(c) $\lambda = 2$, $A = \begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix}$.

ML.4. Consider a living organism that can live to a maximum age of two years and whose Leslie matrix is

$$\begin{bmatrix} 0.2 & 0.8 & 0.3 \\ 0.9 & 0 & 0 \\ 0 & 0.7 & 0 \end{bmatrix}.$$

Find a stable age distribution.

8.2 DIAGONALIZATION

In this section we show how to find the eigenvalues and associated eigenvectors of a given matrix A by finding the eigenvalues and eigenvectors of a related matrix B that has the same eigenvalues and eigenvectors as A. The matrix B has the helpful property that its eigenvalues are easily obtained. Thus we will have found the eigenvalues of A. In Section 8.3, this equivalent approach will shed much light on the eigenvalue-eigenvector problem.

Similar Matrices

DEFINITION A matrix B is said to be **similar** to a matrix A if there is a nonsingular matrix P such that

$$B = P^{-1}AP.$$

EXAMPLE 1 Let

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

be the matrix of Example 4 in Section 8.1. Let

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

Then

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

and

$$B = P^{-1}AP = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}.$$

Thus B is similar to A.

We shall let the reader (Exercise T.1) show that the following elementary properties hold for similarity:

1. A is similar to A.
2. If B is similar to A, then A is similar to B.
3. If A is similar to B and B is similar to C, then A is similar to C.

By property 2 we replace the statements "A is similar to B" and "B is similar to A" by "A and B are similar."

DEFINITION We shall say that the matrix A is **diagonalizable** if it is similar to a diagonal matrix. In this case we also say that A can be diagonaliz

EXAMPLE 2 If A and B are as in Example 1, then A is diagonalizable, since it is similar to B. ■

THEOREM 8.3 *Similar matrices have the same eigenvalues.*

Proof Let A and B be similar. Then $B = P^{-1}AP$, for some nonsingular matrix P. We prove that A and B have the same characteristic polynomials, $f_A(\lambda)$ and $f_B(\lambda)$, respectively. We have

$$\begin{aligned} f_B(\lambda) &= \det(\lambda I_n - B) = \det(\lambda I_n - P^{-1}AP) \\ &= \det(P^{-1}\lambda I_n P - P^{-1}AP) = \det(P^{-1}(\lambda I_n - A)P) \\ &= \det(P^{-1})\det(\lambda I_n - A)\det(P) \\ &= \det(P^{-1})\det(P)\det(\lambda I_n - A) \\ &= \det(\lambda I_n - A) = f_A(\lambda). \end{aligned} \tag{1}$$

Since $f_A(\lambda) = f_B(\lambda)$, it follows that A and B have the same eigenvalues. ■

It follows from Exercise T.3 in Section 8.1 that the eigenvalues of a diagonal matrix are the entries on its main diagonal. The following theorem establishes when a matrix is diagonalizable.

THEOREM 8.4 *An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. In this case A is similar to a diagonal matrix D, with $P^{-1}AP = D$, whose diagonal elements are the eigenvalues of A, while P is a matrix whose columns are respectively the n linearly independent eigenvectors of A.*

Proof Suppose that A is similar to D. Then

$$P^{-1}AP = D,$$

so that

$$AP = PD. \tag{2}$$

Let

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix},$$

and let $\mathbf{x}_j$, $j = 1, 2, \ldots, n$, be the jth column of P. From Exercise T.9 in Section 1.3, it follows that the jth column of the matrix AP is $A\mathbf{x}_j$, and the jth column of PD is $\lambda_j\mathbf{x}_j$.

Thus from (2) we have

$$A\mathbf{x}_j = \lambda_j\mathbf{x}_j. \tag{3}$$

Since P is a nonsingular matrix, by Theorem 6.13 in Section 6.6 its columns are linearly independent and so are all nonzero. Hence λ_j is an eigenvalue of A and $\mathbf{x}_j$ is a corresponding eigenvector.

Conversely, suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are n eigenvalues of A and that the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are linearly independent. Let $P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$ be the matrix whose jth column is $\mathbf{x}_j$. Since the columns of P are linearly independent, it follows from Theorem 6.13 in Section 6.6 that P is nonsingular. From (3) we obtain (2), which implies that A is diagonalizable. This completes the proof. ■

Observe that in Theorem 8.4 the order of the columns of P de… the order of the diagonal entries in D.

EXAMPLE 3

Let A be as in Example 1. The eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 3$ … Example 4 in Section 8.1.) The corresponding eigenvectors

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

are linearly independent. Hence A is diagonalizable. Here

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}.$$

Thus, as in Example 1,

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

On the other hand, if we let $\lambda_1 = 3$ and $\lambda_2 = 2$, then

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

Then

$$P = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} -1 & 1 \\ 2 & -1 \end{bmatrix}.$$

Hence

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

EXAMPLE 4

Let

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

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 1$. Eigenvectors associated w… and λ_2 are vectors of the form

$$\begin{bmatrix} r \\ 0 \end{bmatrix},$$

where r is any nonzero real number. Since A does not have two li… independent eigenvectors, we conclude that A is not diagonalizable.

The following is a useful theorem because it identifies a large cl… matrices that can be diagonalized.

THEOREM 8.5

A matrix A is diagonalizable if all the roots of its characteristic pol… are real and distinct.

Proof Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the distinct eigenvalues of A and let $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ be a set of associated eigenvectors. We wish to show that S is linearly independent.

Suppose that S is linearly dependent. Then Theorem 6.4 of Section 6.3 implies that some vector $\mathbf{x}_j$ is a linear combination of the preceding vectors in S. We can assume that $S_1 = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{j-1}\}$ is linearly independent, for otherwise one of the vectors in S_1 is a linear combination of the preceding ones, and we can choose a new set S_2, and so on. We thus have that S_1 is linearly independent and that

$$\mathbf{x}_j = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_{j-1}\mathbf{x}_{j-1}, \tag{4}$$

where $c_1, c_2, \ldots, c_{j-1}$ are real numbers. Premultiplying (multiplying on the left) both sides of Equation (4) by A, we obtain

$$\begin{aligned} A\mathbf{x}_j &= A(c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_{j-1}\mathbf{x}_{j-1}) \\ &= c_1A\mathbf{x}_1 + c_2A\mathbf{x}_2 + \cdots + c_{j-1}A\mathbf{x}_{j-1}. \end{aligned} \tag{5}$$

Since $\lambda_1, \lambda_2, \ldots, \lambda_j$ are eigenvalues of A and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j$, its associated eigenvectors, we know that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ for $i = 1, 2, \ldots, j$. Substituting in (5), we have

$$\lambda_j\mathbf{x}_j = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_{j-1}\lambda_{j-1}\mathbf{x}_{j-1}. \tag{6}$$

Multiplying (4) by λ_j, we obtain

$$\lambda_j\mathbf{x}_j = \lambda_jc_1\mathbf{x}_1 + \lambda_jc_2\mathbf{x}_2 + \cdots + \lambda_jc_{j-1}\mathbf{x}_{j-1}. \tag{7}$$

Subtracting (7) from (6), we have

$$\begin{aligned} \mathbf{0} &= \lambda_j\mathbf{x}_j - \lambda_j\mathbf{x}_j \\ &= c_1(\lambda_1 - \lambda_j)\mathbf{x}_1 + c_2(\lambda_2 - \lambda_j)\mathbf{x}_2 + \cdots + c_{j-1}(\lambda_{j-1} - \lambda_j)\mathbf{x}_{j-1}. \end{aligned}$$

Since S_1 is linearly independent, we must have

$$c_1(\lambda_1 - \lambda_j) = 0, \quad c_2(\lambda_2 - \lambda_j) = 0, \ldots, \quad c_{j-1}(\lambda_{j-1} - \lambda_j) = 0.$$

Now

$$\lambda_1 - \lambda_j \neq 0, \quad \lambda_2 - \lambda_j \neq 0, \ldots, \quad \lambda_{j-1} - \lambda_j \neq 0$$

(because the λ's are distinct), which implies that

$$c_1 = c_2 = \cdots = c_{j-1} = 0.$$

From (4) we conclude that $\mathbf{x}_j = \mathbf{0}$, which is impossible if $\mathbf{x}_j$ is an eigenvector. Hence S is linearly independent, and from Theorem 8.4 it follows that A is diagonalizable. ■

Remark In the proof of Theorem 8.5, we have actually established the following somewhat stronger result: Let A be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be k distinct eigenvalues of A with associated eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$. *Then* $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linearly independent (Exercise T.11).

If all the roots of the characteristic polynomial of A are real and not distinct, then A may or may not be diagonalizable. The characteristic polynomial of A can be written as the product of n factors, each of the form $\lambda - \lambda_j$, where λ_j is a root of the characteristic polynomial and the eigenvalues of A are the real roots of the characteristic polynomial of A. Thus the characteristic polynomial can be written as

$$(\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}\cdots(\lambda - \lambda_r)^{k_r},$$

where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are the distinct eigenvalues of A, and $k_1, k_2, \ldots, k_r$ are integers whose sum is n. The integer k_i is called the **multiplicity** of λ_i. Thus in Example 4, $\lambda = 1$ is an eigenvalue of

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

of multiplicity 2. It can be shown that if the roots of the characteristic polynomial of A are all real, then A can be diagonalized if and only if for each eigenvalue λ_j of multiplicity k_j we can find k_j linearly independent eigenvectors. This means that the solution space of the linear system $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$ has dimension k_j. It can also be shown that if λ_j is an eigenvalue of A of multiplicity k_j, then we can never find more than k_j linearly independent eigenvectors associated with λ_j. We consider the following examples.

EXAMPLE 5 Let

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

The characteristic polynomial of A is $f(\lambda) = \lambda(\lambda - 1)^2$, so the eigenvalues of A are $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 1$. Thus $\lambda_2 = 1$ is an eigenvalue of multiplicity 2. We now consider the eigenvectors associated with the eigenvalues $\lambda_2 = \lambda_3 = 1$. They are obtained by solving the linear system $(1I_3 - A)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A solution is any vector of the form

$$\begin{bmatrix} 0 \\ r \\ 0 \end{bmatrix},$$

where r is any real number, so the dimension of the solution space of the linear system $(1I_3 - A)\mathbf{x} = \mathbf{0}$ is 1. There do not exist two linearly independent eigenvectors associated with $\lambda_2 = 1$. Thus A cannot be diagonalized.

EXAMPLE 6 Let

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

The characteristic polynomial of A is $f(\lambda) = \lambda(\lambda - 1)^2$, so the eigenvalues of A are $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 1$; $\lambda_2 = 1$ is again an eigenvalue of multiplicity 2. Now we consider the solution space of $(1I_3 - A)\mathbf{x} = \mathbf{0}$, that is, of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A solution is any vector of the form

$$\begin{bmatrix} 0 \\ r \\ s \end{bmatrix}$$

for any real numbers r and s. Thus we can take as eigenvectors $\mathbf{x}_2$ and $\mathbf{x}_3$ the vectors

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Now we look for an eigenvector associated with $\lambda_1 = 0$. We have to solve $(0I_3 - A)\mathbf{x} = \mathbf{0}$, or

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

A solution is any vector of the form

$$\begin{bmatrix} t \\ 0 \\ -t \end{bmatrix}$$

for any real number t. Thus

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

is an eigenvector associated with $\lambda_1 = 0$. Since $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are linearly independent, A can be diagonalized. ■

Thus an $n \times n$ matrix may fail to be diagonalizable either because not all the roots of its characteristic polynomial are real numbers, or because it does not have n linearly independent eigenvectors.

The procedure for diagonalizing a matrix A is as follows.

Step 1. Form the characteristic polynomial $f(\lambda) = \det(\lambda I_n - A)$ of A.

Step 2. Find the roots of the characteristic polynomial of A. If the roots are not all real, then A cannot be diagonalized.

Step 3. For each eigenvalue λ_j of A of multiplicity k_j, find a basis for the solution space of $(\lambda_j I_n - A)\mathbf{x} = \mathbf{0}$ (the eigenspace associated with λ_j). If the dimension of the eigenspace is less than k_j, then A is not diagonalizable. We thus determine n linearly independent eigenvectors of A. In Section 6.5 we solved the problem of finding a basis for the solution space of a homogeneous system.

Step 4. Let P be the matrix whose columns are the n linearly independent eigenvectors determined in Step 3. Then $P^{-1}AP = D$, a diagonal matrix whose diagonal elements are the eigenvalues of A that correspond to the columns of P.