Chapter 5: Eigenvalues and Eigenvectors

Based on Elementary Linear Algebra (11th Edition)

5.1 Eigenvalues and Eigenvectors

Learning Objectives:

- Define eigenvalues and eigenvectors.
- Understand the characteristic equation.
- Compute eigenvalues for 2×2 and 3×3 matrices.

5.1 Eigenvalues and Eigenvectors

DEFINITION 1 If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an *eigenvector* of A (or of the matrix operator T_A) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is,

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar λ. The scalar λ is called an *eigenvalue* of A (or of T_A), and $\mathbf{x}$ is said to be an *eigenvector corresponding to* λ.

5.1 Eigenvalues and Eigenvectors

THEOREM 5.1.1 *If A is an $n \times n$ matrix, then λ is an eigenvalue of A if and only if it satisfies the equation*

$$\det(\lambda I - A) = 0 \tag{1}$$

This is called the ***characteristic equation*** *of A.*

▶ EXAMPLE 2 Finding Eigenvalues

In Example 1 we observed that $\lambda = 3$ is an eigenvalue of the matrix

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

but we did not explain how we found it. Use the characteristic equation to find all eigenvalues of this matrix.

5.1 Eigenvalues and Eigenvectors

▶ **EXAMPLE 1 Eigenvector of a 2 × 2 Matrix**

The vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

corresponding to the eigenvalue $\lambda = 3$, since

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$

Geometrically, multiplication by A has stretched the vector $\mathbf{x}$ by a factor of 3 (Figure 5.1.2). ◀

5.1 Eigenvalues and Eigenvectors

EXAMPLE 3 Eigenvalues of a 3 x 3 Matrix

Find the eigenvalues of

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

Solution The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4$$

The eigenvalues of A must therefore satisfy the cubic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$$

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0$$

Thus, the remaining solutions of (5) satisfy the quadratic equation

$$\lambda^2 - 4\lambda + 1 = 0$$

which can be solved by the quadratic formula. Thus, the eigenvalues of A are

$$\lambda = 4, \quad \lambda = 2 + \sqrt{3}, \quad \text{and} \quad \lambda = 2 - \sqrt{3}$$

5.1 Eigenvalues and Eigenvectors

Solution It follows from Formula (1) that the eigenvalues of A are the solutions of the equation $\det(\lambda I - A) = 0$, which we can write as

$$\begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

from which we obtain

$$(\lambda - 3)(\lambda + 1) = 0 \tag{2}$$

This shows that the eigenvalues of A are $\lambda = 3$ and $\lambda = -1$. Thus, in addition to the eigenvalue $\lambda = 3$ noted in Example 1, we have discovered a second eigenvalue $\lambda = -1$. ◀

When the determinant $\det(\lambda I - A)$ in (1) is expanded, the characteristic equation of A takes the form

$$\lambda^n + c_1\lambda^{n-1} + \cdots + c_n = 0 \tag{3}$$

where the left side of this equation is a polynomial of degree n in which the coefficient of λ^n is 1 (Exercise 37). The polynomial

$$p(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_n \tag{4}$$

is called the ***characteristic polynomial*** of A. For example, it follows from (2) that the characteristic polynomial of the 2×2 matrix in Example 2 is

$$p(\lambda) = (\lambda - 3)(\lambda + 1) = \lambda^2 - 2\lambda - 3$$

5.1 Eigenvalues and Eigenvectors

EXAMPLE 6 Bases for Eigenspaces

Find bases for the eigenspaces of the matrix

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

Solution The characteristic equation of A is

$$\begin{vmatrix} \lambda + 1 & -3 \\ -2 & \lambda \end{vmatrix} = \lambda(\lambda + 1) - 6 = (\lambda - 2)(\lambda + 3) = 0$$

so the eigenvalues of A are $\lambda = 2$ and $\lambda = -3$. Thus, there are two eigenspaces of A, one for each eigenvalue.

By definition,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an eigenvector of A corresponding to an eigenvalue λ if and only if $(\lambda I - A)\mathbf{x} = \mathbf{0}$, that is,

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

In the case where $\lambda = 2$ this equation becomes

$$\begin{bmatrix} 3 & -3 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

whose general solution is

$$x_1 = t, \quad x_2 = t$$

5.1 Eigenvalues and Eigenvectors

whose general solution is

$$x_1 = t, \quad x_2 = t$$

(verify). Since this can be written in matrix form as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

it follows that

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 2$. We leave it for you to follow the pattern of these computations and show that

$$\begin{bmatrix} -\frac{3}{2} \\ 1 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = -3$. ◀

Figure 5.1.3 illustrates the geometric effect of multiplication by the matrix A in Example 6. The eigenspace corresponding to $\lambda = 2$ is the line L_1 through the origin and the point $(1, 1)$, and the eigenspace corresponding to $\lambda = 3$ is the line L_2 through the origin and the point $(-\frac{3}{2}, 1)$. As indicated in the figure, multiplication by A maps each vector in L_1 back into L_1, scaling it by a factor of 2, and it maps each vector in L_2 back into L_2, scaling it by a factor of -3.

Definition:

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.

5.2 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

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

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

$$\begin{aligned}\det(B) &= \det(P^{-1}AP) = \det(P^{-1})\det(A)\det(P)\\ &= \frac{1}{\det(P)}\det(A)\det(P) = \det(A)\end{aligned}$$

5.2 Diagonalization

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

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.*

Proof (*a*) $\Rightarrow$ (*b*) Since A is assumed to be diagonalizable, it follows that there exist an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ or, equivalently,

$$AP = PD \tag{1}$$

If we denote the column vectors of P by $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$, and if we assume that the diagonal entries of D are $\lambda_1, \lambda_2, \dots, \lambda_n$, then by Formula (6) of Section 1.3 the left side of (1) can be expressed as

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

and, as noted in the comment following Example 1 of Section 1.7, the right side of (1) can be expressed as

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

Thus, it follows from (1) that

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \dots, \quad A\mathbf{p}_n = \lambda_n\mathbf{p}_n \tag{2}$$

Since P is invertible, we know from Theorem 5.1.5 that its column vectors $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$ are linearly independent (and hence nonzero). Thus, it follows from (2) that these n column vectors are eigenvectors of A.

Proof (*b*) $\Rightarrow$ (*a*) Assume that A has n linearly independent eigenvectors, $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n$, and that $\lambda_1, \lambda_2, \dots, \lambda_n$ are the corresponding eigenvalues. If we let

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

and if we let D be the diagonal matrix that has $\lambda_1, \lambda_2, \dots, \lambda_n$ as its successive diagonal entries, then

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

Proof

5.2 Diagonalization

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

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

5.2 Diagonalization

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

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

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

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

5.2 Diagonalization

▶ **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.

5.2 Diagonalization

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

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

▶ 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

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

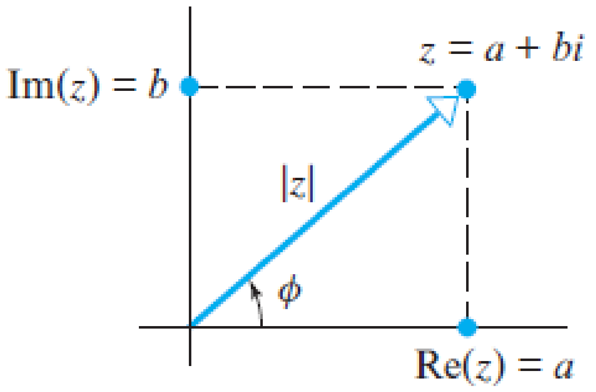
$$= \begin{bmatrix} -8190 & 0 & -16382 \\ 8191 & 8192 & 8191 \\ 8191 & 0 & 16383 \end{bmatrix} ◀$$

5.3 Complex Vector Spaces

Review of Complex Numbers

Recall that if $z = a + bi$ is a complex number, then:

- $\operatorname{Re}(z) = a$ and $\operatorname{Im}(z) = b$ are called the ***real part*** of z and the ***imaginary part*** of z, respectively,
- $|z| = \sqrt{a^2 + b^2}$ is called the ***modulus*** (or ***absolute value***) of z,
- $\overline{z} = a - bi$ is called the ***complex conjugate*** of z,



▲ Figure 5.3.1

- $z\overline{z} = a^2 + b^2 = |z|^2$,
- the angle ϕ in Figure 5.3.1 is called an ***argument*** of z,
- $\operatorname{Re}(z) = |z|\cos\phi$
- $\operatorname{Im}(z) = |z|\sin\phi$
- $z = |z|(\cos\phi + i\sin\phi)$ is called the ***polar form*** of z.

5.3 Complex Vector Spaces

Complex Eigenvalues

In Formula (3) of Section 5.1 we observed that the characteristic equation of a general $n \times n$ matrix A has the form

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

in which the highest power of λ has a coefficient of 1. Up to now we have limited our discussion to matrices in which the solutions of (1) are real numbers. However, it is possible for the characteristic equation of a matrix A with real entries to have imaginary solutions; for example, the characteristic equation of the matrix

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

is

$$\begin{vmatrix} \lambda + 2 & 1 \\ -5 & \lambda - 2 \end{vmatrix} = \lambda^2 + 1 = 0$$

which has the imaginary solutions $\lambda = i$ and $\lambda = -i$. To deal with this case we will need to explore the notion of a complex vector space and some related ideas.

5.3 Complex Vector Spaces

As you might expect, if A is a complex matrix, then A and $\overline{A}$ can be expressed in terms of $\mathrm{Re}(A)$ and $\mathrm{Im}(A)$ as

$$A = \mathrm{Re}(A) + i\,\mathrm{Im}(A)$$
$$\overline{A} = \mathrm{Re}(A) - i\,\mathrm{Im}(A)$$

▶ **EXAMPLE 1 Real and Imaginary Parts of Vectors and Matrices**

Let

$$\mathbf{v} = (3+i, -2i, 5) \quad \text{and} \quad A = \begin{bmatrix} 1+i & -i \\ 4 & 6-2i \end{bmatrix}$$

Then

$$\overline{\mathbf{v}} = (3-i, 2i, 5), \quad \mathrm{Re}(\mathbf{v}) = (3, 0, 5), \quad \mathrm{Im}(\mathbf{v}) = (1, -2, 0)$$

$$\overline{A} = \begin{bmatrix} 1-i & i \\ 4 & 6+2i \end{bmatrix}, \quad \mathrm{Re}(A) = \begin{bmatrix} 1 & 0 \\ 4 & 6 \end{bmatrix}, \quad \mathrm{Im}(A) = \begin{bmatrix} 1 & -1 \\ 0 & -2 \end{bmatrix}$$

$$\det(A) = \begin{vmatrix} 1+i & -i \\ 4 & 6-2i \end{vmatrix} = (1+i)(6-2i) - (-i)(4) = 8 + 8i \quad ◀$$

Complex Eigen

Symmetric Matrices Have Real Eigenvalues

Our next result, which is concerned with the eigenvalues of real symmetric matrices, is important in a wide variety of applications. The key to its proof is to think of a real symmetric matrix as a complex matrix whose entries have an imaginary part of zero.

THEOREM 5.3.6 *If A is a real symmetric matrix, then A has real eigenvalues.*

Proof Suppose that λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, where we allow for the possibility that λ is complex and $\mathbf{x}$ is in C^n. Thus,

$$A\mathbf{x} = \lambda\mathbf{x}$$

where $\mathbf{x} \neq \mathbf{0}$. If we multiply both sides of this equation by $\overline{\mathbf{x}}^T$ and use the fact that

$$\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T(\lambda\mathbf{x}) = \lambda(\overline{\mathbf{x}}^T\mathbf{x}) = \lambda(\mathbf{x}\cdot\mathbf{x}) = \lambda\|\mathbf{x}\|^2$$

then we obtain

$$\lambda = \frac{\overline{\mathbf{x}}^T A\mathbf{x}}{\|\mathbf{x}\|^2}$$

Since the denominator in this expression is real, we can prove that λ is real by showing that

$$\overline{\overline{\mathbf{x}}^T A\mathbf{x}} = \overline{\mathbf{x}}^T A\mathbf{x} \tag{14}$$

But A is symmetric and has real entries, so it follows from the second equality in (5) and properties of the conjugate that

$$\overline{\overline{\mathbf{x}}^T A\mathbf{x}} = \overline{\overline{\mathbf{x}}}^T\,\overline{A\mathbf{x}} = \mathbf{x}^T\,\overline{A\mathbf{x}} = (\overline{A\mathbf{x}})^T\mathbf{x} = (\overline{A}\,\overline{\mathbf{x}})^T\mathbf{x} = (A\overline{\mathbf{x}})^T\mathbf{x} = \overline{\mathbf{x}}^T A^T\mathbf{x} = \overline{\mathbf{x}}^T A\mathbf{x} \quad \blacktriangleleft$$

Checking Attendance:

2. Find the eigenvalues of

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

Exercise 1

19. Let

$$A = \begin{bmatrix} -1 & 7 & -1 \\ 0 & 1 & 0 \\ 0 & 15 & -2 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 5 \end{bmatrix}$$

Confirm that P diagonalizes A, and then compute A^{11}.

Chapter 6: Inner Product Spaces

Based on Elementary Linear Algebra (11th Edition)

DEFINITION 1 An ***inner product*** on a real vector space V is a function that associates a real number $\langle \mathbf{u}, \mathbf{v} \rangle$ with each pair of vectors in V in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V and all scalars k.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ **[Symmetry axiom]**
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$ **[Additivity axiom]**
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$ **[Homogeneity axiom]**
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ **[Positivity axiom]**

A real vector space with an inner product is called a ***real inner product space***.

6.1 Inner Products

Key Concepts and Examples

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

DEFINITION 2 If V is a real inner product space, then the ***norm*** (or ***length***) of a vector $\mathbf{v}$ in V is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the ***distance*** between two vectors is denoted by $d(\mathbf{u}, \mathbf{v})$ and is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}$$

A vector of norm 1 is called a ***unit vector***.

6.1 Inner Products

Key Concepts and Examples

6.1 Inner Products

Example:

For $u = (1,2,3)$ and $v = (2,-1,2)$:

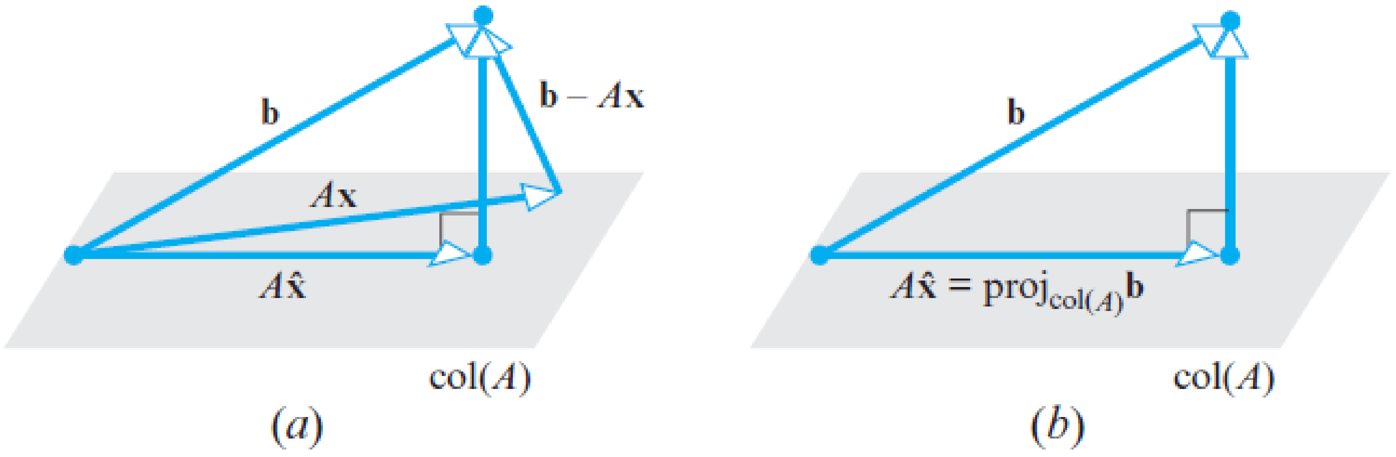
$\langle u,v \rangle = 1\cdot 2 + 2\cdot(-1) + 3\cdot 2 = 6.$

$\|u\| = \sqrt{14}$, $\|v\| = \sqrt{9} = 3.$

6.4 Best Approximation and Least Squares

To explain the terminology in this problem, suppose that the column form of $\mathbf{b} - A\mathbf{x}$ is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$



THEOREM 6.4.1 Best Approximation Theorem

*If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V, then $\text{proj}_W\,\mathbf{b}$ is the **best approximation** to $\mathbf{b}$ from W in the sense that*

$$\|\mathbf{b} - \text{proj}_W\,\mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ in W that is different from $\text{proj}_W\,\mathbf{b}$.

6.4 Best Approximation and Least Squares

6.4 Best Approximation and Least Squares

Proof For every vector $\mathbf{w}$ in W, we can write

$$\mathbf{b} - \mathbf{w} = (\mathbf{b} - \text{proj}_W \mathbf{b}) + (\text{proj}_W \mathbf{b} - \mathbf{w}) \tag{1}$$

But $\text{proj}_W \mathbf{b} - \mathbf{w}$, being a difference of vectors in W, is itself in W; and since $\mathbf{b} - \text{proj}_W \mathbf{b}$ is orthogonal to W, the two terms on the right side of (1) are orthogonal. Thus, it follows from the Theorem of Pythagoras (Theorem 6.2.3) that

$$\|\mathbf{b} - \mathbf{w}\|^2 = \|\mathbf{b} - \text{proj}_W \mathbf{b}\|^2 + \|\text{proj}_W \mathbf{b} - \mathbf{w}\|^2$$

If $\mathbf{w} \neq \text{proj}_W \mathbf{b}$, it follows that the second term in this sum is positive, and hence that

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\|^2 < \|\mathbf{b} - \mathbf{w}\|^2$$

Since norms are nonnegative, it follows (from a property of inequalities) that

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\| \quad \blacktriangleleft$$

It follows from Theorem 6.4.1 that if $V = R^n$ and $W = \text{col}(A)$, then the best approximation to $\mathbf{b}$ from $\text{col}(A)$ is $\text{proj}_{\text{col}(A)}\mathbf{b}$. But every vector in the column space of A is expressible in the form $A\mathbf{x}$ for some vector $\mathbf{x}$, so there is at least one vector $\hat{\mathbf{x}}$ in $\text{col}(A)$ for which $A\hat{\mathbf{x}} = \text{proj}_{\text{col}(A)}\mathbf{b}$. Each such vector is a least squares solution of $A\mathbf{x} = \mathbf{b}$. Note, however, that although there may be more than one least squares solution of $A\mathbf{x} = \mathbf{b}$, each such solution $\hat{\mathbf{x}}$ has the same error vector $\mathbf{b} - A\hat{\mathbf{x}}$.

6.4 Best Approximation and Least Squares

Finding Least Squares Solutions

One way to find a least squares solution of $A\mathbf{x} = \mathbf{b}$ is to calculate the orthogonal projection $\text{proj}_W \mathbf{b}$ on the column space W of A and then solve the equation

$$A\mathbf{x} = \text{proj}_W \mathbf{b} \tag{2}$$

However, we can avoid calculating the projection by rewriting (2) as

$$\mathbf{b} - A\mathbf{x} = \mathbf{b} - \text{proj}_W \mathbf{b}$$

and then multiplying both sides of this equation by A^T to obtain

$$A^T(\mathbf{b} - A\mathbf{x}) = A^T(\mathbf{b} - \text{proj}_W \mathbf{b}) \tag{3}$$

6.4 Best Approximation and Least Squares

Since $\mathbf{b} - \text{proj}_W \mathbf{b}$ is the component of $\mathbf{b}$ that is orthogonal to the column space of A, it follows from Theorem 4.8.7(b) that this vector lies in the null space of A^T, and hence that

$$A^T(\mathbf{b} - \text{proj}_W \mathbf{b}) = \mathbf{0}$$

Thus, (3) simplifies to

$$A^T(\mathbf{b} - A\mathbf{x}) = \mathbf{0}$$

which we can rewrite as

$$A^TA\mathbf{x} = A^T\mathbf{b} \tag{4}$$

This is called the ***normal equation*** or the ***normal system*** associated with $A\mathbf{x} = \mathbf{b}$. When viewed as a linear system, the individual equations are called the ***normal equations*** associated with $A\mathbf{x} = \mathbf{b}$.

In summary, we have established the following result.

6.4 Best Approximation and Least Squares

▶ **EXAMPLE 1 Unique Least Squares Solution**

Find the least squares solution, the least squares error vector, and the least squares error of the linear system

$$\begin{aligned} x_1 - x_2 &= 4 \\ 3x_1 + 2x_2 &= 1 \\ -2x_1 + 4x_2 &= 3 \end{aligned}$$

Solution It will be convenient to express the system in the matrix form $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \tag{7}$$

It follows that

$$A^TA = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix} \tag{8}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

6.4 Best Approximation and Least Squares

so the normal system $A^TA\mathbf{x} = A^T\mathbf{b}$ is

$$\begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$$

Solving this system yields a unique least squares solution, namely,

$$x_1 = \tfrac{17}{95}, \quad x_2 = \tfrac{143}{285}$$

The least squares error vector is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} \frac{17}{95} \\ \frac{143}{285} \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{95}{57} \end{bmatrix} = \begin{bmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{bmatrix}$$

and the least squares error is

$$\|\mathbf{b} - A\mathbf{x}\| \approx 4.556 \blacktriangleleft$$

6.4 Best Approximation and Least Squares

EXAMPLE 2 Infinitely Many Least Squares Solutions

Find the least squares solutions, the least squares error vector, and the least squares error of the linear system

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 2 \\ x_1 - 4x_2 + 3x_3 &= -2 \\ x_1 + 10x_2 - 7x_3 &= 1 \end{aligned}$$

Solution The matrix form of the system is $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

It follows that

$$A^TA = \begin{bmatrix} 11 & 12 & -7 \\ 12 & 120 & -84 \\ -7 & -84 & 59 \end{bmatrix} \quad \text{and} \quad A^T\mathbf{b} = \begin{bmatrix} 5 \\ 22 \\ -15 \end{bmatrix}$$

6.4 Best Approximation and Least Squares

so the augmented matrix for the normal system $A^TA\mathbf{x} = A^T\mathbf{b}$ is

$$\left[\begin{array}{rrr|r} 11 & 12 & -7 & 5 \\ 12 & 120 & -84 & 22 \\ -7 & -84 & 59 & -15 \end{array}\right]$$

The reduced row echelon form of this matrix is

$$\left[\begin{array}{rrr|r} 1 & 0 & \frac{1}{7} & \frac{2}{7} \\ 0 & 1 & -\frac{5}{7} & \frac{13}{84} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

from which it follows that there are infinitely many least squares solutions, and that they are given by the parametric equations

$$\begin{aligned} x_1 &= \tfrac{2}{7} - \tfrac{1}{7}t \\ x_2 &= \tfrac{13}{84} + \tfrac{5}{7}t \\ x_3 &= t \end{aligned}$$

As a check, let us verify that all least squares solutions produce the same least squares error vector and the same least squares error. To see that this is so, we first compute

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 & -1 \\ 1 & -4 & 3 \\ 1 & 10 & -7 \end{bmatrix} \begin{bmatrix} \frac{2}{7} - \frac{1}{7}t \\ \frac{13}{84} + \frac{5}{7}t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{7}{6} \\ -\frac{1}{3} \\ \frac{11}{6} \end{bmatrix} = \begin{bmatrix} \frac{5}{6} \\ -\frac{5}{3} \\ -\frac{5}{6} \end{bmatrix}$$

Since $\mathbf{b} - A\mathbf{x}$ does not depend on t, all least squares solutions produce the same error

6.4 Application of Least Square

THEOREM 6.5.1 Uniqueness of the Least Squares Solution

Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ *be a set of two or more data points, not all lying on a vertical line, and let*

$$M = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \textit{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{4}$$

Then there is a unique least squares straight line fit

$$y = a^* + b^*x \tag{5}$$

to the data points. Moreover,

$$\mathbf{v}^* = \begin{bmatrix} a^* \\ b^* \end{bmatrix} \tag{6}$$

is given by the formula

$$\mathbf{v}^* = (M^TM)^{-1}M^T\mathbf{y} \tag{7}$$

which expresses the fact that $\mathbf{v} = \mathbf{v}^*$ *is the unique solution of the normal equation*

$$M^TM\mathbf{v} = M^T\mathbf{y} \tag{8}$$

EXAMPLE 3 Fitting a Quadratic Curve to Data

According to Newton's second law of motion, a body near the Earth's surface falls vertically downward in accordance with the equation

$$s = s_0 + v_0 t + \tfrac{1}{2} g t^2 \tag{13}$$

where

s = vertical displacement downward relative to some reference point

s_0 = displacement from the reference point at time $t = 0$

v_0 = velocity at time $t = 0$

g = acceleration of gravity at the Earth's surface

Suppose that a laboratory experiment is performed to approximate g by measuring the displacement s relative to a fixed reference point of a falling weight at various times. Use the experimental results shown in the following table to approximate g.

Time t (sec)	.1	.2	.3	.4	.5
Displacement s (ft)	−0.18	0.31	1.03	2.48	3.73

Example

Solution For notational simplicity, let $a_0 = s_0$, $a_1 = v_0$, and $a_2 = \frac{1}{2}g$ in (13), so our mathematical problem is to fit a quadratic curve

$$s = a_0 + a_1 t + a_2 t^2 \tag{14}$$

to the five data points:

$$(.1, -0.18), \quad (.2, 0.31), \quad (.3, 1.03), \quad (.4, 2.48), \quad (.5, 3.73)$$

With the appropriate adjustments in notation, the matrices M and $\mathbf{y}$ in (11) are

$$M = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} = \begin{bmatrix} 1 & .1 & .01 \\ 1 & .2 & .04 \\ 1 & .3 & .09 \\ 1 & .4 & .16 \\ 1 & .5 & .25 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} = \begin{bmatrix} -0.18 \\ 0.31 \\ 1.03 \\ 2.48 \\ 3.73 \end{bmatrix}$$

Thus, from (12),

$$\mathbf{v}^* = \begin{bmatrix} a_0^* \\ a_1^* \\ a_2^* \end{bmatrix} = (M^T M)^{-1} M^T \mathbf{y} \approx \begin{bmatrix} -0.40 \\ 0.35 \\ 16.1 \end{bmatrix}$$

so the least squares quadratic fit is

$$s = -0.40 + 0.35t + 16.1t^2$$

From this equation we estimate that $\frac{1}{2}g = 16.1$ and hence that $g = 32.2$ ft/sec^2. Note that this equation also provides the following estimates of the initial displacement and velocity of the weight:

$$s_0 = a_0^* = -0.40 \text{ ft}$$
$$v_0 = a_1^* = 0.35 \text{ ft/sec}$$

In Figure 6.5.5 we have plotted the data points and the approximating polynomial. ◀

Example

T7. (a) Show that making the substitution $X = \ln x$ in the equation $y = a + b \ln x$ produces the equation $y = a + bX$ whose graph in the Xy-plane is a line of slope b and y-intercept a.

(b) Part (a) suggests that a curve of the form $y = a + b \ln x$ can be fitted to n data points (x_i, y_i) by letting $X_i = \ln x_i$ and then fitting a straight line to the transformed data points (X_i, y_i) by least squares to find b and a. Use this method to fit a logarithmic model to the following data, and graph the curve and data in the same coordinate system.

x	2	3	4	5	6	7	8	9
y	4.07	5.30	6.21	6.79	7.32	7.91	8.23	8.51

Exercise 2:

Chapter 7: Diagonalization and Quadratic Forms

Based on Elementary Linear Algebra (11th Edition)

7.1 Orthogonal Matrices

Key Concepts and Examples

DEFINITION 1 A square matrix A is said to be ***orthogonal*** if its transpose is the same as its inverse, that is, if

$$A^{-1} = A^T$$

or, equivalently, if

$$AA^T = A^TA = I \tag{1}$$

EXAMPLE 1 A 3 x 3 Orthogonal Matrix

The matrix

$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$

is orthogonal since

$$A^TA = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix} \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

7.1 Orthogonal Matrices

EXAMPLE 2 Rotation and Reflection Matrices Are Orthogonal

Recall from Table 5 of Section 4.9 that the standard matrix for the counterclockwise rotation of R^2 through an angle θ is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This matrix is orthogonal for all choices of θ since

$$A^T A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We leave it for you to verify that the reflection matrices in Tables 1 and 2 and the rotation matrices in Table 6 of Section 4.9 are all orthogonal.

7.2 Orthogonal Diagonalization

DEFINITION 1 If A and B are square matrices, then we say that B is ***orthogonally similar*** to A if there is an orthogonal matrix P such that $B = P^TAP$.

Note that if B is orthogonally similar to A, then it is also true that A is orthogonally similar to B since we can express A as $A = Q^TBQ$ by taking $Q = P^T$ (verify). This being the case we will say that A and B are ***orthogonally similar matrices*** if either is orthogonally similar to the other.

If A is orthogonally similar to some diagonal matrix, say

$$P^TAP = D$$

7.2 Orthogonal Diagonalization

Our first goal in this section is to determine what conditions a matrix must satisfy to be orthogonally diagonalizable. As an initial step, observe that there is no hope of orthogonally diagonalizing a matrix that is not symmetric. To see why this is so, suppose that

$$P^TAP = D \tag{1}$$

where P is an orthogonal matrix and D is a diagonal matrix. Multiplying the left side of (1) by P, the right side by P^T, and then using the fact that $PP^T = P^TP = I$, we can rewrite this equation as

$$A = PDP^T \tag{2}$$

Now transposing both sides of this equation and using the fact that a diagonal matrix is the same as its transpose we obtain

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

so A must be symmetric if it is orthogonally diagonalizable.

7.2 Orthogonal Diagonalization

Orthogonally Diagonalizing an $n \times n$ Symmetric Matrix

Step 1. Find a basis for each eigenspace of A.

Step 2. Apply the Gram–Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

Step 3. Form the matrix P whose columns are the vectors constructed in Step 2. This matrix will orthogonally diagonalize A, and the eigenvalues on the diagonal of $D = P^TAP$ will be in the same order as their corresponding eigenvectors in P.

7.2 Orthogonal Diagonalization

EXAMPLE 1 Orthogonally Diagonalizing a Symmetric Matrix

Find an orthogonal matrix P that diagonalizes

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

Solution We leave it for you to verify that the characteristic equation of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 4 & -2 & -2 \\ -2 & \lambda - 4 & -2 \\ -2 & -2 & \lambda - 4 \end{bmatrix} = (\lambda - 2)^2(\lambda - 8) = 0$$

Thus, the distinct eigenvalues of A are $\lambda = 2$ and $\lambda = 8$. By the method used in Example 7 of Section 5.1, it can be shown that

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \tag{5}$$

form a basis for the eigenspace corresponding to $\lambda = 2$. Applying the Gram–Schmidt process to $\{\mathbf{u}_1, \mathbf{u}_2\}$ yields the following orthonormal eigenvectors (verify):

$$\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \tag{6}$$

7.2 Orthogonal Diagonalization

The eigenspace corresponding to $\lambda = 8$ has

$$\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

as a basis. Applying the Gram–Schmidt process to $\{\mathbf{u}_3\}$ (i.e., normalizing $\mathbf{u}_3$) yields

$$\mathbf{v}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

Finally, using $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ as column vectors, we obtain

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

which orthogonally diagonalizes A. As a check, we leave it for you to confirm that

$$P^TAP = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$ ◀

7.2 Exercise 2: find P and A^{20}

5. Find a matrix P that orthogonally diagonalizes

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

and determine the diagonal matrix $D = P^TAP$.