Section 8.2 Similarity and Diagonalizability

Definition 8.2.1 If A and C are square matrices with the same size, then we say that ***C is similar to A*** if there is an invertible matrix P such that $C = P^{-1}AP$.

REMARK If C is similar to A, then it is also true that A is similar to C. You can see this by letting $Q = P^{-1}$ and rewriting the equation $C = P^{-1}AP$ as

$$A = PCP^{-1} = (P^{-1})^{-1}C(P^{-1}) = Q^{-1}CQ$$

When we want to emphasize that similarity goes both ways, we will say that ***A and C are similar***.

Theorem 8.2.2 *Two square matrices are similar if and only if there exist bases with respect to which the matrices represent the same linear operator.*

Proof We will show first that if A and C are similar $n \times n$ matrices, then there exist bases with respect to which they represent the same linear operator on R^n. For this purpose, let $T: R^n \to R^n$ denote multiplication by A; that is,

$$A = [T] \tag{1}$$

Since A and C are similar, there exists an invertible matrix P such that $C = P^{-1}AP$, so it follows from (1) that

$$C = P^{-1}[T]P \tag{2}$$

If we assume that the column-vector form of P is

$$P = [\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n]$$

If P can be inverted, its column vectors are LI.

then the invertibility of P and Theorem 7.4.4 imply that $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for R^n. It now follows from Formula (2) above and Formula (20) of Section 8.1 that

$$C = P^{-1}[T]P = [T]_B$$

Thus we have shown that A is the matrix for T with respect to the standard basis, and C is the matrix for T with respect to the basis B, so this part of the proof is complete.

Conversely, assume that C represents the linear operator $T: R^n \to R^n$ with respect to some basis B, and A represents the same operator with respect to a basis B'; that is,

$$C = [T]_B \quad \text{and} \quad A = [T]_{B'}$$

If we let $P = P_{B \to B'}$, then it follows from Formula (12) in Theorem 8.1.2 that

$$[T]_{B'} = P[T]_B P^{-1} \quad \text{or, equivalently,} \quad A = PCP^{-1}$$

Rewriting the last equation as $C = P^{-1}AP$ shows that A and C are similar. ■

SIMILARITY INVARIANTS

There are a number of basic properties of matrices that are shared by similar matrices. For example, if $C = P^{-1}AP$, then

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

which shows that similar matrices have the same determinant.

In general, any property that is shared by similar matrices is said to be a ***similarity invariant***. The following theorem lists some of the most important similarity invariants.

Theorem 8.2.3

(*a*) *Similar matrices have the same determinant.*

(*b*) *Similar matrices have the same rank.*

(*c*) *Similar matrices have the same nullity.*

(*d*) *Similar matrices have the same trace.*

(*e*) *Similar matrices have the same characteristic polynomial and hence have the same eigenvalues with the same algebraic multiplicities.*

We have already proved part (a). We will prove part (e) and leave the proofs of the other three parts as exercises.

***Proof* (e)** We want to prove that if A and C are similar matrices, then

$$\det(\lambda I - C) = \det(\lambda I - A) \tag{3}$$

As a first step we will show that if A and C are similar matrices, then so are $\lambda I - A$ and $\lambda I - C$ for any scalar λ. To see this, suppose that $C = P^{-1}AP$ and write

$$\begin{aligned}\lambda I - C = \lambda I - P^{-1}AP = \lambda P^{-1}P - P^{-1}AP &= P^{-1}(\lambda P - AP)\\ = P^{-1}(\lambda I P - AP) &= P^{-1}(\lambda I - A)P\end{aligned}$$

This shows that $\lambda I - A$ and $\lambda I - C$ are similar, so (3) now follows from part (a). ■

EXAMPLE 1

Similarity

Show that there do not exist bases for R^2 with respect to which the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

represent the same linear operator.

Solution For A and C to represent the same linear operator, the two matrices would have to be similar by Theorem 8.2.2. But this cannot be, since $\operatorname{tr}(A) = 7$ and $\operatorname{tr}(C) = 5$, contradicting the fact that the trace is a similarity invariant. ■

EIGENVECTORS AND EIGENVALUES OF SIMILAR MATRICES

Recall that the solution space of

$$(\lambda_0 I - A)\mathbf{x} = \mathbf{0}$$

is called the *eigenspace* of A corresponding to λ_0. We call the dimension of this solution space the ***geometric multiplicity*** of λ_0.

Do not confuse this with the *algebraic multiplicity* of λ_0, which, as you may recall, is the number of repetitions of the factor $\lambda - \lambda_0$ in the complete factorization of the characteristic polynomial of A.

EXAMPLE 2 Algebraic and Geometric Multiplicities

Find the algebraic and geometric multiplicities of the eigenvalues of

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

Solution Since A is triangular its characteristic polynomial is

$$p(\lambda) = (\lambda - 2)(\lambda - 3)(\lambda - 3) = (\lambda - 2)(\lambda - 3)^2$$ ← det(λI – A)

This implies that the distinct eigenvalues are $\lambda = 2$ and $\lambda = 3$ and that

$\lambda = 2$ has algebraic multiplicity 1

$\lambda = 3$ has algebraic multiplicity 2

One way to find the geometric multiplicities of the eigenvalues is to find bases for the eigenspaces and then determine the dimensions of those spaces from the number of basis vectors.

Let us do this. By definition, the eigenspace corresponding to an eigenvalue λ is the solution space of $(\lambda I - A)\mathbf{x} = \mathbf{0}$, which in this case is

$$\begin{bmatrix} \lambda - 2 & 0 & 0 \\ -1 & \lambda - 3 & 0 \\ 3 & -5 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

If $\lambda = 2$, this system becomes

$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & -5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{5}$$

We leave it for you to show that a general solution of this system is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{8}t \\ -\frac{1}{8}t \\ t \end{bmatrix} = t \begin{bmatrix} \frac{1}{8} \\ -\frac{1}{8} \\ 1 \end{bmatrix} \tag{6}$$

which shows that the eigenspace corresponding to $\lambda = 2$ has dimension 1 and that the column vector on the right side of (6) is a basis for this eigenspace.

Similarly, it follows from (4) that the eigenspace corresponding to $\lambda = 3$ is the solution space of

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 3 & -5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{7}$$

We leave it for you to show that a general solution of this system is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{8}$$

which shows that the eigenspace corresponding to $\lambda = 3$ has dimension 1 and that the column vector on the right side of (8) is a basis for this eigenspace. Since both eigenspaces have dimension 1, we have shown that

$\lambda = 2$ has geometric multiplicity 1

$\lambda = 3$ has geometric multiplicity 1 ■

EXAMPLE 3 Algebraic and Geometric Multiplicities

Find the algebraic and geometric multiplicities of the eigenvalues of

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

Solution We leave it for you to confirm that the characteristic polynomial of A is

$$p(\lambda) = \det(\lambda I - A) = \lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2$$

This implies that the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$ and that

$\lambda = 1$ has algebraic multiplicity 1
$\lambda = 2$ has algebraic multiplicity 2

By definition, the eigenspace corresponding to an eigenvalue λ is the solution space of the system $(\lambda I - A)\mathbf{x} = \mathbf{0}$, which in this case is

$$\begin{bmatrix} \lambda & 0 & 2 \\ -1 & \lambda - 2 & -1 \\ -1 & 0 & \lambda - 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{9}$$

We leave it for you to show that a general solution of this system for $\lambda = 1$ is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2t \\ t \\ t \end{bmatrix} = t\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \tag{10}$$

and that a general solution for $\lambda = 2$ is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -s \\ t \\ s \end{bmatrix} = \begin{bmatrix} -s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} 0 \\ t \\ 0 \end{bmatrix} = s\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \tag{11}$$

This shows that the eigenspace corresponding to $\lambda = 1$ has dimension 1 and that the column vector on the right side of (10) is a basis for this eigenspace, and it shows that the eigenspace corresponding to $\lambda = 2$ has dimension 2 and that the column vectors on the right side of (11) are a basis for this eigenspace. Thus,

$\lambda = 1$ has geometric multiplicity 1

$\lambda = 2$ has geometric multiplicity 2 ■

REMARK It is not essential to find bases for the eigenspaces to determine the geometric multiplicities of the eigenvalues.

For example, to find the dimensions of the eigenspaces in Example 2 we could have calculated the ranks of the coefficient matrices in (5) and (7) by row reduction and then used the relationship rank $+$ nullity $= 3$ to determine the nullities.

The next theorem shows that eigenvalues and their multiplicities are similarity invariants.

Theorem 8.2.4 *Similar matrices have the same eigenvalues and those eigenvalues have the same algebraic and geometric multiplicities for both matrices.*

Proof Let us assume first that A and C are similar matrices. Since similar matrices have the same characteristic polynomial, it follows that A and C have the same eigenvalues with the same algebraic multiplicities. To show that an eigenvalue λ has the same geometric multiplicity for both matrices, we must show that the solution spaces of

$$(\lambda I - A)\mathbf{x} = \mathbf{0} \quad \text{and} \quad (\lambda I - C)\mathbf{x} = \mathbf{0}$$

have the same dimension, or equivalently, that the matrices

$$\lambda I - A \quad \text{and} \quad \lambda I - C \tag{12}$$

have the same nullity. But we showed in the proof of Theorem 8.2.3 that the similarity of A and C implies the similarity of the matrices in (12). Thus, these matrices have the same nullity by part (c) of Theorem 8.2.3. ■

Do not read more into Theorem 8.2.4 than it actually says; the theorem states that similar matrices have the same eigenvalues with the same algebraic and geometric multiplicities, but it does not say that similar matrices have the same eigenspaces.

The following theorem establishes the relationship between the eigenspaces of similar matrices.

Theorem 8.2.5 *Suppose that $C = P^{-1}AP$ and that λ is an eigenvalue of A and C.*

(a) *If $\mathbf{x}$ is an eigenvector of C corresponding to λ, then $P\mathbf{x}$ is an eigenvector of A corresponding to λ.*

(b) *If $\mathbf{x}$ is an eigenvector of A corresponding to λ, then $P^{-1}\mathbf{x}$ is an eigenvector of C corresponding to λ.*

We will prove part (*a*) and leave the proof of part (*b*) as an exercise.

Proof (a) Assume that $\mathbf{x}$ is an eigenvector of C corresponding to λ, so $\mathbf{x} \neq \mathbf{0}$ and $C\mathbf{x} = \lambda\mathbf{x}$. If we substitute $P^{-1}AP$ for C, we obtain

$$P^{-1}AP\mathbf{x} = \lambda\mathbf{x}$$

which we can rewrite as

$$AP\mathbf{x} = P\lambda\mathbf{x} \quad \text{or equivalently,} \quad A(P\mathbf{x}) = \lambda(P\mathbf{x}) \tag{13}$$

Since P is invertible and $\mathbf{x} \neq \mathbf{0}$, it follows that $P\mathbf{x} \neq \mathbf{0}$. Thus, the second equation in (13) implies that $P\mathbf{x}$ is an eigenvector of A corresponding to λ. ■

Recall that since P is invertible, its column vectors are LI.

DIAGONALIZATION

Diagonal matrices play an important role in many applications because, in many respects, they represent the simplest kinds of linear operators. For example, suppose that $T : R^n \to R^n$ is a linear operator whose matrix with respect to a basis $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is

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

If $\mathbf{w}$ is a vector in R^n, and if $\mathbf{x} = [\mathbf{w}]_B$ is the coordinate matrix for $\mathbf{w}$ with respect to B, then

$$D\mathbf{x} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \\ \vdots \\ d_n x_n \end{bmatrix}$$

Thus, multiplying $\mathbf{x}$ by D has the effect of "scaling" each coordinate of $\mathbf{w}$ (with a sign reversal for negative d's).

In particular, the effect of T on a vector that is parallel to one of the basis vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is to contract or dilate that vector (with a possible reversal of direction) (Figure 8.2.1).

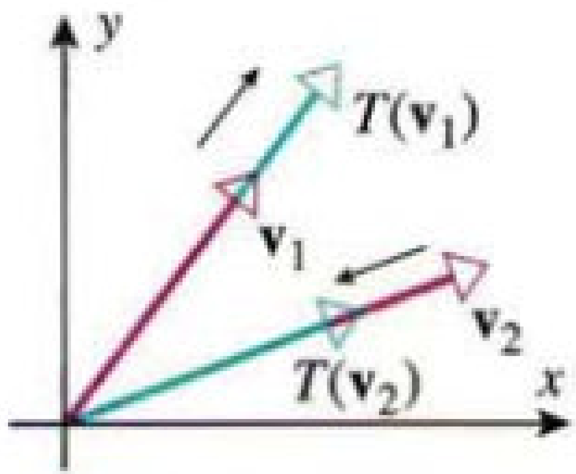

Figure 8.2.1

If T is represented by a diagonal matrix with respect to the basis $B = \{\mathbf{v}_1, \mathbf{v}_2\}$, then T contracts or dilates vectors that are parallel to $\mathbf{v}_1$ or $\mathbf{v}_2$ (with possible reversals of direction).

We will now consider the problem of determining conditions under which a linear operator can be represented by a diagonal matrix with respect to some basis. Since we will generally know the standard matrix for a linear operator, we will consider the following form of this problem.

The Diagonalization Problem Given a square matrix A, does there exist an invertible matrix P for which $P^{-1}AP$ is a diagonal matrix, and if so, how does one find such a P? If such a matrix P exists, then A is said to be ***diagonalizable***, and P is said to ***diagonalize*** A.

Theorem 8.2.6 *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.*

Proof We will show first that if the matrix A is diagonalizable, then it has n linearly independent eigenvectors. The diagonalizability of A implies that there is an invertible matrix P and a diagonal matrix D, say

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{14}$$

such that $P^{-1}AP = D$. If we rewrite this as $AP = PD$ and substitute (14), we obtain

$$AP = PD$$

$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix} \tag{15}$$

Thus, if we denote the column vectors of P by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, then the left side of (15) can be expressed as

$$AP = A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n] = [A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \cdots \quad A\mathbf{p}_n] \tag{16}$$

and the right side of (15) as

$$[\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n] \tag{17}$$

It follows from (16) and (17) that

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \ldots, \quad A\mathbf{p}_n = \lambda_n\mathbf{p}_n$$

and it follows from the invertibility of P that $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are nonzero, so we have shown that $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are eigenvectors of A corresponding to $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively.

Moreover, the invertibility of P also implies that $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are linearly independent (Theorem 7.4.4 applied to P), so the column vectors of P form a set of n linearly independent eigenvectors of A.

Conversely, assume that A has n linearly independent eigenvectors, $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, and that the corresponding eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_n$, so

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \ldots, \quad A\mathbf{p}_n = \lambda_1\mathbf{p}_n$$ ← It is $A\mathbf{p}_n = \lambda_n\mathbf{p}_n$, and not $\lambda_1\mathbf{p}_n$.

If we now form the matrices

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n] = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

then we obtain

$$\begin{aligned} AP = A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n] &= [A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \cdots \quad A\mathbf{p}_n] \\ &= [\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n] \end{aligned}$$

$$AP = A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n] = [A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \cdots \quad A\mathbf{p}_n]$$

$$= [\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n]$$

$$= \begin{bmatrix} \lambda_1 p_{11} & \lambda_2 p_{12} & \cdots & \lambda_n p_{1n} \\ \lambda_1 p_{21} & \lambda_2 p_{22} & \cdots & \lambda_n p_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda_1 p_{n1} & \lambda_2 p_{n2} & \cdots & \lambda_n p_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = PD$$

However, the matrix P is invertible, since its column vectors are linearly independent, so it follows from this computation that $D = P^{-1}AP$, which shows that A is diagonalizable. ■

REMARK Keeping in mind that a set of n linearly independent vectors in R^n must be a basis for R^n, Theorem 8.2.6 is equivalent to saying that an $n \times n$ *matrix* A *is diagonalizable if and only if there is a basis for* R^n *consisting of eigenvectors of* A.

A METHOD FOR DIAGONALIZING A MATRIX

Theorem 8.2.6 guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and its proof provides the following method for diagonalizing A in that case.

Diagonalizing an $n \times n$ Matrix with n Linearly Independent Eigenvectors

Step 1. Find n linearly independent eigenvectors of A, say $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$.

Step 2. Form the matrix $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]$.

Step 3. The matrix $P^{-1}AP$ will be diagonal and will have the eigenvalues corresponding to $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, respectively, as its successive diagonal entries.

EXAMPLE 4 Diagonalizing a Matrix

We showed in Example 3 that the matrix

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

has eigenvalues $\lambda = 1$ and $\lambda = 2$ and that basis vectors for these eigenspaces are

$$\underbrace{\mathbf{p}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}}_{\lambda = 1} \quad \text{and} \quad \underbrace{\mathbf{p}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\lambda = 2}$$

It is a straightforward matter to show that these three vectors are linearly independent, so A is diagonalizable and is diagonalized by

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

As a check, we leave it for you to verify that

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

REMARK There is no preferred order for the columns of a diagonalizing matrix P—the only effect of changing the order of the columns is to change the order in which the eigenvalues appear along the main diagonal of $D = P^{-1}AP$.

For example, had we written the column vectors of P in Example 4 in the order

$$P = [\mathbf{p}_3 \quad \mathbf{p}_1 \quad \mathbf{p}_2] = \begin{bmatrix} 0 & -2 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

then the resulting diagonal matrix would have been

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

EXAMPLE 5 A Matrix That Is Not Diagonalizable

We showed in Example 2 that the matrix

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

has eigenvalues $\lambda = 2$ and $\lambda = 3$ and that bases for the corresponding eigenspaces are

$$\mathbf{p}_1 = \begin{bmatrix} \frac{1}{8} \\ -\frac{1}{8} \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda = 2 \qquad\qquad \lambda = 3$$

These eigenvectors are linearly independent, since they are not scalar multiples of one another, but it is impossible to produce a third linearly independent eigenvector since all other eigenvectors must be scalar multiples of one of these two. Thus, A is not diagonalizable. ∎

LINEAR INDEPENDENCE OF EIGENVECTORS

Theorem 8.2.7 *If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are eigenvectors of a matrix A that correspond to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$, then the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent.*

Proof We will assume that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent and obtain a contradiction. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent, then some vector in this sequence must be a linear combination of predecessors (Theorem 7.1.2).

C r

If we let $\mathbf{v}_{r+1}$ be the *first* vector in the sequence that is a linear combination of predecessors, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are linearly independent, and there exist scalars $c_1, c_2, \ldots, c_k$ such that

$$\mathbf{v}_{r+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r \tag{18}$$

Multiplying both sides of (18) by A and using the fact that $A\mathbf{v}_j = \lambda_j\mathbf{v}_j$ for each j yields

$$\lambda_{r+1}\mathbf{v}_{r+1} = c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_r\lambda_r\mathbf{v}_r \tag{19}$$

Now multiplying (18) by λ_{r+1} and subtracting from (19) yields

$$\mathbf{0} = c_1(\lambda_1 - \lambda_{r+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{r+1})\mathbf{v}_2 + \cdots + c_r(\lambda_r - \lambda_{r+1})\mathbf{v}_r \tag{20}$$

Since $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ are linearly independent, it follows that all of the coefficients on the right side of (20) are zero. However, the eigenvalues are all distinct, so it must be that

$$c_1 = c_2 = \cdots = c_r = 0$$

But this and (18) imply that $\mathbf{v}_{r+1} = \mathbf{0}$, which is impossible since eigenvectors are nonzero. Thus, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ must be linearly independent. ■

REMARK If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of a matrix A, then Theorem 8.2.7 tells us that a linearly independent set is produced by choosing one eigenvector from each of the corresponding eigenspaces.

More generally, it can be proved that if one chooses linearly independent *sets* of eigenvectors from distinct eigenspaces and combines them into a single set, then that combined set will be linearly independent.

For example, for the matrix A in Example 4 we had an eigenvector $\mathbf{p}_1$ from the eigenspace corresponding to $\lambda = 1$ and two linearly independent eigenvectors $\mathbf{p}_2$ and $\mathbf{p}_3$ from the eigenspace corresponding to $\lambda = 2$, so we are guaranteed without any computations that $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is a linearly independent set.

It follows from Theorems 8.2.6 and 8.2.7 that an $n \times n$ matrix with n distinct real eigenvalues must be diagonalizable, since we can produce a set of n linearly independent eigenvectors by choosing one eigenvector from each eigenspace.

Theorem 8.2.8 *An $n \times n$ matrix with n distinct real eigenvalues is diagonalizable.*

EXAMPLE 6 Diagonalizable Matrix with Distinct Eigenvalues

The 3×3 matrix

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

is diagonalizable, since it has three distinct eigenvalues, $\lambda = 2$, $\lambda = 3$, and $\lambda = 4$. ■

The converse of Theorem 8.2.8 is *false*; that is, it is possible for an $n \times n$ matrix to be diagonalizable without having n distinct eigenvalues. For example, the matrix A in Example 4 was seen to be diagonalizable, even though it had only two distinct eigenvalues, $\lambda = 1$ and $\lambda = 2$.

The diagonalizability was a consequence of the fact that the eigenspaces had dimensions 1 and 2, respectively, thereby allowing us to produce three linearly independent eigenvectors.

Thus, we see that the key to diagonalizability rests with the dimensions of the eigenspaces.

Theorem 8.2.9 *An $n \times n$ matrix A is diagonalizable if and only if the sum of the geometric multiplicities of its eigenvalues is n.*

Proof Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of A, let $E_1, E_2, \ldots, E_k$ denote the corresponding eigenspaces, let $B_1, B_2, \ldots, B_k$ be any bases for these eigenspaces, and let B be the linearly independent set that results when the bases are merged into a single set (i.e., B is the union of the bases).

If the sum of the geometric multiplicities is n, then B is a set of n linearly independent eigenvectors, so A is diagonalizable by Theorem 8.2.6.

The proof of the converse is left for more advanced courses. ■

EXAMPLE 7 Diagonalizability and Geometric Multiplicity

We showed in Example 2 that the matrix

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

has eigenvalues $\lambda = 2$ and $\lambda = 3$, both with geometric multiplicity 1. Since the sum of the geometric multiplicities is less than 3, the matrix is not diagonalizable. Also, we showed in Example 3 that the matrix

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

has eigenvalues $\lambda = 1$ and $\lambda = 2$ with geometric multiplicities 1 and 2, respectively. Since the sum of the geometric multiplicities is 3, the matrix is diagonalizable (see Example 4). ■

RELATIONSHIP BETWEEN ALGEBRAIC AND GEOMETRIC MULTIPLICITY

A full excursion into the study of diagonalizability will be left for more advanced courses, but we will mention one result that is important for a full understanding of the diagonalizability question:

➡ It can be proved that the geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity.

For example, if the characteristic polynomial of some 6×6 matrix A is

$$p(\lambda) = (\lambda - 3)(\lambda - 5)^2(\lambda - 6)^3$$

then, depending on the particular matrix A, the eigenspace corresponding to $\lambda = 6$ might have dimension 1, 2, or 3, the eigenspace corresponding to $\lambda = 5$ might have dimension 1 or 2, and the eigenspace corresponding to $\lambda = 3$ must have dimension 1.

For the matrix A to be diagonalizable there would have to be six linearly independent eigenvectors, and this will only occur if the geometric and algebraic multiplicities are the same;

that is, if the eigenspace corresponding to $\lambda = 6$ has dimension 3, the eigenspace corresponding to $\lambda = 5$ has dimension 2, and the eigenspace corresponding to $\lambda = 3$ has dimension 1. The following theorem, whose proof is outlined in the exercises, summarizes these ideas.

Theorem 8.2.10 *If A is a square matrix, then:*

(*a*) *The geometric multiplicity of an eigenvalue of A is less than or equal to its algebraic multiplicity.*

(*b*) *A is diagonalizable if and only if the geometric multiplicity of each eigenvalue of A is the same as its algebraic multiplicity.*