

5

Eigenvalues and Eigenvectors

INTRODUCTORY EXAMPLE

Dynamical Systems and Spotted Owls

In 1990, the northern spotted owl became the center of a nationwide controversy over the use and misuse of the majestic forests in the Pacific Northwest. Environmentalists convinced the federal government that the owl was threatened with extinction if logging continued in the old-growth forests (with trees over 200 years old), where the owls prefer to live. The timber industry, anticipating the loss of 30,000 to 100,000 jobs as a result of new government restrictions on logging, argued that the owl should not be classified as a “threatened species” and cited a number of published scientific reports to support its case.¹

Caught in the crossfire of the two lobbying groups, mathematical ecologists intensified their drive to understand the population dynamics of the spotted owl. The life cycle of a spotted owl divides naturally into three stages: juvenile (up to 1 year old), subadult (1 to 2 years), and adult (over 2 years). The owls mate for life during the subadult and adult stages, begin to breed as adults, and live for up to 20 years. Each owl pair requires about 1000 hectares (4 square miles) for its own home territory. A critical time in the life cycle is when the juveniles leave the nest. To survive and become a subadult, a juvenile must successfully find a new home range (and usually a mate).

¹ “The Great Spotted Owl War,” *Reader’s Digest*, November 1992, pp. 91–95.



A first step in studying the population dynamics is to model the population at yearly intervals, at times denoted by $k = 0, 1, 2, \dots$. Usually, one assumes that there is a 1:1 ratio of males to females in each life stage and counts only the females. The population at year k can be described by a vector $\mathbf{x}_k = (j_k, s_k, a_k)$, where j_k , s_k , and a_k are the numbers of females in the juvenile, subadult, and adult stages, respectively.

Using actual field data from demographic studies, R. Lamberson and co-workers considered the following *stage-matrix model*:²

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

Here the number of new juvenile females in year $k + 1$ is .33 times the number of adult females in year k (based on the average birth rate per owl pair). Also, 18% of the juveniles survive to become subadults, and 71% of the subadults and 94% of the adults survive to be counted as adults.

The stage-matrix model is a difference equation of the form $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Such an equation is often called a

² R. H. Lamberson, R. McKelvey, B. R. Noon, and C. Voss, “A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment,” *Conservation Biology* 6 (1992), 505–512. Also, a private communication from Professor Lamberson, 1993.

dynamical system (or a **discrete linear dynamical system**) because it describes the changes in a system as time passes.

The 18% juvenile survival rate in the Lamberson stage matrix is the entry affected most by the amount of old-growth forest available. Actually, 60% of the juveniles normally survive to leave the nest, but in the Willow Creek region of California studied by Lamberson and his colleagues, only 30% of the juveniles that left the nest were able to find new home ranges. The rest perished during the search process.

A significant reason for the failure of owls to find new home ranges is the increasing fragmentation of old-growth timber stands due to clear-cutting of scattered areas on the old-growth land. When an owl leaves the protective canopy of the forest and crosses a clear-cut area, the risk of attack by predators increases dramatically. Section 5.6 will show that the model described above predicts the eventual demise of the spotted owl, but that if 50% of the juveniles who survive to leave the nest also find new home ranges, then the owl population will thrive.

WEB

The goal of this chapter is to dissect the action of a linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ into elements that are easily visualized. Except for a brief digression in Section 5.4, all matrices in the chapter are square. The main applications described here are to discrete dynamical systems, including the spotted owls discussed above. However, the basic concepts—eigenvectors and eigenvalues—are useful throughout pure and applied mathematics, and they appear in settings far more general than we consider here. Eigenvalues are also used to study differential equations and *continuous* dynamical systems, they provide critical information in engineering design, and they arise naturally in fields such as physics and chemistry.

5.1 EIGENVECTORS AND EIGENVALUES

Although a transformation $\mathbf{x} \mapsto A\mathbf{x}$ may move vectors in a variety of directions, it often happens that there are special vectors on which the action of A is quite simple.

EXAMPLE 1 Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The images of \mathbf{u} and \mathbf{v} under multiplication by A are shown in Fig. 1. In fact, $A\mathbf{v}$ is just $2\mathbf{v}$. So A only “stretches,” or dilates, \mathbf{v} . ■

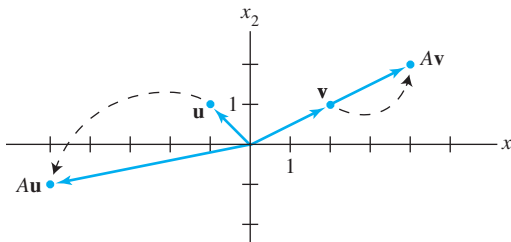


FIGURE 1 Effects of multiplication by A .

As another example, readers of Section 4.9 will recall that if A is a stochastic matrix, then the steady-state vector \mathbf{q} for A satisfies the equation $A\mathbf{x} = \mathbf{x}$. That is, $A\mathbf{q} = 1 \cdot \mathbf{q}$.

This section studies equations such as

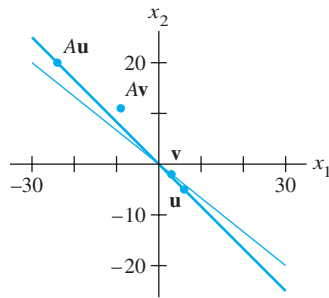
$$A\mathbf{x} = 2\mathbf{x} \quad \text{or} \quad A\mathbf{x} = -4\mathbf{x}$$

where special vectors are transformed by A into scalar multiples of themselves.

DEFINITION

An **eigenvector** of an $n \times n$ matrix A is a nonzero vector \mathbf{x} such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ . A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution \mathbf{x} of $A\mathbf{x} = \lambda\mathbf{x}$; such an \mathbf{x} is called an *eigenvector corresponding to λ* .¹

It is easy to determine if a given vector is an eigenvector of a matrix. It is also easy to decide if a specified scalar is an eigenvalue.



$A\mathbf{u} = -4\mathbf{u}$, but $A\mathbf{v} \neq \lambda\mathbf{v}$.

EXAMPLE 2 Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Are \mathbf{u} and \mathbf{v} eigenvectors of A ?

SOLUTION

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Thus \mathbf{u} is an eigenvector corresponding to an eigenvalue (-4) , but \mathbf{v} is not an eigenvector of A , because $A\mathbf{v}$ is not a multiple of \mathbf{v} . ■

EXAMPLE 3 Show that 7 is an eigenvalue of matrix A in Example 2, and find the corresponding eigenvectors.

SOLUTION The scalar 7 is an eigenvalue of A if and only if the equation

$$A\mathbf{x} = 7\mathbf{x} \tag{1}$$

has a nontrivial solution. But (1) is equivalent to $A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$, or

$$(A - 7I)\mathbf{x} = \mathbf{0} \tag{2}$$

To solve this homogeneous equation, form the matrix

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

The columns of $A - 7I$ are obviously linearly dependent, so (2) has nontrivial solutions. Thus 7 is an eigenvalue of A . To find the corresponding eigenvectors, use row operations:

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

The general solution has the form $x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$. ■

¹Note that an eigenvector must be *nonzero*, by definition, but an eigenvalue may be zero. The case in which the number 0 is an eigenvalue is discussed after Example 5.

Warning: Although row reduction was used in Example 3 to find eigenvectors, it cannot be used to find eigenvalues. An echelon form of a matrix A usually does *not* display the eigenvalues of A .

The equivalence of equations (1) and (2) obviously holds for any λ in place of $\lambda = 7$. Thus λ is an eigenvalue of an $n \times n$ matrix A if and only if the equation

$$(A - \lambda I)\mathbf{x} = \mathbf{0} \quad (3)$$

has a nontrivial solution. The set of *all* solutions of (3) is just the null space of the matrix $A - \lambda I$. So this set is a *subspace* of \mathbb{R}^n and is called the **eigenspace** of A corresponding to λ . The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ .

Example 3 shows that for matrix A in Example 2, the eigenspace corresponding to $\lambda = 7$ consists of *all* multiples of $(1, 1)$, which is the line through $(1, 1)$ and the origin. From Example 2, you can check that the eigenspace corresponding to $\lambda = -4$ is the line through $(6, -5)$. These eigenspaces are shown in Fig. 2, along with eigenvectors $(1, 1)$ and $(3/2, -5/4)$ and the geometric action of the transformation $\mathbf{x} \mapsto A\mathbf{x}$ on each eigenspace.

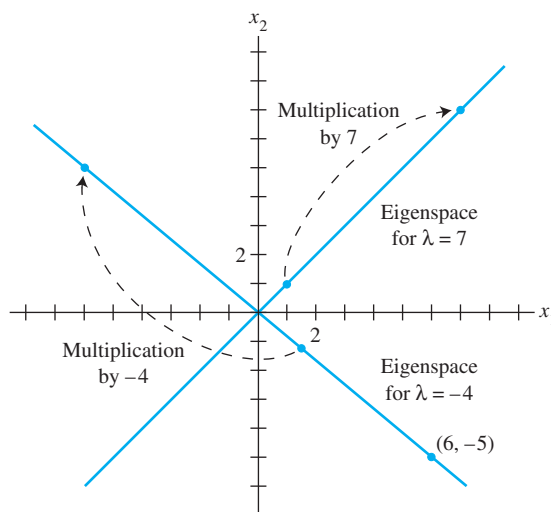


FIGURE 2 Eigenspaces for $\lambda = -4$ and $\lambda = 7$.

EXAMPLE 4 Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is 2. Find a basis for the corresponding eigenspace.

SOLUTION Form

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

and row reduce the augmented matrix for $(A - 2I)\mathbf{x} = \mathbf{0}$:

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

At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A - 2I)\mathbf{x} = \mathbf{0}$ has free variables. The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ and } x_3 \text{ free}$$

The eigenspace, shown in Fig. 3, is a two-dimensional subspace of \mathbb{R}^3 . A basis is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

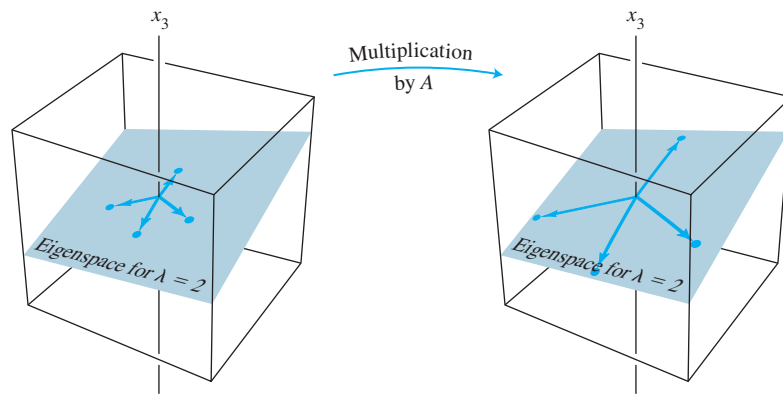


FIGURE 3 A acts as a dilation on the eigenspace.

NUMERICAL NOTE

Example 4 shows a good method for manual computation of eigenvectors in simple cases when an eigenvalue is known. Using a matrix program and row reduction to find an eigenspace (for a specified eigenvalue) usually works, too, but this is not entirely reliable. Roundoff error can lead occasionally to a reduced echelon form with the wrong number of pivots. The best computer programs compute approximations for eigenvalues and eigenvectors simultaneously, to any desired degree of accuracy, for matrices that are not too large. The size of matrices that can be analyzed increases each year as computing power and software improve.

The following theorem describes one of the few special cases in which eigenvalues can be found precisely. Calculation of eigenvalues will also be discussed in Section 5.2.

THEOREM 1

The eigenvalues of a triangular matrix are the entries on its main diagonal.

PROOF For simplicity, consider the 3×3 case. If A is upper triangular, then $A - \lambda I$

has the form

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix} \end{aligned}$$

The scalar λ is an eigenvalue of A if and only if the equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution, that is, if and only if the equation has a free variable. Because of the zero entries in $A - \lambda I$, it is easy to see that $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a free variable if and only if at least one of the entries on the diagonal of $A - \lambda I$ is zero. This happens if and only if λ equals one of the entries a_{11}, a_{22}, a_{33} in A . For the case in which A is lower triangular, see Exercise 28. ■

EXAMPLE 5 Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. The eigenvalues of A are 3, 0, and 2. The eigenvalues of B are 4 and 1. ■

What does it mean for a matrix A to have an eigenvalue of 0, such as in Example 5? This happens if and only if the equation

$$A\mathbf{x} = 0\mathbf{x} \tag{4}$$

has a nontrivial solution. But (4) is equivalent to $A\mathbf{x} = \mathbf{0}$, which has a nontrivial solution if and only if A is not invertible. Thus 0 is an eigenvalue of A if and only if A is not invertible. This fact will be added to the Invertible Matrix Theorem in Section 5.2.

The following important theorem will be needed later. Its proof illustrates a typical calculation with eigenvectors.

THEOREM 2

If $\mathbf{v}_1, \dots, \mathbf{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \dots, \lambda_r$ of an $n \times n$ matrix A , then the set $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is linearly independent.

PROOF Suppose $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ is linearly dependent. Since \mathbf{v}_1 is nonzero, Theorem 7 in Section 1.7 says that one of the vectors in the set is a linear combination of the preceding vectors. Let p be the least index such that \mathbf{v}_{p+1} is a linear combination of the preceding (linearly independent) vectors. Then there exist scalars c_1, \dots, c_p such that

$$c_1\mathbf{v}_1 + \dots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{5}$$

Multiplying both sides of (5) by A and using the fact that $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$ for each k , we obtain

$$\begin{aligned} c_1A\mathbf{v}_1 + \dots + c_pA\mathbf{v}_p &= A\mathbf{v}_{p+1} \\ c_1\lambda_1\mathbf{v}_1 + \dots + c_p\lambda_p\mathbf{v}_p &= \lambda_{p+1}\mathbf{v}_{p+1} \end{aligned} \tag{6}$$

Multiplying both sides of (5) by λ_{p+1} and subtracting the result from (6), we have

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \dots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0} \tag{7}$$

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is linearly independent, the weights in (7) are all zero. But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct. Hence $c_i = 0$ for $i = 1, \dots, p$. But then (5) says that $\mathbf{v}_{p+1} = \mathbf{0}$, which is impossible. Hence $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ cannot be linearly dependent and therefore must be linearly independent. ■

Eigenvectors and Difference Equations

This section concludes by showing how to construct solutions of the first-order difference equation discussed in the chapter introductory example:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad (k = 0, 1, 2, \dots) \quad (8)$$

If A is an $n \times n$ matrix, then (8) is a *recursive* description of a sequence $\{\mathbf{x}_k\}$ in \mathbb{R}^n . A **solution** of (8) is an explicit description of $\{\mathbf{x}_k\}$ whose formula for each \mathbf{x}_k does not depend directly on A or on the preceding terms in the sequence other than the initial term \mathbf{x}_0 .

The simplest way to build a solution of (8) is to take an eigenvector \mathbf{x}_0 and its corresponding eigenvalue λ and let

$$\mathbf{x}_k = \lambda^k \mathbf{x}_0 \quad (k = 1, 2, \dots) \quad (9)$$

This sequence is a solution because

$$A\mathbf{x}_k = A(\lambda^k \mathbf{x}_0) = \lambda^k (A\mathbf{x}_0) = \lambda^k (\lambda \mathbf{x}_0) = \lambda^{k+1} \mathbf{x}_0 = \mathbf{x}_{k+1}$$

Linear combinations of solutions in the form of equation (9) are solutions, too! See Exercise 33.

PRACTICE PROBLEMS

- Is 5 an eigenvalue of $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?
- If \mathbf{x} is an eigenvector of A corresponding to λ , what is $A^3\mathbf{x}$?
- Suppose that \mathbf{b}_1 and \mathbf{b}_2 are eigenvectors corresponding to distinct eigenvalues λ_1 and λ_2 , respectively, and suppose that \mathbf{b}_3 and \mathbf{b}_4 are linearly independent eigenvectors corresponding to a third distinct eigenvalue λ_3 . Does it necessarily follow that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ is a linearly independent set? [*Hint*: Consider the equation $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$.]

5.1 EXERCISES

- Is $\lambda = 2$ an eigenvalue of $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? Why or why not?
 - Is $\lambda = -3$ an eigenvalue of $\begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$? Why or why not?
 - Is $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 1 & -1 \\ 6 & -4 \end{bmatrix}$? If so, find the eigenvalue.
 - Is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 5 & 2 \\ 3 & 6 \end{bmatrix}$? If so, find the eigenvalue.
 - Is $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} -4 & 3 & 3 \\ 2 & -3 & -2 \\ -1 & 0 & -2 \end{bmatrix}$? If so, find the eigenvalue.
 - Is $\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 2 & 7 \\ 5 & 6 & 4 \end{bmatrix}$? If so, find the eigenvalue.
 - Is $\lambda = 4$ an eigenvalue of $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? If so, find one corresponding eigenvector.
 - Is $\lambda = 1$ an eigenvalue of $\begin{bmatrix} 4 & -2 & 3 \\ 0 & -1 & 3 \\ -1 & 2 & -2 \end{bmatrix}$? If so, find one corresponding eigenvector.
- In Exercises 9–16, find a basis for the eigenspace corresponding to each listed eigenvalue.
- $A = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}, \lambda = 1, 3$

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

11. $A = \begin{bmatrix} 1 & -3 \\ -4 & 5 \end{bmatrix}, \lambda = -1, 7$

12. $A = \begin{bmatrix} 4 & 1 \\ 3 & 6 \end{bmatrix}, \lambda = 3, 7$

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

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

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

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

Find the eigenvalues of the matrices in Exercises 17 and 18.

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

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

19. For $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, find one eigenvalue, with no calculation. Justify your answer.

20. Without calculation, find one eigenvalue and two linearly independent eigenvectors of $A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$. Justify your answer.

In Exercises 21 and 22, A is an $n \times n$ matrix. Mark each statement True or False. Justify each answer.

21. a. If $A\mathbf{x} = \lambda\mathbf{x}$ for some vector \mathbf{x} , then λ is an eigenvalue of A .
 b. A matrix A is not invertible if and only if 0 is an eigenvalue of A .
 c. A number c is an eigenvalue of A if and only if the equation $(A - cI)\mathbf{x} = \mathbf{0}$ has a nontrivial solution.
 d. Finding an eigenvector of A may be difficult, but checking whether a given vector is in fact an eigenvector is easy.
 e. To find the eigenvalues of A , reduce A to echelon form.
22. a. If $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ , then \mathbf{x} is an eigenvector of A .
 b. If \mathbf{v}_1 and \mathbf{v}_2 are linearly independent eigenvectors, then they correspond to distinct eigenvalues.

c. A steady-state vector for a stochastic matrix is actually an eigenvector.

d. The eigenvalues of a matrix are on its main diagonal.

e. An eigenspace of A is a null space of a certain matrix.

23. Explain why a 2×2 matrix can have at most two distinct eigenvalues. Explain why an $n \times n$ matrix can have at most n distinct eigenvalues.

24. Construct an example of a 2×2 matrix with only one distinct eigenvalue.

25. Let λ be an eigenvalue of an invertible matrix A . Show that λ^{-1} is an eigenvalue of A^{-1} . [Hint: Suppose a nonzero \mathbf{x} satisfies $A\mathbf{x} = \lambda\mathbf{x}$.]

26. Show that if A^2 is the zero matrix, then the only eigenvalue of A is 0.

27. Show that λ is an eigenvalue of A if and only if λ is an eigenvalue of A^T . [Hint: Find out how $A - \lambda I$ and $A^T - \lambda I$ are related.]

28. Use Exercise 27 to complete the proof of Theorem 1 for the case in which A is lower triangular.

29. Consider an $n \times n$ matrix A with the property that the row sums all equal the same number s . Show that s is an eigenvalue of A . [Hint: Find an eigenvector.]

30. Consider an $n \times n$ matrix A with the property that the column sums all equal the same number s . Show that s is an eigenvalue of A . [Hint: Use Exercises 27 and 29.]

In Exercises 31 and 32, let A be the matrix of the linear transformation T . Without writing A , find an eigenvalue of A and describe the eigenspace.

31. T is the transformation on \mathbb{R}^2 that reflects points across some line through the origin.

32. T is the transformation on \mathbb{R}^3 that rotates points about some line through the origin.

33. Let \mathbf{u} and \mathbf{v} be eigenvectors of a matrix A , with corresponding eigenvalues λ and μ , and let c_1 and c_2 be scalars. Define

$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \dots)$$

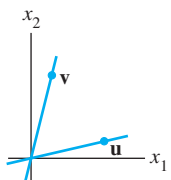
a. What is \mathbf{x}_{k+1} , by definition?

b. Compute $A\mathbf{x}_k$ from the formula for \mathbf{x}_k , and show that $A\mathbf{x}_k = \mathbf{x}_{k+1}$. This calculation will prove that the sequence $\{\mathbf{x}_k\}$ defined above satisfies the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \dots$).

34. Describe how you might try to build a solution of a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \dots$) if you were given the initial \mathbf{x}_0 and this vector did not happen to be an eigenvector of A . [Hint: How might you relate \mathbf{x}_0 to eigenvectors of A ?

35. Let \mathbf{u} and \mathbf{v} be the vectors shown in the figure, and suppose \mathbf{u} and \mathbf{v} are eigenvectors of a 2×2 matrix A that correspond to eigenvalues 2 and 3, respectively. Let $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be the linear transformation given by $T(\mathbf{x}) = A\mathbf{x}$ for each \mathbf{x} in \mathbb{R}^2 , and let $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Make a copy of the figure, and on

the same coordinate system, carefully plot the vectors $T(\mathbf{u})$, $T(\mathbf{v})$, and $T(\mathbf{w})$.



36. Repeat Exercise 35, assuming \mathbf{u} and \mathbf{v} are eigenvectors of A that correspond to eigenvalues -1 and 3 , respectively.

[M] In Exercises 37–40, use a matrix program to find the eigenvalues of the matrix. Then use the method of Example 4 with a row reduction routine to produce a basis for each eigenspace.

$$37. \begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix} \quad 38. \begin{bmatrix} 5 & -2 & 2 & -4 \\ 7 & -4 & 2 & -4 \\ 4 & -4 & 2 & 0 \\ 3 & -1 & 1 & -3 \end{bmatrix}$$

$$39. \begin{bmatrix} 12 & -90 & 30 & 30 & 30 \\ 8 & -49 & 15 & 15 & 15 \\ 16 & -52 & 12 & 0 & 20 \\ 0 & -30 & 10 & 22 & 10 \\ 8 & -41 & 15 & 15 & 7 \end{bmatrix}$$

$$40. \begin{bmatrix} -23 & 57 & -9 & -15 & -59 \\ -10 & 12 & -10 & 2 & -22 \\ 11 & 5 & -3 & -19 & -15 \\ -27 & 31 & -27 & 25 & -37 \\ -5 & -15 & -5 & 1 & 31 \end{bmatrix}$$

SOLUTIONS TO PRACTICE PROBLEMS

1. The number 5 is an eigenvalue of A if and only if the equation $(A - 5I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution. Form

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

and row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

At this point, it is clear that the homogeneous system has no free variables. Thus $A - 5I$ is an invertible matrix, which means that 5 is *not* an eigenvalue of A .

2. If \mathbf{x} is an eigenvector of A corresponding to λ , then $A\mathbf{x} = \lambda\mathbf{x}$ and so

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Again, $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. The general pattern, $A^k\mathbf{x} = \lambda^k\mathbf{x}$, is proved by induction.

3. Yes. Suppose $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 = \mathbf{0}$. Since any linear combination of eigenvectors from the same eigenvalue is again an eigenvector for that eigenvalue, $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ is an eigenvector for λ_3 . By Theorem 2, the vectors \mathbf{b}_1 , \mathbf{b}_2 , and $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ are linearly independent, so

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$$

implies $c_1 = c_2 = 0$. But then, c_3 and c_4 must also be zero since \mathbf{b}_3 and \mathbf{b}_4 are linearly independent. Hence all the coefficients in the original equation must be zero, and the vectors \mathbf{b}_1 , \mathbf{b}_2 , \mathbf{b}_3 , and \mathbf{b}_4 are linearly independent.

5.2 THE CHARACTERISTIC EQUATION

Useful information about the eigenvalues of a square matrix A is encoded in a special scalar equation called the characteristic equation of A . A simple example will lead to the general case.

EXAMPLE 1 Find the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

SOLUTION We must find all scalars λ such that the matrix equation

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

has a nontrivial solution. By the Invertible Matrix Theorem in Section 2.3, this problem is equivalent to finding all λ such that the matrix $A - \lambda I$ is *not* invertible, where

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix}$$

By Theorem 4 in Section 2.2, this matrix fails to be invertible precisely when its determinant is zero. So the eigenvalues of A are the solutions of the equation

$$\det(A - \lambda I) = \det \begin{bmatrix} 2 - \lambda & 3 \\ 3 & -6 - \lambda \end{bmatrix} = 0$$

Recall that

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

So

$$\begin{aligned} \det(A - \lambda I) &= (2 - \lambda)(-6 - \lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21 \\ &= (\lambda - 3)(\lambda + 7) \end{aligned}$$

If $\det(A - \lambda I) = 0$, then $\lambda = 3$ or $\lambda = -7$. So the eigenvalues of A are 3 and -7 . ■

The determinant in Example 1 transformed the matrix equation $(A - \lambda I)\mathbf{x} = \mathbf{0}$, which involves *two* unknowns (λ and \mathbf{x}), into the scalar equation $\lambda^2 + 4\lambda - 21 = 0$, which involves only one unknown. The same idea works for $n \times n$ matrices. However, before turning to larger matrices, we summarize the properties of determinants needed to study eigenvalues.

Determinants

Let A be an $n \times n$ matrix, let U be any echelon form obtained from A by row replacements and row interchanges (without scaling), and let r be the number of such row interchanges. Then the **determinant** of A , written as $\det A$, is $(-1)^r$ times the product of the diagonal entries u_{11}, \dots, u_{nn} in U . If A is invertible, then u_{11}, \dots, u_{nn} are all *pivots* (because $A \sim I_n$ and the u_{ii} have not been scaled to 1's). Otherwise, at least u_{nn} is zero, and the product $u_{11} \cdots u_{nn}$ is zero. Thus¹

$$\det A = \begin{cases} (-1)^r \cdot \left(\text{product of} \right. \\ \left. \text{pivots in } U \right), & \text{when } A \text{ is invertible} \\ 0, & \text{when } A \text{ is not invertible} \end{cases} \quad (1)$$

¹Formula (1) was derived in Section 3.2. Readers who have not studied Chapter 3 may use this formula as the definition of $\det A$. It is a remarkable and nontrivial fact that any echelon form U obtained from A without scaling gives the same value for $\det A$.

EXAMPLE 2 Compute $\det A$ for $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

SOLUTION The following row reduction uses one row interchange:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U_1$$

So $\det A$ equals $(-1)^1(1)(-2)(-1) = -2$. The following alternative row reduction avoids the row interchange and produces a different echelon form. The last step adds $-1/3$ times row 2 to row 3:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & 0 & 1/3 \end{bmatrix} = U_2$$

This time $\det A$ is $(-1)^0(1)(-6)(1/3) = -2$, the same as before. ■

Formula (1) for the determinant shows that A is invertible if and only if $\det A$ is nonzero. This fact, and the characterization of invertibility found in Section 5.1, can be added to the Invertible Matrix Theorem.

THEOREM

The Invertible Matrix Theorem (continued)

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

- s. The number 0 is *not* an eigenvalue of A .
- t. The determinant of A is *not* zero.

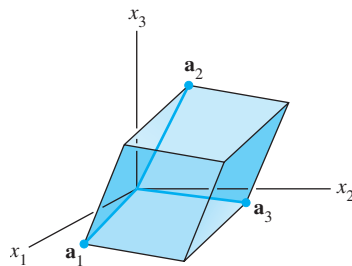


FIGURE 1

When A is a 3×3 matrix, $|\det A|$ turns out to be the *volume* of the parallelepiped determined by the columns \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 of A , as in Fig. 1. (See Section 3.3 for details.) This volume is *nonzero* if and only if the vectors \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{a}_3 are linearly independent, in which case the matrix A is invertible. (If the vectors are nonzero and linearly dependent, they lie in a plane or along a line.)

The next theorem lists facts needed from Sections 3.1 and 3.2. Part (a) is included here for convenient reference.

THEOREM 3

Properties of Determinants

Let A and B be $n \times n$ matrices.

- a. A is invertible if and only if $\det A \neq 0$.
- b. $\det AB = (\det A)(\det B)$.
- c. $\det A^T = \det A$.
- d. If A is triangular, then $\det A$ is the product of the entries on the main diagonal of A .
- e. A row replacement operation on A does not change the determinant. A row interchange changes the sign of the determinant. A row scaling also scales the determinant by the same scalar factor.

The Characteristic Equation

Theorem 3(a) shows how to determine when a matrix of the form $A - \lambda I$ is *not* invertible. The scalar equation $\det(A - \lambda I) = 0$ is called the **characteristic equation** of A , and the argument in Example 1 justifies the following fact.

A scalar λ is an eigenvalue of an $n \times n$ matrix A if and only if λ satisfies the characteristic equation

$$\det(A - \lambda I) = 0$$

EXAMPLE 3 Find the characteristic equation of

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

SOLUTION Form $A - \lambda I$, and use Theorem 3(d):

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 5 - \lambda & -2 & 6 & -1 \\ 0 & 3 - \lambda & -8 & 0 \\ 0 & 0 & 5 - \lambda & 4 \\ 0 & 0 & 0 & 1 - \lambda \end{bmatrix} \\ &= (5 - \lambda)(3 - \lambda)(5 - \lambda)(1 - \lambda) \end{aligned}$$

The characteristic equation is

$$(5 - \lambda)^2(3 - \lambda)(1 - \lambda) = 0$$

or

$$(\lambda - 5)^2(\lambda - 3)(\lambda - 1) = 0$$

Expanding the product, we can also write

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0 \quad \blacksquare$$

In Examples 1 and 3, $\det(A - \lambda I)$ is a polynomial in λ . It can be shown that if A is an $n \times n$ matrix, then $\det(A - \lambda I)$ is a polynomial of degree n called the **characteristic polynomial** of A .

The eigenvalue 5 in Example 3 is said to have *multiplicity* 2 because $(\lambda - 5)$ occurs two times as a factor of the characteristic polynomial. In general, the **(algebraic) multiplicity** of an eigenvalue λ is its multiplicity as a root of the characteristic equation.

EXAMPLE 4 The characteristic polynomial of a 6×6 matrix is $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Find the eigenvalues and their multiplicities.

SOLUTION Factor the polynomial

$$\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12) = \lambda^4(\lambda - 6)(\lambda + 2)$$

The eigenvalues are 0 (multiplicity 4), 6 (multiplicity 1), and -2 (multiplicity 1). \blacksquare

We could also list the eigenvalues in Example 4 as 0, 0, 0, 0, 6, and -2 , so that the eigenvalues are repeated according to their multiplicities.

Because the characteristic equation for an $n \times n$ matrix involves an n th-degree polynomial, the equation has exactly n roots, counting multiplicities, provided complex roots are allowed. Such complex roots, called *complex eigenvalues*, will be discussed in Section 5.5. Until then, we consider only real eigenvalues, and scalars will continue to be real numbers.

The characteristic equation is important for theoretical purposes. In practical work, however, eigenvalues of any matrix larger than 2×2 should be found by a computer, unless the matrix is triangular or has other special properties. Although a 3×3 characteristic polynomial is easy to compute by hand, factoring it can be difficult (unless the matrix is carefully chosen). See the Numerical Notes at the end of this section.

SG

Factoring a
Polynomial 5-8

Similarity

The next theorem illustrates one use of the characteristic polynomial, and it provides the foundation for several iterative methods that *approximate* eigenvalues. If A and B are $n \times n$ matrices, then A is **similar to** B if there is an invertible matrix P such that $P^{-1}AP = B$, or, equivalently, $A = PBP^{-1}$. Writing Q for P^{-1} , we have $Q^{-1}BQ = A$. So B is also similar to A , and we say simply that A and B are **similar**. Changing A into $P^{-1}AP$ is called a **similarity transformation**.

THEOREM 4

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

PROOF If $B = P^{-1}AP$, then

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Using the multiplicative property (b) in Theorem 3, we compute

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \end{aligned} \quad (2)$$

Since $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, we see from equation (2) that $\det(B - \lambda I) = \det(A - \lambda I)$. ■

WARNINGS:

1. The matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

are not similar even though they have the same eigenvalues.

2. Similarity is not the same as row equivalence. (If A is row equivalent to B , then $B = EA$ for some invertible matrix E .) Row operations on a matrix usually change its eigenvalues.

Application to Dynamical Systems

Eigenvalues and eigenvectors hold the key to the discrete evolution of a dynamical system, as mentioned in the chapter introduction.

EXAMPLE 5 Let $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analyze the long-term behavior of the dynamical system defined by $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \dots$), with $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

SOLUTION The first step is to find the eigenvalues of A and a basis for each eigenspace. The characteristic equation for A is

$$\begin{aligned} 0 &= \det \begin{bmatrix} .95 - \lambda & .03 \\ .05 & .97 - \lambda \end{bmatrix} = (.95 - \lambda)(.97 - \lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

By the quadratic formula

$$\begin{aligned} \lambda &= \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} \\ &= \frac{1.92 \pm .08}{2} = 1 \quad \text{or} \quad .92 \end{aligned}$$

It is readily checked that eigenvectors corresponding to $\lambda = 1$ and $\lambda = .92$ are multiples of

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectively.

The next step is to write the given \mathbf{x}_0 in terms of \mathbf{v}_1 and \mathbf{v}_2 . This can be done because $\{\mathbf{v}_1, \mathbf{v}_2\}$ is obviously a basis for \mathbb{R}^2 . (Why?) So there exist weights c_1 and c_2 such that

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad (3)$$

In fact,

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix}^{-1} \mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .60 \\ .40 \end{bmatrix} \\ &= \frac{1}{-8} \begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix} \begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned} \quad (4)$$

Because \mathbf{v}_1 and \mathbf{v}_2 in (3) are eigenvectors of A , with $A\mathbf{v}_1 = \mathbf{v}_1$ and $A\mathbf{v}_2 = .92\mathbf{v}_2$, we easily compute each \mathbf{x}_k :

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 && \text{Using linearity of } \mathbf{x} \mapsto A\mathbf{x} \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 && \mathbf{v}_1 \text{ and } \mathbf{v}_2 \text{ are eigenvectors.} \\ \mathbf{x}_2 &= A\mathbf{x}_1 = c_1A\mathbf{v}_1 + c_2(.92)A\mathbf{v}_2 \\ &= c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2 \end{aligned}$$

and so on. In general,

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \dots)$$

Using c_1 and c_2 from (4),

$$\mathbf{x}_k = .125 \begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \dots) \quad (5)$$

This explicit formula for \mathbf{x}_k gives the solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. As $k \rightarrow \infty$, $(.92)^k$ tends to zero and \mathbf{x}_k tends to $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$. ■

The calculations in Example 5 have an interesting application to a Markov chain discussed in Section 4.9. Those who read that section may recognize that matrix A in Example 5 above is the same as the migration matrix M in Section 4.9, \mathbf{x}_0 is the initial population distribution between city and suburbs, and \mathbf{x}_k represents the population distribution after k years.

Theorem 18 in Section 4.9 stated that for a matrix such as A , the sequence \mathbf{x}_k tends to a steady-state vector. Now we know *why* the \mathbf{x}_k behave this way, at least for the migration matrix. The steady-state vector is $.125\mathbf{v}_1$, a multiple of the eigenvector \mathbf{v}_1 , and formula (5) for \mathbf{x}_k shows precisely why $\mathbf{x}_k \rightarrow .125\mathbf{v}_1$.

NUMERICAL NOTES

1. Computer software such as Mathematica and Maple can use symbolic calculations to find the characteristic polynomial of a moderate-sized matrix. But there is no formula or finite algorithm to solve the characteristic equation of a general $n \times n$ matrix for $n \geq 5$.
2. The best numerical methods for finding eigenvalues avoid the characteristic polynomial entirely. In fact, MATLAB finds the characteristic polynomial of a matrix A by first computing the eigenvalues $\lambda_1, \dots, \lambda_n$ of A and then expanding the product $(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$.
3. Several common algorithms for estimating the eigenvalues of a matrix A are based on Theorem 4. The powerful *QR algorithm* is discussed in the exercises. Another technique, called *Jacobi's method*, works when $A = A^T$ and computes a sequence of matrices of the form

$$A_1 = A \quad \text{and} \quad A_{k+1} = P_k^{-1}A_kP_k \quad (k = 1, 2, \dots)$$

Each matrix in the sequence is similar to A and so has the same eigenvalues as A . The nondiagonal entries of A_{k+1} tend to zero as k increases, and the diagonal entries tend to approach the eigenvalues of A .

4. Other methods of estimating eigenvalues are discussed in Section 5.8.

PRACTICE PROBLEM

Find the characteristic equation and eigenvalues of $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.2 EXERCISES

Find the characteristic polynomial and the real eigenvalues of the matrices in Exercises 1–8.

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

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

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

4. $\begin{bmatrix} 8 & 2 \\ 3 & 3 \end{bmatrix}$

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

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

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

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

Exercises 9–14 require techniques from Section 3.1. Find the characteristic polynomial of each matrix, using either a cofactor expansion or the special formula for 3×3 determinants described

prior to Exercises 15–18 in Section 3.1. [Note: Finding the characteristic polynomial of a 3×3 matrix is not easy to do with just row operations, because the variable λ is involved.]

$$\begin{array}{ll} 9. \begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \\ 1 & 0 & 2 \end{bmatrix} & 10. \begin{bmatrix} 3 & 1 & 1 \\ 0 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix} \\ 11. \begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 4 \\ 1 & 0 & 4 \end{bmatrix} & 12. \begin{bmatrix} -1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \\ 13. \begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix} & 14. \begin{bmatrix} 4 & 0 & -1 \\ -1 & 0 & 4 \\ 0 & 2 & 3 \end{bmatrix} \end{array}$$

For the matrices in Exercises 15–17, list the real eigenvalues, repeated according to their multiplicities.

$$15. \begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix} \quad 16. \begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 0 & 3 & 6 & 0 \\ 2 & 3 & 3 & -5 \end{bmatrix}$$

$$17. \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$$

18. It can be shown that the algebraic multiplicity of an eigenvalue λ is always greater than or equal to the dimension of the eigenspace corresponding to λ . Find h in the matrix A below such that the eigenspace for $\lambda = 4$ is two-dimensional:

$$A = \begin{bmatrix} 4 & 2 & 3 & 3 \\ 0 & 2 & h & 3 \\ 0 & 0 & 4 & 14 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

19. Let A be an $n \times n$ matrix, and suppose A has n real eigenvalues, $\lambda_1, \dots, \lambda_n$, repeated according to multiplicities, so that
- $$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$$

Explain why $\det A$ is the product of the n eigenvalues of A . (This result is true for any square matrix when complex eigenvalues are considered.)

20. Use a property of determinants to show that A and A^T have the same characteristic polynomial.

In Exercises 21 and 22, A and B are $n \times n$ matrices. Mark each statement True or False. Justify each answer.

21. a. The determinant of A is the product of the diagonal entries in A .
 b. An elementary row operation on A does not change the determinant.
 c. $(\det A)(\det B) = \det AB$
 d. If $\lambda + 5$ is a factor of the characteristic polynomial of A , then 5 is an eigenvalue of A .

22. a. If A is 3×3 , with columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, then $\det A$ equals the volume of the parallelepiped determined by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.
 b. $\det A^T = (-1) \det A$.
 c. The multiplicity of a root r of the characteristic equation of A is called the algebraic multiplicity of r as an eigenvalue of A .
 d. A row replacement operation on A does not change the eigenvalues.

A widely used method for estimating eigenvalues of a general matrix A is the *QR algorithm*. Under suitable conditions, this algorithm produces a sequence of matrices, all similar to A , that become almost upper triangular, with diagonal entries that approach the eigenvalues of A . The main idea is to factor A (or another matrix similar to A) in the form $A = Q_1 R_1$, where $Q_1^T = Q_1^{-1}$ and R_1 is upper triangular. The factors are interchanged to form $A_1 = R_1 Q_1$, which is again factored as $A_1 = Q_2 R_2$; then to form $A_2 = R_2 Q_2$, and so on. The similarity of A, A_1, \dots follows from the more general result in Exercise 23.

23. Show that if $A = QR$ with Q invertible, then A is similar to $A_1 = RQ$.

24. Show that if A and B are similar, then $\det A = \det B$.

25. Let $A = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$, and $\mathbf{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. [Note: A is the stochastic matrix studied in Example 5 in Section 4.9.]

- a. Find a basis for \mathbb{R}^2 consisting of \mathbf{v}_1 and another eigenvector \mathbf{v}_2 of A .
 b. Verify that \mathbf{x}_0 may be written in the form $\mathbf{x}_0 = \mathbf{v}_1 + c\mathbf{v}_2$.
 c. For $k = 1, 2, \dots$, define $\mathbf{x}_k = A^k \mathbf{x}_0$. Compute \mathbf{x}_1 and \mathbf{x}_2 , and write a formula for \mathbf{x}_k . Then show that $\mathbf{x}_k \rightarrow \mathbf{v}_1$ as k increases.

26. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Use formula (1) for a determinant (given before Example 2) to show that $\det A = ad - bc$. Consider two cases: $a \neq 0$ and $a = 0$.

$$27. \text{ Let } A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix},$$

$$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and } \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

- a. Show that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are eigenvectors of A . [Note: A is the stochastic matrix studied in Example 3 of Section 4.9.]
 b. Let \mathbf{x}_0 be any vector in \mathbb{R}^3 with nonnegative entries whose sum is 1. (In Section 4.9, \mathbf{x}_0 was called a probability vector.) Explain why there are constants c_1, c_2, c_3 such that $\mathbf{x}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3$. Compute $\mathbf{w}^T \mathbf{x}_0$, and deduce that $c_1 = 1$.
 c. For $k = 1, 2, \dots$, define $\mathbf{x}_k = A^k \mathbf{x}_0$, with \mathbf{x}_0 as in part (b). Show that $\mathbf{x}_k \rightarrow \mathbf{v}_1$ as k increases.

28. [M] Construct a random integer-valued 4×4 matrix A , and verify that A and A^T have the same characteristic polynomial (the same eigenvalues with the same multiplicities). Do A and A^T have the same eigenvectors? Make the same analysis of a 5×5 matrix. Report the matrices and your conclusions.
29. [M] Construct a random integer-valued 4×4 matrix A .
- Reduce A to echelon form U with no row scaling, and use U in formula (1) (before Example 2) to compute $\det A$. (If A happens to be singular, start over with a new random matrix.)
 - Compute the eigenvalues of A and the product of these eigenvalues (as accurately as possible).
- c. List the matrix A , and, to four decimal places, list the pivots in U and the eigenvalues of A . Compute $\det A$ with your matrix program, and compare it with the products you found in (a) and (b).
30. [M] Let $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. For each value of a in the set $\{32, 31.9, 31.8, 32.1, 32.2\}$, compute the characteristic polynomial of A and the eigenvalues. In each case, create a graph of the characteristic polynomial $p(t) = \det(A - tI)$ for $0 \leq t \leq 3$. If possible, construct all graphs on one coordinate system. Describe how the graphs reveal the changes in the eigenvalues as a changes.

SOLUTION TO PRACTICE PROBLEM

The characteristic equation is

$$\begin{aligned} 0 &= \det(A - \lambda I) = \det \begin{bmatrix} 1 - \lambda & -4 \\ 4 & 2 - \lambda \end{bmatrix} \\ &= (1 - \lambda)(2 - \lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18 \end{aligned}$$

From the quadratic formula,

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} = \frac{3 \pm \sqrt{-63}}{2}$$

It is clear that the characteristic equation has no real solutions, so A has no real eigenvalues. The matrix A is acting on the real vector space \mathbb{R}^2 , and there is no nonzero vector \mathbf{v} in \mathbb{R}^2 such that $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ .

5.3 DIAGONALIZATION

In many cases, the eigenvalue–eigenvector information contained within a matrix A can be displayed in a useful factorization of the form $A = PDP^{-1}$ where D is a diagonal matrix. In this section, the factorization enables us to compute A^k quickly for large values of k , a fundamental idea in several applications of linear algebra. Later, in Sections 5.6 and 5.7, the factorization will be used to analyze (and *decouple*) dynamical systems.

The following example illustrates that powers of a diagonal matrix are easy to compute.

EXAMPLE 1 If $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, then $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$ and

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

In general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{for } k \geq 1 \quad \blacksquare$$

If $A = PDP^{-1}$ for some invertible P and diagonal D , then A^k is also easy to compute, as the next example shows.

EXAMPLE 2 Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for A^k , given that $A = PDP^{-1}$, where

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

SOLUTION The standard formula for the inverse of a 2×2 matrix yields

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Then, by associativity of matrix multiplication,

$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = PD \underbrace{(P^{-1}P)}_I DP^{-1} = PDDP^{-1} \\ &= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \end{aligned}$$

Again,

$$A^3 = (PDP^{-1})A^2 = (PDP^{-1}) \underbrace{PD^2P^{-1}}_I = PDD^2P^{-1} = PD^3P^{-1}$$

In general, for $k \geq 1$,

$$\begin{aligned} A^k &= PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix} \quad \blacksquare \end{aligned}$$

A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D . The next theorem gives a characterization of diagonalizable matrices and tells how to construct a suitable factorization.

THEOREM 5

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A . In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P .

In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of \mathbb{R}^n . We call such a basis an **eigenvector basis** of \mathbb{R}^n .

PROOF First, observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$, and if D is any diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$, then

$$AP = A[\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n] = [A\mathbf{v}_1 \ A\mathbf{v}_2 \ \cdots \ A\mathbf{v}_n] \quad (1)$$

while

$$PD = P \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1\mathbf{v}_1 \ \lambda_2\mathbf{v}_2 \ \cdots \ \lambda_n\mathbf{v}_n] \quad (2)$$

Now suppose A is diagonalizable and $A = PDP^{-1}$. Then right-multiplying this relation by P , we have $AP = PD$. In this case, equations (1) and (2) imply that

$$[A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n] = [\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n] \quad (3)$$

Equating columns, we find that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \quad \dots, \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n \quad (4)$$

Since P is invertible, its columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ must be linearly independent. Also, since these columns are nonzero, the equations in (4) show that $\lambda_1, \dots, \lambda_n$ are eigenvalues and $\mathbf{v}_1, \dots, \mathbf{v}_n$ are corresponding eigenvectors. This argument proves the “only if” parts of the first and second statements, along with the third statement, of the theorem.

Finally, given any n eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, use them to construct the columns of P and use corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ to construct D . By equations (1)–(3), $AP = PD$. This is true without any condition on the eigenvectors. If, in fact, the eigenvectors are linearly independent, then P is invertible (by the Invertible Matrix Theorem), and $AP = PD$ implies that $A = PDP^{-1}$. ■

Diagonalizing Matrices

EXAMPLE 3 Diagonalize the following matrix, if possible.

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

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

SOLUTION There are four steps to implement the description in Theorem 5.

Step 1. Find the eigenvalues of A . As mentioned in Section 5.2, the mechanics of this step are appropriate for a computer when the matrix is larger than 2×2 . To avoid unnecessary distractions, the text will usually supply information needed for this step. In the present case, the characteristic equation turns out to involve a cubic polynomial that can be factored:

$$\begin{aligned} 0 = \det(A - \lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$.

Step 2. Find three linearly independent eigenvectors of A . Three vectors are needed because A is a 3×3 matrix. This is the critical step. If it fails, then Theorem 5 says that A cannot be diagonalized. The method in Section 5.1 produces a basis for each eigenspace:

$$\begin{aligned} \text{Basis for } \lambda = 1: \quad \mathbf{v}_1 &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \\ \text{Basis for } \lambda = -2: \quad \mathbf{v}_2 &= \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

You can check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

Step 3. Construct P from the vectors in step 2. The order of the vectors is unimportant. Using the order chosen in step 2, form

$$P = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Step 4. Construct D from the corresponding eigenvalues. In this step, it is essential that the order of the eigenvalues matches the order chosen for the columns of P . Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

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

It is a good idea to check that P and D really work. To avoid computing P^{-1} , simply verify that $AP = PD$. This is equivalent to $A = PDP^{-1}$ when P is invertible. (However, be sure that P is invertible!) Compute

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

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix} \quad \blacksquare$$

EXAMPLE 4 Diagonalize the following matrix, if possible.

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

SOLUTION The characteristic equation of A turns out to be exactly the same as that in Example 3:

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$. However, it is easy to verify that each eigenspace is only one-dimensional:

$$\begin{aligned} \text{Basis for } \lambda = 1: \quad \mathbf{v}_1 &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \\ \text{Basis for } \lambda = -2: \quad \mathbf{v}_2 &= \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}$$

There are no other eigenvalues, and every eigenvector of A is a multiple of either \mathbf{v}_1 or \mathbf{v}_2 . Hence it is impossible to construct a basis of \mathbb{R}^3 using eigenvectors of A . By Theorem 5, A is *not* diagonalizable. \blacksquare

The following theorem provides a *sufficient* condition for a matrix to be diagonalizable.

THEOREM 6

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

PROOF Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of a matrix A . Then $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent, by Theorem 2 in Section 5.1. Hence A is diagonalizable, by Theorem 5. ■

It is not *necessary* for an $n \times n$ matrix to have n distinct eigenvalues in order to be diagonalizable. The 3×3 matrix in Example 3 is diagonalizable even though it has only two distinct eigenvalues.

EXAMPLE 5 Determine if the following matrix is diagonalizable.

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

SOLUTION This is easy! Since the matrix is triangular, its eigenvalues are obviously 5, 0, and -2 . Since A is a 3×3 matrix with three distinct eigenvalues, A is diagonalizable. ■

Matrices Whose Eigenvalues Are Not Distinct

If an $n \times n$ matrix A has n distinct eigenvalues, with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, and if $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, then P is automatically invertible because its columns are linearly independent, by Theorem 2. When A is diagonalizable but has fewer than n distinct eigenvalues, it is still possible to build P in a way that makes P automatically invertible, as the next theorem shows.¹

THEOREM 7

Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \dots, \lambda_p$.

- For $1 \leq k \leq p$, the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k .
- The matrix A is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n , and this happens if and only if (i) the characteristic polynomial factors completely into linear factors and (ii) the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k .
- If A is diagonalizable and \mathcal{B}_k is a basis for the eigenspace corresponding to λ_k for each k , then the total collection of vectors in the sets $\mathcal{B}_1, \dots, \mathcal{B}_p$ forms an eigenvector basis for \mathbb{R}^n .

EXAMPLE 6 Diagonalize the following matrix, if possible.

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

¹The proof of Theorem 7 is somewhat lengthy but not difficult. For instance, see S. Friedberg, A. Insel, and L. Spence, *Linear Algebra*, 4th ed. (Englewood Cliffs, NJ: Prentice-Hall, 2002), Section 5.2.

SOLUTION Since A is a triangular matrix, the eigenvalues are 5 and -3 , each with multiplicity 2. Using the method in Section 5.1, we find a basis for each eigenspace.

$$\begin{aligned} \text{Basis for } \lambda = 5: \quad \mathbf{v}_1 &= \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix} \\ \text{Basis for } \lambda = -3: \quad \mathbf{v}_3 &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

The set $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is linearly independent, by Theorem 7. So the matrix $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_4]$ is invertible, and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \quad \blacksquare$$

PRACTICE PROBLEMS

1. Compute A^8 , where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.
2. Let $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors of A . Use this information to diagonalize A .
3. Let A be a 4×4 matrix with eigenvalues 5, 3, and -2 , and suppose you know that the eigenspace for $\lambda = 3$ is two-dimensional. Do you have enough information to determine if A is diagonalizable?

WEB

5.3 EXERCISES

In Exercises 1 and 2, let $A = PDP^{-1}$ and compute A^4 .

$$1. \quad P = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$2. \quad P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

In Exercises 3 and 4, use the factorization $A = PDP^{-1}$ to compute A^k , where k represents an arbitrary positive integer.

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

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

In Exercises 5 and 6, the matrix A is factored in the form PDP^{-1} . Use the Diagonalization Theorem to find the eigenvalues of A and a basis for each eigenspace.

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

$$\begin{aligned} 6. \quad A &= \begin{bmatrix} 3 & 0 & 0 \\ -3 & 4 & 9 \\ 0 & 0 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ -3 & 1 & 9 \\ -1 & 0 & 3 \end{bmatrix} \end{aligned}$$

Diagonalize the matrices in Exercises 7–20, if possible. The real eigenvalues for Exercises 11–16 and 18 are included below the matrix.

$$7. \quad \begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$$

$$8. \quad \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$$

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

10.
$$\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$$

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

12.
$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

$\lambda = -1, 5$

$\lambda = 2, 5$

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

14.
$$\begin{bmatrix} 2 & 0 & -2 \\ 1 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$\lambda = 1, 5$

$\lambda = 2, 3$

15.
$$\begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$$

16.
$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -2 \\ 1 & 3 & 1 \end{bmatrix}$$

$\lambda = 0, 1$

$\lambda = 0$

17.
$$\begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$$

18.
$$\begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix}$$

$\lambda = -2, -1, 0$

19.
$$\begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

20.
$$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$$

In Exercises 21 and 22, A , B , P , and D are $n \times n$ matrices. Mark each statement True or False. Justify each answer. (Study Theorems 5 and 6 and the examples in this section carefully before you try these exercises.)

21. a. A is diagonalizable if $A = PDP^{-1}$ for some matrix D and some invertible matrix P .
 b. If \mathbb{R}^n has a basis of eigenvectors of A , then A is diagonalizable.
 c. A is diagonalizable if and only if A has n eigenvalues, counting multiplicities.
 d. If A is diagonalizable, then A is invertible.
22. a. A is diagonalizable if A has n eigenvectors.
 b. If A is diagonalizable, then A has n distinct eigenvalues.
 c. If $AP = PD$, with D diagonal, then the nonzero columns of P must be eigenvectors of A .
 d. If A is invertible, then A is diagonalizable.
23. A is a 5×5 matrix with two eigenvalues. One eigenspace is three-dimensional, and the other eigenspace is two-dimensional. Is A diagonalizable? Why?
24. A is a 3×3 matrix with two eigenvalues. Each eigenspace is one-dimensional. Is A diagonalizable? Why?

25. A is a 4×4 matrix with three eigenvalues. One eigenspace is one-dimensional, and one of the other eigenspaces is two-dimensional. Is it possible that A is *not* diagonalizable? Justify your answer.

26. A is a 7×7 matrix with three eigenvalues. One eigenspace is two-dimensional, and one of the other eigenspaces is three-dimensional. Is it possible that A is *not* diagonalizable? Justify your answer.

27. Show that if A is both diagonalizable and invertible, then so is A^{-1} .

28. Show that if A has n linearly independent eigenvectors, then so does A^T . [Hint: Use the Diagonalization Theorem.]

29. A factorization $A = PDP^{-1}$ is not unique. Demonstrate this for the matrix A in Example 2. With $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, use the information in Example 2 to find a matrix P_1 such that $A = P_1 D_1 P_1^{-1}$.

30. With A and D as in Example 2, find an invertible P_2 unequal to the P in Example 2, such that $A = P_2 D P_2^{-1}$.

31. Construct a nonzero 2×2 matrix that is invertible but not diagonalizable.

32. Construct a nondiagonal 2×2 matrix that is diagonalizable but not invertible.

[M] Diagonalize the matrices in Exercises 33–36. Use your matrix program's eigenvalue command to find the eigenvalues, and then compute bases for the eigenspaces as in Section 5.1.

33.
$$\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$$

34.
$$\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$$

35.
$$\begin{bmatrix} 13 & -12 & 9 & -15 & 9 \\ 6 & -5 & 9 & -15 & 9 \\ 6 & -12 & -5 & 6 & 9 \\ 6 & -12 & 9 & -8 & 9 \\ -6 & 12 & 12 & -6 & -2 \end{bmatrix}$$

36.
$$\begin{bmatrix} 24 & -6 & 2 & 6 & 2 \\ 72 & 51 & 9 & -99 & 9 \\ 0 & -63 & 15 & 63 & 63 \\ 72 & 15 & 9 & -63 & 9 \\ 0 & 63 & 21 & -63 & -27 \end{bmatrix}$$

SOLUTIONS TO PRACTICE PROBLEMS

1. $\det(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. The eigenvalues are 2 and 1, and the corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Next, form

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Since $A = PDP^{-1}$,

$$\begin{aligned} A^8 &= PD^8P^{-1} = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

2. Compute $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1$, and

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

So, \mathbf{v}_1 and \mathbf{v}_2 are eigenvectors for the eigenvalues 1 and 3, respectively. Thus

$$A = PDP^{-1}, \quad \text{where} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

3. Yes, A is diagonalizable. There is a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the eigenspace corresponding to $\lambda = 3$. In addition, there will be at least one eigenvector for $\lambda = 5$ and one for $\lambda = -2$. Call them \mathbf{v}_3 and \mathbf{v}_4 . Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent by Theorem 2 and Practice Problem 3 in Section 5.1. There can be no additional eigenvectors that are linearly independent from $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, because the vectors are all in \mathbb{R}^4 . Hence the eigenspaces for $\lambda = 5$ and $\lambda = -2$ are both one-dimensional. It follows that A is diagonalizable by Theorem 7(b).

SG

Mastering: Eigenvalue and Eigenspace 5-14

5.4 EIGENVECTORS AND LINEAR TRANSFORMATIONS

The goal of this section is to understand the matrix factorization $A = PDP^{-1}$ as a statement about linear transformations. We shall see that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is essentially the same as the very simple mapping $\mathbf{u} \mapsto D\mathbf{u}$, when viewed from the proper perspective. A similar interpretation will apply to A and D even when D is not a diagonal matrix.

Recall from Section 1.9 that any linear transformation T from \mathbb{R}^n to \mathbb{R}^m can be implemented via left-multiplication by a matrix A , called the *standard matrix* of T . Now we need the same sort of representation for any linear transformation between two finite-dimensional vector spaces.

The Matrix of a Linear Transformation

Let V be an n -dimensional vector space, let W be an m -dimensional vector space, and let T be any linear transformation from V to W . To associate a matrix with T , choose (ordered) bases \mathcal{B} and \mathcal{C} for V and W , respectively.

Given any \mathbf{x} in V , the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ is in \mathbb{R}^n and the coordinate vector of its image, $[T(\mathbf{x})]_{\mathcal{C}}$, is in \mathbb{R}^m , as shown in Fig. 1.

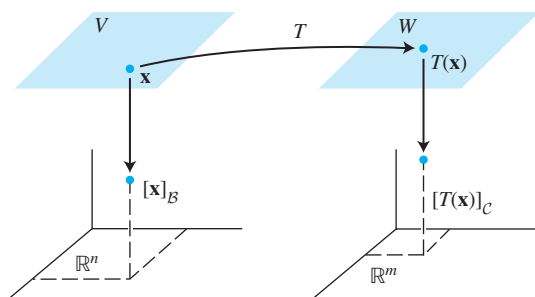


FIGURE 1 A linear transformation from V to W .

The connection between $[\mathbf{x}]_{\mathcal{B}}$ and $[T(\mathbf{x})]_{\mathcal{C}}$ is easy to find. Let $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be the basis \mathcal{B} for V . If $\mathbf{x} = r_1\mathbf{b}_1 + \dots + r_n\mathbf{b}_n$, then

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

and

$$T(\mathbf{x}) = T(r_1\mathbf{b}_1 + \dots + r_n\mathbf{b}_n) = r_1T(\mathbf{b}_1) + \dots + r_nT(\mathbf{b}_n) \quad (1)$$

because T is linear. Now, since the coordinate mapping from W to \mathbb{R}^m is linear (Theorem 8 in Section 4.4), equation (1) leads to

$$[T(\mathbf{x})]_{\mathcal{C}} = r_1[T(\mathbf{b}_1)]_{\mathcal{C}} + \dots + r_n[T(\mathbf{b}_n)]_{\mathcal{C}} \quad (2)$$

Since \mathcal{C} -coordinate vectors are in \mathbb{R}^m , the vector equation (2) can be written as a matrix equation, namely,

$$[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}} \quad (3)$$

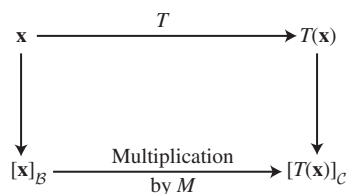


FIGURE 2

where

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix} \quad (4)$$

The matrix M is a matrix representation of T , called the **matrix for T relative to the bases \mathcal{B} and \mathcal{C}** . See Fig. 2.

Equation (3) says that, so far as coordinate vectors are concerned, the action of T on \mathbf{x} may be viewed as left-multiplication by M .

EXAMPLE 1 Suppose $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ is a basis for V and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ is a basis for W . Let $T : V \rightarrow W$ be a linear transformation with the property that

$$T(\mathbf{b}_1) = 3\mathbf{c}_1 - 2\mathbf{c}_2 + 5\mathbf{c}_3 \quad \text{and} \quad T(\mathbf{b}_2) = 4\mathbf{c}_1 + 7\mathbf{c}_2 - \mathbf{c}_3$$

Find the matrix M for T relative to \mathcal{B} and \mathcal{C} .

SOLUTION The \mathcal{C} -coordinate vectors of the *images* of \mathbf{b}_1 and \mathbf{b}_2 are

$$[T(\mathbf{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \text{and} \quad [T(\mathbf{b}_2)]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Hence

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

If \mathcal{B} and \mathcal{C} are bases for the same space V and if T is the identity transformation $T(\mathbf{x}) = \mathbf{x}$ for \mathbf{x} in V , then matrix M in (4) is just a change-of-coordinates matrix (see Section 4.7).

Linear Transformations from V into V

In the common case where W is the same as V and the basis \mathcal{C} is the same as \mathcal{B} , the matrix M in (4) is called the **matrix for T relative to \mathcal{B}** , or simply the **\mathcal{B} -matrix for T** , and is denoted by $[T]_{\mathcal{B}}$. See Fig. 3.

The \mathcal{B} -matrix for $T : V \rightarrow V$ satisfies

$$[T(\mathbf{x})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}, \quad \text{for all } \mathbf{x} \text{ in } V \tag{5}$$

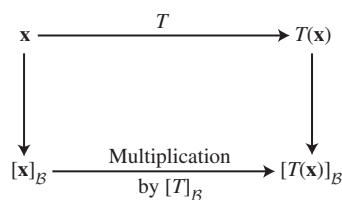


FIGURE 3

EXAMPLE 2 The mapping $T : \mathbb{P}_2 \rightarrow \mathbb{P}_2$ defined by

$$T(a_0 + a_1t + a_2t^2) = a_1 + 2a_2t$$

is a linear transformation. (Calculus students will recognize T as the differentiation operator.)

- Find the \mathcal{B} -matrix for T , when \mathcal{B} is the basis $\{1, t, t^2\}$.
- Verify that $[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}}$ for each \mathbf{p} in \mathbb{P}_2 .

SOLUTION

- Compute the images of the basis vectors:

$$\begin{aligned} T(1) &= 0 && \text{The zero polynomial} \\ T(t) &= 1 && \text{The polynomial whose value is always 1} \\ T(t^2) &= 2t \end{aligned}$$

Then write the \mathcal{B} -coordinate vectors of $T(1)$, $T(t)$, and $T(t^2)$ (which are found by inspection in this example) and place them together as the \mathcal{B} -matrix for T :

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(t^2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

b. For a general $\mathbf{p}(t) = a_0 + a_1t + a_2t^2$,

$$\begin{aligned} [T(\mathbf{p})]_{\mathcal{B}} &= [a_1 + 2a_2t]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}} \end{aligned}$$

See Fig. 4.

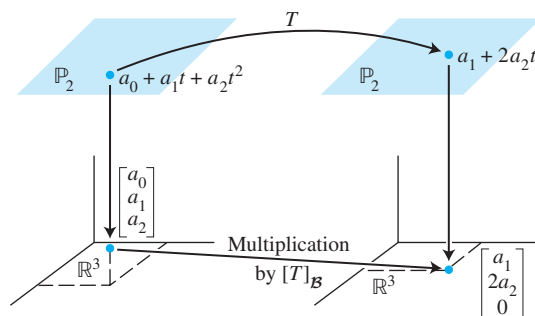


FIGURE 4 Matrix representation of a linear transformation.

WEB

Linear Transformations on \mathbb{R}^n

In an applied problem involving \mathbb{R}^n , a linear transformation T usually appears first as a matrix transformation, $\mathbf{x} \mapsto A\mathbf{x}$. If A is diagonalizable, then there is a basis \mathcal{B} for \mathbb{R}^n consisting of eigenvectors of A . Theorem 8 below shows that, in this case, the \mathcal{B} -matrix for T is diagonal. Diagonalizing A amounts to finding a diagonal matrix representation of $\mathbf{x} \mapsto A\mathbf{x}$.

THEOREM 8

Diagonal Matrix Representation

Suppose $A = PDP^{-1}$, where D is a diagonal $n \times n$ matrix. If \mathcal{B} is the basis for \mathbb{R}^n formed from the columns of P , then D is the \mathcal{B} -matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

PROOF Denote the columns of P by $\mathbf{b}_1, \dots, \mathbf{b}_n$, so that $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $P = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$. In this case, P is the change-of-coordinates matrix $P_{\mathcal{B}}$ discussed in Section 4.4, where

$$P[\mathbf{x}]_{\mathcal{B}} = \mathbf{x} \quad \text{and} \quad [\mathbf{x}]_{\mathcal{B}} = P^{-1}\mathbf{x}$$

If $T(\mathbf{x}) = A\mathbf{x}$ for \mathbf{x} in \mathbb{R}^n , then

$$\begin{aligned} [T]_{\mathcal{B}} &= [[T(\mathbf{b}_1)]_{\mathcal{B}} \quad \cdots \quad [T(\mathbf{b}_n)]_{\mathcal{B}}] && \text{Definition of } [T]_{\mathcal{B}} \\ &= [[A\mathbf{b}_1]_{\mathcal{B}} \quad \cdots \quad [A\mathbf{b}_n]_{\mathcal{B}}] && \text{Since } T(\mathbf{x}) = A\mathbf{x} \\ &= [P^{-1}A\mathbf{b}_1 \quad \cdots \quad P^{-1}A\mathbf{b}_n] && \text{Change of coordinates} \\ &= P^{-1}A[\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n] && \text{Matrix multiplication} \\ &= P^{-1}AP \end{aligned}$$

(6)

Since $A = PDP^{-1}$, we have $[T]_{\mathcal{B}} = P^{-1}AP = D$. ■

EXAMPLE 3 Define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$, where $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a basis \mathcal{B} for \mathbb{R}^2 with the property that the \mathcal{B} -matrix for T is a diagonal matrix.

SOLUTION From Example 2 in Section 5.3, we know that $A = PDP^{-1}$, where

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

The columns of P , call them \mathbf{b}_1 and \mathbf{b}_2 , are eigenvectors of A . By Theorem 8, D is the \mathcal{B} -matrix for T when $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. The mappings $\mathbf{x} \mapsto A\mathbf{x}$ and $\mathbf{u} \mapsto D\mathbf{u}$ describe the same linear transformation, relative to different bases. ■

Similarity of Matrix Representations

The proof of Theorem 8 did not use the information that D was diagonal. Hence, if A is similar to a matrix C , with $A = PCP^{-1}$, then C is the \mathcal{B} -matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ when the basis \mathcal{B} is formed from the columns of P . The factorization $A = PCP^{-1}$ is shown in Fig. 5.

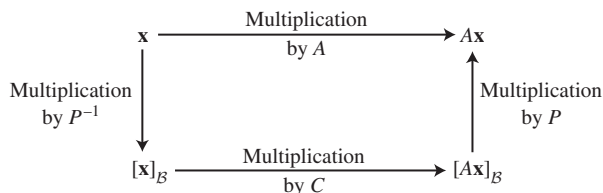


FIGURE 5 Similarity of two matrix representations:
 $A = PCP^{-1}$.

Conversely, if $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined by $T(\mathbf{x}) = A\mathbf{x}$, and if \mathcal{B} is any basis for \mathbb{R}^n , then the \mathcal{B} -matrix for T is similar to A . In fact, the calculations in the proof of Theorem 8 show that if P is the matrix whose columns come from the vectors in \mathcal{B} , then $[T]_{\mathcal{B}} = P^{-1}AP$. Thus, the set of all matrices similar to a matrix A coincides with the set of all matrix representations of the transformation $\mathbf{x} \mapsto A\mathbf{x}$.

EXAMPLE 4 Let $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, and $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The characteristic polynomial of A is $(\lambda + 2)^2$, but the eigenspace for the eigenvalue -2 is only one-dimensional; so A is not diagonalizable. However, the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ has the property that the \mathcal{B} -matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is a triangular matrix called the *Jordan form* of A .¹ Find this \mathcal{B} -matrix.

SOLUTION If $P = [\mathbf{b}_1 \quad \mathbf{b}_2]$, then the \mathcal{B} -matrix is $P^{-1}AP$. Compute

$$\begin{aligned} AP &= \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} \\ P^{-1}AP &= \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \end{aligned}$$

Notice that the eigenvalue of A is on the diagonal. ■

¹Every square matrix A is similar to a matrix in Jordan form. The basis used to produce a Jordan form consists of eigenvectors and so-called “generalized eigenvectors” of A . See Chapter 9 of *Applied Linear Algebra*, 3rd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1988), by B. Noble and J. W. Daniel.

NUMERICAL NOTE

An efficient way to compute a \mathcal{B} -matrix $P^{-1}AP$ is to compute AP and then to row reduce the augmented matrix $[P \ AP]$ to $[I \ P^{-1}AP]$. A separate computation of P^{-1} is unnecessary. See Exercise 15 in Section 2.2.

PRACTICE PROBLEMS

1. Find $T(a_0 + a_1t + a_2t^2)$, if T is the linear transformation from \mathbb{P}_2 to \mathbb{P}_2 whose matrix relative to $\mathcal{B} = \{1, t, t^2\}$ is

$$[T]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

2. Let A , B , and C be $n \times n$ matrices. The text has shown that if A is similar to B , then B is similar to A . This property, together with the statements below, shows that “similar to” is an *equivalence relation*. (Row equivalence is another example of an equivalence relation.) Verify parts (a) and (b).
- A is similar to A .
 - If A is similar to B and B is similar to C , then A is similar to C .

5.4 EXERCISES

- Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ be bases for vector spaces V and W , respectively. Let $T : V \rightarrow W$ be a linear transformation with the property that

$$T(\mathbf{b}_1) = 3\mathbf{d}_1 - 5\mathbf{d}_2, \quad T(\mathbf{b}_2) = -\mathbf{d}_1 + 6\mathbf{d}_2, \quad T(\mathbf{b}_3) = 4\mathbf{d}_2$$
 Find the matrix for T relative to \mathcal{B} and \mathcal{D} .
- Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be bases for vector spaces V and W , respectively. Let $T : V \rightarrow W$ be a linear transformation with the property that

$$T(\mathbf{d}_1) = 3\mathbf{b}_1 - 3\mathbf{b}_2, \quad T(\mathbf{d}_2) = -2\mathbf{b}_1 + 5\mathbf{b}_2$$
 Find the matrix for T relative to \mathcal{D} and \mathcal{B} .
- Let $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the standard basis for \mathbb{R}^3 , let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V , and let $T : \mathbb{R}^3 \rightarrow V$ be a linear transformation with the property that

$$T(x_1, x_2, x_3) = (2x_3 - x_2)\mathbf{b}_1 - (2x_2)\mathbf{b}_2 + (x_1 + 3x_3)\mathbf{b}_3$$
 - Compute $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$.
 - Compute $[T(\mathbf{e}_1)]_{\mathcal{B}}$, $[T(\mathbf{e}_2)]_{\mathcal{B}}$, and $[T(\mathbf{e}_3)]_{\mathcal{B}}$.
 - Find the matrix for T relative to \mathcal{E} and \mathcal{B} .
- Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V and let $T : V \rightarrow \mathbb{R}^2$ be a linear transformation with the property that

$$T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3) = \begin{bmatrix} 2x_1 - 3x_2 + x_3 \\ -2x_1 + 5x_3 \end{bmatrix}$$
 Find the matrix for T relative to \mathcal{B} and the standard basis for \mathbb{R}^2 .
- Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_3$ be the transformation that maps a polynomial $\mathbf{p}(t)$ into the polynomial $(t + 3)\mathbf{p}(t)$.
 - Find the image of $\mathbf{p}(t) = 3 - 2t + t^2$.
 - Show that T is a linear transformation.
 - Find the matrix for T relative to the bases $\{1, t, t^2\}$ and $\{1, t, t^2, t^3\}$.
- Let $T : \mathbb{P}_2 \rightarrow \mathbb{P}_4$ be the transformation that maps a polynomial $\mathbf{p}(t)$ into the polynomial $\mathbf{p}(t) + 2t^2\mathbf{p}(t)$.
 - Find the image of $\mathbf{p}(t) = 3 - 2t + t^2$.
 - Show that T is a linear transformation.
 - Find the matrix for T relative to the bases $\{1, t, t^2\}$ and $\{1, t, t^2, t^3, t^4\}$.
- Assume the mapping $T : \mathbb{P}_2 \rightarrow \mathbb{P}_2$ defined by

$$T(a_0 + a_1t + a_2t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$$
 is linear. Find the matrix representation of T relative to the basis $\mathcal{B} = \{1, t, t^2\}$.
- Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V . Find $T(4\mathbf{b}_1 - 3\mathbf{b}_2)$ when T is a linear transformation from V to V whose matrix relative to \mathcal{B} is

$$[T]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$$

9. Define $T : \mathbb{P}_2 \rightarrow \mathbb{R}^3$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$.
- Find the image under T of $\mathbf{p}(t) = 5 + 3t$.
 - Show that T is a linear transformation.
 - Find the matrix for T relative to the basis $\{1, t, t^2\}$ for \mathbb{P}_2 and the standard basis for \mathbb{R}^3 .
10. Define $T : \mathbb{P}_3 \rightarrow \mathbb{R}^4$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-2) \\ \mathbf{p}(3) \\ \mathbf{p}(1) \\ \mathbf{p}(0) \end{bmatrix}$.
- Show that T is a linear transformation.
 - Find the matrix for T relative to the basis $\{1, t, t^2, t^3\}$ for \mathbb{P}_3 and the standard basis for \mathbb{R}^4 .

In Exercises 11 and 12, find the \mathcal{B} -matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$, where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

11. $A = \begin{bmatrix} -4 & -1 \\ 6 & 1 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
12. $A = \begin{bmatrix} -6 & -2 \\ 4 & 0 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

In Exercises 13–16, define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$. Find a basis \mathcal{B} for \mathbb{R}^2 with the property that $[T]_{\mathcal{B}}$ is diagonal.

13. $A = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$ 14. $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$
15. $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$ 16. $A = \begin{bmatrix} 4 & -2 \\ -1 & 5 \end{bmatrix}$

17. Let $A = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, for $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Define $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$.
- Verify that \mathbf{b}_1 is an eigenvector of A but that A is not diagonalizable.
 - Find the \mathcal{B} -matrix for T .

18. Define $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, where A is a 3×3 matrix with eigenvalues 5, 5, and -2 . Does there exist a basis \mathcal{B} for \mathbb{R}^3 such that the \mathcal{B} -matrix for T is a diagonal matrix? Discuss.

Verify the statements in Exercises 19–24. The matrices are square.

19. If A is invertible and similar to B , then B is invertible and A^{-1} is similar to B^{-1} . [Hint: $P^{-1}AP = B$ for some invertible P . Explain why B is invertible. Then find an invertible Q such that $Q^{-1}A^{-1}Q = B^{-1}$.]
20. If A is similar to B , then A^2 is similar to B^2 .
21. If B is similar to A and C is similar to A , then B is similar to C .

22. If A is diagonalizable and B is similar to A , then B is also diagonalizable.
23. If $B = P^{-1}AP$ and \mathbf{x} is an eigenvector of A corresponding to an eigenvalue λ , then $P^{-1}\mathbf{x}$ is an eigenvector of B corresponding also to λ .
24. If A and B are similar, then they have the same rank. [Hint: Refer to Supplementary Exercises 13 and 14 in Chapter 4.]
25. The *trace* of a square matrix A is the sum of the diagonal entries in A and is denoted by $\text{tr } A$. It can be verified that $\text{tr}(FG) = \text{tr}(GF)$ for any two $n \times n$ matrices F and G . Show that if A and B are similar, then $\text{tr } A = \text{tr } B$.
26. It can be shown that the trace of a matrix A equals the sum of the eigenvalues of A . Verify this statement for the case when A is diagonalizable.
27. Let V be \mathbb{R}^n with a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$; let W be \mathbb{R}^n with the standard basis, denoted here by \mathcal{E} ; and consider the identity transformation $I : \mathbb{R}^n \rightarrow \mathbb{R}^n$, where $I(\mathbf{x}) = \mathbf{x}$. Find the matrix for I relative to \mathcal{B} and \mathcal{E} . What was this matrix called in Section 4.4?
28. Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, let W be the same space V with a basis $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$, and let I be the identity transformation $I : V \rightarrow W$. Find the matrix for I relative to \mathcal{B} and \mathcal{C} . What was this matrix called in Section 4.7?
29. Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. Find the \mathcal{B} -matrix for the identity transformation $I : V \rightarrow V$.

[M] In Exercises 30 and 31, find the \mathcal{B} -matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

30. $A = \begin{bmatrix} 6 & -2 & -2 \\ 3 & 1 & -2 \\ 2 & -2 & 2 \end{bmatrix}$,
 $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$
31. $A = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}$,
 $\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. [M] Let T be the transformation whose standard matrix is given below. Find a basis for \mathbb{R}^4 with the property that $[T]_{\mathcal{B}}$ is diagonal.

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

SOLUTIONS TO PRACTICE PROBLEMS

1. Let $\mathbf{p}(t) = a_0 + a_1t + a_2t^2$ and compute

$$[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

So $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. a. $A = (I)^{-1}AI$, so A is similar to A .
 b. By hypothesis, there exist invertible matrices P and Q with the property that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Substitute the formula for B into the formula for C , and use a fact about the inverse of a product:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

This equation has the proper form to show that A is similar to C .

5.5 COMPLEX EIGENVALUES

Since the characteristic equation of an $n \times n$ matrix involves a polynomial of degree n , the equation always has exactly n roots, counting multiplicities, *provided that possibly complex roots are included*. This section shows that if the characteristic equation of a real matrix A has some complex roots, then these roots provide critical information about A . The key is to let A act on the space \mathbb{C}^n of n -tuples of complex numbers.¹

Our interest in \mathbb{C}^n does not arise from a desire to “generalize” the results of the earlier chapters, although that would in fact open up significant new applications of linear algebra.² Rather, this study of complex eigenvalues is essential in order to uncover “hidden” information about certain matrices with real entries that arise in a variety of real-life problems. Such problems include many real dynamical systems that involve periodic motion, vibration, or some type of rotation in space.

The matrix eigenvalue–eigenvector theory already developed for \mathbb{R}^n applies equally well to \mathbb{C}^n . So a complex scalar λ satisfies $\det(A - \lambda I) = 0$ if and only if there is a nonzero vector \mathbf{x} in \mathbb{C}^n such that $A\mathbf{x} = \lambda\mathbf{x}$. We call λ a **(complex) eigenvalue** and \mathbf{x} a **(complex) eigenvector** corresponding to λ .

EXAMPLE 1 If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ on \mathbb{R}^2 rotates the plane counterclockwise through a quarter-turn. The action of A is periodic, since after four quarter-turns, a vector is back where it started. Obviously, no nonzero vector is mapped into a multiple of itself, so A has no eigenvectors in \mathbb{R}^2 and hence no real eigenvalues. In fact, the characteristic equation of A is

$$\lambda^2 + 1 = 0$$

¹Refer to Appendix B for a brief discussion of complex numbers. Matrix algebra and concepts about real vector spaces carry over to the case with complex entries and scalars. In particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, for A an $m \times n$ matrix with complex entries, \mathbf{x}, \mathbf{y} in \mathbb{C}^n , and c, d in \mathbb{C} .

²A second course in linear algebra often discusses such topics. They are of particular importance in electrical engineering.

The only roots are complex: $\lambda = i$ and $\lambda = -i$. However, if we permit A to act on \mathbb{C}^2 , then

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

Thus i and $-i$ are eigenvalues, with $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$ as corresponding eigenvectors. (A method for *finding* complex eigenvectors is discussed in Example 2.) ■

The main focus of this section will be on the matrix in the next example.

EXAMPLE 2 Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Find the eigenvalues of A , and find a basis for each eigenspace.

SOLUTION The characteristic equation of A is

$$\begin{aligned} 0 = \det \begin{bmatrix} .5 - \lambda & -.6 \\ .75 & 1.1 - \lambda \end{bmatrix} &= (.5 - \lambda)(1.1 - \lambda) - (-.6)(.75) \\ &= \lambda^2 - 1.6\lambda + 1 \end{aligned}$$

From the quadratic formula, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = .8 \pm .6i$. For the eigenvalue $\lambda = .8 - .6i$, construct

$$\begin{aligned} A - (.8 - .6i)I &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8 - .6i & 0 \\ 0 & .8 - .6i \end{bmatrix} \\ &= \begin{bmatrix} -.3 + .6i & -.6 \\ .75 & .3 + .6i \end{bmatrix} \end{aligned} \quad (1)$$

Row reduction of the usual augmented matrix is quite unpleasant by hand because of the complex arithmetic. However, here is a nice observation that really simplifies matters: Since $.8 - .6i$ is an eigenvalue, the system

$$\begin{aligned} (-.3 + .6i)x_1 - .6x_2 &= 0 \\ .75x_1 + (.3 + .6i)x_2 &= 0 \end{aligned} \quad (2)$$

has a nontrivial solution (with x_1 and x_2 possibly complex numbers). Therefore, *both equations in (2) determine the same relationship between x_1 and x_2 , and either equation can be used to express one variable in terms of the other.*³

The second equation in (2) leads to

$$\begin{aligned} .75x_1 &= (-.3 - .6i)x_2 \\ x_1 &= (-.4 - .8i)x_2 \end{aligned}$$

Choose $x_2 = 5$ to eliminate the decimals, and obtain $x_1 = -2 - 4i$. A basis for the eigenspace corresponding to $\lambda = .8 - .6i$ is

$$\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$$

³Another way to see this is to realize that the matrix in equation (1) is not invertible, so its rows are linearly dependent (as vectors in \mathbb{C}^2), and hence one row is a (complex) multiple of the other.

Analogous calculations for $\lambda = .8 + .6i$ produce the eigenvector

$$\mathbf{v}_2 = \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix}$$

As a check on the work, compute

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} \begin{bmatrix} -2 + 4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4 + 2i \\ 4 + 3i \end{bmatrix} = (.8 + .6i)\mathbf{v}_2 \quad \blacksquare$$

Surprisingly, the matrix A in Example 2 determines a transformation $\mathbf{x} \mapsto A\mathbf{x}$ that is essentially a rotation. This fact becomes evident when appropriate points are plotted.

EXAMPLE 3 One way to see how multiplication by the matrix A in Example 2 affects points is to plot an arbitrary initial point—say, $\mathbf{x}_0 = (2, 0)$ —and then to plot successive images of this point under repeated multiplications by A . That is, plot

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} \\ \mathbf{x}_2 &= A\mathbf{x}_1 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix} \\ \mathbf{x}_3 &= A\mathbf{x}_2, \dots \end{aligned}$$

Figure 1 shows $\mathbf{x}_0, \dots, \mathbf{x}_8$ as larger dots. The smaller dots are the locations of $\mathbf{x}_9, \dots, \mathbf{x}_{100}$. The sequence lies along an elliptical orbit. \blacksquare

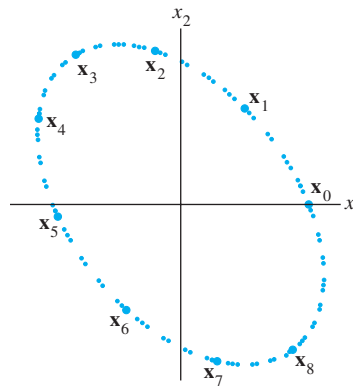


FIGURE 1 Iterates of a point \mathbf{x}_0 under the action of a matrix with a complex eigenvalue.

Of course, Fig. 1 does not explain *why* the rotation occurs. The secret to the rotation is hidden in the real and imaginary parts of a complex eigenvector.

Real and Imaginary Parts of Vectors

The complex conjugate of a complex vector \mathbf{x} in \mathbb{C}^n is the vector $\bar{\mathbf{x}}$ in \mathbb{C}^n whose entries are the complex conjugates of the entries in \mathbf{x} . The **real** and **imaginary parts** of a complex vector \mathbf{x} are the vectors $\text{Re } \mathbf{x}$ and $\text{Im } \mathbf{x}$ in \mathbb{R}^n formed from the real and imaginary parts of the entries of \mathbf{x} .

EXAMPLE 4 If $\mathbf{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$, then

$$\operatorname{Re} \mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \operatorname{Im} \mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}, \quad \text{and} \quad \bar{\mathbf{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix} \quad \blacksquare$$

If B is an $m \times n$ matrix with possibly complex entries, then \bar{B} denotes the matrix whose entries are the complex conjugates of the entries in B . Properties of conjugates for complex numbers carry over to complex matrix algebra:

$$\overline{r\mathbf{x}} = \bar{r}\bar{\mathbf{x}}, \quad \overline{B\mathbf{x}} = \bar{B}\bar{\mathbf{x}}, \quad \overline{BC} = \bar{B}\bar{C}, \quad \text{and} \quad \overline{rB} = \bar{r}\bar{B}$$

Eigenvalues and Eigenvectors of a Real Matrix That Acts on \mathbb{C}^n

Let A be an $n \times n$ matrix whose entries are real. Then $\overline{A\mathbf{x}} = \overline{A}\bar{\mathbf{x}} = A\bar{\mathbf{x}}$. If λ is an eigenvalue of A and \mathbf{x} is a corresponding eigenvector in \mathbb{C}^n , then

$$A\bar{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$$

Hence $\bar{\lambda}$ is also an eigenvalue of A , with $\bar{\mathbf{x}}$ a corresponding eigenvector. This shows that *when A is real, its complex eigenvalues occur in conjugate pairs.* (Here and elsewhere, we use the term *complex eigenvalue* to refer to an eigenvalue $\lambda = a + bi$, with $b \neq 0$.)

EXAMPLE 5 The eigenvalues of the real matrix in Example 2 are complex conjugates, namely, $.8 - .6i$ and $.8 + .6i$. The corresponding eigenvectors found in Example 2 are also conjugates:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \bar{\mathbf{v}}_1 \quad \blacksquare$$

The next example provides the basic “building block” for all real 2×2 matrices with complex eigenvalues.

EXAMPLE 6 If $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, where a and b are real and not both zero, then the eigenvalues of C are $\lambda = a \pm bi$. (See the Practice Problem at the end of this section.) Also, if $r = |\lambda| = \sqrt{a^2 + b^2}$, then

$$C = r \begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix} \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix}$$

where φ is the angle between the positive x -axis and the ray from $(0, 0)$ through (a, b) . See Fig. 2 and Appendix B. The angle φ is called the *argument* of $\lambda = a + bi$. Thus the transformation $\mathbf{x} \mapsto C\mathbf{x}$ may be viewed as the composition of a rotation through the angle φ and a scaling by $|\lambda|$ (see Fig. 3). ■

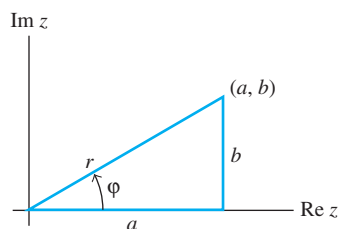


FIGURE 2

Finally, we are ready to uncover the rotation that is hidden within a real matrix having a complex eigenvalue.

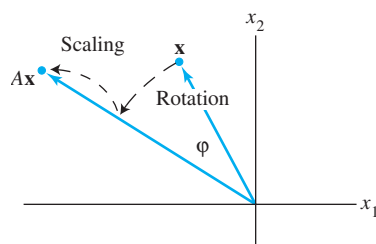


FIGURE 3 A rotation followed by a scaling.

EXAMPLE 7 Let $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$, and $\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$, as in Example 2. Also, let P be the 2×2 real matrix

$$P = [\operatorname{Re} \mathbf{v}_1 \quad \operatorname{Im} \mathbf{v}_1] = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

and let

$$C = P^{-1}AP = \frac{1}{20} \begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix} \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix}$$

By Example 6, C is a pure rotation because $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$. From $C = P^{-1}AP$, we obtain

$$A = PCP^{-1} = P \begin{bmatrix} .8 & -.6 \\ .6 & .8 \end{bmatrix} P^{-1}$$

Here is the rotation “inside” A ! The matrix P provides a change of variable, say, $\mathbf{x} = P\mathbf{u}$. The action of A amounts to a change of variable from \mathbf{x} to \mathbf{u} , followed by a rotation, and then a return to the original variable. See Fig. 4. The rotation produces an ellipse, as in Fig. 1, instead of a circle, because the coordinate system determined by the columns of P is not rectangular and does not have equal unit lengths on the two axes. ■

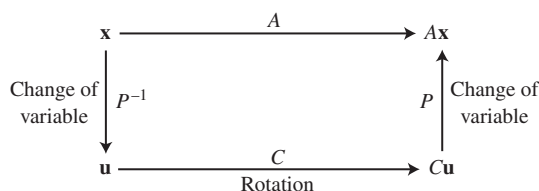


FIGURE 4 Rotation due to a complex eigenvalue.

The next theorem shows that the calculations in Example 7 can be carried out for any 2×2 real matrix A having a complex eigenvalue λ . The proof uses the fact that if the entries in A are real, then $A(\operatorname{Re} \mathbf{x}) = \operatorname{Re} A\mathbf{x}$ and $A(\operatorname{Im} \mathbf{x}) = \operatorname{Im} A\mathbf{x}$, and if \mathbf{x} is an eigenvector for a complex eigenvalue, then $\operatorname{Re} \mathbf{x}$ and $\operatorname{Im} \mathbf{x}$ are linearly independent in \mathbb{R}^2 . (See Exercises 25 and 26.) The details are omitted.

THEOREM 9

Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and an associated eigenvector \mathbf{v} in \mathbb{C}^2 . Then

$$A = PCP^{-1}, \quad \text{where } P = [\operatorname{Re} \mathbf{v} \quad \operatorname{Im} \mathbf{v}] \quad \text{and} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

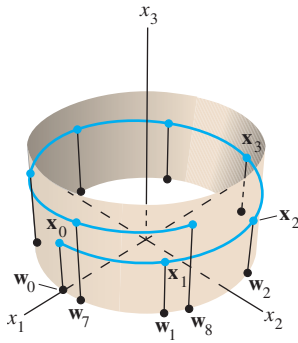


FIGURE 5
Iterates of two points under the action of a 3×3 matrix with a complex eigenvalue.

The phenomenon displayed in Example 7 persists in higher dimensions. For instance, if A is a 3×3 matrix with a complex eigenvalue, then there is a plane in \mathbb{R}^3 on which A acts as a rotation (possibly combined with scaling). Every vector in that plane is rotated into another point on the same plane. We say that the plane is **invariant** under A .

EXAMPLE 8 The matrix $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ has eigenvalues $.8 \pm .6i$ and

1.07. Any vector \mathbf{w}_0 in the x_1x_2 -plane (with third coordinate 0) is rotated by A into another point in the plane. Any vector \mathbf{x}_0 not in the plane has its x_3 -coordinate multiplied by 1.07. The iterates of the points $\mathbf{w}_0 = (2, 0, 0)$ and $\mathbf{x}_0 = (2, 0, 1)$ under multiplication by A are shown in Fig. 5. ■

PRACTICE PROBLEM

Show that if a and b are real, then the eigenvalues of $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are $a \pm bi$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

5.5 EXERCISES

Let each matrix in Exercises 1–6 act on \mathbb{C}^2 . Find the eigenvalues and a basis for each eigenspace in \mathbb{C}^2 .

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

In Exercises 7–12, use Example 6 to list the eigenvalues of A . In each case, the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the composition of a rotation and a scaling. Give the angle φ of the rotation, where $-\pi < \varphi \leq \pi$, and give the scale factor r .

- 7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$
- 8. $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$
- 9. $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$
- 10. $\begin{bmatrix} 0 & .5 \\ -5 & 0 \end{bmatrix}$
- 11. $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$
- 12. $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

In Exercises 13–20, find an invertible matrix P and a matrix C of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ such that the given matrix has the form $A = PCP^{-1}$.

- 13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$
- 14. $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

- 15. $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$
- 16. $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$
- 17. $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$
- 18. $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$
- 19. $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$
- 20. $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

- 21. In Example 2, solve the first equation in (2) for x_2 in terms of x_1 , and from that produce the eigenvector $\mathbf{y} = \begin{bmatrix} 2 \\ -1 + 2i \end{bmatrix}$ for the matrix A . Show that this \mathbf{y} is a (complex) multiple of the vector \mathbf{v}_1 used in Example 2.
- 22. Let A be a complex (or real) $n \times n$ matrix, and let \mathbf{x} in \mathbb{C}^n be an eigenvector corresponding to an eigenvalue λ in \mathbb{C} . Show that for each nonzero complex scalar μ , the vector $\mu\mathbf{x}$ is an eigenvector of A .

Chapter 7 will focus on matrices A with the property that $A^T = A$. Exercises 23 and 24 show that every eigenvalue of such a matrix is necessarily real.

- 23. Let A be an $n \times n$ real matrix with the property that $A^T = A$, let \mathbf{x} be any vector in \mathbb{C}^n , and let $q = \overline{\mathbf{x}}^T A \mathbf{x}$. The equalities below show that q is a real number by verifying that $\overline{q} = q$. Give a reason for each step.

$$\overline{q} = \overline{\overline{\mathbf{x}}^T A \mathbf{x}} = \mathbf{x}^T \overline{A \mathbf{x}} = \mathbf{x}^T A \overline{\mathbf{x}} = (\mathbf{x}^T A \overline{\mathbf{x}})^T = \overline{\mathbf{x}}^T A^T \mathbf{x} = q$$

(a) (b) (c) (d) (e)

24. Let A be an $n \times n$ real matrix with the property that $A^T = A$. Show that if $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero vector \mathbf{x} in \mathbb{C}^n , then, in fact, λ is real and the real part of \mathbf{x} is an eigenvector of A . [Hint: Compute $\bar{\mathbf{x}}^T A\mathbf{x}$, and use Exercise 23. Also, examine the real and imaginary parts of $A\mathbf{x}$.]
25. Let A be a real $n \times n$ matrix, and let \mathbf{x} be a vector in \mathbb{C}^n . Show that $\operatorname{Re}(A\mathbf{x}) = A(\operatorname{Re} \mathbf{x})$ and $\operatorname{Im}(A\mathbf{x}) = A(\operatorname{Im} \mathbf{x})$.
26. Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and an associated eigenvector \mathbf{v} in \mathbb{C}^2 .
- Show that $A(\operatorname{Re} \mathbf{v}) = a \operatorname{Re} \mathbf{v} + b \operatorname{Im} \mathbf{v}$ and $A(\operatorname{Im} \mathbf{v}) = -b \operatorname{Re} \mathbf{v} + a \operatorname{Im} \mathbf{v}$. [Hint: Write $\mathbf{v} = \operatorname{Re} \mathbf{v} + i \operatorname{Im} \mathbf{v}$, and compute $A\mathbf{v}$.]
 - Verify that if P and C are given as in Theorem 9, then $AP = PC$.
- [M] In Exercises 27 and 28, find a factorization of the given matrix A in the form $A = PCP^{-1}$, where C is a block-diagonal matrix with 2×2 blocks of the form shown in Example 6. (For each conjugate pair of eigenvalues, use the real and imaginary parts of one eigenvector in \mathbb{C}^4 to create two columns of P .)
27. $A = \begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$
28. $A = \begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

SOLUTION TO PRACTICE PROBLEM

Remember that it is easy to test whether a vector is an eigenvector. There is no need to examine the characteristic equation. Compute

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a + bi \\ b - ai \end{bmatrix} = (a + bi) \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Thus $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to $\lambda = a + bi$. From the discussion in this section, $\begin{bmatrix} 1 \\ i \end{bmatrix}$ must be an eigenvector corresponding to $\bar{\lambda} = a - bi$.

5.6 DISCRETE DYNAMICAL SYSTEMS

Eigenvalues and eigenvectors provide the key to understanding the long-term behavior, or *evolution*, of a dynamical system described by a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Such an equation was used to model population movement in Section 1.10, various Markov chains in Section 4.9, and the spotted owl population in the introductory example for this chapter. The vectors \mathbf{x}_k give information about the system as time (denoted by k) passes. In the spotted owl example, for instance, \mathbf{x}_k listed the numbers of owls in three age classes at time k .

The applications in this section focus on ecological problems because they are easier to state and explain than, say, problems in physics or engineering. However, dynamical systems arise in many scientific fields. For instance, standard undergraduate courses in control systems discuss several aspects of dynamical systems. The modern *state-space* design method in such courses relies heavily on matrix algebra.¹ The *steady-state response* of a control system is the engineering equivalent of what we call here the “long-term behavior” of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

¹See G. F. Franklin, J. D. Powell, and A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 5th ed. (Upper Saddle River, NJ: Prentice-Hall, 2006). This undergraduate text has a nice introduction to dynamic models (Chapter 2). State-space design is covered in Chapters 7 and 8.

Until Example 6, we assume that A is diagonalizable, with n linearly independent eigenvectors, $\mathbf{v}_1, \dots, \mathbf{v}_n$, and corresponding eigenvalues, $\lambda_1, \dots, \lambda_n$. For convenience, assume the eigenvectors are arranged so that $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$. Since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for \mathbb{R}^n , any initial vector \mathbf{x}_0 can be written uniquely as

$$\mathbf{x}_0 = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n \quad (1)$$

This *eigenvector decomposition* of \mathbf{x}_0 determines what happens to the sequence $\{\mathbf{x}_k\}$. The next calculation generalizes the simple case examined in Example 5 of Section 5.2. Since the \mathbf{v}_i are eigenvectors,

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = c_1 A\mathbf{v}_1 + \dots + c_n A\mathbf{v}_n \\ &= c_1 \lambda_1 \mathbf{v}_1 + \dots + c_n \lambda_n \mathbf{v}_n \end{aligned}$$

In general,

$$\mathbf{x}_k = c_1 (\lambda_1)^k \mathbf{v}_1 + \dots + c_n (\lambda_n)^k \mathbf{v}_n \quad (k = 0, 1, 2, \dots) \quad (2)$$

The examples that follow illustrate what can happen in (2) as $k \rightarrow \infty$.

A Predator–Prey System

Deep in the redwood forests of California, dusky-footed wood rats provide up to 80% of the diet for the spotted owl, the main predator of the wood rat. Example 1 uses a linear dynamical system to model the physical system of the owls and the rats. (Admittedly, the model is unrealistic in several respects, but it can provide a starting point for the study of more complicated nonlinear models used by environmental scientists.)

EXAMPLE 1 Denote the owl and wood rat populations at time k by $\mathbf{x}_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, where k is the time in months, O_k is the number of owls in the region studied, and R_k is the number of rats (measured in thousands). Suppose

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned} \quad (3)$$

where p is a positive parameter to be specified. The $(.5)O_k$ in the first equation says that with no wood rats for food, only half of the owls will survive each month, while the $(.4)R_k$ in the second equation says that with no owls as predators, the rat population will grow by 10% per month. If rats are plentiful, the $(.4)R_k$ will tend to make the owl population rise, while the negative term $-p \cdot O_k$ measures the deaths of rats due to predation by owls. (In fact, $1000p$ is the average number of rats eaten by one owl in one month.) Determine the evolution of this system when the predation parameter p is .104.

SOLUTION When $p = .104$, the eigenvalues of the coefficient matrix A for the equations in (3) turn out to be $\lambda_1 = 1.02$ and $\lambda_2 = .58$. Corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

An initial \mathbf{x}_0 can be written as $\mathbf{x}_0 = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2$. Then, for $k \geq 0$,

$$\begin{aligned} \mathbf{x}_k &= c_1 (1.02)^k \mathbf{v}_1 + c_2 (.58)^k \mathbf{v}_2 \\ &= c_1 (1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} + c_2 (.58)^k \begin{bmatrix} 5 \\ 1 \end{bmatrix} \end{aligned}$$



As $k \rightarrow \infty$, $(.58)^k$ rapidly approaches zero. Assume $c_1 > 0$. Then, for all sufficiently large k , \mathbf{x}_k is approximately the same as $c_1(1.02)^k \mathbf{v}_1$, and we write

$$\mathbf{x}_k \approx c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \quad (4)$$

The approximation in (4) improves as k increases, and so for large k ,

$$\mathbf{x}_{k+1} \approx c_1(1.02)^{k+1} \begin{bmatrix} 10 \\ 13 \end{bmatrix} = (1.02)c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \approx 1.02\mathbf{x}_k \quad (5)$$

The approximation in (5) says that eventually both entries of \mathbf{x}_k (the numbers of owls and rats) grow by a factor of almost 1.02 each month, a 2% monthly growth rate. By (4), \mathbf{x}_k is approximately a multiple of $(10, 13)$, so the entries in \mathbf{x}_k are nearly in the same ratio as 10 to 13. That is, for every 10 owls there are about 13 thousand rats. ■

Example 1 illustrates two general facts about a dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ in which A is $n \times n$, its eigenvalues satisfy $|\lambda_1| \geq 1$ and $1 > |\lambda_j|$ for $j = 2, \dots, n$, and \mathbf{v}_1 is an eigenvector corresponding to λ_1 . If \mathbf{x}_0 is given by equation (1), with $c_1 \neq 0$, then for all sufficiently large k ,

$$\mathbf{x}_{k+1} \approx \lambda_1 \mathbf{x}_k \quad (6)$$

and

$$\mathbf{x}_k \approx c_1(\lambda_1)^k \mathbf{v}_1 \quad (7)$$

The approximations in (6) and (7) can be made as close as desired by taking k sufficiently large. By (6), the \mathbf{x}_k eventually grow almost by a factor of λ_1 each time, so λ_1 determines the eventual growth rate of the system. Also, by (7), the ratio of any two entries in \mathbf{x}_k (for large k) is nearly the same as the ratio of the corresponding entries in \mathbf{v}_1 . The case in which $\lambda_1 = 1$ is illustrated in Example 5 in Section 5.2.

Graphical Description of Solutions

When A is 2×2 , algebraic calculations can be supplemented by a geometric description of a system's evolution. We can view the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ as a description of what happens to an initial point \mathbf{x}_0 in \mathbb{R}^2 as it is transformed repeatedly by the mapping $\mathbf{x} \mapsto A\mathbf{x}$. The graph of $\mathbf{x}_0, \mathbf{x}_1, \dots$ is called a **trajectory** of the dynamical system.

EXAMPLE 2 Plot several trajectories of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, when

$$A = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

SOLUTION The eigenvalues of A are .8 and .64, with eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and

$\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then

$$\mathbf{x}_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Of course, \mathbf{x}_k tends to $\mathbf{0}$ because $(.8)^k$ and $(.64)^k$ both approach 0 as $k \rightarrow \infty$. But *the way* \mathbf{x}_k goes toward $\mathbf{0}$ is interesting. Figure 1 (on page 304) shows the first few terms of several trajectories that begin at points on the boundary of the box with corners at $(\pm 3, \pm 3)$. The points on each trajectory are connected by a thin curve, to make the trajectory easier to see. ■

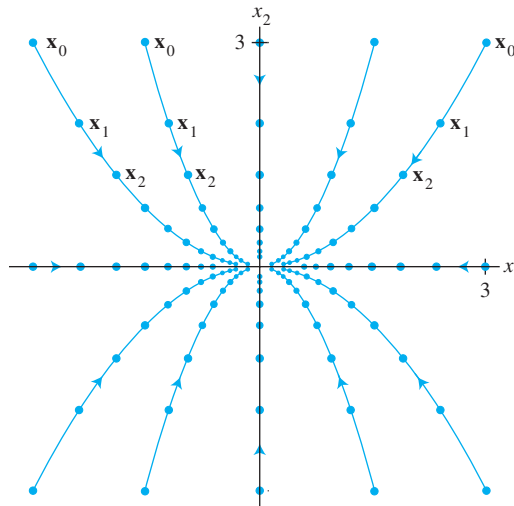


FIGURE 1 The origin as an attractor.

In Example 2, the origin is called an **attractor** of the dynamical system because all trajectories tend toward $\mathbf{0}$. This occurs whenever both eigenvalues are less than 1 in magnitude. The direction of greatest attraction is along the line through $\mathbf{0}$ and the eigenvector \mathbf{v}_2 for the eigenvalue of smaller magnitude.

In the next example, both eigenvalues of A are larger than 1 in magnitude, and $\mathbf{0}$ is called a **repeller** of the dynamical system. All solutions of $\mathbf{x}_{k+1} = A\mathbf{x}_k$ except the (constant) zero solution are unbounded and tend away from the origin.²

EXAMPLE 3 Plot several typical solutions of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

SOLUTION The eigenvalues of A are 1.44 and 1.2. If $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$\mathbf{x}_k = c_1(1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Both terms grow in size, but the first term grows faster. So the direction of greatest repulsion is the line through $\mathbf{0}$ and the eigenvector for the eigenvalue of larger magnitude. Figure 2 shows several trajectories that begin at points quite close to $\mathbf{0}$. ■

In the next example, $\mathbf{0}$ is called a **saddle point** because the origin attracts solutions from some directions and repels them in other directions. This occurs whenever one eigenvalue is greater than 1 in magnitude and the other is less than 1 in magnitude. The direction of greatest attraction is determined by an eigenvector for the eigenvalue of smaller magnitude. The direction of greatest repulsion is determined by an eigenvector for the eigenvalue of greater magnitude.

²The origin is the only possible attractor or repeller in a *linear* dynamical system, but there can be multiple attractors and repellers in a more general dynamical system for which the mapping $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ is not linear. In such a system, attractors and repellers are defined in terms of the eigenvalues of a special matrix (with variable entries) called the *Jacobian matrix* of the system.

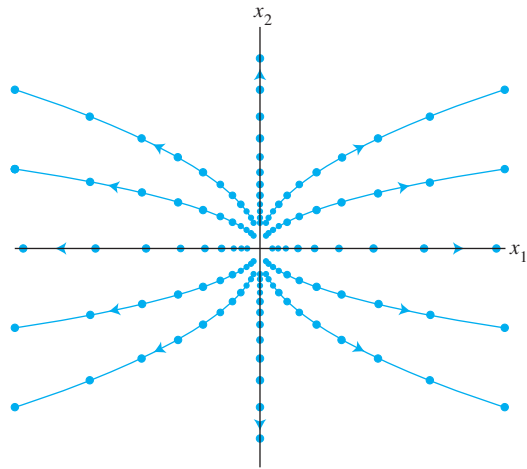


FIGURE 2 The origin as a repeller.

EXAMPLE 4 Plot several typical solutions of the equation $\mathbf{y}_{k+1} = D\mathbf{y}_k$, where

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

(We write D and \mathbf{y} here instead of A and \mathbf{x} because this example will be used later.) Show that a solution $\{\mathbf{y}_k\}$ is unbounded if its initial point is not on the x_2 -axis.

SOLUTION The eigenvalues of D are 2 and .5. If $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$\mathbf{y}_k = c_1 2^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 (.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (8)$$

If \mathbf{y}_0 is on the x_2 -axis, then $c_1 = 0$ and $\mathbf{y}_k \rightarrow \mathbf{0}$ as $k \rightarrow \infty$. But if \mathbf{y}_0 is not on the x_2 -axis, then the first term in the sum for \mathbf{y}_k becomes arbitrarily large, and so $\{\mathbf{y}_k\}$ is unbounded. Figure 3 shows ten trajectories that begin near or on the x_2 -axis. ■

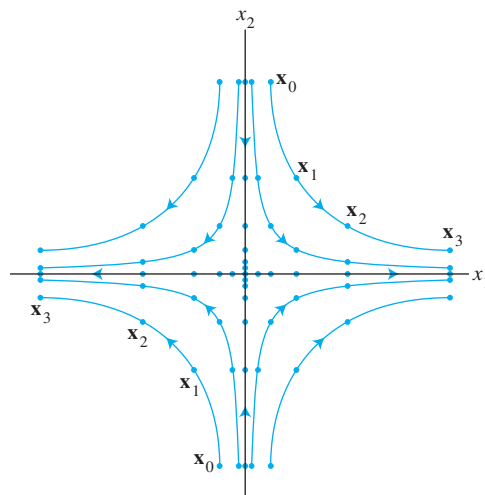


FIGURE 3 The origin as a saddle point.

Change of Variable

The preceding three examples involved diagonal matrices. To handle the nondiagonal case, we return for a moment to the $n \times n$ case in which eigenvectors of A form a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for \mathbb{R}^n . Let $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$, and let D be the diagonal matrix with the corresponding eigenvalues on the diagonal. Given a sequence $\{\mathbf{x}_k\}$ satisfying $\mathbf{x}_{k+1} = A\mathbf{x}_k$, define a new sequence $\{\mathbf{y}_k\}$ by

$$\mathbf{y}_k = P^{-1}\mathbf{x}_k, \quad \text{or equivalently,} \quad \mathbf{x}_k = P\mathbf{y}_k$$

Substituting these relations into the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ and using the fact that $A = PDP^{-1}$, we find that

$$P\mathbf{y}_{k+1} = AP\mathbf{y}_k = (PDP^{-1})P\mathbf{y}_k = PD\mathbf{y}_k$$

Left-multiplying both sides by P^{-1} , we obtain

$$\mathbf{y}_{k+1} = D\mathbf{y}_k$$

If we write \mathbf{y}_k as $\mathbf{y}(k)$ and denote the entries in $\mathbf{y}(k)$ by $y_1(k), \dots, y_n(k)$, then

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

The change of variable from \mathbf{x}_k to \mathbf{y}_k has *decoupled* the system of difference equations. The evolution of $y_1(k)$, for example, is unaffected by what happens to $y_2(k), \dots, y_n(k)$, because $y_1(k+1) = \lambda_1 \cdot y_1(k)$ for each k .

The equation $\mathbf{x}_k = P\mathbf{y}_k$ says that \mathbf{y}_k is the coordinate vector of \mathbf{x}_k with respect to the eigenvector basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. We can decouple the system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ by making calculations in the new eigenvector coordinate system. When $n = 2$, this amounts to using graph paper with axes in the directions of the two eigenvectors.

EXAMPLE 5 Show that the origin is a saddle point for solutions of $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$A = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

Find the directions of greatest attraction and greatest repulsion.

SOLUTION Using standard techniques, we find that A has eigenvalues 2 and .5, with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively. Since $|2| > 1$ and $|.5| < 1$, the origin is a saddle point of the dynamical system. If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then

$$\mathbf{x}_k = c_1 2^k \mathbf{v}_1 + c_2 (.5)^k \mathbf{v}_2 \quad (9)$$

This equation looks just like equation (8) in Example 4, with \mathbf{v}_1 and \mathbf{v}_2 in place of the standard basis.

On graph paper, draw axes through $\mathbf{0}$ and the eigenvectors \mathbf{v}_1 and \mathbf{v}_2 . See Fig. 4. Movement along these axes corresponds to movement along the standard axes in Fig. 3. In Fig. 4, the direction of greatest *repulsion* is the line through $\mathbf{0}$ and the eigenvector \mathbf{v}_1 whose eigenvalue is greater than 1 in magnitude. If \mathbf{x}_0 is on this line, the c_2 in (9) is zero and \mathbf{x}_k moves quickly away from $\mathbf{0}$. The direction of greatest *attraction* is determined by the eigenvector \mathbf{v}_2 whose eigenvalue is less than 1 in magnitude.

A number of trajectories are shown in Fig. 4. When this graph is viewed in terms of the eigenvector axes, the picture “looks” essentially the same as the picture in Fig. 3. ■

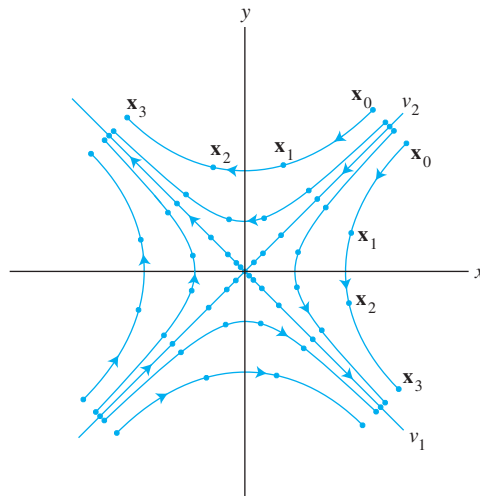


FIGURE 4 The origin as a saddle point.

Complex Eigenvalues

When a real 2×2 matrix A has complex eigenvalues, A is not diagonalizable (when acting on \mathbb{R}^2), but the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ is easy to describe. Example 3 of Section 5.5 illustrated the case in which the eigenvalues have absolute value 1. The iterates of a point \mathbf{x}_0 spiraled around the origin along an elliptical trajectory.

If A has two complex eigenvalues whose absolute value is greater than 1, then $\mathbf{0}$ is a repeller and iterates of \mathbf{x}_0 will spiral outward around the origin. If the absolute values of the complex eigenvalues are less than 1, then the origin is an attractor and the iterates of \mathbf{x}_0 spiral inward toward the origin, as in the following example.

EXAMPLE 6 It can be verified that the matrix

$$A = \begin{bmatrix} .8 & .5 \\ -.1 & 1.0 \end{bmatrix}$$

has eigenvalues $.9 \pm .2i$, with eigenvectors $\begin{bmatrix} 1 \mp 2i \\ 1 \end{bmatrix}$. Figure 5 (on page 308) shows

three trajectories of the system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, with initial vectors $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$. ■

Survival of the Spotted Owls

Recall from this chapter's introductory example that the spotted owl population in the Willow Creek area of California was modeled by a dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ in which the entries in $\mathbf{x}_k = (j_k, s_k, a_k)$ listed the numbers of females (at time k) in the juvenile, subadult, and adult life stages, respectively, and A is the stage-matrix

$$A = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \quad (10)$$

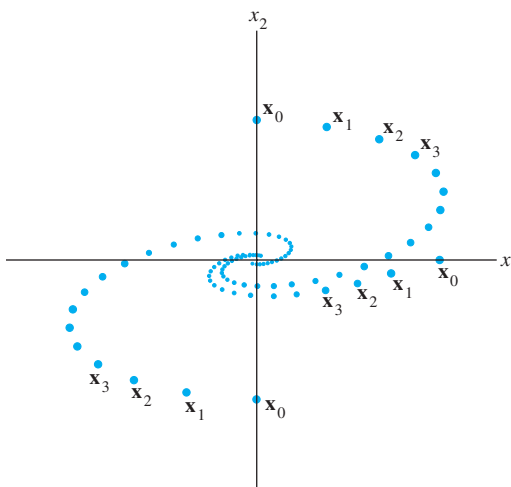


FIGURE 5 Rotation associated with complex eigenvalues.

MATLAB shows that the eigenvalues of A are approximately $\lambda_1 = .98$, $\lambda_2 = -.02 + .21i$, and $\lambda_3 = -.02 - .21i$. Observe that all three eigenvalues are less than 1 in magnitude, because $|\lambda_2|^2 = |\lambda_3|^2 = (-.02)^2 + (.21)^2 = .0445$.

For the moment, let A act on the complex vector space \mathbb{C}^3 . Then, because A has three distinct eigenvalues, the three corresponding eigenvectors are linearly independent and form a basis for \mathbb{C}^3 . Denote the eigenvectors by \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Then the general solution of $\mathbf{x}_{k+1} = A\mathbf{x}_k$ (using vectors in \mathbb{C}^3) has the form

$$\mathbf{x}_k = c_1(\lambda_1)^k \mathbf{v}_1 + c_2(\lambda_2)^k \mathbf{v}_2 + c_3(\lambda_3)^k \mathbf{v}_3 \quad (11)$$

If \mathbf{x}_0 is a real initial vector, then $\mathbf{x}_1 = A\mathbf{x}_0$ is real because A is real. Similarly, the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ shows that each \mathbf{x}_k on the left side of (11) is real, even though it is expressed as a sum of complex vectors. However, each term on the right side of (11) is approaching the zero vector, because the eigenvalues are all less than 1 in magnitude. Therefore the real sequence \mathbf{x}_k approaches the zero vector, too. Sadly, this model predicts that the spotted owls will eventually all perish.

Is there hope for the spotted owl? Recall from the introductory example that the 18% entry in the matrix A in (10) comes from the fact that although 60% of the juvenile owls live long enough to leave the nest and search for new home territories, only 30% of that group survive the search and find new home ranges. Search survival is strongly influenced by the number of clear-cut areas in the forest, which make the search more difficult and dangerous.

Some owl populations live in areas with few or no clear-cut areas. It may be that a larger percentage of the juvenile owls there survive and find new home ranges. Of course, the problem of the spotted owl is more complex than we have described, but the final example provides a happy ending to the story.

EXAMPLE 7 Suppose the search survival rate of the juvenile owls is 50%, so the (2, 1)-entry in the stage-matrix A in (10) is .3 instead of .18. What does the stage-matrix model predict about this spotted owl population?

SOLUTION Now the eigenvalues of A turn out to be approximately $\lambda_1 = 1.01$, $\lambda_2 = -.03 + .26i$, and $\lambda_3 = -.03 - .26i$. An eigenvector for λ_1 is approximately $\mathbf{v}_1 = (10, 3, 31)$. Let \mathbf{v}_2 and \mathbf{v}_3 be (complex) eigenvectors for λ_2 and λ_3 . In this case, equation

(11) becomes

$$\mathbf{x}_k = c_1(1.01)^k \mathbf{v}_1 + c_2(-.03 + .26i)^k \mathbf{v}_2 + c_3(-.03 - .26i)^k \mathbf{v}_3$$

As $k \rightarrow \infty$, the second two vectors tend to zero. So \mathbf{x}_k becomes more and more like the (real) vector $c_1(1.01)^k \mathbf{v}_1$. The approximations in equations (6) and (7), following Example 1, apply here. Also, it can be shown that the constant c_1 in the initial decomposition of \mathbf{x}_0 is positive when the entries in \mathbf{x}_0 are nonnegative. Thus the owl population will grow slowly, with a long-term growth rate of 1.01. The eigenvector \mathbf{v}_1 describes the eventual distribution of the owls by life stages: for every 31 adults, there will be about 10 juveniles and 3 subadults. ■

Further Reading

Franklin, G. F., J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems—Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Philip. *Management of a Buffalo Herd*, UMAP Module 207. Lexington, MA: COMAP, 1980.

PRACTICE PROBLEMS

1. The matrix A below has eigenvalues 1, $\frac{2}{3}$, and $\frac{1}{3}$, with corresponding eigenvectors \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 :

$$A = \frac{1}{9} \begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

Find the general solution of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ if $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 11 \\ -2 \end{bmatrix}$.

2. What happens to the sequence $\{\mathbf{x}_k\}$ in Practice Problem 1 as $k \rightarrow \infty$?

5.6 EXERCISES

- Let A be a 2×2 matrix with eigenvalues 3 and $1/3$ and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Let $\{\mathbf{x}_k\}$ be a solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $\mathbf{x}_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$.
 - Compute $\mathbf{x}_1 = A\mathbf{x}_0$. [Hint: You do not need to know A itself.]
 - Find a formula for \mathbf{x}_k involving k and the eigenvectors \mathbf{v}_1 and \mathbf{v}_2 .
- Suppose the eigenvalues of a 3×3 matrix A are 3, $4/5$, and $3/5$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -3 \\ 7 \end{bmatrix}$. Let $\mathbf{x}_0 = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$. Find the solution of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for the specified \mathbf{x}_0 , and describe what happens as $k \rightarrow \infty$.

In Exercises 3–6, assume that any initial vector \mathbf{x}_0 has an eigenvector decomposition such that the coefficient c_1 in equation (1) of this section is positive.³

- Determine the evolution of the dynamical system in Example 1 when the predation parameter p is .2 in equation (3). (Give a formula for \mathbf{x}_k .) Does the owl population grow or decline? What about the wood rat population?
- Determine the evolution of the dynamical system in Example 1 when the predation parameter p is .125. (Give a formula for \mathbf{x}_k .) As time passes, what happens to the sizes of the owl and wood rat populations? The system tends toward what is sometimes called an unstable equilibrium. What do you think might happen to the system if some aspect of the model (such as birth rates or the predation rate) were to change slightly?
- In old-growth forests of Douglas fir, the spotted owl dines mainly on flying squirrels. Suppose the predator–prey matrix for these two populations is $A = \begin{bmatrix} .4 & .3 \\ -p & 1.2 \end{bmatrix}$. Show that if the predation parameter p is .325, both populations grow. Estimate the long-term growth rate and the eventual ratio of owls to flying squirrels.
- Show that if the predation parameter p in Exercise 5 is .5, both the owls and the squirrels will eventually perish. Find a value of p for which populations of both owls and squirrels tend toward constant levels. What are the relative population sizes in this case?
- Let A have the properties described in Exercise 1.
 - Is the origin an attractor, a repeller, or a saddle point of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$?
 - Find the directions of greatest attraction and/or repulsion for this dynamical system.
 - Make a graphical description of the system, showing the directions of greatest attraction or repulsion. Include a rough sketch of several typical trajectories (without computing specific points).
- Determine the nature of the origin (attractor, repeller, or saddle point) for the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ if A has the properties described in Exercise 2. Find the directions of greatest attraction or repulsion.

In Exercises 9–14, classify the origin as an attractor, repeller, or saddle point of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Find the directions of greatest attraction and/or repulsion.

$$9. A = \begin{bmatrix} 1.7 & -.3 \\ -1.2 & .8 \end{bmatrix} \quad 10. A = \begin{bmatrix} .3 & .4 \\ -.3 & 1.1 \end{bmatrix}$$

³ One of the limitations of the model in Example 1 is that there always exist initial population vectors \mathbf{x}_0 with positive entries such that the coefficient c_1 is negative. The approximation (7) is still valid, but the entries in \mathbf{x}_k eventually become negative.

- $A = \begin{bmatrix} .4 & .5 \\ -.4 & 1.3 \end{bmatrix}$
- $A = \begin{bmatrix} .5 & .6 \\ -.3 & 1.4 \end{bmatrix}$
- $A = \begin{bmatrix} .8 & .3 \\ -.4 & 1.5 \end{bmatrix}$
- $A = \begin{bmatrix} 1.7 & .6 \\ -.4 & .7 \end{bmatrix}$
- Let $A = \begin{bmatrix} .4 & 0 & .2 \\ .3 & .8 & .3 \\ .3 & .2 & .5 \end{bmatrix}$. The vector $\mathbf{v}_1 = \begin{bmatrix} .1 \\ .6 \\ .3 \end{bmatrix}$ is an eigenvector for A , and two eigenvalues are .5 and .2. Construct the solution of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ that satisfies $\mathbf{x}_0 = (0, .3, .7)$. What happens to \mathbf{x}_k as $k \rightarrow \infty$?
- [M] Produce the general solution of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ when A is the stochastic matrix for the Hertz Rent A Car model in Exercise 16 of Section 4.9.
- Construct a stage-matrix model for an animal species that has two life stages: juvenile (up to 1 year old) and adult. Suppose the female adults give birth each year to an average of 1.6 female juveniles. Each year, 30% of the juveniles survive to become adults and 80% of the adults survive. For $k \geq 0$, let $\mathbf{x}_k = (j_k, a_k)$, where the entries in \mathbf{x}_k are the numbers of female juveniles and female adults in year k .
 - Construct the stage-matrix A such that $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for $k \geq 0$.
 - Show that the population is growing, compute the eventual growth rate of the population, and give the eventual ratio of juveniles to adults.
 - [M] Suppose that initially there are 15 juveniles and 10 adults in the population. Produce four graphs that show how the population changes over eight years: (a) the number of juveniles, (b) the number of adults, (c) the total population, and (d) the ratio of juveniles to adults (each year). When does the ratio in (d) seem to stabilize? Include a listing of the program or keystrokes used to produce the graphs for (c) and (d).
- A herd of American buffalo (bison) can be modeled by a stage matrix similar to that for the spotted owls. The females can be divided into calves (up to 1 year old), yearlings (1 to 2 years), and adults. Suppose an average of 42 female calves are born each year per 100 adult females. (Only adults produce offspring.) Each year, about 60% of the calves survive, 75% of the yearlings survive, and 95% of the adults survive. For $k \geq 0$, let $\mathbf{x}_k = (c_k, y_k, a_k)$, where the entries in \mathbf{x}_k are the numbers of females in each life stage at year k .
 - Construct the stage-matrix A for the buffalo herd, such that $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for $k \geq 0$.
 - [M] Show that the buffalo herd is growing, determine the expected growth rate after many years, and give the expected numbers of calves and yearlings present per 100 adults.

SOLUTIONS TO PRACTICE PROBLEMS

1. The first step is to write \mathbf{x}_0 as a linear combination of \mathbf{v}_1 , \mathbf{v}_2 , and \mathbf{v}_3 . Row reduction of $[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \mathbf{x}_0]$ produces the weights $c_1 = 2$, $c_2 = 1$, and $c_3 = 3$, so that

$$\mathbf{x}_0 = 2\mathbf{v}_1 + 1\mathbf{v}_2 + 3\mathbf{v}_3$$

Since the eigenvalues are 1 , $\frac{2}{3}$, and $\frac{1}{3}$, the general solution is

$$\begin{aligned} \mathbf{x}_k &= 2 \cdot 1^k \mathbf{v}_1 + 1 \cdot \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3 \cdot \left(\frac{1}{3}\right)^k \mathbf{v}_3 \\ &= 2 \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + \left(\frac{2}{3}\right)^k \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 3 \cdot \left(\frac{1}{3}\right)^k \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \end{aligned} \quad (12)$$

2. As $k \rightarrow \infty$, the second and third terms in (12) tend to the zero vector, and

$$\mathbf{x}_k = 2\mathbf{v}_1 + \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3 \left(\frac{1}{3}\right)^k \mathbf{v}_3 \rightarrow 2\mathbf{v}_1 = \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$$

5.7 APPLICATIONS TO DIFFERENTIAL EQUATIONS

This section describes continuous analogues of the difference equations studied in Section 5.6. In many applied problems, several quantities are varying continuously in time, and they are related by a system of differential equations:

$$\begin{aligned} x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

Here x_1, \dots, x_n are differentiable functions of t , with derivatives x_1', \dots, x_n' , and the a_{ij} are constants. The crucial feature of this system is that it is *linear*. To see this, write the system as a matrix differential equation

$$\mathbf{x}'(t) = A\mathbf{x}(t) \quad (1)$$

where

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

A **solution** of equation (1) is a vector-valued function that satisfies (1) for all t in some interval of real numbers, such as $t \geq 0$.

Equation (1) is *linear* because both differentiation of functions and multiplication of vectors by a matrix are linear transformations. Thus, if \mathbf{u} and \mathbf{v} are solutions of $\mathbf{x}' = A\mathbf{x}$, then $c\mathbf{u} + d\mathbf{v}$ is also a solution, because

$$\begin{aligned} (c\mathbf{u} + d\mathbf{v})' &= c\mathbf{u}' + d\mathbf{v}' \\ &= cA\mathbf{u} + dA\mathbf{v} = A(c\mathbf{u} + d\mathbf{v}) \end{aligned}$$

(Engineers call this property *superposition* of solutions.) Also, the identically zero function is a (trivial) solution of (1). In the terminology of Chapter 4, the set of all solutions of (1) is a *subspace* of the set of all continuous functions with values in \mathbb{R}^n .

Standard texts on differential equations show that there always exists what is called a **fundamental set of solutions** to (1). If A is $n \times n$, then there are n linearly independent functions in a fundamental set, and each solution of (1) is a unique linear combination of these n functions. That is, a fundamental set of solutions is a *basis* for the set of all solutions of (1), and the solution set is an n -dimensional vector space of functions. If a vector \mathbf{x}_0 is specified, then the **initial value problem** is to construct the (unique) function \mathbf{x} such that $\mathbf{x}' = A\mathbf{x}$ and $\mathbf{x}(0) = \mathbf{x}_0$.

When A is a diagonal matrix, the solutions of (1) can be produced by elementary calculus. For instance, consider

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

that is,

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \tag{3}$$

The system (2) is said to be *decoupled* because each derivative of a function depends only on the function itself, not on some combination or “coupling” of both $x_1(t)$ and $x_2(t)$. From calculus, the solutions of (3) are $x_1(t) = c_1e^{3t}$ and $x_2(t) = c_2e^{-5t}$, for any constants c_1 and c_2 . Each solution of equation (2) can be written in the form

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1e^{3t} \\ c_2e^{-5t} \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{3t} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^{-5t}$$

This example suggests that for the general equation $\mathbf{x}' = A\mathbf{x}$, a solution might be a linear combination of functions of the form

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t} \tag{4}$$

for some scalar λ and some fixed nonzero vector \mathbf{v} . [If $\mathbf{v} = \mathbf{0}$, the function $\mathbf{x}(t)$ is identically zero and hence satisfies $\mathbf{x}' = A\mathbf{x}$.] Observe that

$$\mathbf{x}'(t) = \lambda \mathbf{v}e^{\lambda t} \quad \text{By calculus, since } \mathbf{v} \text{ is a constant vector}$$

$$A\mathbf{x}(t) = A\mathbf{v}e^{\lambda t} \quad \text{Multiplying both sides of (4) by } A$$

Since $e^{\lambda t}$ is never zero, $\mathbf{x}'(t)$ will equal $A\mathbf{x}(t)$ if and only if $\lambda \mathbf{v} = A\mathbf{v}$, that is, if and only if λ is an eigenvalue of A and \mathbf{v} is a corresponding eigenvector. Thus each eigenvalue–eigenvector pair provides a solution (4) of $\mathbf{x}' = A\mathbf{x}$. Such solutions are sometimes called *eigenfunctions* of the differential equation. Eigenfunctions provide the key to solving systems of differential equations.

EXAMPLE 1 The circuit in Fig. 1 can be described by the differential equation

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

where $x_1(t)$ and $x_2(t)$ are the voltages across the two capacitors at time t . Suppose resistor R_1 is 1 ohm, R_2 is 2 ohms, capacitor C_1 is 1 farad, and C_2 is .5 farad, and suppose there is an initial charge of 5 volts on capacitor C_1 and 4 volts on capacitor C_2 . Find formulas for $x_1(t)$ and $x_2(t)$ that describe how the voltages change over time.

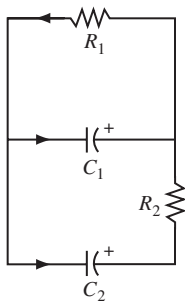


FIGURE 1

SOLUTION Let A denote the matrix displayed above, and let $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. For the data given, $A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}$, and $\mathbf{x}(0) = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. The eigenvalues of A are $\lambda_1 = -.5$ and $\lambda_2 = -2$, with corresponding eigenvectors

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

The eigenfunctions $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda_1 t}$ and $\mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda_2 t}$ both satisfy $\mathbf{x}' = A\mathbf{x}$, and so does any linear combination of \mathbf{x}_1 and \mathbf{x}_2 . Set

$$\mathbf{x}(t) = c_1 \mathbf{v}_1 e^{\lambda_1 t} + c_2 \mathbf{v}_2 e^{\lambda_2 t} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

and note that $\mathbf{x}(0) = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2$. Since \mathbf{v}_1 and \mathbf{v}_2 are obviously linearly independent and hence span \mathbb{R}^2 , c_1 and c_2 can be found to make $\mathbf{x}(0)$ equal to \mathbf{x}_0 . In fact, the equation

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{x}_0 \end{matrix}$

leads easily to $c_1 = 3$ and $c_2 = -2$. Thus the desired solution of the differential equation $\mathbf{x}' = A\mathbf{x}$ is

$$\mathbf{x}(t) = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

or

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 3e^{-.5t} + 2e^{-2t} \\ 6e^{-.5t} - 2e^{-2t} \end{bmatrix}$$

Figure 2 shows the graph, or *trajectory*, of $\mathbf{x}(t)$, for $t \geq 0$, along with trajectories for some other initial points. The trajectories of the two eigenfunctions \mathbf{x}_1 and \mathbf{x}_2 lie in the eigenspaces of A .

The functions \mathbf{x}_1 and \mathbf{x}_2 both decay to zero as $t \rightarrow \infty$, but the values of \mathbf{x}_2 decay faster because its exponent is more negative. The entries in the corresponding eigenvector \mathbf{v}_2 show that the voltages across the capacitors will decay to zero as rapidly as possible if the initial voltages are equal in magnitude but opposite in sign. ■

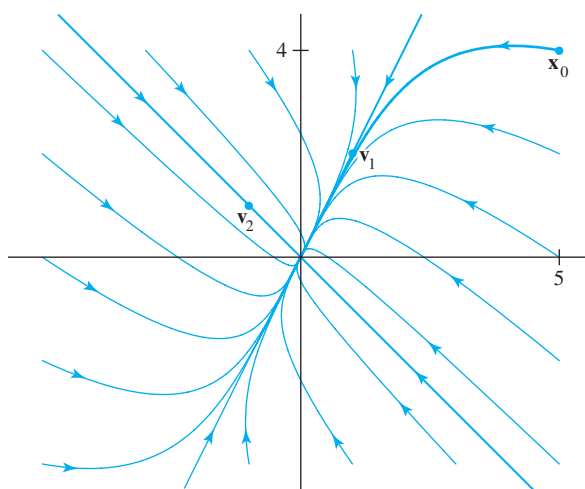


FIGURE 2 The origin as an attractor.

In Fig. 2, the origin is called an **attractor**, or **sink**, of the dynamical system because all trajectories are drawn into the origin. The direction of greatest attraction is along the trajectory of the eigenfunction \mathbf{x}_2 (along the line through $\mathbf{0}$ and \mathbf{v}_2) corresponding to the more negative eigenvalue, $\lambda = -2$. Trajectories that begin at points not on this line become asymptotic to the line through $\mathbf{0}$ and \mathbf{v}_1 because their components in the \mathbf{v}_2 direction decay so rapidly.

If the eigenvalues in Example 1 were positive instead of negative, the corresponding trajectories would be similar in shape, but the trajectories would be traversed *away* from the origin. In such a case, the origin is called a **repeller**, or **source**, of the dynamical system, and the direction of greatest repulsion is the line containing the trajectory of the eigenfunction corresponding to the more positive eigenvalue.

EXAMPLE 2 Suppose a particle is moving in a planar force field and its position vector \mathbf{x} satisfies $\mathbf{x}' = A\mathbf{x}$ and $\mathbf{x}(0) = \mathbf{x}_0$, where

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix}, \quad \mathbf{x}_0 = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Solve this initial value problem for $t \geq 0$, and sketch the trajectory of the particle.

SOLUTION The eigenvalues of A turn out to be $\lambda_1 = 6$ and $\lambda_2 = -1$, with corresponding eigenvectors $\mathbf{v}_1 = (-5, 2)$ and $\mathbf{v}_2 = (1, 1)$. For any constants c_1 and c_2 , the function

$$\mathbf{x}(t) = c_1 \mathbf{v}_1 e^{\lambda_1 t} + c_2 \mathbf{v}_2 e^{\lambda_2 t} = c_1 \begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

is a solution of $\mathbf{x}' = A\mathbf{x}$. We want c_1 and c_2 to satisfy $\mathbf{x}(0) = \mathbf{x}_0$, that is,

$$c_1 \begin{bmatrix} -5 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -5 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Calculations show that $c_1 = -3/70$ and $c_2 = 188/70$, and so the desired function is

$$\mathbf{x}(t) = \frac{-3}{70} \begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + \frac{188}{70} \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

Trajectories of \mathbf{x} and other solutions are shown in Fig. 3. ■

In Fig. 3, the origin is called a **saddle point** of the dynamical system because some trajectories approach the origin at first and then change direction and move away from the origin. A saddle point arises whenever the matrix A has both positive and negative eigenvalues. The direction of greatest repulsion is the line through \mathbf{v}_1 and $\mathbf{0}$, corresponding to the positive eigenvalue. The direction of greatest attraction is the line through \mathbf{v}_2 and $\mathbf{0}$, corresponding to the negative eigenvalue.

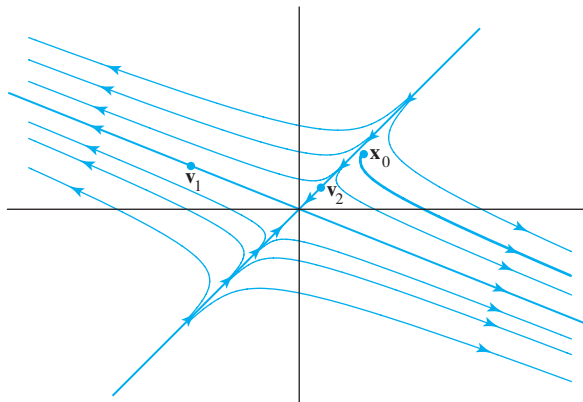


FIGURE 3 The origin as a saddle point.

Decoupling a Dynamical System

The following discussion shows that the method of Examples 1 and 2 produces a fundamental set of solutions for any dynamical system described by $\mathbf{x}' = A\mathbf{x}$ when A is $n \times n$ and has n linearly independent eigenvectors, that is, when A is diagonalizable. Suppose the eigenfunctions for A are

$$\mathbf{v}_1 e^{\lambda_1 t}, \quad \dots, \quad \mathbf{v}_n e^{\lambda_n t}$$

with $\mathbf{v}_1, \dots, \mathbf{v}_n$ linearly independent eigenvectors. Let $P = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$, and let D be the diagonal matrix with entries $\lambda_1, \dots, \lambda_n$, so that $A = PDP^{-1}$. Now make a *change of variable*, defining a new function \mathbf{y} by

$$\mathbf{y}(t) = P^{-1}\mathbf{x}(t) \quad \text{or, equivalently,} \quad \mathbf{x}(t) = P\mathbf{y}(t)$$

The equation $\mathbf{x}(t) = P\mathbf{y}(t)$ says that $\mathbf{y}(t)$ is the coordinate vector of $\mathbf{x}(t)$ relative to the eigenvector basis. Substitution of $P\mathbf{y}$ for \mathbf{x} in the equation $\mathbf{x}' = A\mathbf{x}$ gives

$$\frac{d}{dt}(P\mathbf{y}) = A(P\mathbf{y}) = (PDP^{-1})P\mathbf{y} = PD\mathbf{y} \quad (5)$$

Since P is a constant matrix, the left side of (5) is $P\mathbf{y}'$. Left-multiply both sides of (5) by P^{-1} and obtain $\mathbf{y}' = D\mathbf{y}$, or

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

The change of variable from \mathbf{x} to \mathbf{y} has *decoupled* the system of differential equations, because the derivative of each scalar function y_k depends only on y_k . (Review the analogous change of variables in Section 5.6.) Since $y_1' = \lambda_1 y_1$, we have $y_1(t) = c_1 e^{\lambda_1 t}$, with similar formulas for y_2, \dots, y_n . Thus

$$\mathbf{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \quad \text{where} \quad \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{y}(0) = P^{-1}\mathbf{x}(0) = P^{-1}\mathbf{x}_0$$

To obtain the general solution \mathbf{x} of the original system, compute

$$\begin{aligned} \mathbf{x}(t) &= P\mathbf{y}(t) = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]\mathbf{y}(t) \\ &= c_1 \mathbf{v}_1 e^{\lambda_1 t} + \dots + c_n \mathbf{v}_n e^{\lambda_n t} \end{aligned}$$

This is the eigenfunction expansion constructed as in Example 1.

Complex Eigenvalues

In the next example, a real matrix A has a pair of complex eigenvalues λ and $\bar{\lambda}$, with associated complex eigenvectors \mathbf{v} and $\bar{\mathbf{v}}$. (Recall from Section 5.5 that for a real matrix, complex eigenvalues and associated eigenvectors come in conjugate pairs.) So two solutions of $\mathbf{x}' = A\mathbf{x}$ are

$$\mathbf{x}_1(t) = \mathbf{v}e^{\lambda t} \quad \text{and} \quad \mathbf{x}_2(t) = \bar{\mathbf{v}}e^{\bar{\lambda}t} \quad (6)$$

It can be shown that $\mathbf{x}_2(t) = \overline{\mathbf{x}_1(t)}$ by using a power series representation for the complex exponential function. Although the complex eigenfunctions \mathbf{x}_1 and \mathbf{x}_2 are convenient for some calculations (particularly in electrical engineering), real functions

are more appropriate for many purposes. Fortunately, the real and imaginary parts of \mathbf{x}_1 are (real) solutions of $\mathbf{x}' = A\mathbf{x}$, because they are linear combinations of the solutions in (6):

$$\operatorname{Re}(\mathbf{v}e^{\lambda t}) = \frac{1}{2}[\mathbf{x}_1(t) + \overline{\mathbf{x}_1(t)}], \quad \operatorname{Im}(\mathbf{v}e^{\lambda t}) = \frac{1}{2i}[\mathbf{x}_1(t) - \overline{\mathbf{x}_1(t)}]$$

To understand the nature of $\operatorname{Re}(\mathbf{v}e^{\lambda t})$, recall from calculus that for any number x , the exponential function e^x can be computed from the power series:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

This series can be used to define $e^{\lambda t}$ when λ is complex:

$$e^{\lambda t} = 1 + (\lambda t) + \frac{1}{2!}(\lambda t)^2 + \cdots + \frac{1}{n!}(\lambda t)^n + \cdots$$

By writing $\lambda = a + bi$ (with a and b real), and using similar power series for the cosine and sine functions, one can show that

$$e^{(a+bi)t} = e^{at} \cdot e^{ibt} = e^{at}(\cos bt + i \sin bt) \quad (7)$$

Hence

$$\begin{aligned} \mathbf{v}e^{\lambda t} &= (\operatorname{Re} \mathbf{v} + i \operatorname{Im} \mathbf{v}) \cdot e^{at}(\cos bt + i \sin bt) \\ &= [(\operatorname{Re} \mathbf{v}) \cos bt - (\operatorname{Im} \mathbf{v}) \sin bt] e^{at} \\ &\quad + i [(\operatorname{Re} \mathbf{v}) \sin bt + (\operatorname{Im} \mathbf{v}) \cos bt] e^{at} \end{aligned}$$

So two real solutions of $\mathbf{x}' = A\mathbf{x}$ are

$$\begin{aligned} \mathbf{y}_1(t) &= \operatorname{Re} \mathbf{x}_1(t) = [(\operatorname{Re} \mathbf{v}) \cos bt - (\operatorname{Im} \mathbf{v}) \sin bt] e^{at} \\ \mathbf{y}_2(t) &= \operatorname{Im} \mathbf{x}_1(t) = [(\operatorname{Re} \mathbf{v}) \sin bt + (\operatorname{Im} \mathbf{v}) \cos bt] e^{at} \end{aligned}$$

It can be shown that \mathbf{y}_1 and \mathbf{y}_2 are linearly independent functions (when $b \neq 0$).¹

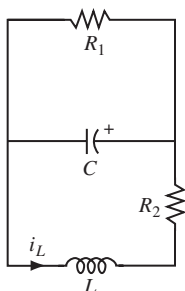


FIGURE 4

EXAMPLE 3 The circuit in Fig. 4 can be described by the equation

$$\begin{bmatrix} i_L' \\ v_C' \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1 C) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

where i_L is the current passing through the inductor L and v_C is the voltage drop across the capacitor C . Suppose R_1 is 5 ohms, R_2 is .8 ohm, C is .1 farad, and L is .4 henry. Find formulas for i_L and v_C , if the initial current through the inductor is 3 amperes and the initial voltage across the capacitor is 3 volts.

SOLUTION For the data given, $A = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$ and $\mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. The method discussed in Section 5.5 produces the eigenvalue $\lambda = -2 + 5i$ and the corresponding eigenvector $\mathbf{v}_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$. The complex solutions of $\mathbf{x}' = A\mathbf{x}$ are complex linear combinations of

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{(-2+5i)t} \quad \text{and} \quad \mathbf{x}_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix} e^{(-2-5i)t}$$

¹Since $\mathbf{x}_2(t)$ is the complex conjugate of $\mathbf{x}_1(t)$, the real and imaginary parts of $\mathbf{x}_2(t)$ are $\mathbf{y}_1(t)$ and $-\mathbf{y}_2(t)$, respectively. Thus one can use either $\mathbf{x}_1(t)$ or $\mathbf{x}_2(t)$, but not both, to produce two real linearly independent solutions of $\mathbf{x}' = A\mathbf{x}$.

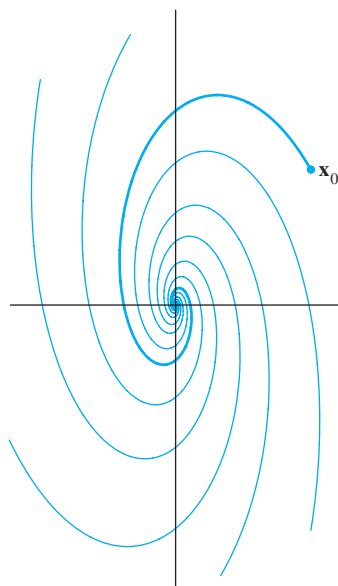


FIGURE 5
The origin as a spiral point.

Next, use equation (7) to write

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t} (\cos 5t + i \sin 5t)$$

The real and imaginary parts of \mathbf{x}_1 provide real solutions:

$$\mathbf{y}_1(t) = \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t}, \quad \mathbf{y}_2(t) = \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t}$$

Since \mathbf{y}_1 and \mathbf{y}_2 are linearly independent functions, they form a basis for the two-dimensional real vector space of solutions of $\mathbf{x}' = A\mathbf{x}$. Thus the general solution is

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t}$$

To satisfy $\mathbf{x}(0) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, we need $c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, which leads to $c_1 = 1.5$ and $c_2 = 3$. Thus

$$\mathbf{x}(t) = 1.5 \begin{bmatrix} -\sin 5t \\ 2 \cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2 \sin 5t \end{bmatrix} e^{-2t}$$

or

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5 \sin 5t + 3 \cos 5t \\ 3 \cos 5t + 6 \sin 5t \end{bmatrix} e^{-2t}$$

See Fig. 5. ■

In Fig. 5, the origin is called a **spiral point** of the dynamical system. The rotation is caused by the sine and cosine functions that arise from a complex eigenvalue. The trajectories spiral inward because the factor e^{-2t} tends to zero. Recall that -2 is the real part of the eigenvalue in Example 3. When A has a complex eigenvalue with positive real part, the trajectories spiral outward. If the real part of the eigenvalue is zero, the trajectories form ellipses around the origin.

PRACTICE PROBLEMS

A real 3×3 matrix A has eigenvalues $-.5$, $.2 + .3i$, and $.2 - .3i$, with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 + 2i \\ 4i \\ 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} 1 - 2i \\ -4i \\ 2 \end{bmatrix}$$

1. Is A diagonalizable as $A = PDP^{-1}$, using complex matrices?
2. Write the general solution of $\mathbf{x}' = A\mathbf{x}$ using complex eigenfunctions, and then find the general real solution.
3. Describe the shapes of typical trajectories.

5.7 EXERCISES

1. A particle moving in a planar force field has a position vector \mathbf{x} that satisfies $\mathbf{x}' = A\mathbf{x}$. The 2×2 matrix A has eigenvalues 4 and 2, with corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and

$\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Find the position of the particle at time t , assuming that $\mathbf{x}(0) = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$.

2. Let A be a 2×2 matrix with eigenvalues -3 and -1 and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Let $\mathbf{x}(t)$ be the position of a particle at time t . Solve the initial value problem $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

In Exercises 3–6, solve the initial value problem $\mathbf{x}'(t) = A\mathbf{x}(t)$ for $t \geq 0$, with $\mathbf{x}(0) = (3, 2)$. Classify the nature of the origin as an attractor, repeller, or saddle point of the dynamical system described by $\mathbf{x}' = A\mathbf{x}$. Find the directions of greatest attraction and/or repulsion. When the origin is a saddle point, sketch typical trajectories.

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

In Exercises 7 and 8, make a change of variable that decouples the equation $\mathbf{x}' = A\mathbf{x}$. Write the equation $\mathbf{x}(t) = P\mathbf{y}(t)$ and show the calculation that leads to the uncoupled system $\mathbf{y}' = D\mathbf{y}$, specifying P and D .

7. A as in Exercise 5 8. A as in Exercise 6

In Exercises 9–18, construct the general solution of $\mathbf{x}' = A\mathbf{x}$ involving complex eigenfunctions and then obtain the general real solution. Describe the shapes of typical trajectories.

9. $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$ 10. $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$
 11. $A = \begin{bmatrix} -3 & -9 \\ 2 & 3 \end{bmatrix}$ 12. $A = \begin{bmatrix} -7 & 10 \\ -4 & 5 \end{bmatrix}$
 13. $A = \begin{bmatrix} 4 & -3 \\ 6 & -2 \end{bmatrix}$ 14. $A = \begin{bmatrix} -2 & 1 \\ -8 & 2 \end{bmatrix}$
 15. [M] $A = \begin{bmatrix} -8 & -12 & -6 \\ 2 & 1 & 2 \\ 7 & 12 & 5 \end{bmatrix}$

16. [M] $A = \begin{bmatrix} -6 & -11 & 16 \\ 2 & 5 & -4 \\ -4 & -5 & 10 \end{bmatrix}$

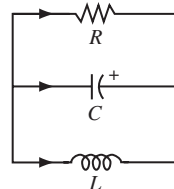
17. [M] $A = \begin{bmatrix} 30 & 64 & 23 \\ -11 & -23 & -9 \\ 6 & 15 & 4 \end{bmatrix}$

18. [M] $A = \begin{bmatrix} 53 & -30 & -2 \\ 90 & -52 & -3 \\ 20 & -10 & 2 \end{bmatrix}$

19. [M] Find formulas for the voltages v_1 and v_2 (as functions of time t) for the circuit in Example 1, assuming that $R_1 = 1/5$ ohm, $R_2 = 1/3$ ohm, $C_1 = 4$ farads, $C_2 = 3$ farads, and the initial charge on each capacitor is 4 volts.
20. [M] Find formulas for the voltages v_1 and v_2 for the circuit in Example 1, assuming that $R_1 = 1/15$ ohm, $R_2 = 1/3$ ohm, $C_1 = 9$ farads, $C_2 = 2$ farads, and the initial charge on each capacitor is 3 volts.
21. [M] Find formulas for the current i_L and the voltage v_C for the circuit in Example 3, assuming that $R_1 = 1$ ohm, $R_2 = .125$ ohm, $C = .2$ farad, $L = .125$ henry, the initial current is 0 amp, and the initial voltage is 15 volts.
22. [M] The circuit in the figure is described by the equation

$$\begin{bmatrix} i_L' \\ v_C' \end{bmatrix} = \begin{bmatrix} 0 & 1/L \\ -1/C & -1/(RC) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

where i_L is the current through the inductor L and v_C is the voltage drop across the capacitor C . Find formulas for i_L and v_C when $R = .5$ ohm, $C = 2.5$ farads, $L = .5$ henry, the initial current is 0 amp, and the initial voltage is 12 volts.



SOLUTIONS TO PRACTICE PROBLEMS

- Yes, the 3×3 matrix is diagonalizable because it has three distinct eigenvalues. Theorem 2 in Section 5.1 and Theorem 5 in Section 5.3 are valid when complex scalars are used. (The proofs are essentially the same as for real scalars.)
- The general solution has the form

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-5t} + c_2 \begin{bmatrix} 1 + 2i \\ 4i \\ 2 \end{bmatrix} e^{(.2+.3i)t} + c_3 \begin{bmatrix} 1 - 2i \\ -4i \\ 2 \end{bmatrix} e^{(.2-.3i)t}$$

The scalars c_1, c_2, c_3 here can be any complex numbers. The first term in $\mathbf{x}(t)$ is real. Two more real solutions can be produced using the real and imaginary parts of the

second term in $\mathbf{x}(t)$:

$$\begin{bmatrix} 1 + 2i \\ 4i \\ 2 \end{bmatrix} e^{2t} (\cos .3t + i \sin .3t)$$

The general real solution has the following form, with *real* scalars c_1, c_2, c_3 :

$$c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} \cos .3t - 2 \sin .3t \\ -4 \sin .3t \\ 2 \cos .3t \end{bmatrix} e^{2t} + c_3 \begin{bmatrix} \sin .3t + 2 \cos .3t \\ 4 \cos .3t \\ 2 \sin .3t \end{bmatrix} e^{-2t}$$

3. Any solution with $c_2 = c_3 = 0$ is attracted to the origin because of the negative exponential factor. Other solutions have components that grow without bound, and the trajectories spiral outward.

Be careful not to mistake this problem for one in Section 5.6. There the condition for attraction toward $\mathbf{0}$ was that an eigenvalue be less than 1 in magnitude, to make $|\lambda|^k \rightarrow 0$. Here the real part of the eigenvalue must be negative, to make $e^{\lambda t} \rightarrow 0$.

5.8 ITERATIVE ESTIMATES FOR EIGENVALUES

In scientific applications of linear algebra, eigenvalues are seldom known precisely. Fortunately, a close numerical approximation is usually quite satisfactory. In fact, some applications require only a rough approximation to the largest eigenvalue. The first algorithm described below can work well for this case. Also, it provides a foundation for a more powerful method that can give fast estimates for other eigenvalues as well.

The Power Method

The power method applies to an $n \times n$ matrix A with a **strictly dominant eigenvalue** λ_1 , which means that λ_1 must be larger in absolute value than all the other eigenvalues. In this case, the power method produces a scalar sequence that approaches λ_1 and a vector sequence that approaches a corresponding eigenvector. The background for the method rests on the eigenvector decomposition used at the beginning of Section 5.6.

Assume for simplicity that A is diagonalizable and \mathbb{R}^n has a basis of eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$, arranged so their corresponding eigenvalues $\lambda_1, \dots, \lambda_n$ decrease in size, with the strictly dominant eigenvalue first. That is,

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n| \quad (1)$$

↑ Strictly larger

As we saw in equation (2) of Section 5.6, if \mathbf{x} in \mathbb{R}^n is written as $\mathbf{x} = c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n$, then

$$A^k \mathbf{x} = c_1 (\lambda_1)^k \mathbf{v}_1 + c_2 (\lambda_2)^k \mathbf{v}_2 + \dots + c_n (\lambda_n)^k \mathbf{v}_n \quad (k = 1, 2, \dots)$$

Assume $c_1 \neq 0$. Then, dividing by $(\lambda_1)^k$,

$$\frac{1}{(\lambda_1)^k} A^k \mathbf{x} = c_1 \mathbf{v}_1 + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{v}_2 + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{v}_n \quad (k = 1, 2, \dots) \quad (2)$$

From inequality (1), the fractions $\lambda_2/\lambda_1, \dots, \lambda_n/\lambda_1$ are all less than 1 in magnitude and so their powers go to zero. Hence

$$(\lambda_1)^{-k} A^k \mathbf{x} \rightarrow c_1 \mathbf{v}_1 \quad \text{as } k \rightarrow \infty \quad (3)$$

Thus, for large k , a scalar multiple of $A^k \mathbf{x}$ determines almost the same *direction* as the eigenvector $c_1 \mathbf{v}_1$. Since positive scalar multiples do not change the direction of a vector, $A^k \mathbf{x}$ itself points almost in the same direction as \mathbf{v}_1 or $-\mathbf{v}_1$, provided $c_1 \neq 0$.

EXAMPLE 1 Let $A = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -.5 \\ 1 \end{bmatrix}$. Then A has eigenvalues 2 and 1, and the eigenspace for $\lambda_1 = 2$ is the line through $\mathbf{0}$ and \mathbf{v}_1 . For $k = 0, \dots, 8$, compute $A^k \mathbf{x}$ and construct the line through $\mathbf{0}$ and $A^k \mathbf{x}$. What happens as k increases?

SOLUTION The first three calculations are

$$\begin{aligned} A\mathbf{x} &= \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix} \begin{bmatrix} -.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -.1 \\ 1.1 \end{bmatrix} \\ A^2\mathbf{x} &= A(A\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix} \begin{bmatrix} -.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} .7 \\ 1.3 \end{bmatrix} \\ A^3\mathbf{x} &= A(A^2\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix} \begin{bmatrix} .7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix} \end{aligned}$$

Analogous calculations complete Table 1.

TABLE 1 Iterates of a Vector

k	0	1	2	3	4	5	6	7	8
$A^k \mathbf{x}$	$\begin{bmatrix} -.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

The vectors $\mathbf{x}, A\mathbf{x}, \dots, A^4\mathbf{x}$ are shown in Fig. 1. The other vectors are growing too long to display. However, line segments are drawn showing the directions of those vectors. In fact, the directions of the vectors are what we really want to see, not the vectors themselves. The lines seem to be approaching the line representing the eigenspace spanned by \mathbf{v}_1 . More precisely, the angle between the line (subspace) determined by $A^k \mathbf{x}$ and the line (eigenspace) determined by \mathbf{v}_1 goes to zero as $k \rightarrow \infty$. ■

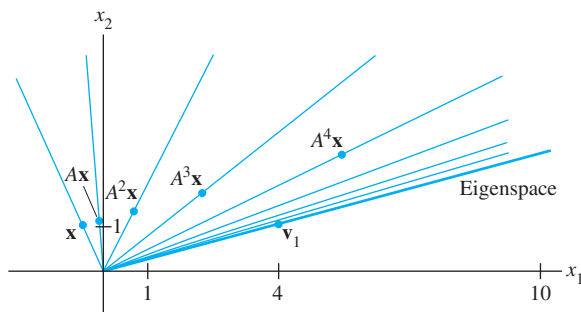


FIGURE 1 Directions determined by $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \dots, A^7\mathbf{x}$.

The vectors $(\lambda_1)^{-k} A^k \mathbf{x}$ in (3) are scaled to make them converge to $c_1 \mathbf{v}_1$, provided $c_1 \neq 0$. We cannot scale $A^k \mathbf{x}$ in this way because we do not know λ_1 . But we can scale each $A^k \mathbf{x}$ to make its largest entry a 1. It turns out that the resulting sequence $\{\mathbf{x}_k\}$ will converge to a multiple of \mathbf{v}_1 whose largest entry is 1. Figure 2 shows the scaled sequence

for Example 1. The eigenvalue λ_1 can be estimated from the sequence $\{\mathbf{x}_k\}$, too. When \mathbf{x}_k is close to an eigenvector for λ_1 , the vector $A\mathbf{x}_k$ is close to $\lambda_1\mathbf{x}_k$, with each entry in $A\mathbf{x}_k$ approximately λ_1 times the corresponding entry in \mathbf{x}_k . Because the largest entry in \mathbf{x}_k is 1, the largest entry in $A\mathbf{x}_k$ is close to λ_1 . (Careful proofs of these statements are omitted.)

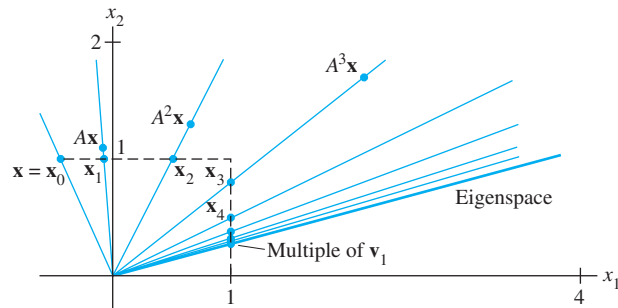


FIGURE 2 Scaled multiples of \mathbf{x} , $A\mathbf{x}$, $A^2\mathbf{x}$, \dots , $A^7\mathbf{x}$.

THE POWER METHOD FOR ESTIMATING A STRICTLY DOMINANT EIGENVALUE

1. Select an initial vector \mathbf{x}_0 whose largest entry is 1.
2. For $k = 0, 1, \dots$,
 - a. Compute $A\mathbf{x}_k$.
 - b. Let μ_k be an entry in $A\mathbf{x}_k$ whose absolute value is as large as possible.
 - c. Compute $\mathbf{x}_{k+1} = (1/\mu_k)A\mathbf{x}_k$.
3. For almost all choices of \mathbf{x}_0 , the sequence $\{\mu_k\}$ approaches the dominant eigenvalue, and the sequence $\{\mathbf{x}_k\}$ approaches a corresponding eigenvector.

EXAMPLE 2 Apply the power method to $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ with $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Stop when $k = 5$, and estimate the dominant eigenvalue and a corresponding eigenvector of A .

SOLUTION Calculations in this example and the next were made with MATLAB, which computes with 16-digit accuracy, although we show only a few significant figures here. To begin, compute $A\mathbf{x}_0$ and identify the largest entry μ_0 in $A\mathbf{x}_0$:

$$A\mathbf{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad \mu_0 = 5$$

Scale $A\mathbf{x}_0$ by $1/\mu_0$ to get \mathbf{x}_1 , compute $A\mathbf{x}_1$, and identify the largest entry in $A\mathbf{x}_1$:

$$\mathbf{x}_1 = \frac{1}{\mu_0}A\mathbf{x}_0 = \frac{1}{5} \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$A\mathbf{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \quad \mu_1 = 8$$

Scale $A\mathbf{x}_1$ by $1/\mu_1$ to get \mathbf{x}_2 , compute $A\mathbf{x}_2$, and identify the largest entry in $A\mathbf{x}_2$:

$$\mathbf{x}_2 = \frac{1}{\mu_1}A\mathbf{x}_1 = \frac{1}{8} \begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \quad \mu_2 = 7.125$$

Scale $A\mathbf{x}_2$ by $1/\mu_2$ to get \mathbf{x}_3 , and so on. The results of MATLAB calculations for the first five iterations are arranged in Table 2.

TABLE 2 The Power Method for Example 2

k	0	1	2	3	4	5
\mathbf{x}_k	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
$A\mathbf{x}_k$	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

The evidence from Table 2 strongly suggests that $\{\mathbf{x}_k\}$ approaches $(1, .2)$ and $\{\mu_k\}$ approaches 7. If so, then $(1, .2)$ is an eigenvector and 7 is the dominant eigenvalue. This is easily verified by computing

$$A \begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ .2 \end{bmatrix} \quad \blacksquare$$

The sequence $\{\mu_k\}$ in Example 2 converged quickly to $\lambda_1 = 7$ because the second eigenvalue of A was much smaller. (In fact, $\lambda_2 = 1$.) In general, the rate of convergence depends on the ratio $|\lambda_2/\lambda_1|$, because the vector $c_2(\lambda_2/\lambda_1)^k \mathbf{v}_2$ in equation (2) is the main source of error when using a scaled version of $A^k \mathbf{x}$ as an estimate of $c_1 \mathbf{v}_1$. (The other fractions λ_j/λ_1 are likely to be smaller.) If $|\lambda_2/\lambda_1|$ is close to 1, then $\{\mu_k\}$ and $\{\mathbf{x}_k\}$ can converge very slowly, and other approximation methods may be preferred.

With the power method, there is a slight chance that the chosen initial vector \mathbf{x} will have no component in the \mathbf{v}_1 direction (when $c_1 = 0$). But computer rounding errors during the calculations of the \mathbf{x}_k are likely to create a vector with at least a small component in the direction of \mathbf{v}_1 . If that occurs, the \mathbf{x}_k will start to converge to a multiple of \mathbf{v}_1 .

The Inverse Power Method

This method provides an approximation for *any* eigenvalue, provided a good initial estimate α of the eigenvalue λ is known. In this case, we let $B = (A - \alpha I)^{-1}$ and apply the power method to B . It can be shown that if the eigenvalues of A are $\lambda_1, \dots, \lambda_n$, then the eigenvalues of B are

$$\frac{1}{\lambda_1 - \alpha}, \quad \frac{1}{\lambda_2 - \alpha}, \quad \dots, \quad \frac{1}{\lambda_n - \alpha}$$

and the corresponding eigenvectors are the same as those for A . (See Exercises 15 and 16.)

Suppose, for example, that α is closer to λ_2 than to the other eigenvalues of A . Then $1/(\lambda_2 - \alpha)$ will be a strictly dominant eigenvalue of B . If α is really close to λ_2 , then $1/(\lambda_2 - \alpha)$ is *much* larger than the other eigenvalues of B , and the inverse power method produces a very rapid approximation to λ_2 for almost all choices of \mathbf{x}_0 . The following algorithm gives the details.

THE INVERSE POWER METHOD FOR ESTIMATING AN EIGENVALUE λ OF A

1. Select an initial estimate α sufficiently close to λ .
2. Select an initial vector \mathbf{x}_0 whose largest entry is 1.
3. For $k = 0, 1, \dots$,
 - a. Solve $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ for \mathbf{y}_k .
 - b. Let μ_k be an entry in \mathbf{y}_k whose absolute value is as large as possible.
 - c. Compute $v_k = \alpha + (1/\mu_k)$.
 - d. Compute $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$.
4. For almost all choices of \mathbf{x}_0 , the sequence $\{v_k\}$ approaches the eigenvalue λ of A , and the sequence $\{\mathbf{x}_k\}$ approaches a corresponding eigenvector.

Notice that B , or rather $(A - \alpha I)^{-1}$, does not appear in the algorithm. Instead of computing $(A - \alpha I)^{-1}\mathbf{x}_k$ to get the next vector in the sequence, it is better to *solve* the equation $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ for \mathbf{y}_k (and then scale \mathbf{y}_k to produce \mathbf{x}_{k+1}). Since this equation for \mathbf{y}_k must be solved for each k , an LU factorization of $A - \alpha I$ will speed up the process.

EXAMPLE 3 It is not uncommon in some applications to need to know the smallest eigenvalue of a matrix A and to have at hand rough estimates of the eigenvalues. Suppose 21, 3.3, and 1.9 are estimates for the eigenvalues of the matrix A below. Find the smallest eigenvalue, accurate to six decimal places.

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

SOLUTION The two smallest eigenvalues seem close together, so we use the inverse power method for $A - 1.9I$. Results of a MATLAB calculation are shown in Table 3. Here \mathbf{x}_0 was chosen arbitrarily, $\mathbf{y}_k = (A - 1.9I)^{-1}\mathbf{x}_k$, μ_k is the largest entry in \mathbf{y}_k , $v_k = 1.9 + 1/\mu_k$, and $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$. As it turns out, the initial eigenvalue estimate was fairly good, and the inverse power sequence converged quickly. The smallest eigenvalue is exactly 2. ■

TABLE 3 The Inverse Power Method

k	0	1	2	3	4
\mathbf{x}_k	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5736 \\ .0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5054 \\ .0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5004 \\ .0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .50003 \\ .00002 \\ 1 \end{bmatrix}$
\mathbf{y}_k	$\begin{bmatrix} 4.45 \\ .50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ .0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ .0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ .0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ .000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
v_k	2.03	2.0008	2.00005	2.000004	2.0000002

If an estimate for the smallest eigenvalue of a matrix is not available, one can simply take $\alpha = 0$ in the inverse power method. This choice of α works reasonably well if the smallest eigenvalue is much closer to zero than to the other eigenvalues.

The two algorithms presented in this section are practical tools for many simple situations, and they provide an introduction to the problem of eigenvalue estimation. A more robust and widely used iterative method is the QR algorithm. For instance, it is the heart of the MATLAB command `eig(A)`, which rapidly computes eigenvalues and eigenvectors of A . A brief description of the QR algorithm was given in the exercises for Section 5.2. Further details are presented in most modern numerical analysis texts.

PRACTICE PROBLEM

How can you tell if a given vector \mathbf{x} is a good approximation to an eigenvector of a matrix A ? If it is, how would you estimate the corresponding eigenvalue? Experiment with

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

5.8 EXERCISES

In Exercises 1–4, the matrix A is followed by a sequence $\{\mathbf{x}_k\}$ produced by the power method. Use these data to estimate the largest eigenvalue of A , and give a corresponding eigenvector.

$$1. \quad A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}; \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .25 \end{bmatrix}, \begin{bmatrix} 1 \\ .3158 \end{bmatrix}, \begin{bmatrix} 1 \\ .3298 \end{bmatrix}, \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$$

$$2. \quad A = \begin{bmatrix} 1.8 & -.8 \\ -3.2 & 4.2 \end{bmatrix}; \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -.5625 \end{bmatrix}, \begin{bmatrix} 1 \\ -.3021 \end{bmatrix}, \begin{bmatrix} 1 \\ -.2601 \end{bmatrix}, \begin{bmatrix} 1 \\ -.2520 \end{bmatrix}$$

$$3. \quad A = \begin{bmatrix} .5 & .2 \\ .4 & .7 \end{bmatrix}; \\ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .8 \end{bmatrix}, \begin{bmatrix} 1 \\ .6875 \end{bmatrix}, \begin{bmatrix} 1 \\ .5577 \end{bmatrix}, \begin{bmatrix} 1 \\ .5188 \end{bmatrix}$$

$$4. \quad A = \begin{bmatrix} 4.1 & -6 \\ 3 & -4.4 \end{bmatrix}; \\ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .7368 \end{bmatrix}, \begin{bmatrix} 1 \\ .7541 \end{bmatrix}, \begin{bmatrix} 1 \\ .7490 \end{bmatrix}, \begin{bmatrix} 1 \\ .7502 \end{bmatrix}$$

$$5. \quad \text{Let } A = \begin{bmatrix} 15 & 16 \\ -20 & -21 \end{bmatrix}. \text{ The vectors } \mathbf{x}, \dots, A^5 \mathbf{x} \text{ are } \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \\ \begin{bmatrix} 31 \\ -41 \end{bmatrix}, \begin{bmatrix} -191 \\ 241 \end{bmatrix}, \begin{bmatrix} 991 \\ -1241 \end{bmatrix}, \begin{bmatrix} -4991 \\ 6241 \end{bmatrix}, \begin{bmatrix} 24991 \\ -31241 \end{bmatrix}.$$

Find a vector with a 1 in the second entry that is close to an eigenvector of A . Use four decimal places. Check your estimate, and give an estimate for the dominant eigenvalue of A .

6. Let $A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}$. Repeat Exercise 5, using the following sequence $\mathbf{x}, A\mathbf{x}, \dots, A^5 \mathbf{x}$.

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \begin{bmatrix} -29 \\ 61 \end{bmatrix}, \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \begin{bmatrix} -509 \\ 1021 \end{bmatrix}, \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}$$

[M] Exercises 7–12 require MATLAB or other computational aid. In Exercises 7 and 8, use the power method with the \mathbf{x}_0 given. List $\{\mathbf{x}_k\}$ and $\{\mu_k\}$ for $k = 1, \dots, 5$. In Exercises 9 and 10, list μ_5 and μ_6 .

$$7. \quad A = \begin{bmatrix} 6 & 7 \\ 8 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$8. \quad A = \begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$9. \quad A = \begin{bmatrix} 8 & 0 & 12 \\ 1 & -2 & 1 \\ 0 & 3 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$10. \quad A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 9 \\ 0 & 1 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Another estimate can be made for an eigenvalue when an approximate eigenvector is available. Observe that if $A\mathbf{x} = \lambda\mathbf{x}$, then $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T (\lambda\mathbf{x}) = \lambda(\mathbf{x}^T \mathbf{x})$, and the **Rayleigh quotient**

$$R(\mathbf{x}) = \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T \mathbf{x}}$$

equals λ . If \mathbf{x} is close to an eigenvector for λ , then this quotient is close to λ . When A is a symmetric matrix ($A^T = A$), the Rayleigh quotient $R(\mathbf{x}_k) = (\mathbf{x}_k^T A\mathbf{x}_k) / (\mathbf{x}_k^T \mathbf{x}_k)$ will have roughly twice as many digits of accuracy as the scaling factor μ_k in the power method. Verify this increased accuracy in Exercises 11 and 12 by computing μ_k and $R(\mathbf{x}_k)$ for $k = 1, \dots, 4$.

$$11. A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$12. A = \begin{bmatrix} -3 & 2 \\ 2 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Exercises 13 and 14 apply to a 3×3 matrix A whose eigenvalues are estimated to be 4, -4 , and 3.

13. If the eigenvalues close to 4 and -4 are known to have different absolute values, will the power method work? Is it likely to be useful?
14. Suppose the eigenvalues close to 4 and -4 are known to have exactly the same absolute value. Describe how one might obtain a sequence that estimates the eigenvalue close to 4.
15. Suppose $A\mathbf{x} = \lambda\mathbf{x}$ with $\mathbf{x} \neq \mathbf{0}$. Let α be a scalar different from the eigenvalues of A , and let $B = (A - \alpha I)^{-1}$. Subtract $\alpha\mathbf{x}$ from both sides of the equation $A\mathbf{x} = \lambda\mathbf{x}$, and use algebra to show that $1/(\lambda - \alpha)$ is an eigenvalue of B , with \mathbf{x} a corresponding eigenvector.
16. Suppose μ is an eigenvalue of the B in Exercise 15, and that \mathbf{x} is a corresponding eigenvector, so that $(A - \alpha I)^{-1}\mathbf{x} = \mu\mathbf{x}$. Use this equation to find an eigenvalue of A in terms of μ and α . [Note: $\mu \neq 0$ because B is invertible.]
17. [M] Use the inverse power method to estimate the middle eigenvalue of the A in Example 3, with accuracy to four decimal places. Set $\mathbf{x}_0 = (1, 0, 0)$.

18. [M] Let A be as in Exercise 9. Use the inverse power method with $\mathbf{x}_0 = (1, 0, 0)$ to estimate the eigenvalue of A near $\alpha = -1.4$, with an accuracy to four decimal places.

[M] In Exercises 19 and 20, find (a) the largest eigenvalue and (b) the eigenvalue closest to zero. In each case, set $\mathbf{x}_0 = (1, 0, 0, 0)$ and carry out approximations until the approximating sequence seems accurate to four decimal places. Include the approximate eigenvector.

$$19. A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$$

$$20. A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 12 & 13 & 11 \\ -2 & 3 & 0 & 2 \\ 4 & 5 & 7 & 2 \end{bmatrix}$$

21. A common misconception is that if A has a strictly dominant eigenvalue, then, for any sufficiently large value of k , the vector $A^k\mathbf{x}$ is approximately equal to an eigenvector of A . For the three matrices below, study what happens to $A^k\mathbf{x}$ when $\mathbf{x} = (.5, .5)$, and try to draw general conclusions (for a 2×2 matrix).

$$a. A = \begin{bmatrix} .8 & 0 \\ 0 & .2 \end{bmatrix} \quad b. A = \begin{bmatrix} 1 & 0 \\ 0 & .8 \end{bmatrix}$$

$$c. A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$$

SOLUTION TO PRACTICE PROBLEM

For the given A and \mathbf{x} ,

$$A\mathbf{x} = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1.00 \\ -4.30 \\ 8.10 \end{bmatrix} = \begin{bmatrix} 3.00 \\ -13.00 \\ 24.50 \end{bmatrix}$$

If $A\mathbf{x}$ is nearly a multiple of \mathbf{x} , then the ratios of corresponding entries in the two vectors should be nearly constant. So compute:

$$\begin{array}{r} \{\text{entry in } A\mathbf{x}\} \div \{\text{entry in } \mathbf{x}\} = \{\text{ratio}\} \\ \begin{array}{ccc} 3.00 & 1.00 & 3.000 \\ -13.00 & -4.30 & 3.023 \\ 24.50 & 8.10 & 3.025 \end{array} \end{array}$$

Each entry in $A\mathbf{x}$ is about 3 times the corresponding entry in \mathbf{x} , so \mathbf{x} is close to an eigenvector. Any of the ratios above is an estimate for the eigenvalue. (To five decimal places, the eigenvalue is 3.02409.)

WEB

CHAPTER 5 SUPPLEMENTARY EXERCISES

Throughout these supplementary exercises, A and B represent square matrices of appropriate sizes.

1. Mark each statement as True or False. Justify each answer.
 - a. If A is invertible and 1 is an eigenvalue for A , then 1 is also an eigenvalue of A^{-1} .
 - b. If A is row equivalent to the identity matrix I , then A is diagonalizable.
 - c. If A contains a row or column of zeros, then 0 is an eigenvalue of A .
 - d. Each eigenvalue of A is also an eigenvalue of A^2 .
 - e. Each eigenvector of A is also an eigenvector of A^2 .
 - f. Each eigenvector of an invertible matrix A is also an eigenvector of A^{-1} .
 - g. Eigenvalues must be nonzero scalars.
 - h. Eigenvectors must be nonzero vectors.
 - i. Two eigenvectors corresponding to the same eigenvalue are always linearly dependent.
 - j. Similar matrices always have exactly the same eigenvalues.
 - k. Similar matrices always have exactly the same eigenvectors.
 - l. The sum of two eigenvectors of a matrix A is also an eigenvector of A .
 - m. The eigenvalues of an upper triangular matrix A are exactly the nonzero entries on the diagonal of A .
 - n. The matrices A and A^T have the same eigenvalues, counting multiplicities.
 - o. If a 5×5 matrix A has fewer than 5 distinct eigenvalues, then A is not diagonalizable.
 - p. There exists a 2×2 matrix that has no eigenvectors in \mathbb{R}^2 .
 - q. If A is diagonalizable, then the columns of A are linearly independent.
 - r. A nonzero vector cannot correspond to two different eigenvalues of A .
 - s. A (square) matrix A is invertible if and only if there is a coordinate system in which the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is represented by a diagonal matrix.
 - t. If each vector \mathbf{e}_j in the standard basis for \mathbb{R}^n is an eigenvector of A , then A is a diagonal matrix.
 - u. If A is similar to a diagonalizable matrix B , then A is also diagonalizable.
 - v. If A and B are invertible $n \times n$ matrices, then AB is similar to BA .
 - w. An $n \times n$ matrix with n linearly independent eigenvectors is invertible.

x. If A is an $n \times n$ diagonalizable matrix, then each vector in \mathbb{R}^n can be written as a linear combination of eigenvectors of A .

2. Show that if \mathbf{x} is an eigenvector of the matrix product AB and $B\mathbf{x} \neq \mathbf{0}$, then $B\mathbf{x}$ is an eigenvector of BA .
3. Suppose \mathbf{x} is an eigenvector of A corresponding to an eigenvalue λ .
 - a. Show that \mathbf{x} is an eigenvector of $5I - A$. What is the corresponding eigenvalue?
 - b. Show that \mathbf{x} is an eigenvector of $5I - 3A + A^2$. What is the corresponding eigenvalue?
4. Use mathematical induction to show that if λ is an eigenvalue of an $n \times n$ matrix A , with \mathbf{x} a corresponding eigenvector, then, for each positive integer m , λ^m is an eigenvalue of A^m , with \mathbf{x} a corresponding eigenvector.
5. If $p(t) = c_0 + c_1t + c_2t^2 + \cdots + c_nt^n$, define $p(A)$ to be the matrix formed by replacing each power of t in $p(t)$ by the corresponding power of A (with $A^0 = I$). That is,

$$p(A) = c_0I + c_1A + c_2A^2 + \cdots + c_nA^n$$
 Show that if λ is an eigenvalue of A , then one eigenvalue of $p(A)$ is $p(\lambda)$.
6. Suppose $A = PDP^{-1}$, where P is 2×2 and $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.
 - a. Let $B = 5I - 3A + A^2$. Show that B is diagonalizable by finding a suitable factorization of B .
 - b. Given $p(t)$ and $p(A)$ as in Exercise 5, show that $p(A)$ is diagonalizable.
7. Suppose A is diagonalizable and $p(t)$ is the characteristic polynomial of A . Define $p(A)$ as in Exercise 5, and show that $p(A)$ is the zero matrix. This fact, which is also true for any square matrix, is called the *Cayley–Hamilton theorem*.
8.
 - a. Let A be a diagonalizable $n \times n$ matrix. Show that if the multiplicity of an eigenvalue λ is n , then $A = \lambda I$.
 - b. Use part (a) to show that the matrix $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ is not diagonalizable.
9. Show that $I - A$ is invertible when all the eigenvalues of A are less than 1 in magnitude. [*Hint*: What would be true if $I - A$ were not invertible?]
10. Show that if A is diagonalizable, with all eigenvalues less than 1 in magnitude, then A^k tends to the zero matrix as $k \rightarrow \infty$. [*Hint*: Consider $A^k\mathbf{x}$ where \mathbf{x} represents any one of the columns of I .]
11. Let \mathbf{u} be an eigenvector of A corresponding to an eigenvalue λ , and let H be the line in \mathbb{R}^n through \mathbf{u} and the origin.
 - a. Explain why H is invariant under A in the sense that $A\mathbf{x}$ is in H whenever \mathbf{x} is in H .

b. Let K be a one-dimensional subspace of \mathbb{R}^n that is invariant under A . Explain why K contains an eigenvector of A .

12. Let $G = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$. Use formula (1) for the determinant in Section 5.2 to explain why $\det G = (\det A)(\det B)$. From this, deduce that the characteristic polynomial of G is the product of the characteristic polynomials of A and B .

Use Exercise 12 to find the eigenvalues of the matrices in Exercises 13 and 14.

13. $A = \begin{bmatrix} 3 & -2 & 8 \\ 0 & 5 & -2 \\ 0 & -4 & 3 \end{bmatrix}$

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

15. Let J be the $n \times n$ matrix of all 1's, and consider $A = (a - b)I + bJ$; that is,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Use the results of Exercise 16 in the Supplementary Exercises for Chapter 3 to show that the eigenvalues of A are $a - b$ and $a + (n - 1)b$. What are the multiplicities of these eigenvalues?

16. Apply the result of Exercise 15 to find the eigenvalues of the

matrices $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ and $\begin{bmatrix} 7 & 3 & 3 & 3 & 3 \\ 3 & 7 & 3 & 3 & 3 \\ 3 & 3 & 7 & 3 & 3 \\ 3 & 3 & 3 & 7 & 3 \\ 3 & 3 & 3 & 3 & 7 \end{bmatrix}$.

17. Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Recall from Exercise 25 in Section 5.4 that $\operatorname{tr} A$ (the trace of A) is the sum of the diagonal entries in A . Show that the characteristic polynomial of A is

$$\lambda^2 - (\operatorname{tr} A)\lambda + \det A$$

Then show that the eigenvalues of a 2×2 matrix A are both real if and only if $\det A \leq \left(\frac{\operatorname{tr} A}{2}\right)^2$.

18. Let $A = \begin{bmatrix} .4 & -.3 \\ .4 & 1.2 \end{bmatrix}$. Explain why A^k approaches $\begin{bmatrix} -.5 & -.75 \\ 1.0 & 1.50 \end{bmatrix}$ as $k \rightarrow \infty$.

Exercises 19–23 concern the polynomial

$$p(t) = a_0 + a_1t + \cdots + a_{n-1}t^{n-1} + t^n$$

and an $n \times n$ matrix C_p called the **companion matrix** of p :

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

19. Write the companion matrix C_p for $p(t) = 6 - 5t + t^2$, and then find the characteristic polynomial of C_p .

20. Let $p(t) = (t - 2)(t - 3)(t - 4) = -24 + 26t - 9t^2 + t^3$. Write the companion matrix for $p(t)$, and use techniques from Chapter 3 to find its characteristic polynomial.

21. Use mathematical induction to prove that for $n \geq 2$,

$$\det(C_p - \lambda I) = (-1)^n(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) = (-1)^n p(\lambda)$$

[Hint: Expanding by cofactors down the first column, show that $\det(C_p - \lambda I)$ has the form $(-\lambda)B + (-1)^n a_0$, where B is a certain polynomial (by the induction assumption).]

22. Let $p(t) = a_0 + a_1t + a_2t^2 + t^3$, and let λ be a zero of p .

a. Write the companion matrix for p .

b. Explain why $\lambda^3 = -a_0 - a_1\lambda - a_2\lambda^2$, and show that $(1, \lambda, \lambda^2)$ is an eigenvector of the companion matrix for p .

23. Let p be the polynomial in Exercise 22, and suppose the equation $p(t) = 0$ has distinct roots $\lambda_1, \lambda_2, \lambda_3$. Let V be the Vandermonde matrix

$$V = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

(The transpose of V was considered in Supplementary Exercise 11 in Chapter 2.) Use Exercise 22 and a theorem from this chapter to deduce that V is invertible (but do not compute V^{-1}). Then explain why $V^{-1}C_pV$ is a diagonal matrix.

24. [M] The MATLAB command `roots(p)` computes the roots of the polynomial equation $p(t) = 0$. Read a MATLAB manual, and then describe the basic idea behind the algorithm for the `roots` command.

25. [M] Use a matrix program to diagonalize

$$A = \begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$$

if possible. Use the eigenvalue command to create the diagonal matrix D . If the program has a command that produces eigenvectors, use it to create an invertible matrix P . Then compute $AP - PD$ and PDP^{-1} . Discuss your results.

26. [M] Repeat Exercise 25 for $A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$.

This page intentionally left blank