

CHAPTER 8

Linear Algebra:

Matrix Eigenvalue Problems

A **matrix eigenvalue problem** considers the vector equation

$$(1) \quad \mathbf{Ax} = \lambda \mathbf{x}.$$

Here \mathbf{A} is a given square matrix, λ an unknown scalar, and \mathbf{x} an unknown vector. In a matrix eigenvalue problem, the task is to determine λ 's and \mathbf{x} 's that satisfy (1).

Since $\mathbf{x} = \mathbf{0}$ is always a solution for any λ and thus not interesting, we only admit solutions with $\mathbf{x} \neq \mathbf{0}$.

The solutions to (1) are given the following names: The λ 's that satisfy (1) are called **eigenvalues of \mathbf{A}** and the corresponding nonzero \mathbf{x} 's that also satisfy (1) are called **eigenvectors of \mathbf{A}** .

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

We formalize our observation. Let $\mathbf{A} = [a_{jk}]$ be a given nonzero square matrix of dimension $n \times n$. Consider the following vector equation:

$$(1) \quad \mathbf{Ax} = \lambda \mathbf{x}.$$

The problem of finding nonzero \mathbf{x} 's and λ 's that satisfy equation (1) is called **an eigenvalue problem**.

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

A value of λ for which (1) has a solution $\mathbf{x} \neq \mathbf{0}$ is called an **eigenvalue** or *characteristic value* of the matrix \mathbf{A} .

The corresponding solutions $\mathbf{x} \neq \mathbf{0}$ of (1) are called the **eigenvectors** or *characteristic vectors* of \mathbf{A} corresponding to that eigenvalue λ .

The set of all the eigenvalues of \mathbf{A} is called the **spectrum of \mathbf{A}** . We shall see that the spectrum consists of at least one eigenvalue and at most of n numerically different eigenvalues.

The largest of the absolute values of the eigenvalues of \mathbf{A} is called the spectral radius of \mathbf{A} , a name to be motivated later.

How to Find Eigenvalues and Eigenvectors

EXAMPLE 1

Determination of Eigenvalues and Eigenvectors

We illustrate all the steps in terms of the matrix

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

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 1 (continued 1)

Determination of Eigenvalues and Eigenvectors

Solution.

(a) *Eigenvalues.* These must be determined *first*.

Equation (1) is

$$\mathbf{Ax} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

in components

$$-5x_1 + 2x_2 = \lambda x_1$$

$$2x_1 - 2x_2 = \lambda x_2.$$

EXAMPLE 1 (continued 2)

Determination of Eigenvalues and Eigenvectors

Solution. (continued 1)

(a) *Eigenvalues. (continued 1)*

Transferring the terms on the right to the left, we get

$$(2^*) \quad \begin{aligned} (-5 - \lambda)x_1 + 2x_2 &= 0 \\ 2x_1 + (-2 - \lambda)x_2 &= 0 \end{aligned}$$

This can be written in matrix notation

$$(3^*) \quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

Because (1) is $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{Ax} - \lambda\mathbf{Ix} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, which gives (3*).

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 1 (continued 3)

Determination of Eigenvalues and Eigenvectors

Solution. (continued 2)

(a) *Eigenvalues. (continued 2)*

We see that this is a *homogeneous* linear system. It has a nontrivial solution (an eigenvector of \mathbf{A} we are looking for) if and only if its coefficient determinant is zero, that is,

$$\begin{aligned} D(\lambda) &= \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -5 - \lambda & 2 \\ 2 & -2 - \lambda \end{vmatrix} \\ (4^*) \quad &= (-5 - \lambda)(-2 - \lambda) - 4 = \lambda^2 + 7\lambda + 6 = 0. \end{aligned}$$

EXAMPLE 1 (continued 4)

Determination of Eigenvalues and Eigenvectors

Solution. (continued 3)

(a) *Eigenvalues.* (continued 3)

We call $D(\lambda)$ the **characteristic determinant** or, if expanded, the **characteristic polynomial**, and $D(\lambda) = 0$ the **characteristic equation of \mathbf{A}** . The solutions of this quadratic equation are $\lambda_1 = -1$ and $\lambda_2 = -6$. These are the eigenvalues of \mathbf{A} .

(b₁) *Eigenvector of \mathbf{A} corresponding to λ_1* . This vector is obtained from (2*) with $\lambda = \lambda_1 = -1$, that is,

$$-4x_1 + 2x_2 = 0$$

$$2x_1 - x_2 = 0.$$

EXAMPLE 1 (continued 5)

Determination of Eigenvalues and Eigenvectors

Solution. (continued 4)

(b₁) *Eigenvector of A corresponding to λ_1 . (continued)*

A solution is $x_2 = 2x_1$, as we see from either of the two equations, so that we need only one of them. This determines an eigenvector corresponding to $\lambda_1 = -1$ up to a scalar multiple. If we choose $x_1 = 1$, we obtain the eigenvector

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \text{Check: } \mathbf{A}\mathbf{x}_1 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = (-1)\mathbf{x}_1 = \lambda_1\mathbf{x}_1.$$

EXAMPLE 1 (continued 6)

Determination of Eigenvalues and Eigenvectors

Solution. (continued 5)

(b₂) Eigenvector of A corresponding to λ_2 .

For $\lambda = \lambda_2 = -6$, equation (2*) becomes

$$x_1 + 2x_2 = 0$$

$$2x_1 + 4x_2 = 0.$$

A solution is $x_2 = -x_1/2$ with arbitrary x_1 . If we choose $x_1 = 2$, we get $x_2 = -1$. Thus an eigenvector of **A** corresponding to $\lambda_2 = -6$ is

$$\mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \text{Check: } \mathbf{A}\mathbf{x}_2 = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -12 \\ 6 \end{bmatrix} = (-6)\mathbf{x}_2 = \lambda_2\mathbf{x}_2.$$

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

This example illustrates the general case as follows.

Equation (1) written in components is

$$a_{11}x_1 + \cdots + a_{1n}x_n = \lambda x_1$$

$$a_{21}x_1 + \cdots + a_{2n}x_n = \lambda x_2$$

.....

$$a_{n1}x_1 + \cdots + a_{nn}x_n = \lambda x_n.$$

Transferring the terms on the right side to the left side, we have

$$(a_{11} - \lambda)x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$

$$a_{21}x_1 + (a_{22} - \lambda)x_2 + \cdots + a_{2n}x_n = 0$$

.....

$$a_{n1}x_1 + a_{n2}x_2 + \cdots + (a_{nn} - \lambda)x_n = 0.$$

(2)

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

In matrix notation,

$$(3) \quad (\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}.$$

By **Cramer's theorem** in Sec. 7.7, this homogeneous linear system of equations has a nontrivial solution if and only if the corresponding determinant of the coefficients is zero:

$$(4) \quad D(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

$\mathbf{A} - \lambda\mathbf{I}$ is called the **characteristic matrix** and $D(\lambda)$ the **characteristic determinant** of \mathbf{A} . Equation (4) is called the **characteristic equation** of \mathbf{A} . By developing $D(\lambda)$ we obtain a polynomial of n th degree in λ . This is called the **characteristic polynomial** of \mathbf{A} .

Theorem 1

Eigenvalues

The eigenvalues of a square matrix \mathbf{A} are the roots of the characteristic equation (4) of \mathbf{A} .

Hence an $n \times n$ matrix has at least one eigenvalue and at most n numerically different eigenvalues.

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

The eigenvalues must be determined first.

Once these are known, corresponding **eigenvectors** are obtained from the system (2), for instance, **by the Gauss elimination**, where λ is the eigenvalue for which an eigenvector is wanted.

Theorem 2

Eigenvectors, Eigenspace

If \mathbf{w} and \mathbf{x} are eigenvectors of a matrix \mathbf{A} corresponding to the same eigenvalue λ , so are $\mathbf{w} + \mathbf{x}$ (provided $\mathbf{x} \neq -\mathbf{w}$) and $k\mathbf{x}$ for any $k \neq 0$.

Hence the eigenvectors corresponding to one and the same eigenvalue λ of \mathbf{A} , together with $\mathbf{0}$, form a vector space, called the **eigenspace** of \mathbf{A} corresponding to that λ .

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

In particular, *an eigenvector \mathbf{x} is determined only up to a constant factor.*

Hence we can **normalize** \mathbf{x} , that is, multiply it by a scalar to get a unit vector.

EXAMPLE 2 Multiple Eigenvalues

Find the eigenvalues and eigenvectors of

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

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 2 Multiple Eigenvalues (continued 1)

Solution.

For our matrix, the characteristic determinant gives the characteristic equation

$$-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

The roots (eigenvalues of \mathbf{A}) are $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$.

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 2 Multiple Eigenvalues (continued 2)

Solution. (continued 1)

To find eigenvectors, we apply the **Gauss elimination** to the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, first with $\lambda = 5$ and then with $\lambda = -3$. For $\lambda = 5$ the characteristic matrix is

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} - 5\mathbf{I} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}.$$

It row-reduces to $\begin{bmatrix} -7 & 2 & -3 \\ 0 & -24/7 & -48/7 \\ 0 & 0 & 0 \end{bmatrix}$.

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 2 Multiple Eigenvalues (continued 3)

Solution. (continued 2)

Hence it has rank 2. Choosing $x_3 = -1$ we have $x_2 = 2$ from $-\frac{24}{7}x_2 - \frac{48}{7}x_3 = 0$ and then $x_1 = 1$ from $-7x_1 + 2x_2 - 3x_3 = 0$.

Hence an eigenvector of \mathbf{A} corresponding to $\lambda = 5$ is $\mathbf{x}_1 = [1 \ 2 \ -1]^T$.

For $\lambda = -3$ the characteristic matrix

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}$$

row-reduces to

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

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

EXAMPLE 2 Multiple Eigenvalues (continued 4)

Solution. (continued 3)

Hence it has rank 1 and there are 2 free variables.

From $x_1 + 2x_2 - 3x_3 = 0$ we have $x_1 = -2x_2 + 3x_3$.

Choosing (1) $x_2 = 1, x_3 = 0$ and (2) $x_2 = 0, x_3 = 1$, we obtain two linearly independent eigenvectors of \mathbf{A} corresponding to $\lambda = -3$

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

- The order M_λ of an eigenvalue λ as a root of the characteristic polynomial is called the **algebraic multiplicity** of λ .
- The number m_λ of linearly independent eigenvectors corresponding to λ is called the **geometric multiplicity** of λ .
Thus m_λ is the dimension of the eigenspace corresponding to this λ .
- Since the characteristic polynomial has degree n , *the sum of all the algebraic multiplicities must equal n .*
- In Example 2 for $\lambda = -3$ we have $m_\lambda = M_\lambda = 2$. Generally speaking, $m_\lambda \leq M_\lambda$, as can be shown. The difference $\Delta_\lambda = M_\lambda - m_\lambda$ is called the **defect of λ** . Thus $\Delta_{-3} = 0$ in Example 2, but positive defects Δ_λ can easily occur.

Theorem 3

Eigenvalues of the Transpose

The transpose \mathbf{A}^T of a square matrix \mathbf{A} has the same eigenvalues as \mathbf{A} .

8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

Definitions

Symmetric, Skew-Symmetric, and Orthogonal Matrices

A *real* square matrix $\mathbf{A} = [a_{jk}]$ is called
symmetric if transposition leaves it unchanged,

$$(1) \quad \mathbf{A}^T = \mathbf{A}, \quad \text{thus} \quad a_{kj} = a_{jk}$$

skew-symmetric if transposition gives the negative of \mathbf{A} ,

$$(2) \quad \mathbf{A}^T = -\mathbf{A}, \quad \text{thus} \quad a_{kj} = -a_{jk}$$

orthogonal if transposition gives the inverse of \mathbf{A} ,

$$(3) \quad \mathbf{A}^T = \mathbf{A}^{-1}.$$

8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

Any real square matrix \mathbf{A} may be written as the sum of a symmetric matrix \mathbf{R} and a skew-symmetric matrix \mathbf{S} , where

$$(4) \quad \mathbf{R} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) \quad \text{and} \quad \mathbf{S} = \frac{1}{2}(\mathbf{A} - \mathbf{A}^T).$$

Theorem 1

Eigenvalues of Symmetric and Skew-Symmetric Matrices

- (a) *The eigenvalues of a symmetric matrix are real.*
- (b) *The eigenvalues of a skew-symmetric matrix are pure imaginary or zero.*

Orthogonal Transformations and Orthogonal Matrices

Orthogonal transformations are transformations

$$(5) \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{where } \mathbf{A} \text{ is an } \text{orthogonal matrix}.$$

With each vector \mathbf{x} in R^n such a transformation assigns a vector \mathbf{y} in R^n .

For instance, the plane rotation through an angle θ

$$(6) \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is an orthogonal transformation.

Theorem 2

Invariance of Inner Product

An orthogonal transformation preserves the value of the **inner product** of vectors \mathbf{a} and \mathbf{b} in R^n , defined by

$$(7) \quad \mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}.$$

That is, for any \mathbf{a} and \mathbf{b} in R^n , orthogonal $n \times n$ matrix \mathbf{A} , and $\mathbf{u} = \mathbf{A}\mathbf{a}$, $\mathbf{v} = \mathbf{A}\mathbf{b}$ we have $\mathbf{u} \cdot \mathbf{v} = \mathbf{a} \cdot \mathbf{b}$.

Hence the transformation also preserves the **length** or **norm** of any vector \mathbf{a} in R^n given by

$$(8) \quad \|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\mathbf{a}^T \mathbf{a}}.$$

Theorem 3

Orthonormality of Column and Row Vectors

A real square matrix is orthogonal if and only if its column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$ (and also its row vectors) form an **orthonormal system**, that is,

$$(10) \quad \mathbf{a}_j \cdot \mathbf{a}_k = \mathbf{a}_j^T \mathbf{a}_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k. \end{cases}$$

Theorem 4

Determinant of an Orthogonal Matrix

The determinant of an orthogonal matrix has the value +1 or -1.

Theorem 5

Eigenvalues of an Orthogonal Matrix

The eigenvalues of an orthogonal matrix \mathbf{A} are real or complex conjugates in pairs and have absolute value 1.

8.4 Eigenbases. Diagonalization. Quadratic Forms

Eigenvectors of an $n \times n$ matrix \mathbf{A} **may (or may not!)** form a basis for R^n .

If we are interested in a transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, such an “**eigenbasis**” (basis of eigenvectors)—if it exists—is of great advantage because then we can represent any \mathbf{x} in R^n uniquely as a linear combination of the eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, say,

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \dots + c_n\mathbf{x}_n.$$

And, denoting the corresponding (not necessarily distinct) eigenvalues of the matrix \mathbf{A} by $\lambda_1, \dots, \lambda_n$, we have $\mathbf{A}\mathbf{x}_j = \lambda_j\mathbf{x}_j$, so that we simply obtain

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x} = \mathbf{A}(c_1\mathbf{x}_1 + \dots + c_n\mathbf{x}_n) \\ (1) \quad &= c_1\mathbf{A}\mathbf{x}_1 + \dots + c_n\mathbf{A}\mathbf{x}_n \\ &= c_1\lambda_1\mathbf{x}_1 + \dots + c_n\lambda_n\mathbf{x}_n. \end{aligned}$$

This shows that we have decomposed
the complicated action of \mathbf{A} on an arbitrary vector \mathbf{x}
into
a sum of simple actions along the eigenvectors of \mathbf{A} .

This is the point of an eigenbasis.

Theorem 1

Basis of Eigenvectors

*If an $n \times n$ matrix \mathbf{A} has n **distinct** eigenvalues, then \mathbf{A} has a basis of eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ for \mathbb{R}^n .*

Theorem 2

Symmetric Matrices

*A symmetric matrix has an **orthonormal basis** of eigenvectors for R^n .*

Similarity of Matrices. Diagonalization

DEFINITION

Similar Matrices. Similarity Transformation

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix \mathbf{A} if

$$(4) \quad \hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

for some (nonsingular!) $n \times n$ matrix \mathbf{P} .

This transformation, which gives $\hat{\mathbf{A}}$ from \mathbf{A} , is called a **similarity transformation**.

Theorem 3

Eigenvalues and Eigenvectors of Similar Matrices

If $\hat{\mathbf{A}}$ is similar to \mathbf{A} , then $\hat{\mathbf{A}}$ has the same eigenvalues as \mathbf{A} .

Furthermore, if \mathbf{x} is an eigenvector of \mathbf{A} , then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue.

Theorem 4

Diagonalization of a Matrix

If an $n \times n$ matrix \mathbf{A} has a basis of eigenvectors, then

$$(5) \quad \mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

is diagonal, with the eigenvalues of \mathbf{A} as the entries on the main diagonal.

Here \mathbf{X} is the matrix with these eigenvectors as column vectors.

Also,

$$(5^*) \quad \mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X} \quad (m = 2, 3, \dots).$$

EXAMPLE 4 Diagonalization

Diagonalize

$$\mathbf{A} = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}.$$

Solution.

- The characteristic determinant gives the characteristic equation $-\lambda^3 - \lambda^2 + 12\lambda = 0$. The roots (eigenvalues of \mathbf{A}) are $\lambda_1 = 3$, $\lambda_2 = -4$, $\lambda_3 = 0$.
- By the Gauss elimination applied to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1, \lambda_2, \lambda_3$ we find eigenvectors to form \mathbf{X} and then find \mathbf{X}^{-1} by the Gauss–Jordan elimination.

EXAMPLE 4 (continued 1) **Diagonalization**

Solution. (continued 1)

The results are

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix},$$

$$\mathbf{X}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

EXAMPLE 4 (continued 2) **Diagonalization***Solution.* (continued 2)

Calculating \mathbf{AX} and multiplying by \mathbf{X}^{-1} from the left, we thus obtain

$$\begin{aligned} \mathbf{D} &= \mathbf{X}^{-1} \mathbf{AX} = \mathbf{X}^{-1} \mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 \end{bmatrix} \\ &= \mathbf{X}^{-1} \begin{bmatrix} \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \lambda_3 \mathbf{x}_3 \end{bmatrix} \\ &= \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} \begin{bmatrix} -3 & -4 & 0 \\ 9 & 4 & 0 \\ -3 & -12 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Quadratic Forms.

Transformation to Principal Axes

By definition, a **quadratic form** Q in the components x_1, \dots, x_n of a vector \mathbf{x} is a sum n^2 of terms, namely,

$$\begin{aligned}
 (7) \quad Q &= \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k \\
 &= a_{11} x_1^2 + a_{12} x_1 x_2 + \cdots + a_{1n} x_1 x_n \\
 &\quad + a_{21} x_2 x_1 + a_{22} x_2^2 + \cdots + a_{2n} x_2 x_n \\
 &\quad + \cdots \cdots \cdots \\
 &\quad + a_{n1} x_n x_1 + a_{n2} x_n x_2 + \cdots + a_{nn} x_n^2.
 \end{aligned}$$

$\mathbf{A} = [a_{jk}]$ is called the **coefficient matrix** of the form.

We may assume that \mathbf{A} is *symmetric*, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms.

EXAMPLE 5 Quadratic Form. Symmetric Coefficient Matrix

Let

$$\begin{aligned}\mathbf{x}^T \mathbf{A} \mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= 3x_1^2 + 4x_1x_2 + 6x_2x_1 + 2x_2^2 \\ &= 3x_1^2 + 10x_1x_2 + 2x_2^2.\end{aligned}$$

Here $4 + 6 = 10 = 5 + 5$.

EXAMPLE 5 (continued) **Quadratic Form.**
Symmetric Coefficient Matrix

From the corresponding *symmetric* matrix $\mathbf{C} = [c_{jk}]$ where $c_{jk} = \frac{1}{2}(a_{jk} + a_{kj})$, thus $c_{11} = 3$, $c_{12} = c_{21} = 5$, $c_{22} = 2$, we get the same result; indeed,

$$\begin{aligned}\mathbf{x}^T \mathbf{C} \mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 2x_2^2 \\ &= 3x_1^2 + 10x_1x_2 + 2x_2^2.\end{aligned}$$

By Theorem 2, the *symmetric* coefficient matrix \mathbf{A} of (7) has an orthonormal basis of eigenvectors.

Hence if we take these as column vectors, we obtain a **matrix \mathbf{X} that is orthogonal**, so that $\mathbf{X}^{-1} = \mathbf{X}^T$.

From (5) we thus have $\mathbf{A} = \mathbf{XDX}^{-1} = \mathbf{XDX}^T$. Substitution into (7) gives

$$(8) \quad Q = \mathbf{x}^T \mathbf{XDX}^T \mathbf{x}.$$

If we set $\mathbf{X}^T \mathbf{x} = \mathbf{y}$, then, since $\mathbf{X}^{-1} = \mathbf{X}^T$, we have $\mathbf{X}^{-1} \mathbf{x} = \mathbf{y}$ and thus obtain

$$(9) \quad \mathbf{x} = \mathbf{Xy}.$$

Furthermore, in (8) we have $\mathbf{x}^T \mathbf{X} = (\mathbf{X}^T \mathbf{x})^T = \mathbf{y}^T$ and $\mathbf{X}^T \mathbf{x} = \mathbf{y}$, so that Q becomes simply

$$(10) \quad Q = \mathbf{y}^T \mathbf{Dy} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2.$$

Theorem 5

Principal Axes Theorem

The substitution (9) transforms a quadratic form

$$Q = \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k \quad (a_{kj} = a_{jk})$$

*to the **principal axes form** or **canonical form** (10), where $\lambda_1, \dots, \lambda_n$ are the (not necessarily distinct) eigenvalues of the (symmetric!) matrix \mathbf{A} , and \mathbf{X} is an orthogonal matrix with corresponding eigenvectors $\mathbf{x}_1, \dots, \mathbf{x}_n$, respectively, as column vectors.*

EXAMPLE 6 Transformation to Principal Axes. Conic Sections

Find out what type of conic section the following quadratic form represents and transform it to principal axes:

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128.$$

Solution. We have $Q = \mathbf{x}^T \mathbf{A} \mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

EXAMPLE 6 (continued) Transformation to Principal Axes. Conic Sections

Solution. (continued 1)

This gives the characteristic equation $(17 - \lambda)^2 - 15^2 = 0$. It has the roots $\lambda_1 = 2$, $\lambda_2 = 32$. Hence (10) becomes

$$Q = 2y_1^2 + 32y_2^2.$$

We see that $Q = 128$ represents the ellipse $2y_1^2 + 32y_2^2 = 128$, that is,

$$\frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1.$$

EXAMPLE 6 (continued) Transformation to Principal Axes. Conic Sections

Solution. (continued 2)

If we want to know the direction of the principal axes in the x_1x_2 -coordinates, we have to determine normalized eigenvectors from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1 = 2$ and $\lambda = \lambda_2 = 32$ and then use (9). We get

$$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

EXAMPLE 6 (continued) Transformation to Principal Axes. Conic Sections

Solution. (continued 3)

hence

$$\mathbf{x} = \mathbf{X}\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad \begin{aligned} x_1 &= y_1/\sqrt{2} - y_2/\sqrt{2} \\ x_2 &= y_1/\sqrt{2} + y_2/\sqrt{2}. \end{aligned}$$

This is a 45° rotation.

8.5 Complex Matrices and Forms.

Optional

Notations

$\bar{\mathbf{A}} = [\bar{a}_{jk}]$ is obtained from $\mathbf{A} = [a_{jk}]$ by replacing each entry $a_{jk} = \alpha + i\beta$ (α, β real) with its complex conjugate $\bar{a}_{jk} = \alpha - i\beta$. Also, $\bar{\mathbf{A}}^T = [\bar{a}_{kj}]$ is the transpose of $\bar{\mathbf{A}}$, hence the conjugate transpose of \mathbf{A} .

DEFINITION**Hermitian, Skew-Hermitian, and Unitary Matrices**

A square matrix $\mathbf{A} = [a_{kj}]$ is called

Hermitian	if $\bar{\mathbf{A}}^T = \mathbf{A}$,	that is,	$\bar{a}_{kj} = a_{jk}$
skew-Hermitian	if $\bar{\mathbf{A}}^T = -\mathbf{A}$,	that is,	$\bar{a}_{kj} = -a_{jk}$
unitary	if $\bar{\mathbf{A}}^T = \mathbf{A}^{-1}$.		

Eigenvalues

It is quite remarkable that the matrices under consideration have spectra (sets of eigenvalues; see Sec. 8.1) that can be characterized in a general way as follows (see Fig. 163).

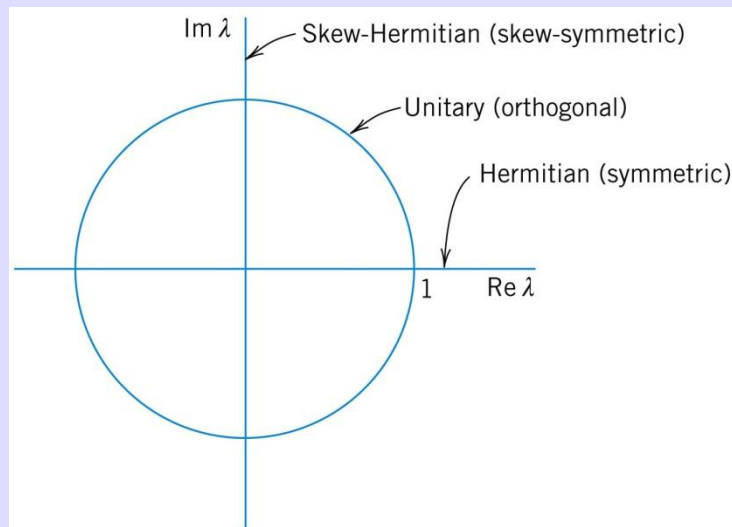


Fig. 163. Location of the eigenvalues of Hermitian, skew-Hermitian, and unitary matrices in the complex λ -plane

Theorem 1

Eigenvalues

- (a) The eigenvalues of a Hermitian matrix (and thus of a symmetric matrix) are real.*
- (b) The eigenvalues of a skew-Hermitian matrix (and thus of a skew-symmetric matrix) are pure imaginary or zero.*
- (c) The eigenvalues of a unitary matrix (and thus of an orthogonal matrix) have absolute value 1.*

Theorem 2

Invariance of Inner Product

A **unitary transformation**, that is, $\mathbf{y} = \mathbf{A}\mathbf{x}$ with a unitary matrix \mathbf{A} , preserves the value of the inner product (4), hence also the norm (5).

DEFINITION

Unitary System

A *unitary system* is a set of complex vectors satisfying the relationships

$$(6) \quad \mathbf{a}_j \cdot \mathbf{a}_k = \bar{\mathbf{a}}_j^T \mathbf{a}_k = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k. \end{cases}$$

Theorem 4

Determinant of a Unitary Matrix

Let \mathbf{A} be a unitary matrix. Then its determinant has absolute value one, that is, $|\det \mathbf{A}| = 1$.

Theorem 5

Basis of Eigenvectors

A Hermitian, skew-Hermitian, or unitary matrix has a basis of eigenvectors for C^n that is a unitary system.

Hermitian and Skew-Hermitian Forms

The concept of a quadratic form (Sec. 8.4) can be extended to complex. We call the numerator $\bar{\mathbf{x}}^T \mathbf{A} \mathbf{x}$ in (1) a **form** in the components x_1, \dots, x_n of \mathbf{x} , which may now be complex. This form is again a sum of n^2 terms

$$\begin{aligned}
 (7) \quad \bar{\mathbf{x}}^T \mathbf{A} \mathbf{x} &= \sum_{j=1}^n \sum_{k=1}^n a_{jk} \bar{x}_j x_k \\
 &= a_{11} \bar{x}_1 x_1 + \cdots + a_{1n} \bar{x}_1 x_n \\
 &\quad + a_{21} \bar{x}_2 x_1 + \cdots + a_{2n} \bar{x}_2 x_n \\
 &\quad + \cdots \cdots \cdots \\
 &\quad + a_{n1} \bar{x}_n x_1 + \cdots + a_{nn} \bar{x}_n x_n.
 \end{aligned}$$

Hermitian and Skew-Hermitian Forms

(continued)

A is called its **coefficient matrix**. The form is called a **Hermitian** or **skew-Hermitian form** if **A** is Hermitian or skew-Hermitian, respectively. *The value of a Hermitian form is real, and that of a skew-Hermitian form is pure imaginary or zero.*

SUMMARY OF CHAPTER 8
Linear Algebra:
Matrix Eigenvalue Problems

Linear Algebra: Matrix Eigenvalue Problems

The practical importance of matrix eigenvalue problems can hardly be overrated. The problems are defined by the vector equation

$$(1) \quad \mathbf{Ax} = \lambda \mathbf{x}.$$

\mathbf{A} is a given square matrix. All matrices in this chapter are *square*. λ is a scalar. To *solve* the problem (1) means to determine values of λ , called **eigenvalues** (or **characteristic values**) of \mathbf{A} , such that (1) has a nontrivial solution \mathbf{x} (that is, $\mathbf{x} \neq \mathbf{0}$), called an **eigenvector** of \mathbf{A} corresponding to that λ . An $n \times n$ matrix has at least one and at most n numerically different eigenvalues.

*(continued 1)***Linear Algebra: Matrix Eigenvalue Problems**

These are the solutions of the **characteristic equation**
(Sec. 8.1)

$$(2) \quad D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

$D(\lambda)$ is called the **characteristic determinant** of \mathbf{A} . By expanding it we get the **characteristic polynomial** of \mathbf{A} , which is of degree n in λ . Some typical applications are shown in Sec. 8.2.

(continued 2)

Linear Algebra: Matrix Eigenvalue Problems

Section 8.3 is devoted to eigenvalue problems for **symmetric** ($\mathbf{A}^T = \mathbf{A}$), **skew-symmetric** ($\mathbf{A}^T = -\mathbf{A}$), and **orthogonal matrices** ($\mathbf{A}^T = \mathbf{A}^{-1}$). Section 8.4 concerns the diagonalization of matrices and the transformation of quadratic forms to principal axes and its relation to eigenvalues.

Section 8.5 extends Sec. 8.3 to the complex analogs of those real matrices, called **Hermitian** ($\mathbf{A}^T = \mathbf{A}$), **skew-Hermitian** ($\mathbf{A}^T = -\mathbf{A}$), and **unitary matrices** $\bar{\mathbf{A}}^T = \mathbf{A}^{-1}$.

All the eigenvalues of a Hermitian matrix (and a symmetric one) are real. For a skew-Hermitian (and a skew-symmetric) matrix they are pure imaginary or zero. For a unitary (and an orthogonal) matrix they have absolute value 1.