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).

A first step in studying the population dynamics is to model the population at yearly intervals, at times denoted by $k = 0, 1, 2, \ldots$. 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

[1] "The Great Spotted Owl War," *Reader's Digest*, November 1992, pp. 91–95.

[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, $\mathbf{v}$. ■

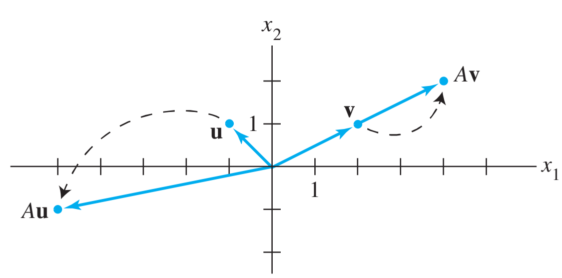


FIGURE 1 Effects of multiplication by A.

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

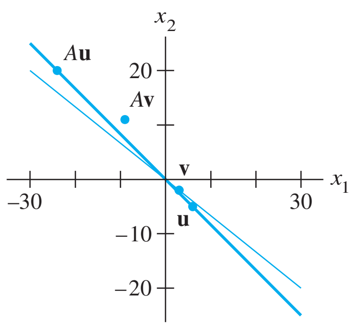
$$A\mathbf{x} = 2\mathbf{x} \quad \text{or} \quad A\mathbf{x} = -4\mathbf{x}$$

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 $\lambda = 7$. ■

---

[1] Note 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 eigen*vectors*, it cannot be used to find eigen*values*. 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

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

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.

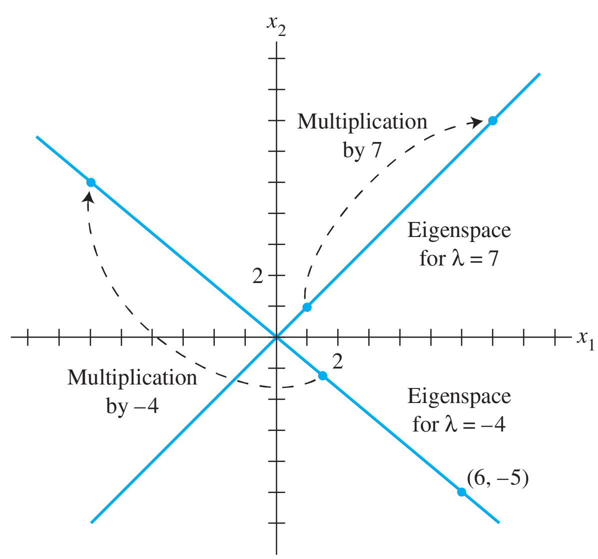


FIGURE 2 Eigenspaces for $\lambda = -4$ and $\lambda = 7$.

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

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

■

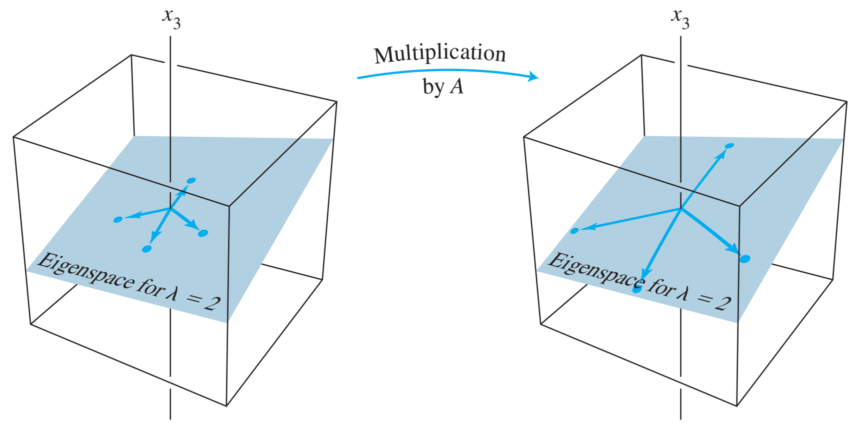


FIGURE 3 A acts as a dilation on the eigenspace.

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

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

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

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

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

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

$$c_1A\mathbf{v}_1 + \cdots + c_pA\mathbf{v}_p = A\mathbf{v}_{p+1}$$

$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p = \lambda_{p+1}\mathbf{v}_{p+1} \tag{6}$$

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

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

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:

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

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

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

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

1. Is 5 an eigenvalue of $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?
2. If $\mathbf{x}$ is an eigenvector of A corresponding to λ, what is $A^3\mathbf{x}$?
3. 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

1. Is $\lambda = 2$ an eigenvalue of $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? Why or why not?
2. Is $\lambda = -3$ an eigenvalue of $\begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$? Why or why not?
3. Is $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 1 & -1 \\ 6 & -4 \end{bmatrix}$? If so, find the eigenvalue.
4. Is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 5 & 2 \\ 3 & 6 \end{bmatrix}$? If so, find the eigenvalue.
5. 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.
6. 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.
7. 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.
8. 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.

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

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

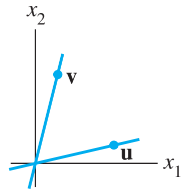
$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \ldots)$$

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 $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$.

34. Describe how you might try to build a solution of a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ 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 $T : \mathbb{R}^2 \to \mathbb{R}^2$ 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 **u** and **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. $\begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix}$

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 **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 $u_{11}, \ldots, u_{nn}$ in U. If A is invertible, then $u_{11}, \ldots, u_{nn}$ 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. Thus[1]

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

[1] Formula (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.

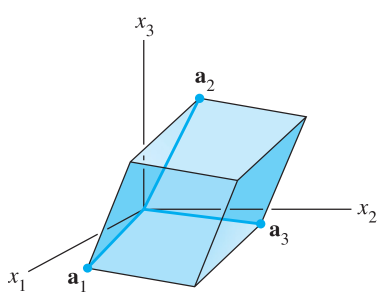


FIGURE 1

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

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

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). ■

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 nth-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

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

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

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

In fact,

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

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,

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

Using c_1 and c_2 from (4),

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

This explicit formula for $\mathbf{x}_k$ gives the solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. As $k \to \infty$, $(.92)^k$ 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 $\mathbf{x}_k \to .125\mathbf{v}_1$.

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 $\lambda_1, \ldots, \lambda_n$ 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
$$A_1 = A \quad \text{and} \quad A_{k+1} = P_k^{-1} A_k P_k \quad (k = 1, 2, \ldots)$$
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.]

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

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

15. $\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ **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 $A = Q_1R_1$, where $Q_1^T = Q_1^{-1}$ and R_1 is upper triangular. The factors are interchanged to form $A_1 = R_1Q_1$, which is again factored as $A_1 = Q_2R_2$; then to form $A_2 = R_2Q_2$, 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 $\mathbf{x}_k \to \mathbf{v}_1$ 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. Let $A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$,

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, 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 $\mathbf{x}_k \to \mathbf{v}_1$ as k increases.