Diagonalization

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 of the form

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

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.

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

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.

Theorem 5.2.1

If A is an $n \times n$ matrix, the following statements are equivalent.

(a) A is diagonalizable.

(b) A has n linearly independent eigenvectors.

Theorem 5.2.2

(a) If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of a matrix A, and if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are corresponding eigenvectors, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a linearly independent set.

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

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

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

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 4.9.2, 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 $\lambda_1 = -1$, $\lambda_2 = 3$, $\lambda_3 = 5$, $\lambda_4 = -2$.

Question:

9. Let

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

a. Find the eigenvalues of A.

b. For each eigenvalue λ, find the rank of the matrix $\lambda I - A$.

c. Is A diagonalizable? Justify your conclusion.

Solution:

(a) Cofactor expansion along the second column yields

$$\det(\lambda I - A) = \begin{vmatrix} \lambda-4 & 0 & -1 \\ -2 & \lambda-3 & -2 \\ -1 & 0 & \lambda-4 \end{vmatrix} = (\lambda-3)\begin{vmatrix} \lambda-4 & -1 \\ -1 & \lambda-4 \end{vmatrix} = (\lambda-3)\left[(\lambda-4)^2-1\right] = (\lambda-3)^2(\lambda-5)$$

therefore A has eigenvalues 3 (with algebraic multiplicity 2) and 5.

(b) The reduced row echelon form of $3I - A$ is $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, consequently $\operatorname{rank}(3I - A) = 1$.

The reduced row echelon form of $5I - A$ is $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$, consequently $\operatorname{rank}(5I - A) = 2$.

(c) Based on part (b), the geometric multiplicities of the eigenvalues $\lambda = 3$ and $\lambda = 5$ are $3-1=2$ and $3-2=1$, respectively. Since these are equal to the corresponding algebraic multiplicities, by Theorem 5.2.4(b) A is diagonalizable.

Question:

In Exercises 11–14, find the geometric and algebraic multiplicity of each eigenvalue of the matrix A, and determine whether A is diagonalizable. If A is diagonalizable, then find a matrix P that diagonalizes A, and find $P^{-1}AP$.

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

Solution:

Cofactor expansion along the second row yields

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

$$=(-3)\left[(-4)(\lambda-3)-(2)(-1)\right]+(\lambda-4)\left[(\lambda+1)(\lambda-3)-(2)(3)\right]=\lambda^3-6\lambda^2+11\lambda-6.$$

Following the procedure described in Example 3 of Section 5.1, we determine that the only possibilities for integer solutions of the characteristic equation are ±1, ±2, ±3, and ±6.

Since $\det(1I - A) = 0$, $\lambda - 1$ must be a factor of the characteristic polynomial. Dividing $\lambda - 1$ into $\lambda^3-6\lambda^2+11\lambda-6$ leads to $\det(\lambda I - A) = (\lambda-1)(\lambda^2-5\lambda+6) = (\lambda-1)(\lambda-2)(\lambda-3)$.

We conclude that the eigenvalues are 1, 2, and 3 - each of them has the algebraic multiplicity 1.

The reduced row echelon form of $1I - A$ is $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ so that the eigenspace corresponding to $\lambda_1 = 1$ contains vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 = t$, $x_2 = t$, $x_3 = t$. A vector $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ forms a basis for this eigenspace. This eigenvalue has geometric multiplicity 1.

The reduced row echelon form of $2I - A$ is $\begin{bmatrix} 1 & 0 & -\frac{2}{3} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ so that the eigenspace corresponding to $\lambda_2 = 2$ contains vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 = \frac{2}{3}t$, $x_2 = t$, $x_3 = t$. A vector $\mathbf{p}_2 = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$ forms a basis for this eigenspace. This eigenvalue has geometric multiplicity 1.

The reduced row echelon form of $3I - A$ is $\begin{bmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & -\frac{3}{4} \\ 0 & 0 & 0 \end{bmatrix}$ so that the eigenspace corresponding to $\lambda_3 = 3$ contains vectors $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ where $x_1 = \frac{1}{4}t$, $x_2 = \frac{3}{4}t$, $x_3 = t$. A vector $\mathbf{p}_3 = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ forms a basis for this eigenspace. This eigenvalue has geometric multiplicity 1.

Since for each eigenvalue the geometric multiplicity matches the algebraic multiplicity, by Theorem 5.2.4(b) A is diagonalizable.

We form a matrix P using the column vectors p_1, p_2, and p_3: $P=\begin{bmatrix}1&2&1\\1&3&3\\1&3&4\end{bmatrix}$. (Note that this answer is not unique. Any nonzero multiples of these columns would also form a valid matrix P. Furthermore, the columns can be interchanged.)

$$P^{-1}AP=\begin{bmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{bmatrix}=\begin{bmatrix}1&0&0\\0&2&0\\0&0&3\end{bmatrix}.$$

Eigenvalues of Powers of a Matrix

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.

Computing Powers 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

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

which we can rewrite as

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

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

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

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

There is some terminology that is related to these ideas. If λ_0 is an eigenvalue of an $n \times n$ matrix A, then the dimension of the eigenspace corresponding to λ_0 is called the ***geometric multiplicity*** of λ_0, and the number of times that $\lambda - \lambda_0$ appears as a factor in the characteristic polynomial of A is called the ***algebraic multiplicity*** of λ_0. The following theorem, which we state without proof, summarizes the preceding discussion.

Theorem 5.2.4

Geometric and Algebraic Multiplicity

If A is a square matrix, then:

(*a*) For every eigenvalue of A, the geometric multiplicity is less than or equal to the algebraic multiplicity.

(*b*) A is diagonalizable if and only if its characteristic polynomial can be expressed as a product of linear factors, and the geometric multiplicity of every eigenvalue is equal to the algebraic multiplicity.

Question:

Let

$$A = \begin{bmatrix} 1 & -2 & 8 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Confirm that P diagonalizes A, and then compute each of the following powers of A.

a. A^{1000} **b.** A^{-1000} **c.** A^{2301} **d.** A^{-2301}

Solution:

After calculating $P^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix}$, we verify that

$$P^{-1}AP = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix}\begin{bmatrix} 1 & -2 & 8 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = D \text{ is a diagonal matrix therefore}$$

P diagonalizes A.

(a) $$A^{1000} = PD^{1000}P^{-1} = \begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} (-1)^{1000} & 0 & 0 \\ 0 & (-1)^{1000} & 0 \\ 0 & 0 & 1^{1000} \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(b) $$A^{-1000} = PD^{-1000}P^{-1} = \begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} (-1)^{-1000} & 0 & 0 \\ 0 & (-1)^{-1000} & 0 \\ 0 & 0 & 1^{-1000} \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(c) $$A^{2301} = PD^{2301}P^{-1} = \begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} (-1)^{2301} & 0 & 0 \\ 0 & (-1)^{2301} & 0 \\ 0 & 0 & 1^{2301} \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 8 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(d) $$A^{-2301} = PD^{-2301}P^{-1} = \begin{bmatrix} 1 & -4 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} (-1)^{-2301} & 0 & 0 \\ 0 & (-1)^{-2301} & 0 \\ 0 & 0 & 1^{-2301} \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 8 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$