sketch the image of the square with vertices $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ under multiplication by

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

Answer:

Rectangle with vertices at (0, 0), $(-3, 0)$, $(0, 1)$, $(-3, 1)$

8. Find the matrix that rotates a point (x, y) about the origin

(a) 45°

(b) 90°

(c) 180°

(d) 270°

(e) $-30°$

9. Find the matrix that shears by

(a) a factor of $k = 4$ in the y-direction.

(b) a factor of $k = -2$ in the x-direction.

Answer:

(a) $\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}$

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

10. Find the matrix that compresses or expands by

(a) a factor of $\frac{1}{3}$ in the y-direction.

(b) a factor of 6 in the x-direction.

11. In each part, describe the geometric effect of multiplication by A.

(a) $A = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix}$

(c) $A = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}$

Answer:

(a) Expansion by a factor of 3 in the x-direction

(b) Expansion by a factor of 5 in the y-direction and reflection about the x-axis

(c) Shearing by a factor of 4 in the x-direction

12. In each part, express the matrix as a product of elementary matrices, and then describe the effect of multiplication by A in terms of compressions, expansions, reflections, and shears.

(a) $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

(b) $A = \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}$

(c) $A = \begin{bmatrix} 0 & -2 \\ 4 & 0 \end{bmatrix}$

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

13. In each part, find a single matrix that performs the indicated succession of operations.

(a) Compresses by a factor of $\frac{1}{2}$ in the x-direction, then expands by a factor of 5 in the y-direction.

(b) Expands by a factor of 5 in the y-direction, then shears by a factor of 2 in the y-direction.

(c) Reflects about $y = x$, then rotates through an angle of 180° about the origin.

Answer:

(a) $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 5 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

14. In each part, find a single matrix that performs the indicated succession of operations.

(a) Reflects about the y-axis, then expands by a factor of 5 in the x-direction, and then reflects about $y = x$.

(b) Rotates through 30° about the origin, then shears by a factor of -2 in the y-direction, and then expands by a factor of 3 in the y-direction.

15. Use matrix inversion to show the following.

(a) The inverse transformation for a reflection about $y = x$ is a reflection about $y = x$.

(b) The inverse transformation for a compression along an axis is an expansion along that axis.

(c) The inverse transformation for a reflection about a coordinate axis is a reflection about that axis.

(d) The inverse transformation for a shear along a coordinate axis is a shear along that axis.

16. Find an equation of the image of the line $y = -4x + 3$ under multiplication by

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

17. In parts (a) through (e), find an equation of the image of the line $y = 2x$ under

(a) a shear of factor 3 in the x-direction.

(b) a compression of factor $\frac{1}{2}$ in the y-direction.

(c) a reflection about $y = x$.

(d) a reflection about the y-axis.

(e) a rotation of 60° about the origin.

Answer:

(a) $y = \frac{2}{7}x$

(b) $y = x$

(c) $y = \frac{1}{2}x$

(d) $y = -2x$

(e) $y = \left(-\frac{8 + 5\sqrt{3}}{11}\right)x$

18. Find the matrix for a shear in the x-direction that transforms the triangle with vertices $(0, 0)$, $(2, 1)$, and $(3, 0)$ into a right triangle with the right angle at the origin.

19. (a) Show that multiplication by

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

maps each point in the plane onto the line $y = 2x$.

(b) It follows from part (a) that the noncollinear points $(1, 0)$, $(0, 1)$, $(-1, 0)$ are mapped onto a line. Does this violate part (*e*) of Theorem 4.11.3?

Answer:

(b) No

20. Prove part (*a*) of Theorem 4.11.3. [*Hint:* A line in the plane has an equation of the form $Ax + By + C = 0$, where *A* and *B* are not both zero. Use the method of Example 6 to show that the image of this line under multiplication by the invertible matrix

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

has the equation $A'x + B'y + C = 0$, where

$$A' = (dA - cB) / (ad - bc)$$

and

$$B' = (-bA + aB) / (ad - bc)$$

Then show that A' and B' are not both zero to conclude that the image is a line.]

21. Use the hint in Exercise 20 to prove parts (*b*) and (*c*) of Theorem 4.11.3.

22. In each part of the accompanying figure, find the standard matrix for the operator described.

Figure Ex-22

23. In R^3 the ***shear in the xy-direction with factor k*** is the matrix transformation that moves each point (x, y, z) parallel to the *xy*-plane to the new position $(x + kz, y + kz, z)$. (See the accompanying figure.)

(a) Find the standard matrix for the shear in the *xy*-direction with factor *k*.

(b) How would you define the shear in the *xz*-direction with factor *k* and the shear in the *yz*-direction with factor *k*?

Find the standard matrices for these matrix transformations.

Figure Ex-23

Answer:

(a) $\begin{bmatrix} 1 & 0 & k \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix}$

(b) Shear in the *xz*-direction with

factor k maps (x, y, z) to $(x + ky, y, z + ky)$: $\begin{bmatrix} 1 & k & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$.

Shear in the *yz*-direction with factor k maps (x, y, z) to $(x, y + kx, z + kx)$: $\begin{bmatrix} 1 & 0 & 0 \\ k & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$.

True-False Exercises

In parts (a)–(g) determine whether the statement is true or false, and justify your answer.

(a) The image of the unit square under a one-to-one matrix operator is a square.

Answer:

False

(b) A 2×2 invertible matrix operator has the geometric effect of a succession of shears, compressions, expansions, and reflections.

Answer:

True

(c) The image of a line under a one-to-one matrix operator is a line.

Answer:

True

(d) Every reflection operator on R^2 is its own inverse.

Answer:

True

(e) The matrix $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ represents reflection about a line.

Answer:

False

(f) The matrix $\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}$ represents a shear.

Answer:

False

(g) The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ represents an expansion.

Answer:

True

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

4.12 Dynamical Systems and Markov Chains

In this optional section we will show how matrix methods can be used to analyze the behavior of physical systems that evolve over time. The methods that we will study here have been applied to problems in business, ecology, demographics, sociology, and most of the physical sciences.

Dynamical Systems

A ***dynamical system*** is a finite set of variables whose values change with time. The value of a variable at a point in time is called the ***state of the variable*** at that time, and the vector formed from these states is called the ***state of the dynamical system*** at that time. Our primary objective in this section is to analyze how the state of a dynamical system changes with time. Let us begin with an example.

EXAMPLE 1 Market Share as a Dynamical System

Suppose that two competing television channels, channel 1 and channel 2, each have 50% of the viewer market at some initial point in time. Assume that over each one-year period channel 1 captures 10% of channel 2's share, and channel 2 captures 20% of channel 1's share (see Figure 4.12.1). What is each channel's market share after one year?

Figure 4.12.1

Solution Let us begin by introducing the time-dependent variables

$$x_1(t) = \text{fraction of the market held by channel 1 at time } t$$
$$x_2(t) = \text{fraction of the market held by channel 2 at time } t$$

and the column vector

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \begin{array}{l} \leftarrow \textbf{Channel 1's fraction of the market at time } t \textbf{ in years} \\ \leftarrow \textbf{Channel 2's fraction of the market at time } t \textbf{ in years} \end{array}$$

The variables $x_1(t)$ and $x_2(t)$ form a dynamical system whose state at time t is the vector $\mathbf{x}(t)$. If we take $t = 0$ to be the starting point at which the two channels had 50% of the market, then the state of the system at that time is

$$\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \begin{array}{l} \leftarrow \textbf{Channel 1's fraction of the market at time } t = 0 \\ \leftarrow \textbf{Channel 2's fraction of the market at time } t = 0 \end{array} \tag{1}$$

Now let us try to find the state of the system at time $t = 1$ (one year later). Over the one-year period, channel 1 retains 80% of its initial 50%, and it gains 10% of channel 2's initial 50%. Thus,

$$\mathbf{x}_1(1) = 0.8(0.5) + 0.1(0.5) = 0.45 \tag{2}$$

Similarly, channel 2 gains 20% of channel 1's initial 50%, and retains 90% of its initial 50%. Thus,

$$x_2(1) = 0.2(0.5) + 0.9(0.5) = 0.55 \tag{3}$$

Therefore, the state of the system at time $t = 1$ is

$$\mathbf{x}\begin{pmatrix} \\ 1 \\ \\ \end{pmatrix} = \begin{bmatrix} x_1(1) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix} \begin{matrix} \leftarrow \textbf{Channel 1's fraction of the market at time } t = 1 \\ \leftarrow \textbf{Channel 2's fraction of the market at time } t = 1 \end{matrix} \tag{4}$$

EXAMPLE 2 Evolution of Market Share over Five Years

Track the market shares of channels 1 and 2 in Example 1 over a five-year period.

Solution To solve this problem suppose that we have already computed the market share of each channel at time $t = k$ and we are interested in using the known values of $x_1(k)$ and $x_2(k)$ to compute the market shares $x_1(k+1)$ and $x_2(k+1)$ one year later. The analysis is exactly the same as that used to obtain Equations 2 and 3. Over the one-year period, channel 1 retains 80% of its starting fraction $x_1(k)$ and gains 10% of channel 2's starting fraction $x_2(k)$. Thus,

$$x_1(k+1) = (0.8)x_1(k) + (0.1)x_2(k) \tag{5}$$

Similarly, channel 2 gains 20% of channel 1's starting fraction $x_1(k)$ and retains 90% of its own starting fraction $x_2(k)$. Thus,

$$x_2(k+1) = (0.2)x_1(k) + (0.9)x_2(k) \tag{6}$$

Equations 5 and 6 can be expressed in matrix form as

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \tag{7}$$

which provides a way of using matrix multiplication to compute the state of the system at time $t = k + 1$ from the state at time $t = k$. For example, using 1 and 7 we obtain

$$\mathbf{x}(1) = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \mathbf{x}(0) = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}$$

which agrees with 4. Similarly,

$$\mathbf{x}(2) = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \mathbf{x}(1) = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix} = \begin{bmatrix} 0.415 \\ 0.585 \end{bmatrix}$$

We can now continue this process, using Formula 7 to compute $\mathbf{x}(3)$ from $\mathbf{x}(2)$, then $\mathbf{x}(4)$ from $\mathbf{x}(3)$, and so on. This yields (verify)

$$\mathbf{x}(3) = \begin{bmatrix} 0.3905 \\ 0.6095 \end{bmatrix}, \quad \mathbf{x}(4) = \begin{bmatrix} 0.37335 \\ 0.62665 \end{bmatrix}, \quad \mathbf{x}(5) = \begin{bmatrix} 0.361345 \\ 0.638655 \end{bmatrix} \tag{8}$$

Thus, after five years, channel 1 will hold about 36% of the market and channel 2 will hold about 64%of the market.

If desired, we can continue the market analysis in the last example beyond the five-year period and explore what happens to the market share over the long term. We did so, using a computer, and obtained the following state vectors (rounded to six decimal places):

$$\mathbf{x}(10) \approx \begin{bmatrix} 0.338041 \\ 0.661959 \end{bmatrix}, \quad \mathbf{x}(20) \approx \begin{bmatrix} 0.333466 \\ 0.666534 \end{bmatrix}, \quad \mathbf{x}(40) \approx \begin{bmatrix} 0.333333 \\ 0.666667 \end{bmatrix} \tag{9}$$

All subsequent state vectors, when rounded to six decimal places, are the same as $\mathbf{x}(40)$, so we see that the market shares eventually stabilize with channel 1 holding about one-third of the market and channel 2 holding about two-thirds. Later in this section, we will explain why this stabilization occurs.

Markov Chains

In many dynamical systems the states of the variables are not known with certainty but can be expressed as probabilities; such dynamical systems are called ***stochastic processes*** (from the Greek word *stokastikos*, meaning "proceeding by guesswork"). A detailed study of stochastic processes requires a precise definition of the term *probability*, which is outside the scope of this course. However, the following interpretation will suffice for our present purposes:

> *Stated informally, the **probability** that an experiment or observation will have a certain outcome is approximately the fraction of the time that the outcome would occur if the experiment were to be repeated many times under constant conditions—the greater the number of repetitions, the more accurately the probability describes the fraction of occurrences.*

For example, when we say that the probability of tossing heads with a fair coin is $\frac{1}{2}$, we mean that if the coin were tossed many times under constant conditions, then we would expect about half of the outcomes to be heads. Probabilities are often expressed as decimals or percentages. Thus, the probability of tossing heads with a fair coin can also be expressed as 0.5 or 50%.

If an experiment or observation has n possible outcomes, then the probabilities of those outcomes must be nonnegative fractions whose sum is 1. The probabilities are nonnegative because each describes the fraction of occurrences of an outcome over the long term, and the sum is 1 because they account for all possible outcomes. For example, if a box containing 10 balls has one red ball, three green balls, and six yellow balls, and if a ball is drawn at random from the box, then the probabilities of the various outcomes are

$$\begin{aligned} p_1 &= \text{prob(red)} = 1 / 10 = 0.1 \\ p_2 &= \text{prob(green)} = 3 / 10 = 0.3 \\ p_3 &= \text{prob(yellow)} = 6 / 10 = 0.6 \end{aligned}$$

Each probability is a nonnegative fraction and

$$p_1 + p_2 + p_3 = 0.1 + 0.3 + 0.6 = 1$$

In a stochastic process with n possible states, the state vector at each time t has the form

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} \begin{matrix} \textbf{Probability that the system is in state 1} \\ \textbf{Probability that the system is in state 2} \\ \vdots \\ \textbf{Probability that the system is in state } \boldsymbol{n} \end{matrix}$$

The entries in this vector must add up to 1 since they account for all n possibilities. In general, a vector with nonnegative entries that add up to 1 is called a ***probability vector***.

EXAMPLE 3 Example 1 Revisited from the Probability Viewpoint

Observe that the state vectors in Example 1 and Example 2 are all probability vectors. This is to be expected since the entries in each state vector are the fractional market shares of the channels, and together they account for the entire market. In practice, it is preferable to interpret the entries in the state vectors as probabilities rather than exact market fractions, since market information is usually obtained by statistical sampling procedures with intrinsic uncertainties. Thus, for example, the state vector

$$\mathbf{x}(1) = \begin{bmatrix} x_1(1) \\ x_2(1) \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix}$$

which we interpreted in Example 1 to mean that channel 1 has 45% of the market and channel 2 has 55%, can also be interpreted to mean that an individual picked at random from the market will be a channel 1 viewer with probability 0.45 and a channel 2 viewer with probability 0.55.

A square matrix, each of whose columns is a probability vector, is called a ***stochastic matrix***. Such matrices commonly occur in formulas that relate successive states of a stochastic process. For example, the state vectors $\mathbf{x}(k+1)$ and $\mathbf{x}(k)$ in 7 are related by an equation of the form $\mathbf{x}(k+1) = P\mathbf{x}(k)$ in which

$$P = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix} \tag{10}$$

is a stochastic matrix. It should not be surprising that the column vectors of P are probability vectors, since the entries in each column provide a breakdown of what happens to each channel's market share over the year—the entries in column 1 convey that each year channel 1 retains 80% of its market share and loses 20%; and the entries in column 2 convey that each year channel 2 retains 90% of its market share and loses 10%. The entries in 10 can also be viewed as probabilities:

$$\begin{aligned} p_{11} &= 0.8 = \text{probability that a channel 1 viewer remains a channel 1 viewer} \\ p_{21} &= 0.2 = \text{probability that a channel 1 viewer becomes a channel 2 viewer} \\ p_{12} &= 0.1 = \text{probability that a channel 2 viewer becomes a channel 1 viewer} \\ p_{22} &= 0.9 = \text{probability that a channel 2 viewer remains a channel 2 viewer} \end{aligned}$$

Example 1 is a special case of a large class of stochastic processes, called *Markov chains*.

Andrei Andreyevich Markov (1856–1922)

Historical Note Markov chains are named in honor of the Russian mathematician A. A. Markov, a lover of poetry, who used them to analyze the alternation of vowels and consonants in the poem *Eugene Onegin* by Pushkin. Markov believed that the only applications of his chains were to the analysis of literary works, so he would be astonished to learn that his discovery is used today in the social sciences, quantum theory, and genetics!
[*Image: wikipedia*]

DEFINITION 1

A ***Markov chain*** is a dynamical system whose state vectors at a succession of time intervals are probability vectors and for which the state vectors at successive time intervals are related by an equation of the form

$$\mathbf{x}(k+1) = P\mathbf{x}(k)$$

in which $P = [p_{ij}]$ is a stochastic matrix and p_{ij} is the probability that the system will be in state i at time $t = k+1$ if it is in state j at time $t = k$. The matrix P is called the ***transition matrix*** for the system.

Remark Note that in this definition the row index i corresponds to the later state and the column index j to the earlier state (Figure 4.12.2).

Figure 4.12.2

EXAMPLE 4 Wildlife Migration as a Markov Chain

Suppose that a tagged lion can migrate over three adjacent game reserves in search of food, reserve 1, reserve 2, and reserve 3. Based on data about the food resources, researchers conclude that the monthly migration pattern of the lion can be modeled by a Markov chain with transition matrix

$$p = \begin{array}{c} \text{Reserve at time } t = k \\ \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.2 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.1 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{l} \\ \\ \text{Reserve at time } t = k + 1 \end{array}$$

(see Figure 4.12.3). That is,

$p_{11} = 0.5 =$ probability that the lion will stay in reserve 1 when it is in reserve 1

$p_{12} = 0.4 =$ probability that the lion will move from reserve 2 to reserve 1

$p_{13} = 0.6 =$ probability that the lion will move from reserve 3 to reserve 1

$p_{21} = 0.2 =$ probability that the lion will move from reserve 1 to reserve 2

$p_{22} = 0.2 =$ probability that the lion will stay in reserve 2 when it is in reserve 2

$p_{23} = 0.3 =$ probability that the lion will move from reserve 3 to reserve 2

$p_{31} = 0.3 =$ probability that the lion will move from reserve 1 to reserve 3

$p_{32} = 0.4 =$ probability that the lion will move from reserve 2 to reserve 3

$p_{33} = 0.1 =$ probability that the lion will stay in reserve 3 when it is in reserve 3

Assuming that t is in months and the lion is released in reserve 2 at time $t = 0$, track its probable locations over a six-month period.

Figure 4.12.3

Solution Let $x_1(k)$, $x_2(k)$, and $x_3(k)$ be the probabilities that the lion is in reserve 1, 2, or 3, respectively, at time $t = k$, and let

$$\mathbf{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

be the state vector at that time. Since we know with certainty that the lion is in reserve 2 at time $t = 0$, the initial state vector is

$$\mathbf{x}(0) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

We leave it for you to show that the state vectors over a six-month period are

$$\mathbf{x}(1) = P\mathbf{x}(0) = \begin{bmatrix} 0.400 \\ 0.200 \\ 0.400 \end{bmatrix}, \mathbf{x}(2) = P\mathbf{x}(1) = \begin{bmatrix} 0.520 \\ 0.240 \\ 0.240 \end{bmatrix}, \mathbf{x}(3) = P\mathbf{x}(2) = \begin{bmatrix} 0.500 \\ 0.224 \\ 0.276 \end{bmatrix}$$

$$\mathbf{x}(4) = P\mathbf{x}(3) \approx \begin{bmatrix} 0.505 \\ 0.228 \\ 0.267 \end{bmatrix}, \mathbf{x}(5) = P\mathbf{x}(4) \approx \begin{bmatrix} 0.504 \\ 0.227 \\ 0.269 \end{bmatrix}, \mathbf{x}(6) = P\mathbf{x}(5) \approx \begin{bmatrix} 0.504 \\ 0.227 \\ 0.269 \end{bmatrix}$$

As in Example 2, the state vectors here seem to stabilize over time with a probability of approximately 0.504 that the lion is in reserve 1, a probability of approximately 0.227 that it is in reserve 2, and a probability of approximately 0.269 that it is in reserve 3.

Markov Chains in Terms of Powers of the Transition Matrix

In a Markov chain with an initial state of $\mathbf{x}(0)$, the successive state vectors are

$$\mathbf{x}(1) = P\mathbf{x}(0), \mathbf{x}(2) = P\mathbf{x}(1), \mathbf{x}(3) = P\mathbf{x}(2), \mathbf{x}(4) = P\mathbf{x}(3), \ldots$$

For brevity, it is common to denote $\mathbf{x}(k)$ by $\mathbf{x}_k$, which allows us to write the successive state vectors more briefly as

$$\mathbf{x}_1 = P\mathbf{x}_0,\ \mathbf{x}_2 = P\mathbf{x}_1,\ \mathbf{x}_3 = P\mathbf{x}_2,\ \mathbf{x}_4 = P\mathbf{x}_3, \ldots \tag{11}$$

Note that Formula 12 makes it possible to compute the state vector $\mathbf{x}_k$ without first computing the earlier state vectors as required in Formula 11.

Alternatively, these state vectors can be expressed in terms of the initial state vector $\mathbf{x}_0$ as

$$\mathbf{x}_1 = P\mathbf{x}_0,\ \mathbf{x}_2 = P\left(P\mathbf{x}_0\right) = P^2\mathbf{x}_0,\ \mathbf{x}_3 = P\left(P^2\mathbf{x}_0\right) = P^3\mathbf{x}_0,\ \mathbf{x}_4 = P\left(P^3\mathbf{x}_0\right) = P^4\mathbf{x}_0, \ldots$$

from which it follows that

$$\mathbf{x}_k = P^k\mathbf{x}_0 \tag{12}$$

EXAMPLE 5 Finding a State Vector Directly from $\mathbf{x}_0$

Use Formula 12 to find the state vector $\mathbf{x}(3)$ in Example 2.

Solution From 1 and 7, the initial state vector and transition matrix are

$$\mathbf{x}_0 = \mathbf{x}\left(0\right) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}$$

We leave it for you to calculate P^3 and show that

$$\mathbf{x}\left(3\right) = \mathbf{x}_3 = P^3\mathbf{x}_0 = \begin{bmatrix} 0.562 & 0.219 \\ 0.438 & 0.781 \end{bmatrix}\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.3905 \\ 0.6095 \end{bmatrix}$$

which agrees with the result in 8.

Long-Term Behavior of a Markov Chain

We have seen two examples of Markov chains in which the state vectors seem to stabilize after a period of time. Thus, it is reasonable to ask whether all Markov chains have this property. The following example shows that this is not the case.

EXAMPLE 6 A Markov Chain That Does Not Stabilize

The matrix

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

is stochastic and hence can be regarded as the transition matrix for a Markov chain. A simple calculation shows that $P^2 = I$, from which it follows that

$$I = P^2 = P^4 = P^6 = \ldots \text{ and } P = P^3 = P^5 = P^7 = \ldots$$

Thus, the successive states in the Markov chain with initial vector $\mathbf{x}_0$ are

$$\mathbf{x}_0,\ P\mathbf{x}_0,\ \mathbf{x}_0,\ P\mathbf{x}_0,\ \mathbf{x}_0, \ldots$$

which oscillate between $\mathbf{x}_0$ and $P\mathbf{x}_0$. Thus, the Markov chain does not stabilize unless both components of $\mathbf{x}_0$ are $\frac{1}{2}$ (verify).

A precise definition of what it means for a sequence of numbers or vectors to stabilize is given in calculus; however, that level of precision will not be needed here. Stated informally, we will say that a sequence of vectors

$$\mathbf{x}_1,\ \mathbf{x}_2, \ldots,\ \mathbf{x}_k, \ldots$$

approaches a ***limit*** $\mathbf{q}$ or that it ***converges*** to $\mathbf{q}$ if all entries in $\mathbf{x}_k$ can be made as close as we like to the corresponding entries in the vector $\mathbf{q}$ by taking k sufficiently large. We denote this by writing $\mathbf{x}_k \to \mathbf{q}$ as $k \to \infty$.

We saw in Example 6 that the state vectors of a Markov chain need not approach a limit in all cases. However, by imposing a mild condition on the transition matrix of a Markov chain, we can guarantee that the state vectors will approach a limit.

DEFINITION 2

A stochastic matrix P is said to be ***regular*** if P or some positive power of P has all positive entries, and a Markov chain whose transition matrix is regular is said to be a ***regular Markov chain***.

EXAMPLE 7 Regular Stochastic Matrices

The transition matrices in Example 2 and Example 4 are regular because their entries are positive. The matrix

$$P = \begin{bmatrix} 0.5 & 1 \\ 0.5 & 0 \end{bmatrix}$$

is regular because

$$P^2 = \begin{bmatrix} 0.75 & 0.5 \\ 0.25 & 0.5 \end{bmatrix}$$

has positive entries. The matrix P in Example 6 is not regular because P and every positive power of P have some zero entries (verify).

The following theorem, which we state without proof, is the fundamental result about the long-term behavior of Markov chains.

THEOREM 4.12.1

If P is the transition matrix for a regular Markov chain, then:

(a) There is a unique probability vector $\mathbf{q}$ such that $P\mathbf{q} = \mathbf{q}$.

(b) For any initial probability vector $\mathbf{x}_0$, the sequence of state vectors

$$\mathbf{x}_0,\ P\mathbf{x}_0, \ldots,\ P^k\mathbf{x}_0, \ldots$$

converges to $\mathbf{q}$.

The vector $\mathbf{q}$ in this theorem is called the ***steady-state*** vector of the Markov chain. It can be found by rewriting the equation in part (a) as

$$(I - P)\mathbf{q} = 0$$

and then solving this equation for $\mathbf{q}$ subject to the requirement that $\mathbf{q}$ be a probability vector. Here are some examples.

EXAMPLE 7 Example 1 and Example 2 Revisited

The transition matrix for the Markov chain in Example 2 is

$$P = \begin{bmatrix} 0.8 & 0.1 \\ 0.2 & 0.9 \end{bmatrix}$$

Since the entries of P are positive, the Markov chain is regular and hence has a unique steady-state vector $\mathbf{q}$. To find $\mathbf{q}$ we will solve the system $(I - P)\mathbf{q} = 0$, which we can write as

$$\begin{bmatrix} 0.2 & -0.1 \\ -0.2 & 0.1 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The general solution of this system is

$$q_1 = 0.5s, \quad q_2 = s$$

(verify), which we can write in vector form as

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} 0.5s \\ s \end{bmatrix} = \begin{bmatrix} \frac{1}{2}s \\ s \end{bmatrix} \tag{13}$$

For $\mathbf{q}$ to be a probability vector, we must have

$$1 = q_1 + q_2 = \frac{3}{2}s$$

which implies that $s = \frac{2}{3}$. Substituting this value in 13 yields the steady-state vector

$$\boldsymbol{q} = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

which is consistent with the numerical results obtained in 9.

EXAMPLE 9 Example 4 Revisited

The transition matrix for the Markov chain in Example 4 is

$$P = \begin{bmatrix} 0.5 & 0.4 & 0.6 \\ 0.2 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.1 \end{bmatrix}$$

Since the entries of P are positive, the Markov chain is regular and hence has a unique steady-state vector $\mathbf{q}$. To find $\mathbf{q}$ we will solve the system $(I - P)\mathbf{q} = 0$, which we can write (using fractions) as

$$\begin{bmatrix} \frac{1}{2} & -\frac{2}{5} & -\frac{3}{5} \\ -\frac{1}{5} & \frac{4}{5} & -\frac{3}{10} \\ -\frac{3}{10} & -\frac{2}{5} & \frac{9}{10} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{14}$$

(We have converted to fractions to avoid roundoff error in this illustrative example.) We leave it for you to confirm that the reduced row echelon form of the coefficient matrix is

$$\begin{bmatrix} 1 & 0 & -\frac{15}{8} \\ 0 & 1 & -\frac{27}{32} \\ 0 & 0 & 0 \end{bmatrix}$$

and that the general solution of 14 is

$$q_1 = \frac{15}{8}s,\ q_2 = \frac{27}{32}s,\ q_3 = s \tag{15}$$

For $\mathbf{q}$ to be a probability vector we must have $q_1 + q_2 + q_3 = 1$, from which it follows that $s = \frac{32}{119}$ (verify). Substituting this value in 15 yields the steady-state vector

$$\mathbf{q} = \begin{bmatrix} \frac{60}{119} \\ \frac{27}{119} \\ \frac{32}{119} \end{bmatrix} \approx \begin{bmatrix} 0.5042 \\ 0.2269 \\ 0.2689 \end{bmatrix}$$

(verify), which is consistent with the results obtained in Example 4.

Concept Review

- Dynamical system
- State of a variable
- State of a dynamical system
- Stochastic process
- Probability
- Probability vector
- Stochastic matrix
- Markov chain
- Transition matrix
- Regular stochastic matrix
- Regular Markov chain
- Steady-state vector

Skills

- Determine whether a matrix is stochastic.
- Compute the state vectors from a transition matrix and an initial state.
- Determine whether a stochastic matrix is regular.
- Determine whether a Markov chain is regular.
- Find the steady-state vector for a regular transition matrix.

Exercise Set 4.12

In Exercises 1–2, determine whether A is a stochastic matrix. If A is not stochastic, then explain why not.

1. (a) $A = \begin{bmatrix} 0.4 & 0.3 \\ 0.6 & 0.7 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix}$

(c)
$$A=\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{3} \end{bmatrix}$$

(d)
$$A=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{3} & -\frac{2}{2} \\ \frac{1}{2} & \frac{1}{3} & 1 \end{bmatrix}$$

Answer:

(a) Stochastic
(b) Not stochastic
(c) Stochastic
(d) Not stochastic

2. (a) $A=\begin{bmatrix} 0.2 & 0.9 \\ 0.8 & 0.1 \end{bmatrix}$

(b) $A=\begin{bmatrix} 0.2 & 0.8 \\ 0.9 & 0.1 \end{bmatrix}$

(c)
$$A=\begin{bmatrix} \frac{1}{12} & \frac{1}{9} & \frac{1}{6} \\ \frac{1}{2} & 0 & \frac{5}{6} \\ \frac{5}{12} & \frac{8}{9} & 0 \end{bmatrix}$$

(d)
$$A=\begin{bmatrix} -1 & \frac{1}{3} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{2} \\ 2 & \frac{1}{3} & 0 \end{bmatrix}$$

In Exercises 3–4, use Formulas 11 and 12 to compute the state vector $\mathbf{x}_4$ in two different ways.

3. $P=\begin{bmatrix} 0.5 & 0.6 \\ 0.5 & 0.4 \end{bmatrix}; \mathbf{x}_0=\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$

Answer:

$$\begin{bmatrix} 0.54545 \\ 0.45455 \end{bmatrix}$$

4. $P=\begin{bmatrix} 0.8 & 0.5 \\ 0.2 & 0.5 \end{bmatrix}; \quad \mathbf{x}_0=\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

In Exercises 5–6, determine whether P is a regular stochastic matrix.

5. (a)

$$P = \begin{bmatrix} \frac{1}{5} & \frac{1}{7} \\ \frac{4}{5} & \frac{6}{7} \end{bmatrix}$$

(b)

$$P = \begin{bmatrix} \frac{1}{5} & 0 \\ \frac{4}{5} & 1 \end{bmatrix}$$

(c)

$$P = \begin{bmatrix} \frac{1}{5} & 1 \\ \frac{4}{5} & 0 \end{bmatrix}$$

Answer:

(a) Regular

(b) Not regular

(c) Regular

6. (a)

$$P = \begin{bmatrix} \frac{1}{2} & 1 \\ \frac{1}{2} & 0 \end{bmatrix}$$

(b)

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

(c)

$$P = \begin{bmatrix} \frac{3}{4} & \frac{1}{3} \\ \frac{1}{4} & \frac{2}{3} \end{bmatrix}$$

In Exercises 7–10, verify that P is a regular stochastic matrix, and find the steady-state vector for the associated Markov chain.

7.

$$P = \begin{bmatrix} \frac{1}{4} & \frac{2}{3} \\ \frac{3}{4} & \frac{1}{3} \end{bmatrix}$$

Answer:

$$\begin{bmatrix} \frac{8}{17} \\ \frac{9}{17} \end{bmatrix}$$

8. $P = \begin{bmatrix} 0.2 & 0.6 \\ 0.8 & 0.4 \end{bmatrix}$

9.

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{4} & 0 & \frac{2}{3} \end{bmatrix}$$

Answer:

$$\begin{bmatrix} \frac{4}{11} \\ \frac{4}{11} \\ \frac{3}{11} \end{bmatrix}$$

10.

$$P = \begin{bmatrix} \frac{1}{3} & \frac{1}{4} & \frac{2}{5} \\ 0 & \frac{3}{4} & \frac{2}{5} \\ \frac{2}{3} & 0 & \frac{1}{5} \end{bmatrix}$$

11. Consider a Markov process with transition matrix

$$\begin{array}{cc} & \begin{array}{cc} \textbf{State 1} & \textbf{State 2} \end{array} \\ \begin{array}{c} \textbf{State 1} \\ \textbf{State 2} \end{array} & \begin{bmatrix} 0.2 & 0.1 \\ 0.8 & 0.9 \end{bmatrix} \end{array}$$

(a) What does the entry 0.2 represent?

(b) What does the entry 0.1 represent?

(c) If the system is in state 1 initially, what is the probability that it will be in state 2 at the next observation?

(d) If the system has a 50% chance of being in state 1 initially, what is the probability that it will be in state 2 at the next observation?

Answer:

(a) Probability that something in state 1 stays in state 1

(b) Probability that something in state 2 moves to state 1

(c) 0.8

(d) 0.85

12. Consider a Markov process with transition matrix

$$\begin{array}{cc} & \begin{array}{cc} \textbf{State 1} & \textbf{State 2} \end{array} \\ \begin{array}{c} \textbf{State 1} \\ \textbf{State 2} \end{array} & \begin{bmatrix} 0 & \frac{1}{7} \\ 1 & \frac{6}{7} \end{bmatrix} \end{array}$$

(a) What does the entry $\frac{6}{7}$ represent?

(b) What does the entry 0 represent?

(c) If the system is in state 1 initially, what is the probability that it will be in state 1 at the next observation?

(d) If the system has a 50% chance of being in state 1 initially, what is the probability that it will be in state 2 at the next observation?

13. On a given day the air quality in a certain city is either good or bad. Records show that when the air quality is good on one day, then there is a 95% chance that it will be good the next day, and when the air quality is bad on one day, then there is a 45% chance that it will be bad the next day.

(a) Find a transition matrix for this phenomenon.

(b) If the air quality is good today, what is the probability that it will be good two days from now?

(c) If the air quality is bad today, what is the probability that it will be bad three days from now?

(d) If there is a 20% chance that the air quality will be good today, what is the probability that it will be good tomorrow?

Answer:

(a) $\begin{bmatrix} 0.95 & 0.55 \\ 0.05 & 0.45 \end{bmatrix}$

(b) 0.93

(c) 0.142

(d) 0.63

14. In a laboratory experiment, a mouse can choose one of two food types each day, type I or type II. Records show that if the mouse chooses type I on a given day, then there is a 75% chance that it will choose type I the next day, and if it chooses type II on one day, then there is a 50% chance that it will choose type II the next day.

(a) Find a transition matrix for this phenomenon.

(b) If the mouse chooses type I today, what is the probability that it will choose type I two days from now?

(c) If the mouse chooses type II today, what is the probability that it will choose type II three days from now?

(d) If there is a 10% chance that the mouse will choose type I today, what is the probability that it will choose type I tomorrow?

15. Suppose that at some initial point in time 100,000 people live in a certain city and 25,000 people live in its suburbs. The Regional Planning Commission determines that each year 5% of the city population moves to the suburbs and 3% of the suburban population moves to the city.

(a) Assuming that the total population remains constant, make a table that shows the populations of the city and its suburbs over a five-year period (round to the nearest integer).

(b) Over the long term, how will the population be distributed between the city and its suburbs?

Answer:

(a)

Year	**1**	**2**	**3**	**4**	**5**
City	95,750	91,840	88,243	84,933	81,889
Suburbs	29,250	33,160	36,757	40,067	43,111

(b)

City	46,875

Suburbs	78,125

16. Suppose that two competing television stations, station 1 and station 2, each have 50% of the viewer market at some initial point in time. Assume that over each one-year period station 1 captures 5% of station 2's market share and station 2 captures 10% of station 1's market share.

(a) Make a table that shows the market share of each station over a five-year period.

(b) Over the long term, how will the market share be distributed between the two stations?

17. Suppose that a car rental agency has three locations, numbered 1, 2, and 3. A customer may rent a car from any of the three locations and return it to any of the three locations. Records show that cars are rented and returned in accordance with the following probabilities:

		Rented from Location		
		1	**2**	**3**
	1	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{5}$
Returned to Location	**2**	$\frac{4}{5}$	$\frac{3}{10}$	$\frac{1}{5}$
	3	$\frac{1}{10}$	$\frac{1}{2}$	$\frac{1}{5}$

(a) Assuming that a car is rented from location 1, what is the probability that it will be at location 1 after two rentals?

(b) Assuming that this dynamical system can be modeled as a Markov chain, find the steady-state vector.

(c) If the rental agency owns 120 cars, how many parking spaces should it allocate at each location to be reasonably certain that it will have enough spaces for the cars over the long term? Explain your reasoning.

Answer:

(a) $\frac{23}{100}$

(b) $\begin{bmatrix} \frac{46}{159} \\ \frac{22}{53} \\ \frac{47}{159} \end{bmatrix}$

(c) 35, 50, 35

18. Physical traits are determined by the genes that an offspring receives from its parents. In the simplest case a trait in the offspring is determined by one pair of genes, one member of the pair inherited from the male parent and the other from the female parent. Typically, each gene in a pair can assume one of two forms, called ***alleles***, denoted by A and a. This leads to three possible pairings:

$$AA,\ Aa,\ aa$$

called ***genotypes*** (the pairs Aa and aA determine the same trait and hence are not distinguished from one another). It is shown in the study of heredity that if a parent of known genotype is crossed with a random parent of unknown genotype, then the offspring will have the genotype probabilities given in the following table, which can be viewed as a transition matrix for a Markov process:

		Genotype of Parent		
		AA	***Aa***	***aa***
	AA	$\frac{1}{2}$	$\frac{1}{4}$	0
Genotype of Offspring	***Aa***	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	aa	0	$\frac{1}{4}$	$\frac{1}{2}$

Thus, for example, the offspring of a parent of genotype *AA* that is crossed at random with a parent of unknown genotype will have a 50% chance of being *AA*, a 50% chance of being *Aa*, and no chance of being *aa*.

(a) Show that the transition matrix is regular.

(b) Find the steady-state vector, and discuss its physical interpretation.

19. Fill in the missing entries of the stochastic matrix

$$P = \begin{bmatrix} \frac{7}{10} & * & \frac{1}{5} \\ * & \frac{3}{10} & * \\ \frac{1}{10} & \frac{3}{5} & \frac{3}{10} \end{bmatrix}$$

and find its steady-state vector.

Answer:

$$P = \begin{bmatrix} \frac{7}{10} & \frac{1}{10} & \frac{1}{5} \\ \frac{1}{5} & \frac{3}{10} & \frac{1}{2} \\ \frac{1}{10} & \frac{3}{5} & \frac{3}{10} \end{bmatrix}; \quad \mathbf{q} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix}$$

20. If P is an $n \times n$ stochastic matrix, and if M is a $1 \times n$ matrix whose entries are all 1's, then $MP =$ ___

21. If P is a regular stochastic matrix with steady-state vector $\mathbf{q}$, what can you say about the sequence of products

$$P\mathbf{q}, \quad P^2\mathbf{q}, \quad P^3\mathbf{q}, \ldots, \quad P^k\mathbf{q}, \ldots$$

as $k \to \infty$?

Answer:

$P^k\mathbf{q} = \mathbf{q}$ for every positive integer k

22. (a) If P is a regular $n \times n$ stochastic matrix with steady-state vector $\mathbf{q}$, and if $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard unit vectors in column form, what can you say about the behavior of the sequence

$$P\mathbf{e}_i, \quad P^2\mathbf{e}_i, \quad P^3\mathbf{e}_i, \ldots, \quad P^k\mathbf{e}_i, \ldots$$

as $k \to \infty$ for each $i = 1, 2, \ldots, n$?

(b) What does this tell you about the behavior of the column vectors of P^k as $k \to \infty$?

23. Prove that the product of two stochastic matrices is a stochastic matrix. [*Hint:* Write each column of the product as a linear combination of the columns of the first factor.

24. Prove that if P is a stochastic matrix whose entries are all greater than or equal to p, then the entries of P^2 are greater than or equal to p.

True-False Exercises

In parts (a)–(e) determine whether the statement is true or false, and justify your answer.

(a) The vector $\begin{bmatrix} \frac{1}{3} \\ 0 \\ \frac{2}{3} \end{bmatrix}$ is a probability vector.

Answer:

True

(b) The matrix $\begin{bmatrix} 0.2 & 1 \\ 0.8 & 0 \end{bmatrix}$ is a regular stochastic matrix.

Answer:

True

(c) The column vectors of a transition matrix are probability vectors.

Answer:

True

(d) A steady-state vector for a Markov chain with transition matrix P is any solution of the linear system $(I - P)\mathbf{q} = 0$.

Answer:

False

(e) The square of every regular stochastic matrix is stochastic.

Answer:

True

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

Chapter 4 Supplementary Exercises

1. Let V be the set of all ordered pairs of real numbers, and consider the following addition and scalar multiplication operations on $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3), \quad k\mathbf{u} = (k\mathbf{u}_1, 0, 0)$$

(a) Compute $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for $\mathbf{u} = (3, -2, 4)$, $\mathbf{v} = (1, 5, -2)$, and $k = -1$.

(b) In words, explain why V is closed under addition and scalar multiplication.

(c) Since the addition operation on V is the standard addition operation on R^3, certain vector space axioms hold for V because they are known to hold for R^3. Which axioms in Definition 1 of Section 4.1 are they?

(d) Show that Axioms 7, 8, and 9 hold.

(e) Show that Axiom 10 fails for the given operations.

Answer:

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

(c) Axioms 1–5

2. In each part, the solution space of the system is a subspace of R^3 and so must be a line through the origin, a plane through the origin, all of R^3, or the origin only. For each system, determine which is the case. If the subspace is a plane, find an equation for it, and if it is a line, find parametric equations.

(a) $0x + 0y + 0z = 0$

(b)
$$\begin{aligned} 2x - 3y + z &= 0 \\ 6x - 9y + 3z &= 0 \\ -4x + 6y - 2z &= 0 \end{aligned}$$

(c)
$$\begin{aligned} x - 2y + 7z &= 0 \\ -4x + 8y + 5z &= 0 \\ 2x - 4y + 3z &= 0 \end{aligned}$$

(d)
$$\begin{aligned} x + 4y + 8z &= 0 \\ 2x + 5y + 6z &= 0 \\ 3x + y - 4z &= 0 \end{aligned}$$

3. For what values of s is the solution space of

$$\begin{aligned} x_1 + x_2 + sx_3 &= 0 \\ x_1 + sx_2 + x_3 &= 0 \\ sx_1 + x_2 + x_3 &= 0 \end{aligned}$$

the origin only, a line through the origin, a plane through the origin, or all of R^3?

Answer:

If $s \neq 1, \ -2,$ the solution space is the origin. If $s = 1$, the solution space is a plane through the origin. If $s = -2$, the solution space is a line through the origin.

4. (a) Express $(4a, a-b, a+2b)$ as a linear combination of $(4, 1, 1)$ and $(0, -1, 2)$.

(b) Express $(3a+b+3c, \ -a+4b-c, 2a+b+2c)$ as a linear combination of $(3, -1, 2)$ and $(1, 4, 1)$.

(c) Express $(2a-b+4c, 3a-c, 4b+c)$ as a linear combination of three nonzero vectors.

5. Let W be the space spanned by $\mathbf{f} = \sin x$ and $\mathbf{g} = \cos x$.

(a) Show that for any value of θ, $\mathbf{f}_1 = \sin(x+\theta)$ and $\mathbf{g}_1 = \cos(x+\theta)$ are vectors in W.

(b) Show that $\mathbf{f}_1$ and $\mathbf{g}_1$ form a basis for W.

6. (a) Express $\mathbf{v} = (1, 1)$ as a linear combination of $\mathbf{v}_1 = (1, -1)$, $\mathbf{v}_2 = (3, 0)$, and $\mathbf{v}_3 = (2, 1)$ in two different ways.

(b) Explain why this does not violate Theorem 4.4.1.

7. Let A be an $n \times n$ matrix, and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ be linearly independent vectors in R^n expressed as $n \times 1$ matrices. What must be true about A for $A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_n$ to be linearly independent?

Answer:

A must be invertible

8. Must a basis for P_n contain a polynomial of degree k for each $k = 0, 1, 2, \ldots, n$? Justify your answer.

9. For the purpose of this exercise, let us define a "checkerboard matrix" to be a square matrix $A = [a_{ij}]$ such that

$$a_{ij} = \begin{cases} 1 & \text{if } i+j \text{ is even} \\ 0 & \text{if } i+j \text{ is odd} \end{cases}$$

Find the rank and nullity of the following checkerboard matrices.

(a) The 3×3 checkerboard matrix.

(b) The 4×4 checkerboard matrix.

(c) The $n \times n$ checkerboard matrix.

Answer:

(a) $\text{Rank} = 2, \ \text{nullity} = 1$

(b) $\text{Rank} = 2, \ \text{nullity} = 2$

(c) $\text{Rank} = 2, \ \text{nullity} = n-2$

10. For the purpose of this exercise, let us define an "X-matrix" to be a square matrix with an odd number of rows and columns that has 0's everywhere except on the two diagonals where it has 1's. Find the rank and nullity of the following X-matrices.

(a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$

(c) the X-matrix of size $(2n+1) \times (2n+1)$

11. In each part, show that the stated set of polynomials is a subspace of P_n and find a basis for it.

(a) All polynomials in P_n such that $p(-x) = p(x)$.

(b) All polynomials in P_n such that $p(0) = 0$.

Answer:

(a) $\{1,\ x^2,\ x^4,\ \ldots,\ x^{2m}\}$ where $2m = n$ if n is even and $2m = n - 1$ if n is odd.

(b) $\{x,\ x^2,\ x^3,\ \ldots,\ x^n\}$

12. (***Calculus required***) Show that the set of all polynomials in P_n that have a horizontal tangent at $x = 0$ is a subspace of P_n. Find a basis for this subspace.

13. (a) Find a basis for the vector space of all 3×3 symmetric matrices.

(b) Find a basis for the vector space of all 3×3 skew-symmetric matrices.

Answer:

(a) $\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}$

(b) $\left\{ \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \right\}$

14. Various advanced texts in linear algebra prove the following determinant criterion for rank: *The rank of a matrix A is r if and only if A has some $r \times r$ submatrix with a nonzero determinant, and all square submatrices of larger size have determinant zero*. [*Note:* A submatrix of A is any matrix obtained by deleting rows or columns of A. The matrix A itself is also considered to be a submatrix of A.] In each part, use this criterion to find the rank of the matrix.

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

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

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

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

15. Use the result in Exercise 14 above to find the possible ranks for matrices of the form

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & a_{16} \\ 0 & 0 & 0 & 0 & 0 & a_{26} \\ 0 & 0 & 0 & 0 & 0 & a_{36} \\ 0 & 0 & 0 & 0 & 0 & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \end{bmatrix}$$

Answer:

Possible ranks are 2, 1, and 0.

16. Prove: If S is a basis for a vector space V, then for any vectors $\mathbf{u}$ and $\mathbf{v}$ in V and any scalar k, the following relationships hold.

(a) $(\mathbf{u}+\mathbf{v})_S = (\mathbf{u})_S + (\mathbf{v})_S$

(b) $(k\mathbf{u})_S = k(\mathbf{u})_S$

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

CHAPTER

5 Eigenvalues and Eigenvectors

CHAPTER CONTENTS

5.1. Eigenvalues and Eigenvectors

5.2. Diagonalization

5.3. Complex Vector Spaces

5.4. Differential Equations

INTRODUCTION

In this chapter we will focus on classes of scalars and vectors known as "eigenvalues" and "eigenvectors," terms derived from the German word *eigen*, meaning "own," "peculiar to," "characteristic," or "individual." The underlying idea first appeared in the study of rotational motion but was later used to classify various kinds of surfaces and to describe solutions of certain differential equations. In the early 1900s it was applied to matrices and matrix transformations, and today it has applications in such diverse fields as computer graphics, mechanical vibrations, heat flow, population dynamics, quantum mechanics, and economics to name just a few.

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

5.1 Eigenvalues and Eigenvectors

In this section we will define the notions of "eigenvalue" and "eigenvector" and discuss some of their basic properties.

Definition of Eigenvalue and Eigenvector

We begin with the main definition in this section.

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*** λ.

The requirement that an eigenvector be nonzero is imposed to avoid the unimportant case $A0 = \lambda 0$, which holds for every A and λ

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) $\lambda \le 1$ (c) $-1 \le \lambda \le 0$ (d) $\lambda \le -1$

Figure 5.1.1

EXAMPLE 1 Eigenvector of a 2 × 2 Matrix

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 (Figure 5.1.2).

Figure 5.1.2

Computing Eigenvalues and Eigenvectors

Our next objective is to obtain a general procedure for finding eigenvalues and eigenvectors of an $n \times n$ matrix A. We will begin with the problem of finding the eigenvalues of A. Note first that the equation $A\mathbf{x} = \lambda\mathbf{x}$ can be rewritten as $A\mathbf{x} = \lambda I\mathbf{x}$, or equivalently, as

$$(\lambda I - A)\mathbf{x} = 0$$

For λ to be an eigenvalue of A this equation must have a nonzero solution for $\mathbf{x}$. But it follows from parts (*b*) and (*g*) of Theorem 4.10.4 that this is so if and only if the coefficient matrix $\lambda I - A$ has a zero determinant. Thus, we have the following result.

THEOREM 5.1.1

If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation

$$\det(\lambda I - A) = 0 \tag{1}$$

This is called the **characteristic equation** of A.

EXAMPLE 2 Finding Eigenvalues

In Example 1 we observed that $\lambda = 3$ is an eigenvalue of the matrix

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

but we did not explain how we found it. Use the characteristic equation to find all eigenvalues of this matrix.

Solution It follows from Formula 1 that the eigenvalues of A are the solutions of the equation $\det(\lambda I - A) = 0$, which we can write as

$$\begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

from which we obtain

$$(\lambda - 3)(\lambda + 1) = 0 \tag{2}$$

This shows that the eigenvalues of A are $\lambda = 3$ and $\lambda = -1$. Thus, in addition to the eigenvalue $\lambda = 3$ noted in Example 1, we have discovered a second eigenvalue $\lambda = -1$.

When the determinant $\det(\lambda I - A)$ that appears on the left side of 1 is expanded, the result is a polynomial $p(\lambda)$ of degree n that is called the ***characteristic polynomial*** of A. For example, it follows from 2 that the characteristic polynomial of the 2×2 matrix A in Example 2 is

$$p(\lambda) = (\lambda - 3)(\lambda + 1) = \lambda^2 - 2\lambda - 3$$

which is a polynomial of degree 2. In general, the characteristic polynomial of an $n \times n$ matrix has the form

$$p(\lambda) = \lambda^n + c_1\lambda^{n-1} + \ldots + c_n$$

in which the coefficient of λ^n is 1 (Exercise 17). Since a polynomial of degree n has at most n distinct roots, it follows that the equation

$$\lambda^n + c_1\lambda^{n-1} + \ldots + c_n = 0 \tag{3}$$

has at most n distinct solutions and consequently that an $n \times n$ matrix has at most n distinct eigenvalues. Since some of these solutions may be complex numbers, it is possible for a matrix to have complex eigenvalues, even if that matrix itself has real entries. We will discuss this issue in more detail later, but for now we will focus on examples in which the eigenvalues are real numbers.

EXAMPLE 3 Eigenvalues of a 3 × 3 Matrix

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

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0 \tag{4}$$

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*

$$\lambda^n + c_1\lambda^{n-1} + \ldots + c_n = 0$$

In applications involving large matrices it is often not feasible to compute the characteristic equation directly so other methods must be used to find eigenvalues. We will consider such methods in Chapter 9.

must be divisors of the constant term, c_n. Thus, the only possible integer solutions of 4 are the divisors of -4, that is, ± 1, ± 2, ± 4. Successively substituting these values in 4 shows that $\lambda = 4$ is an integer solution. As a consequence, $\lambda - 4$ must be a factor of the left side of 4. Dividing $\lambda - 4$ into $\lambda^3 - 8\lambda^2 + 17\lambda - 4$ shows that 4 can be rewritten as

$$(\lambda - 4)\left(\lambda^2 - 4\lambda + 1\right) = 0$$

Thus, the remaining solutions of 4 satisfy the quadratic equation

$$\lambda^2 - 4\lambda + 1 = 0$$

which can be solved by the quadratic formula. Thus the eigenvalues of A are

$$\lambda = 4, \quad \lambda = 2 + \sqrt{3}, \quad \text{and} \quad \lambda = 2 - \sqrt{3}$$

EXAMPLE 4 Eigenvalues of an Upper Triangular Matrix

Find the eigenvalues of the upper triangular matrix

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

Solution Recalling that the determinant of a triangular matrix is the product of the entries on the main diagonal (Theorem 2.1.2), we obtain

$$\begin{aligned}\det(\lambda I - A) &= \det\begin{bmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & -a_{14} \\ 0 & \lambda - a_{22} & -a_{23} & -a_{24} \\ 0 & 0 & \lambda - a_{33} & -a_{34} \\ 0 & 0 & 0 & \lambda - a_{44} \end{bmatrix} \\ &= (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44})\end{aligned}$$

Thus, the characteristic equation is

$$(\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) = 0$$

and the eigenvalues are

$$\lambda = a_{11}, \quad \lambda = a_{22}, \quad \lambda = a_{33}, \quad \lambda = a_{44}$$

which are precisely the diagonal entries of A.

The following general theorem should be evident from the computations in the preceding example.

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.

EXAMPLE 5 Eigenvalues of a Lower Triangular Matrix

By inspection, 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 $\lambda = -\frac{1}{4}$.

Had Theorem 5.1.2 been available earlier, we could have anticipated the result obtained in Example 2.

THEOREM 5.1.3

If A is an $n \times n$ matrix, the following statements are equivalent.

(a) λ is an eigenvalue of A.

(b) The system of equations $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.

(c) There is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$

(d) λ is a solution of the characteristic equation $\det(\lambda I - A) = 0$

Finding Eigenvectors and Bases for Eigenspaces

Now that we know how to find the eigenvalues of a matrix, we will consider the problem of finding the corresponding eigenvectors. Since the eigenvectors corresponding to an eigenvalue λ of a matrix A are the nonzero vectors that satisfy the equation

$$(\lambda I - A)\mathbf{x} = \mathbf{0}$$

these eigenvectors are the nonzero vectors in the null space of the matrix $\lambda I - A$. We call this null space the ***eigenspace*** of A corresponding to λ. Stated another way, *the eigenspace of A corresponding to the eigenvalue λ is the solution space of the homogeneous system* $(\lambda \mathbf{I} - A)\mathbf{x} = \mathbf{0}$.

Notice that $\mathbf{x} = \mathbf{0}$ is in every eigenspace even though it is not an eigenvector. Thus, it is the *nonzero* vectors in an eigenspace that are the eigenvectors.

EXAMPLE 6 Bases for Eigenspaces

Find bases for the eigenspaces of the matrix

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

Solution In Example 1 we found the characteristic equation of A to be

$$(\lambda - 3)(\lambda + 1) = 0$$

from which we obtained the eigenvalues $\lambda = 3$ and $\lambda = -1$. Thus, there are two eigenspaces of A, one corresponding to each of these eigenvalues.

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 $\mathbf{x}$ is a nontrivial solution of $(\lambda I - A)\mathbf{x} = \mathbf{0}$, that is, of

$$\begin{bmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If $\lambda = 3$, then this equation becomes

$$\begin{bmatrix} 0 & 0 \\ -8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

whose general solution is

$$x_1 = \frac{1}{2}t, \quad x_2 = t$$

(verify) or in matrix form,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

Thus,

$$\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 3$. We leave it as an exercise for you to follow the pattern of these computations and show that

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

is a basis for the eigenspace corresponding to $\lambda = -1$.

Historical Note Methods of linear algebra are used in the emerging field of computerized face recognition. Researchers are working with the idea that every human face in a racial group is a combination of a few dozen primary shapes. For example, by analyzing three-dimensional scans of many faces, researchers at Rockefeller University have produced both an average head shape in the

Caucasian group—dubbed the ***meanhead*** (top row left in the figure to the left)—and a set of standardized variations from that shape, called ***eigenheads*** (15 of which are shown in the picture). These are so named because they are eigenvectors of a certain matrix that stores digitized facial information. Face shapes are represented mathematically as linear combinations of the eigenheads. [*Image: Courtesy Dr. Joseph Atick, Dr. Norman Redlich, and Dr. Paul Griffith*]

EXAMPLE 7 Eigenvectors and Bases for Eigenspaces

Find bases for the eigenspaces of

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

Solution The characteristic equation of A is $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$, or in factored form, $(\lambda - 1)(\lambda - 2)^2 = 0$ (verify). 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,

$$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

In the case where $\lambda = 2$, Formula 5 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 (verify)

$$x_1 = -s, \ x_2 = t, \ x_3 = s$$

Thus, the eigenvectors of A corresponding to $\lambda = 2$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Since

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

are linearly independent (why?), these vectors form a basis for the eigenspace corresponding to $\lambda = 2$.

If $\lambda = 1$, then 5 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 (verify)

$$x_1 = -2s, \ x_2 = s, \ x_3 = s$$

Thus, the eigenvectors corresponding to $\lambda = 1$ are the nonzero vectors of the form

$$\begin{bmatrix} -2s \\ s \\ s \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \text{ so that } \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 1$.

Powers of a Matrix

Once the eigenvalues and eigenvectors of a matrix A are found, it is a simple matter to find the eigenvalues and eigenvectors of any positive integer power of A; for example, if λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, then

$$A^2\mathbf{x} = A(A\mathbf{x}) = A(\lambda\mathbf{x}) = \lambda(A\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$$

which shows that λ^2 is an eigenvalue of A^2 and that $\mathbf{x}$ is a corresponding eigenvector. In general, we have the following result.

THEOREM 5.1.4

If k is a positive integer, λ is an eigenvalue of a matrix A, and $\mathbf{x}$ is a corresponding eigenvector, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

EXAMPLE 8 Powers of a Matrix

In Example 7 we showed that the eigenvalues of

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

are $\lambda = 2$ and $\lambda = 1$, so from Theorem 5.1.4 both $\lambda = 2^7 = 128$ and $\lambda = 1^7 = 1$ are eigenvalues of A^7. We also showed that

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

are eigenvectors of A corresponding to the eigenvalue $\lambda = 2$, so from Theorem 5.1.4 they are also eigenvectors of A^7 corresponding to $\lambda = 2^7 = 128$. Similarly, the eigenvector

$$\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

of A corresponding to the eigenvalue $\lambda = 1$ is also an eigenvector of A^7 corresponding to $\lambda = 1^7 = 1$.

Eigenvalues and Invertibility

The next theorem establishes a relationship between eigenvalues and the invertibility of a matrix.

THEOREM 5.1.5

A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof Assume that A is an $n \times n$ matrix and observe first that $\lambda = 0$ is a solution of the characteristic equation

$$\lambda^n + c_1\lambda^{n-1} + \ldots + c_n = 0$$

if and only if the constant term c_n is zero. Thus, it suffices to prove that A is invertible if and only if $c_n \neq 0$. But

$$\det(\lambda I - A) = \lambda^n + c_1\lambda^{n-1} + \ldots + c_n$$

or, on setting $\lambda = 0$,

$$\det(-A) = c_n \text{ or } (-1)^n \det(A) = c_n$$

It follows from the last equation that $\det(A) = 0$ if and only if $c_n = 0$, and this in turn implies that A is invertible if and only if $c_n \neq 0$.

EXAMPLE 9 Eigenvalues and Invertibility

The matrix A in Example 7 is invertible since it has eigenvalues $\lambda = 1$ and $\lambda = 2$, neither of which is zero. We leave it for you to check this conclusion by showing that $\det(A) \neq 0$.

More on the Equivalence Theorem

As our final result in this section, we will use Theorem 5.1.5 to add one additional part to Theorem 4.10.4.

THEOREM 5.1.6 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix b.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix b.

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

(h) The column vectors of A are linearly independent.

(i) The row vectors of A are linearly independent.

(j) The column vectors of A span R^n.

(k) The row vectors of A span R^n.

(l) The column vectors of A form a basis for R^n.

(m) The row vectors of A form a basis for R^n.

(n) A has $rank\ n$.

(o) A has nullity 0.

(p) The orthogonal complement of the null space of A is R^n.

(q) The orthogonal complement of the row space of A is $\{\mathbf{0}\}$.

(r) The range of T_A is R^n.

(s) T_A is one-to-one.

(t) $\lambda = 0$ is not an eigenvalue of A.

This theorem relates all of the major topics we have studied thus far.

Concept Review

- Eigenvector
- Eigenvalue
- Characteristic equation
- Characteristic polynomial
- Eigenspace
- Equivalence Theorem

Skills

- Find the eigenvalues of a matrix.
- Find bases for the eigenspaces of a matrix.

Exercise Set 5.1

In Exercises 1–2, confirm by multiplication that **x** is an eigenvector of A, and find the corresponding eigenvalue.

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

Answer:

5

2. $A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}; \ \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

3. Find the characteristic equations of the following matrices:

(a) $\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$

(b) $\begin{bmatrix} 10 & -9 \\ 4 & -2 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} -2 & -7 \\ 1 & 2 \end{bmatrix}$

(e) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

(f) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Answer:

(a) $\lambda^2 - 2\lambda - 3 = 0$

(b) $\lambda^2 - 8\lambda + 16 = 0$

(c) $\lambda^2 - 12 = 0$

(d) $\lambda^2 + 3 = 0$

(e) $\lambda^2 = 0$

(f) $\lambda^2 - 2\lambda + 1 = 0$

4. Find the eigenvalues of the matrices in Exercise 3

5. Find bases for the eigenspaces of the matrices in Exercise 3

Answer:

(a) Basis for eigenspace corresponding to $\lambda = 3$: $\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$; basis for eigenspace corresponding to $\lambda = -1$: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(b) Basis for eigenspace corresponding to $\lambda = 4$: $\begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$

(c) Basis for eigenspace corresponding to $\lambda = \sqrt{12}$: $\begin{bmatrix} \frac{3}{\sqrt{12}} \\ 1 \end{bmatrix}$; basis for eigenspace corresponding to $\lambda = -\sqrt{12}$: $\begin{bmatrix} -\frac{3}{\sqrt{12}} \\ 1 \end{bmatrix}$

(d) There are no eigenspaces.

(e) Basis for eigenspace corresponding to $\lambda = 0$: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(f) Basis for eigenspace corresponding to $\lambda = 1$: $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

6. Find the characteristic equations of the following matrices:

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

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

(c) $\begin{bmatrix} -2 & 0 & 1 \\ -6 & -2 & 0 \\ 19 & 5 & -4 \end{bmatrix}$

(d) $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{bmatrix}$

(e) $\begin{bmatrix} 5 & 0 & 1 \\ 1 & 1 & 0 \\ -7 & 1 & 0 \end{bmatrix}$

(f) $\begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}$

7. Find the eigenvalues of the matrices in Exercise 6.

Answer:

(a) 1, 2, 3

(b) $-\sqrt{2}, 0, \sqrt{2}$

(c) -8

(d) 2

(e) 2

(f) $-4, 3$

8. Find bases for the eigenspaces of the matrices in Exercise 6.

9. Find the characteristic equations of the following matrices:

(a) $\begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 10 & -9 & 0 & 0 \\ 4 & -2 & 0 & 0 \\ 0 & 0 & -2 & -7 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

Answer:

(a) $\lambda^4+\lambda^3-3\lambda^2-\lambda+2=0$

(b) $\lambda^4-8\lambda^3+19\lambda^2-24\lambda+48=0$

10. Find the eigenvalues of the matrices in Exercise 9.

11. Find bases for the eigenspaces of the matrices in Exercise 9.

Answer:

(a)

$$\lambda=1:\text{basis}\begin{bmatrix}2\\3\\1\\0\end{bmatrix},\ \begin{bmatrix}0\\0\\0\\1\end{bmatrix};\ \lambda=-2:\text{basis}\begin{bmatrix}-1\\0\\1\\0\end{bmatrix};\ \lambda=-1:\text{basis}\begin{bmatrix}-2\\1\\1\\0\end{bmatrix}$$

(b)

$$\lambda=4:\text{basis}\begin{bmatrix}\frac{3}{2}\\1\\0\\0\end{bmatrix}$$

12. By inspection, find the eigenvalues of the following matrices:

(a) $\begin{bmatrix}-1 & 6\\0 & 5\end{bmatrix}$

(b) $\begin{bmatrix}3 & 0 & 0\\-2 & 7 & 0\\4 & 8 & 1\end{bmatrix}$

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

13. Find the eigenvalues of A^9 for

$$A=\begin{bmatrix}1 & 3 & 7 & 11\\0 & \frac{1}{2} & 3 & 8\\0 & 0 & 0 & 4\\0 & 0 & 0 & 2\end{bmatrix}$$

Answer:

$1,\ \left(\frac{1}{2}\right)^9=\frac{1}{512},\ 2^9=512$

14. Find the eigenvalues and bases for the eigenspaces of A^{25} for

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

15. Let A be a 2×2 matrix, and call a line through the origin of R^2 *invariant* under A if $A\mathbf{x}$ lies on the line when $\mathbf{x}$ does. Find equations for all lines in R^2, if any, that are invariant under the given matrix.

(a) $A = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}$

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

(c) $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

Answer:

(a) $y = x$ and $y = 2x$

(b) No lines

(c) $y = 0$

16. Find $\det(A)$ given that A has $p(\lambda)$ as its characteristic polynomial.

(a) $p(\lambda) = \lambda^3 - 2\lambda^2 + \lambda + 5$

(b) $p(\lambda) = \lambda^4 - \lambda^3 + 7$

[*Hint:* See the proof of Theorem 5.1.5.]

17. Let A be an $n \times n$ matrix.

(a) Prove that the characteristic polynomial of A has degree n.

(b) Prove that the coefficient of λ^n in the characteristic polynomial is 1.

18. Show that the characteristic equation of a 2×2 matrix A can be expressed as $\lambda^2 - \mathrm{tr}(A)\lambda + \det(A) = 0$, where $\mathrm{tr}(A)$ is the trace of A.

19. Use the result in Exercise 18 to show that if

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

then the solutions of the characteristic equation of A are

$$\lambda = \frac{1}{2}\left[(a+d) \pm \sqrt{(a-d)^2 + 4bc}\right]$$

Use this result to show that A has

(a) two distinct real eigenvalues if $(a-d)^2 + 4\mathrm{bc} > 0$.

(b) two repeated real eigenvalues if $(a-d)^2 + 4bc = 0$.

(c) complex conjugate eigenvalues if $(a-d)^2 + 4bc < 0$.

20. Let A be the matrix in Exercise 19. Show that if $b \neq 0$, then

$$\mathbf{x}_1 = \begin{bmatrix} -b \\ a - \lambda_1 \end{bmatrix} \text{ and } \mathbf{x}_2 = \begin{bmatrix} -b \\ a - \lambda_2 \end{bmatrix}$$

are eigenvectors of A that correspond, respectively, to the eigenvalues

$$\lambda_1 = \frac{1}{2}\left[(a+d) + \sqrt{(a-d)^2 + 4bc}\right]$$

and

$$\lambda_2 = \frac{1}{2}\left[(a+d) - \sqrt{(a-d)^2 + 4bc}\right]$$

21. Use the result of Exercise 18 to prove that if $p(\lambda)$ is the characteristic polynomial of a 2×2 matrix A, then $p(A) = 0$.

22. Prove: If a, b, c, and d are integers such that $a + b = c + d$, then

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

has integer eigenvalues—namely, $\lambda_1 = a + b$ and $\lambda_2 = a - c$.

23. Prove: If λ is an eigenvalue of an invertible matrix A, and $\mathbf{x}$ is a corresponding eigenvector, then $1/\lambda$ is an eigenvalue of A^{-1}, and $\mathbf{x}$ is a corresponding eigenvector.

24. Prove: If λ is an eigenvalue of A, $\mathbf{x}$ is a corresponding eigenvector, and s is a scalar, then $\lambda - s$ is an eigenvalue of $A - sI$, and $\mathbf{x}$ is a corresponding eigenvector.

25. Prove: If λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, then $s\lambda$ is an eigenvalue of sA for every scalar s, and $\mathbf{x}$ is a corresponding eigenvector.

26. Find the eigenvalues and bases for the eigenspaces of

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

and then use Exercises 23 and 24 to find the eigenvalues and bases for the eigenspaces of

(a) A^{-1}

(b) $A - 3I$

(c) $A + 2I$

27. (a) Prove that if A is a square matrix, then A and A^T have the same eigenvalues. [*Hint:* Look at the characteristic equation $\det(\lambda I - A) = 0$.]

(b) Show that A and A^T need not have the same eigenspaces. [*Hint:* Use the result in Exercise 20 to find a 2×2 matrix for which A and A^T have different eigenspaces.]

28. Suppose that the characteristic polynomial of some matrix A is found to be $p(\lambda) = (\lambda - 1)(\lambda - 3)^2(\lambda - 4)^3$. In each part, answer the question and explain your reasoning.

(a) What is the size of A?

(b) Is A invertible?

(c) How many eigenspaces does A have?

29. The eigenvectors that we have been studying are sometimes called ***right eigenvectors*** to distinguish them from ***left eigenvectors***, which are $n \times 1$ column matrices $\mathbf{x}$ that satisfy the equation $\mathbf{x}^T A = \mu \mathbf{x}^T$ for some scalar μ. What is the relationship, if any, between the right eigenvectors and corresponding eigenvalues λ of A and the left eigenvectors and corresponding eigenvalues μ of A?

True-False Exercises

In parts (a)–(g) determine whether the statement is true or false, and justify your answer.

(a) If A is a square matrix and $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero scalar λ, then $\mathbf{x}$ is an eigenvector of A.

Answer:

False

(b) If λ is an eigenvalue of a matrix A, then the linear system $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has only the trivial solution.

Answer:

False

(c) If the characteristic polynomial of a matrix A is $p(\lambda) = \lambda^2 + 1$, then A is invertible.

Answer:

True

(d) If λ is an eigenvalue of a matrix A, then the eigenspace of A corresponding to λ is the set of eigenvectors of A corresponding to λ.

Answer:

False

(e) If 0 is an eigenvalue of a matrix A, then A^2 is singular.

Answer:

True

(f) The eigenvalues of a matrix A are the same as the eigenvalues of the reduced row echelon form of A.

Answer:

False

(g) If 0 is an eigenvalue of a matrix A, then the set of columns of A is linearly independent.

Answer:

False

Copyright © 2010 John Wiley & Sons, Inc. All rights reserved.

5.2 Diagonalization

In this section we will be concerned with the problem of finding a basis for R^n that consists of eigenvectors of an $n \times n$ matrix A. Such bases can be used to study geometric properties of A and to simplify various numerical computations. These bases are also of physical significance in a wide variety of applications, some of which will be considered later in this text.

The Matrix Diagonalization Problem

Our first objective in this section is to show that the following two seemingly different problems are equivalent.

Problem 1 Given an $n \times n$ matrix A, does there exist an invertible matrix P such that $P^{-1}AP$ is diagonal?

Problem 2 Given an $n \times n$ matrix A, does A have n linearly independent eigenvectors?

Similarity

The matrix product $P^{-1}AP$ that appears in Problem 1 is called a ***similarity transformation*** of the matrix A. Such products are important in the study of eigenvectors and eigenvalues, so we will begin with some terminology about them.

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 B as $B = Q^{-1}AQ$ 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.

Similarity Invariants

Similar matrices have many properties in common. For example, if $B = P^{-1}AP$, then it follows that A and B have the same determinant, since

$$\begin{aligned}\det(B) &= \det\left(P^{-1}AP\right) = \det\left(P^{-1}\right)\det(A)\det(P) \\ &= \frac{1}{\det(P)}\det(A)\det(P) = \det(A)\end{aligned}$$

In general, any property that is shared by all similar matrices is called a ***similarity invariant*** or is said to be ***invariant under similarity***. Table 1 lists the most important similarity invariants. The proofs of some of these results are given as exercises.

Table 1 Similarity Invariants

Property	**Description**
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A and hence of $P^{-1}AP$, then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ have the same dimension.

Expressed in the language of similarity, Problem 1 posed above is equivalent to asking whether the matrix A is similar to a diagonal matrix. If so, the diagonal matrix will have all of the similarity-invariant properties of A, but will have a simpler form, making it easier to analyze and work with. This important idea has some associated terminology.

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.

The following theorem shows that Problems 1 and 2 posed above are actually two different forms of the same mathematical problem.

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.

Part (*b*) of Theorem 5.2.1 is equivalent to saying that there is a *basis* for R^n consisting of eigenvectors of A. Why?

Proof (a)⇒(b) Since A is assumed to be diagonalizable, it follows that there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ or, equivalently,

$$AP = PD \tag{1}$$

If we denote the column vectors of P by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, and if we assume that the diagonal entries of D are $\lambda_1, \lambda_2, \ldots, \lambda_n$, then by Formula 6 of Section 1.3 the left side of 1 can be expressed as

$$AP = A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \ldots \quad \mathbf{p}_n] = [A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \ldots \quad A\mathbf{p}_n]$$

and, as noted in the comment following Example 1 of Section 1.7, the right side of 1 can be expressed as

$$PD = [\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \ldots \quad \lambda_n\mathbf{p}_n]$$

Thus, it follows from 1 that

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \ldots, \quad A\mathbf{p}_n = \lambda_n\mathbf{p}_n \tag{2}$$

Since P is invertible, we know from Theorem 5.1.6 that its column vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are linearly independent (and hence nonzero). Thus, it follows from 2 that these n column vectors are eigenvectors of A.

Proof (b)⇒(a) Assume that A has n linearly independent eigenvectors, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, and that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the corresponding eigenvalues. If we let

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \ldots \quad \mathbf{p}_n]$$

and if we let D be the diagonal matrix that has $\lambda_1, \lambda_2, \ldots, \lambda_n$ as its successive diagonal entries, then

$$\begin{aligned} AP &= A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \ldots \quad \mathbf{p}_n] = [A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \ldots \quad A\mathbf{p}_n] \\ &= [\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \ldots \quad \lambda_n\mathbf{p}_n] = PD \end{aligned}$$

Since the column vectors of P are linearly independent, it follows from Theorem 5.1.6 that P is invertible, so that this last equation can be rewritten as $P^{-1}AP = D$, which shows that A is diagonalizable.

Procedure for Diagonalizing a Matrix

The preceding theorem guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and the proof suggests the following method for diagonalizing A.

Procedure for Diagonalizing a Matrix

Step 1. Confirm that the matrix is actually diagonalizable by finding n linearly independent eigenvectors. One way to do this is by finding a basis for each eigenspace and merging these basis vectors into a single set S. If this set has fewer than n vectors, then the matrix is not diagonalizable.

Step 2. Form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$ that has the vectors in S as its column vectors.

Step 3. The matrix $P^{-1}AP$ will be diagonal and have the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ corresponding to the eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ as its successive diagonal entries.

EXAMPLE 1 Finding a Matrix *P* That Diagonalizes a Matrix *A*

Find a matrix P that diagonalizes

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

Solution In Example 7 of the preceding section we found the characteristic equation of A to be

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

and we found the following bases for the eigenspaces:

$$\lambda = 2: \ \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \quad \lambda = 1: \ \mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

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, you should verify that

$$P^{-1}AP = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 21 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In general, there is no preferred order for the columns of P. Since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P, changing the order of the columns of P just changes the order of the eigenvalues on the diagonal of $P^{-1}AP$. Thus, had we written

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

in the preceding example, we would have obtained

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

EXAMPLE 2 A Matrix That Is Not Diagonalizable

Find a matrix P that diagonalizes

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

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

$$(\lambda - 1)(\lambda - 2)^2 = 0$$

Thus, the distinct eigenvalues of A are $\lambda = 1$ and $\lambda = 2$. We leave it for you to show that 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.

Alternative Solution If you are concerned only in determining whether a matrix is diagonalizable and not with actually finding a diagonalizing matrix P, then it is not necessary to compute bases for the eigenspaces—it suffices to find the dimensions of the eigenspaces. For this example, the eigenspace corresponding to $\lambda = 1$ is the solution space of the system

$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & -5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since the coefficient matrix has rank 2 (verify), the nullity of this matrix is 1 by Theorem 4.8.2, and hence the eigenspace corresponding to $\lambda = 1$ is one-dimensional.

The eigenspace corresponding to $\lambda = 2$ is the solution space of the system

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 3 & -5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This coefficient matrix also has rank 2 and nullity 1 (verify), so the eigenspace corresponding to $\lambda = 2$ is also one-dimensional. Since the eigenspaces produce a total of two basis vectors, and since three are needed, the matrix A is not diagonalizable.

There is an assumption in Example 1 that the column vectors of P, which are made up of basis vectors from the various eigenspaces of A, are linearly independent. The following theorem, proved at the end of this section, shows that this is so.

THEOREM 5.2.2

If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of a matrix A corresponding to distinct eigenvalues, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

Remark Theorem 5.2.2 is a special case of a more general result: Suppose that $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues and that we choose a linearly independent set in each of the corresponding eigenspaces. If we then merge all these vectors into a single set, the result will still be a linearly independent set. For example, if we choose three linearly independent vectors from one eigenspace and two linearly independent vectors from another eigenspace, then the five vectors together form a linearly independent set. We omit the proof.

As a consequence of Theorem 5.2.2, we obtain the following important result.

THEOREM 5.2.3

If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Proof If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are eigenvectors corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then by Theorem 5.2.2, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are linearly independent. Thus, A is diagonalizable by Theorem 5.2.1.

EXAMPLE 3 Using Theorem 5.2.3

We saw in Example 3 of the preceding section that

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

has three distinct eigenvalues: $\lambda = 4$, $\lambda = 2 + \sqrt{3}$, and $\lambda = 2 - \sqrt{3}$. Therefore, A is diagonalizable and

$$P^{-1}AP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2+\sqrt{3} & 0 \\ 0 & 0 & 2-\sqrt{3} \end{bmatrix}$$

for some invertible matrix P. If needed, the matrix P can be found using the method shown in Example 1 of this section.

EXAMPLE 4 Diagonalizability of Triangular Matrices

From Theorem 5.1.2, the eigenvalues of a triangular matrix are the entries on its main diagonal. Thus, a triangular matrix with distinct entries on the main diagonal is diagonalizable. For example,

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

is a diagonalizable matrix with eigenvalues $\lambda_1 = -1, \lambda_2 = 3, \lambda_3 = 5, \lambda_4 = -2$.

Computing Powers of a Matrix

There are many applications in which it is necessary to compute high powers of a square matrix A. We will show next that if A happens to be diagonalizable, then the computations can be simplified by diagonalizing A.

To start, suppose that A is a diagonalizable $n \times n$ matrix, that P diagonalizes A, and that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} = D$$

Squaring both sides of this equation yields

$$\left(P^{-1}AP^2\right) = \begin{bmatrix} \lambda_1^2 & 0 & \dots & 0 \\ 0 & \lambda_2^2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n^2 \end{bmatrix} = D^2$$

We can rewrite the left side of this equation as

$$\left(P^{-1}AP\right)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P$$

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

$$P^{-1}A^kP = D^k = \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}$$

which we can rewrite as

$$A^k = PD^kP^{-1} = P\begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix}P^{-1} \tag{3}$$

Formula 3 reveals that raising a diagonalizable matrix A to a positive integer power has the effect of raising its eigenvalues to that power.

Note that computing the right side of this formula involves only three matrix multiplications and the powers of the diagonal entries of D. For matrices of large size and high powers of λ, this involves substantially fewer operations than computing A^k directly.

EXAMPLE 5 Power of a Matrix

Use 3 to find A^{13}, where

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

Solution We showed in Example 1 that the matrix A is diagonalized by

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

and that

$$D = P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Thus, it follows from 3 that

$$\begin{aligned} A^{13} = PD^{13}P^{-1} &= \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix}\begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix} \end{aligned} \tag{4}$$

Remark With the method in the preceding example, most of the work is in diagonalizing A. Once that work is done, it can be used to compute any power of A. Thus, to compute A^{1000} we need only change the exponents from 13 to 1000 in 4.

Eigenvalues of Powers of a Matrix

Once the eigenvalues and eigenvectors of any square matrix A are found, it is a simple matter to find the eigenvalues and eigenvectors of any positive integer power of A. For example, if λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, then

$$A^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$$

which shows not only that λ^2 is an eigenvalue of A^2 but that $\mathbf{x}$ is a corresponding eigenvector. In general, we have the following result.

Note that diagonalizability is not a requirement in Theorem 5.2.4.

THEOREM 5.2.4

If λ is an eigenvalue of a square matrix A and $\mathbf{x}$ is a corresponding eigenvector, and if k is any positive integer, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.

Some problems that use this theorem are given in the exercises.

Geometric and Algebraic Multiplicity

Theorem 5.2.3 does not completely settle the diagonalizability question since it only guarantees that a square matrix with n distinct eigenvalues is diagonalizable, but does not preclude the possibility that there may exist diagonalizable matrices with fewer than n distinct eigenvalues. The following example shows that this is indeed the case.

EXAMPLE 6 The Converse of Theorem 5.2.3 Is False

Consider the matrices

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

It follows from Theorem 5.1.2 that both of these matrices have only one distinct eigenvalue, namely $\lambda = 1$, and hence only one eigenspace. We leave it as an exercise for you to solve the characteristic equations

$$(\lambda I - I)\mathbf{x} = 0 \text{ and } (\lambda J - I)\mathbf{x} = 0$$

with $\lambda = 1$ and show that for I the eigenspace is three-dimensional (all of R^3) and for J it is one-dimensional, consisting of all scalar multiples of

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

This shows that the converse of Theorem 5.2.3 is false, since we have produced two 3×3 matrices with fewer than three distinct eigenvalues, one of which is diagonalizable and the other of which is not.

A full excursion into the study of diagonalizability is left for more advanced courses, but we will touch on one theorem that is important to a fuller understanding of diagonalizability. It can be proved that if λ_0 is an eigenvalue of A, then the dimension of the eigenspace corresponding to λ_0 cannot exceed the number of times that $\lambda - \lambda_0$ appears as a factor of the characteristic polynomial of A. For example, in Example 1 and Example 2 the characteristic polynomial is

$$(\lambda - 1)(\lambda - 2)^2$$

Thus, the eigenspace corresponding to $\lambda = 1$ is at most (hence exactly) one-dimensional, and the eigenspace corresponding to $\lambda = 2$ is at most two-dimensional. In Example 1 the eigenspace corresponding to $\lambda = 2$ actually had dimension 2, resulting in diagonalizability, but in Example 2 the eigenspace corresponding to $\lambda = 2$ had only dimension 1, resulting in nondiagonalizability.

There is some terminology that is related to these ideas. If λ_0 is an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the ***geometric multiplicity*** of λ_0, and the number of times that $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the ***algebraic multiplicity*** of λ_0. The following theorem, which we state without proof, summarizes the preceding discussion.

THEOREM 5.2.5 Geometric and Algebraic Multiplicity

If A is a square matrix, then:

(a) For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.

(b) A is diagonalizable if and only if the geometric multiplicity of every eigenvalue is equal to the algebraic multiplicity.

OPTIONAL

We will complete this section with an optional proof of Theorem 5.2.2.

Proof of Theorem 5.2.2 Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. We will assume that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent and obtain a contradiction. We can then conclude that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent.

Since an eigenvector is nonzero by definition, $\{\mathbf{v}_1\}$ is linearly independent. Let r be the largest integer such that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is linearly independent. Since we are assuming that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly dependent, r satisfies $1 \le r < k$. Moreover, by the definition of r, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{r+1}\}$ is linearly dependent. Thus, there are scalars $c_1, c_2, \ldots, c_{r+1}$, not all zero, such that

$$c_1\mathbf{v}_1 + \mathbf{c}_2\mathbf{v}_2 + \ldots + \mathbf{c}_{r+1}\mathbf{v}_{r+1} = 0 \tag{5}$$

Multiplying both sides of 5 by A and using the fact that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad \mathbf{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \ldots, \quad \mathbf{A}\mathbf{v}_{r+1} = \lambda_{r+1}\mathbf{v}_{r+1}$$

we obtain

$$c_1\lambda_1\mathbf{v}_1 + \mathbf{c}_2\lambda_2\mathbf{v}_2 + \ldots + \mathbf{c}_{r+1}\lambda_{r+1}\mathbf{v}_{r+1} = 0 \tag{6}$$

If we now multiply both sides of 5 by λ_{r+1} and subtract the resulting equation from 6 we obtain

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{v}_1 + \mathbf{c}_2(\lambda_2 - \lambda_{r+1})\mathbf{v}_2 + \ldots + \mathbf{c}_r(\lambda_r - \lambda_{r+1})\mathbf{v}_r = 0$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a linearly independent set, this equation implies that

$$c_1(\lambda_1 - \lambda_{r+1}) = c_2(\lambda_2 - \lambda_{r+1}) = \ldots = c_r(\lambda_r - \lambda_{r+1}) = 0$$

and since $\lambda_1, \lambda_2, \ldots, \lambda_{r+1}$ are assumed to be distinct, it follows that

$$c_1 = c_2 = \ldots = c_r = 0 \tag{7}$$

Substituting these values in 5 yields

$$c_{r+1}\mathbf{v}_{r+1} = \mathbf{0}$$

Since the eigenvector $\mathbf{v}_{r+1}$ is nonzero, it follows that

$$c_{r+1} = 0 \tag{8}$$

But equations 7 and 8 contradict the fact that $c_1, c_2, \ldots, c_{r+1}$ are not all zero so the proof is complete.

Concept Review

- Similarity transformation
- Similarity invariant
- Similar matrices
- Diagonalizable matrix
- Geometric multiplicity
- Algebraic multiplicity

Skills

- Determine whether a square matrix A is diagonalizable.
- Diagonalize a square matrix A.
- Find powers of a matrix using similarity.
- Find the geometric multiplicity and the algebraic multiplicity of an eigenvalue.

Exercise Set 5.2

In Exercises 1–4, show that A and B are not similar matrices.

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

Answer:

Possible reason: Determinants are different.

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

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

Answer:

Possible reason: Ranks are different.

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

5. Let A be a 6×6 matrix with characteristic equation $\lambda^2(\lambda - 1)(\lambda - 2)^3 = 0$. What are the possible dimensions for eigenspaces of A?

Answer:

$\lambda = 0: 1 \text{ or } 2;\ \lambda = 1: 1;\ \lambda = 2: 1, 2, \text{ or } 3$

6. Let

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

(a) Find the eigenvalues of A.

(b) For each eigenvalue λ, find the rank of the matrix $\lambda I - A$.

(c) Is A diagonalizable? Justify your conclusion.

In Exercises 7–11, use the method of Exercise 6 to determine whether the matrix is diagonalizable.

7. $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$

Answer:

Not diagonalizable

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

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

Answer:

Not diagonalizable

10. $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{bmatrix}$

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

Answer:

Not diagonalizable

In Exercises 12–15, find a matrix P that diagonalizes A, and compute $P^{-1}AP$.

12. $A = \begin{bmatrix} -14 & 12 \\ -20 & 17 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$

Answer:

$$P = \begin{bmatrix} \frac{1}{3} & 0 \\ 1 & 1 \end{bmatrix};\ P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

14. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

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

Answer:

$$P = \begin{bmatrix} -2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix};\ P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

In Exercises 16–21, find the geometric and algebraic multiplicity of each eigenvalue of the matrix A, and determine whether A is diagonalizable. If A is diagonalizable, then find a matrix P that diagonalizes A, and find $P^{-1}AP$.

16. $A = \begin{bmatrix} 19 & -9 & -6 \\ 25 & -11 & -9 \\ 17 & -9 & -4 \end{bmatrix}$

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

Answer:

$$P = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 4 \end{bmatrix}; \; P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

18. $A = \begin{bmatrix} 5 & 0 & 0 \\ 1 & 5 & 0 \\ 0 & 1 & 5 \end{bmatrix}$

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

Answer:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}; \; P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

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

Answer:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \; P^{-1}AP = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

22. Use the method of Example 5 to compute A^{10}, where

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

23. Use the method of Example 5 to compute A^{11}, where