39. Prove that the intersection of any two distinct eigenspaces of a matrix A is $\{0\}$.

True-False Exercises

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

(a) If A is a square matrix and $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero scalar λ, then $\mathbf{x}$ is an eigenvector of A.

(b) If λ is an eigenvalue of a matrix A, then the linear system $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) If the characteristic polynomial of a matrix A is $p(\lambda) = \lambda^2 + 1$, then A is invertible.

(d) If λ is an eigenvalue of a matrix A, then the eigenspace of A corresponding to λ is the set of eigenvectors of A corresponding to λ.

(e) The eigenvalues of a matrix A are the same as the eigenvalues of the reduced row echelon form of A.

(f) If 0 is an eigenvalue of a matrix A, then the set of columns of A is linearly independent.

Working with Technology

T1. For the given matrix A, find the characteristic polynomial and the eigenvalues, and then use the method of Example 7 to find bases for the eigenspaces.

$$A = \begin{bmatrix} -8 & 33 & 38 & 173 & -30 \\ 0 & 0 & -1 & -4 & 0 \\ 0 & 0 & -5 & -25 & 1 \\ 0 & 0 & 1 & 5 & 0 \\ 4 & -16 & -19 & -86 & 15 \end{bmatrix}$$

T2. The Cayley–Hamilton Theorem states that every square matrix satisfies its characteristic equation; that is, if A is an $n \times n$ matrix whose characteristic equation is

$$\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0$$

then $A^n + c_1A^{n-1} + \cdots + c_n = 0$.

(a) Verify the Cayley–Hamilton Theorem for the matrix

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

(b) Use the result in Exercise 28 to prove the Cayley–Hamilton Theorem for 2×2 matrices.

5.2 Diagonalization

In this section we will be concerned with the problem of finding a basis for R^n that consists of eigenvectors of an $n \times n$ matrix A. Such bases can be used to study geometric properties of A and to simplify various numerical computations. These bases are also of physical significance in a wide variety of applications, some of which will be considered later in this text.

The Matrix Diagonalization Problem

Products of the form $P^{-1}AP$ in which A and P are $n \times n$ matrices and P is invertible will be our main topic of study in this section. There are various ways to think about such products, one of which is to view them as transformations

$$A \rightarrow P^{-1}AP$$

in which the matrix A is mapped into the matrix $P^{-1}AP$. These are called ***similarity transformations***. Such transformations are important because they preserve many properties of the matrix A. For example, if we let $B = P^{-1}AP$, then A and B have the same determinant since

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

$$= \frac{1}{\det(P)}\det(A)\det(P) = \det(A)$$

In general, any property that is preserved by a similarity transformation is called a ***similarity invariant*** and is said to be ***invariant under similarity***. Table 1 lists the most important similarity invariants. The proofs of some of these are given as exercises.

Table 1 Similarity Invariants

Property	Description
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A (and hence of $P^{-1}AP$) then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ have the same dimension.

We will find the following terminology useful in our study of similarity transformations.

DEFINITION 1 If A and B are square matrices, then we say that ***B is similar to A*** if there is an invertible matrix P such that $B = P^{-1}AP$.

Note that if B is similar to A, then it is also true that A is similar to B since we can express A as $A = Q^{-1}BQ$ by taking $Q = P^{-1}$. This bei[illegible] say that A and B are ***similar matrices*** if either is similar to [illegible]

Because diagonal matrices have such a simple forn[illegible]er a given $n \times n$ matrix A is similar to a matrix of this ty[illegible] be the case, and should we be able to actually find a diagona[illegible]ilar to A, then we would be able to ascertain many of the similarity invariant properties of A directly from the diagonal entries of D. For example, the diagonal entries of D will be the eigenvalues of A (Theorem 5.1.2), and the product of the diagonal entries of D will be the determinant of A (Theorem 2.1.2). This leads us to introduce the following terminology.

DEFINITION 2 A square matrix A is said to be ***diagonalizable*** if it is similar to some diagonal matrix; that is, if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal. In this case the matrix P is said to ***diagonalize*** A.

The following theorem and the ideas used in its proof will provide us with a roadmap for devising a technique for determining whether a matrix is diagonalizable and, if so, for finding a matrix P that will perform the diagonalization.

e for
latrix

Theorem 5.2.1 guarantees that an $n \times n$ matrix A with n linearly independent eigenvectors is diagonalizable, and the proof of that theorem together with Theorem 5.2.2 suggests the following procedure for diagonalizing A.

A Procedure for Diagonalizing an $n \times n$ Matrix

Step 1. Determine first whether the matrix is actually diagonalizable by searching for n linearly independent eigenvectors. One way to do this is to find a basis for each eigenspace and count the total number of vectors obtained. If there is a total of n vectors, then the matrix is diagonalizable, and if the total is less than n, then it is not.

Step 2. If you ascertained that the matrix is diagonalizable, then form the matrix $P = [\mathbf{p}_1 \;\; \mathbf{p}_2 \;\; \cdots \;\; \mathbf{p}_n]$ whose column vectors are the n basis vectors you obtained in Step 1.

Step 3. $P^{-1}AP$ will be a diagonal matrix whose successive diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ that correspond to the successive columns of P.

▶ EXAMPLE 1 Finding a Matrix P That Diagonalizes a Matrix A

Find a matrix P that diagonalizes

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

Solution In Example 7 of the preceding section we found the characteristic equation of A to be

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

and we found the following bases for the eigenspaces:

$$\lambda = 2: \quad \mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 1: \quad \mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

There are three basis vectors in total, so the matrix

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

diagonalizes A. As a check, you should verify that

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

In general, there is no preferred order for the columns of P. Since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P, changing the order of the columns of P just changes the order of the eigenvalues on the diagonal of $P^{-1}AP$. Thus, had we written

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

in the preceding example, we would have obtained

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

▶ EXAMPLE 2 A Matrix That Is Not Diagonalizable

Show that the following matrix is not diagonalizable:

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

Solution The characteristic polynomial of A is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & 0 \\ 3 & -5 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2$$

so the characteristic equation is

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

and the distinct eigenvalues of A are $\lambda = 1$ and $\lambda = 2$. We leave it for you to show that bases for the eigenspaces are

$$\lambda = 1: \quad \mathbf{p}_1 = \begin{bmatrix} \frac{1}{8} \\ -\frac{1}{8} \\ 1 \end{bmatrix}; \qquad \lambda = 2: \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Since A is a 3×3 matrix and there are only two basis vectors in total, A is not diagonalizable.

Alternative Solution If you are concerned only in determining whether a matrix is diagonalizable and not with actually finding a diagonalizing matrix P, then it is not necessary to compute bases for the eigenspaces—it suffices to find the dimensions of the eigenspaces. For this example, the eigenspace corresponding to $\lambda = 1$ is the solution space of the system

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

Since the coefficient matrix has rank 2 (verify), the nullity of this matrix is 1 by Theorem 4.8.2, and hence the eigenspace corresponding to $\lambda = 1$ is one-dimensional.

The eigenspace corresponding to $\lambda = 2$ is the solution space of the system

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

This coefficient matrix also has rank 2 and nullity 1 (verify), so the eigenspace corresponding to $\lambda = 2$ is also one-dimensional. Since the eigenspaces produce a total of two basis vectors, and since three are needed, the matrix A is not diagonalizable.

EXAMPLE 3 Recognizing Diagonalizability

We saw in Example 3 of the preceding section that

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}$$

has three distinct eigenvalues: $\lambda = 4$, $\lambda = 2 + \sqrt{3}$, and $\lambda = 2 - \sqrt{3}$. Therefore, A is diagonalizable and

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

for some invertible matrix P. If needed, the matrix P can be found using the method shown in Example 1 of this section.

EXAMPLE 4 Diagonalizability of Triangular Matrices

From Theorem 5.1.2, the eigenvalues of a triangular matrix are the entries on its main diagonal. Thus, a triangular matrix with distinct entries on the main diagonal is diagonalizable. For example,

$$A = \begin{bmatrix} -1 & 2 & 4 & 0 \\ 0 & 3 & 1 & 7 \\ 0 & 0 & 5 & 8 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

is a diagonalizable matrix with eigenvalues $\lambda_1 = -1$, $\lambda_2 = 3$, $\lambda_3 = 5$, $\lambda_4 = -2$.

vers of a Matrix

Since there are many applications in which it is necessary to compute high powers of a square matrix A, we will now turn our attention to that important problem. As we will see, the most efficient way to compute A^k, particularly for large values of k, is to first diagonalize A. But because diagonalizing a matrix A involves finding its eigenvalues and eigenvectors, we will need to know how these quantities are related to those of A^k. As an illustration, suppose that λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector. Then

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

which shows not only that λ^2 is a eigenvalue of A^2 but that $\mathbf{x}$ is a corresponding eigenvector. In general, we have the following result.

ability is
n Theo-

THEOREM 5.2.3 *If k is a positive integer, λ is an eigenvalue of a matrix A, and $\mathbf{x}$ is a corresponding eigenvector, then λ^k is an eigenvalue of A^k and $\mathbf{x}$ is a corresponding eigenvector.*

EXAMPLE 5 Eigenvalues and Eigenvectors of Matrix Powers

In Example 2 we found the eigenvalues and corresponding eigenvectors of the matrix

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

Do the same for A^7.

Solution We know from Example 2 that the eigenvalues of A are $\lambda = 1$ and $\lambda = 2$, so the eigenvalues of A^7 are $\lambda = 1^7 = 1$ and $\lambda = 2^7 = 128$. The eigenvectors $\mathbf{p}_1$ and $\mathbf{p}_2$ obtained in Example 1 corresponding to the eigenvalues $\lambda = 1$ and $\lambda = 2$ of A are also the eigenvectors corresponding to the eigenvalues $\lambda = 1$ and $\lambda = 128$ of A^7. ◀

ers of a Matrix

The problem of computing powers of a matrix is greatly simplified when the matrix is diagonalizable. To see why this is so, suppose that A is a diagonalizable $n \times n$ matrix, that P diagonalizes A, and that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = D$$

Squaring both sides of this equation yields

$$(P^{-1}AP)^2 = \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^2 \end{bmatrix} = D^2$$

We can rewrite the left side of this equation as

$$(P^{-1}AP)^2 = P^{-1}APP^{-1}AP = P^{-1}AIAP = P^{-1}A^2P$$

from which we obtain the relationship $P^{-1}A^2P = D^2$. More generally, if k is a positive integer, then a similar computation will show that

322/802

$$P^{-1}A^kP = D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}$$

which we can rewrite as

$$A^k = PD^kP^{-1} = P \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix} P^{-1} \tag{3}$$

at rais- atrix A ver has eigen-

▶ EXAMPLE 6 Powers of a Matrix

Use (3) to find A^{13}, where

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

Solution We showed in Example 1 that the matrix A is diagonalized by

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

and that

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

Thus, it follows from (3) that

$$A^{13} = PD^{13}P^{-1} = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2^{13} & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 1^{13} \end{bmatrix} \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ -1 & 0 & -1 \end{bmatrix} \tag{4}$$

$$= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix}$$ ◀

Remark With the method in the preceding example, most of the work is in diagonalizing A. Once that work is done, it can be used to compute any power of A. Thus, to compute A^{1000} we need only change the exponents from 13 to 1000 in (4).

Algebraic Multiplicity

Theorem 5.2.2(b) does not completely settle the diagonalizability question since it only guarantees that a square matrix with n distinct eigenvalues is diagonalizable; it does not preclude the possibility that there may exist diagonalizable matrices with fewer than n distinct eigenvalues. The following example shows that this is indeed the case.

▶ EXAMPLE 7 The Converse of Theorem 5.2.2(b) Is False

Consider the matrices

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

It follows from Theorem 5.1.2 that both of these matrices have only one distinct eigenvalue, namely $\lambda = 1$, and hence only one eigenspace. We leave [illegible] you to solve the characteristic equations

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

with $\lambda = 1$ and show that for I the eigenspace is three-dimensional (all of R^3) and for J it is one-dimensional, consisting of all scalar multiples of

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

This shows that the converse of Theorem 5.2.2(b) is false, since we have produced two 3×3 matrices with fewer than three distinct eigenvalues, one of which is diagonalizable and the other of which is not. ◀

A full excursion into the study of diagonalizability is left for more advanced courses, but we will touch on one theorem that is important for a fuller understanding of diagonalizability. It can be proved that if λ_0 is an eigenvalue of A, then the dimension of the eigenspace corresponding to λ_0 cannot exceed the number of times that $\lambda - \lambda_0$ appears as a factor of the characteristic polynomial of A. For example, in Examples 1 and 2 the characteristic polynomial is

$$(\lambda - 1)(\lambda - 2)^2$$

Thus, the eigenspace corresponding to $\lambda = 1$ is at most (hence exactly) one-dimensional, and the eigenspace corresponding to $\lambda = 2$ is at most two-dimensional. In Example 1