Chapter 8:

Linear Algebra:

Matrix Eigenvalue Problems

E. Kreyszig

8.1. The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

A matrix eigenvalue problem considers the vector equation

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}. \tag{1}$$

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 A** and the corresponding nonzero $\mathbf{x}$'s that also satisfy (1) are called **eigenvectors of A**.

Consider multiplying nonzero vectors by a given square matrix, such as

$$\begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 33 \\ 27 \end{bmatrix}, \qquad \begin{bmatrix} 6 & 3 \\ 4 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix}.$$

We want to see what influence the multiplication of the given matrix has on the vectors. In the first case, we get a totally new vector with a different direction and different length when compared to the original vector. This is what usually happens and is of no interest here. In the second case something interesting happens. The multiplication produces a vector $[30 \quad 40]^{\mathsf{T}} = 10\,[3 \quad 4]^{\mathsf{T}}$, which means the new vector has the same direction as the original vector. The scale constant, which we denote by λ is 10. *The problem of systematically finding such* λ's *and nonzero vectors for a given square matrix will be the theme of this chapter*. It is called the *matrix eigenvalue* problem or, more commonly, the *eigenvalue* problem.

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

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

Remark. So $\mathbf{A}$ is a given square (!) matrix, $\mathbf{x}$ is an unknown vector, and λ is an unknown scalar. Our task is to find λ's and nonzero $\mathbf{x}$'s that satisfy (1). Geometrically, we are looking for vectors, $\mathbf{x}$, for which the multiplication by $\mathbf{A}$ has the same effect as the multiplication by a scalar λ; in other words, $\mathbf{Ax}$ should be proportional to $\mathbf{x}$. Thus, the multiplication has the effect of producing, from the original vector $\mathbf{x}$, a new vector $\lambda\mathbf{x}$ that has the same or opposite (minus sign) direction as the original vector.

We introduce more terminology. 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 $\mathbf{A}$. Another term for λ is a *latent* root. ("Eigen" is German and means "proper" or "characteristic."). 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

Eigenvalues have a very large number of applications in diverse fields such as in engineering, geometry, physics, mathematics, biology, environmental science, economics, psychology, and other areas. You will encounter applications for elastic membranes, Markov processes, population models, and others in this chapter.

Since, from the viewpoint of engineering applications, eigenvalue problems are the most important problems in connection with matrices, the student should carefully follow our discussion.

Example 1 demonstrates how to systematically solve a simple eigenvalue problem.

EXAMPLE 1 Determination of Eigenvalues and Eigenvectors

We illustrate all the steps in terms of the matrix

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

Solution. **(a)** ***Eigenvalues.*** These must be determined ***first***. Equation (1) is

$$\mathbf{Ax} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \qquad \text{in components,} \qquad \begin{aligned} -5x_1 + 2x_2 &= \lambda x_1 \\ 2x_1 - 2x_2 &= \lambda x_2. \end{aligned}$$

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

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

This can be written in matrix notation

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

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*). We see that this is a ***homogeneous*** linear system. By Cramer's theorem in Sec. 7.7 it has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$ (an eigenvector of $\mathbf{A}$ we are looking for) if and only if its coefficient determinant is zero, that is,

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

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 $\mathbf{A}$.

($\mathbf{b}_1$) ***Eigenvector of* A *corresponding to*** λ_1. This vector is obtained from (2*) with $\lambda = \lambda_1 = -1$, that is,

$$\begin{aligned} -4x_1 + 2x_2 &= 0 \\ 2x_1 - \quad x_2 &= 0. \end{aligned}$$

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

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

($\mathbf{b}_2$) ***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

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

For the matrix in the intuitive opening example at the start of Sec. 8.1, the characteristic equation is $\lambda^2 - 13\lambda + 30 = (\lambda - 10)(\lambda - 3) = 0$. The eigenvalues are $\{10, \ 3\}$. Corresponding eigenvectors are $[3 \quad 4]^T$ and $[-1 \quad 1]^T$, respectively. The reader may want to verify this. ■

This example illustrates the general case as follows. Equation (1) written in components is

$$
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= \lambda x_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= \lambda x_2 \\
\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\
a_{n1}x_1 + \cdots + a_{nn}x_n &= \lambda x_n.
\end{aligned}
$$

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

$$
\begin{array}{lcccccccl}
(a_{11}-\lambda)x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\
a_{21}x_1 & + & (a_{22}-\lambda)x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & (a_{nn}-\lambda)x_n & = & 0.
\end{array} \tag{2}
$$

In matrix notation,

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

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:

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

$\mathbf{A} - \lambda\mathbf{I}$ 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 nth degree in λ. This is called the **characteristic polynomial** of $\mathbf{A}$.

This proves the following important theorem.

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.*

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* (cf. Sec. 7.4), *called the* **eigenspace** *of* $\mathbf{A}$ *corresponding to that* λ.

PROOF

$\mathbf{Aw} = \lambda\mathbf{w}$ and $\mathbf{Ax} = \lambda\mathbf{x}$ *imply* $\mathbf{A}(\mathbf{w} + \mathbf{x}) = \mathbf{Aw} + \mathbf{Ax} = \lambda\mathbf{w} + \lambda\mathbf{x} = \lambda(\mathbf{w} + \mathbf{x})$ and $\mathbf{A}(k\mathbf{w}) = k(\mathbf{Aw}) = k(\lambda\mathbf{w}) = \lambda(k\mathbf{w})$; hence $\mathbf{A}(k\mathbf{w} + \ell\mathbf{x}) = \lambda(k\mathbf{w} + \ell\mathbf{x})$. ■

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 (see Sec. 7.9). For instance, $\mathbf{x}_1 = [1 \quad 2]^\mathsf{T}$ in Example 1 has the length $\|\mathbf{x}_1\| = \sqrt{1^2 + 2^2} = \sqrt{5}$; hence $[1/\sqrt{5} \quad 2/\sqrt{5}]^\mathsf{T}$ is a normalized eigenvector (a unit eigenvector).

EXAMPLE 2 Multiple Eigenvalues

Find the eigenvalues and eigenvectors of

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

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 $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$. (If you have trouble finding roots, you may want to use a root finding algorithm such as Newton's method (Sec. 19.2). Your CAS or scientific calculator can find roots. However, to really learn and remember this material, you have to do some exercises with paper and pencil.) To find eigenvectors, we apply the Gauss elimination (Sec. 7.3) 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

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} - 5\mathbf{I} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}. \quad \text{It row-reduces to} \quad \begin{bmatrix} -7 & 2 & -3 \\ 0 & -\frac{24}{7} & -\frac{48}{7} \\ 0 & 0 & 0 \end{bmatrix}.$$

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 $\mathbf{x}_1 = [1 \quad 2 \quad -1]^\mathsf{T}$.

For $\lambda = -3$ the characteristic matrix

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix} \quad \text{row-reduces to} \quad \begin{bmatrix} 1 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Hence it has rank 1. From $x_1 + 2x_2 - 3x_3 = 0$ we have $x_1 = -2x_2 + 3x_3$. Choosing $x_2 = 1, x_3 = 0$ and $x_2 = 0, x_3 = 1$, we obtain two linearly independent eigenvectors of $\mathbf{A}$ corresponding to $\lambda = -3$ [as they must exist by (5), Sec. 7.5, with rank $= 1$ and $n = 3$],

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

and

$$\mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

EXAMPLE 4 Real Matrices with Complex Eigenvalues and Eigenvectors

Since real polynomials may have complex roots (which then occur in conjugate pairs), a real matrix may have complex eigenvalues and eigenvectors. For instance, the characteristic equation of the skew-symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \qquad \text{is} \qquad \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = 0.$$

It gives the eigenvalues $\lambda_1 = i\,(=\sqrt{-1})$, $\lambda_2 = -i$. Eigenvectors are obtained from $-ix_1 + x_2 = 0$ and $ix_1 + x_2 = 0$, respectively, and we can choose $x_1 = 1$ to get

$$\begin{bmatrix} 1 \\ i \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

■

THEOREM 3

Eigenvalues of the Transpose

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

8.2. 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) $$\mathbf{A}^\mathsf{T} = \mathbf{A}, \qquad \text{thus} \qquad a_{kj} = a_{jk},$$

skew-symmetric if transposition gives the negative of **A**,

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

orthogonal if transposition gives the inverse of **A**,

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

EXAMPLE 1 Symmetric, Skew-Symmetric, and Orthogonal Matrices

The matrices

$$\begin{bmatrix} -3 & 1 & 5 \\ 1 & 0 & -2 \\ 5 & -2 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

are symmetric, skew-symmetric, and orthogonal, respectively, as you should verify.

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

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

EXAMPLE 2 Illustration of Formula (4)

$$\mathbf{A} = \begin{bmatrix} 9 & 5 & 2 \\ 2 & 3 & -8 \\ 5 & 4 & 3 \end{bmatrix} = \mathbf{R} + \mathbf{S} = \begin{bmatrix} 9.0 & 3.5 & 3.5 \\ 3.5 & 3.0 & -2.0 \\ 3.5 & -2.0 & 3.0 \end{bmatrix} + \begin{bmatrix} 0 & 1.5 & -1.5 \\ -1.5 & 0 & -6.0 \\ 1.5 & 6.0 & 0 \end{bmatrix}$$

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

$$\mathbf{y} = \mathbf{A}\mathbf{x} \qquad (5)$$

where **A** is an 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

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

is an orthogonal transformation. It can be shown that any orthogonal transformation in the plane or in three-dimensional space is a **rotation**

THEOREM 2

Invariance of Inner Product

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

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^\mathsf{T}\mathbf{b} = [a_1 \quad \cdots \quad a_n] \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}. \tag{7}$$

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

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

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\mathbf{a}^\mathsf{T}\mathbf{a}}. \tag{8}$$

PROOF

Let **A** be orthogonal. Let $\mathbf{u} = \mathbf{Aa}$ and $\mathbf{v} = \mathbf{Ab}$. We must show that $\mathbf{u} \cdot \mathbf{v} = \mathbf{a} \cdot \mathbf{b}$. Now $(\mathbf{Aa})^\mathsf{T} = \mathbf{a}^\mathsf{T}\mathbf{A}^\mathsf{T}$ by (10d) in Sec. 7.2 and $\mathbf{A}^\mathsf{T}\mathbf{A} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ by (3). Hence

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\mathsf{T}\mathbf{v} = (\mathbf{Aa})^\mathsf{T}\mathbf{Ab} = \mathbf{a}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{Ab} = \mathbf{a}^\mathsf{T}\mathbf{Ib} = \mathbf{a}^\mathsf{T}\mathbf{b} = \mathbf{a} \cdot \mathbf{b}. \tag{9}$$

THEOREM 3

Orthonormality of Column and Row Vectors

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

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

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.

EXAMPLE 5 Eigenvalues of an Orthogonal Matrix

The orthogonal matrix in Example 1

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

has the characteristic equation

$$-\lambda^3 + \tfrac{2}{3}\lambda^2 + \tfrac{2}{3}\lambda - 1 = 0.$$

eigenvalues

$$\begin{cases} -1 \\ (5 + i\sqrt{11})/6 \\ (5 - i\sqrt{11})/6 \end{cases}$$

8.5 Complex Matrices and Forms

The three classes of matrices in Sec. 8.3 have complex counterparts which are of practical interest in certain applications, for instance, in quantum mechanics.

Notations

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

EXAMPLE 1 Notations

If $\mathbf{A} = \begin{bmatrix} 3+4i & 1-i \\ 6 & 2-5i \end{bmatrix}$, then $\overline{\mathbf{A}} = \begin{bmatrix} 3-4i & 1+i \\ 6 & 2+5i \end{bmatrix}$ and $\overline{\mathbf{A}}^{\mathsf{T}} = \begin{bmatrix} 3-4i & 6 \\ 1+i & 2+5i \end{bmatrix}$.

DEFINITION

Hermitian, Skew-Hermitian, and Unitary Matrices

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

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

From the definitions we see the following. If $\mathbf{A}$ is Hermitian, the entries on the main diagonal must satisfy $\overline{a}_{jj} = a_{jj}$; that is, they are real. Similarly, if $\mathbf{A}$ is skew-Hermitian, then $\overline{a}_{jj} = -a_{jj}$. If we set $a_{jj} = \alpha + i\beta$, this becomes $\alpha - i\beta = -(\alpha + i\beta)$. Hence $\alpha = 0$, so that a_{jj} must be pure imaginary or 0.

EXAMPLE 2 Hermitian, Skew-Hermitian, and Unitary Matrices

$$\mathbf{A} = \begin{bmatrix} 4 & 1-3i \\ 1+3i & 7 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} 3i & 2+i \\ -2+i & -i \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \frac{1}{2}i & \frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2}i \end{bmatrix}$$

If a Hermitian matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = \mathbf{A}$. Hence a real Hermitian matrix is a symmetric matrix (Sec. 8.3).

Similarly, if a skew-Hermitian matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = -\mathbf{A}$. Hence a real skew-Hermitian matrix is a skew-symmetric matrix.

Finally, if a unitary matrix is real, then $\overline{\mathbf{A}}^{\mathsf{T}} = \mathbf{A}^{\mathsf{T}} = \mathbf{A}^{-1}$. Hence a real unitary matrix is an orthogonal matrix.

This shows that *Hermitian, skew-Hermitian, and unitary matrices generalize symmetric, skew-symmetric, and orthogonal matrices, respectively.*

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.

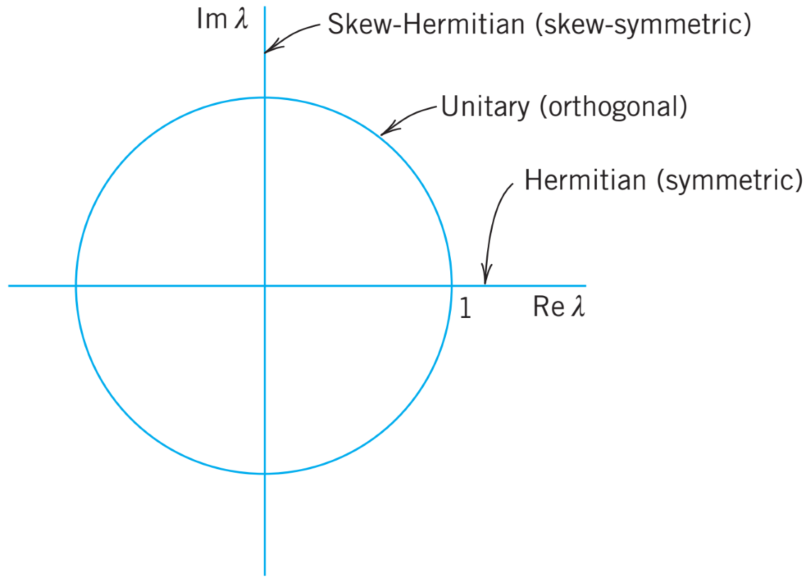

Fig. 163. Location of the eigenvalues of Hermitian, skew-Hermitian and unitary matrices in the complex λ-plane

EXAMPLE 3 Illustration of Theorem 1

For the matrices in Example 2 we find by direct calculation

Matrix		Characteristic Equation	Eigenvalues
A	Hermitian	$\lambda^2 - 11\lambda + 18 = 0$	$9, \quad 2$
B	Skew-Hermitian	$\lambda^2 - 2i\lambda + 8 = 0$	$4i, \quad -2i$
C	Unitary	$\lambda^2 - i\lambda - 1 = 0$	$\frac{1}{2}\sqrt{3} + \frac{1}{2}i, \quad -\frac{1}{2}\sqrt{3} + \frac{1}{2}i$

Key properties of orthogonal matrices (invariance of the inner product, orthonormality of rows and columns; see Sec. 8.3) generalize to unitary matrices in a remarkable way.

To see this, instead of R^n we now use the **complex vector space** C^n of all complex vectors with n complex numbers as components, and complex numbers as scalars. For such complex vectors the **inner product** is defined by (note the overbar for the complex conjugate)

$$\mathbf{a} \bullet \mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{b}. \tag{4}$$

The **length** or **norm** of such a complex vector is a *real* number defined by

$$\|\mathbf{a}\| = \sqrt{\mathbf{a} \bullet \mathbf{a}} = \sqrt{\overline{\mathbf{a}}_j^{\mathsf{T}}\mathbf{a}} = \sqrt{\overline{a}_1 a_1 + \cdots + \overline{a}_n a_n} = \sqrt{|a_1|^2 + \cdots + |a_n|^2}. \tag{5}$$

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).

PROOF $\mathbf{u} \bullet \mathbf{v} = \overline{\mathbf{u}}^{\mathsf{T}}\mathbf{v} = (\overline{\mathbf{A}\mathbf{a}})^{\mathsf{T}}\mathbf{A}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\overline{\mathbf{A}}^{\mathsf{T}}\mathbf{A}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{I}\mathbf{b} = \overline{\mathbf{a}}^{\mathsf{T}}\mathbf{b} = \mathbf{a} \bullet \mathbf{b}.$

DEFINITION

Unitary System

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

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

THEOREM 3

Unitary Systems of Column and Row Vectors

A complex square matrix is unitary if and only if its column vectors (*and also its row vectors*) *form a unitary system.*

THEOREM 4

Determinant of a Unitary Matrix

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

EXAMPLE 4

Unitary Matrix Illustrating Theorems 1c and 2–4

For the vectors $\mathbf{a}^{\mathsf{T}} = [2 \quad -i]$ and $\mathbf{b}^{\mathsf{T}} = [1 + i \quad 4i]$ we get $\bar{\mathbf{a}}^{\mathsf{T}} = [2 \quad i]^{\mathsf{T}}$ and $\bar{\mathbf{a}}^{\mathsf{T}}\mathbf{b} = 2(1 + i) - 4 = -2 + 2i$ and with

$$\mathbf{A} = \begin{bmatrix} 0.8i & 0.6 \\ 0.6 & 0.8i \end{bmatrix} \qquad \text{also} \qquad \mathbf{Aa} = \begin{bmatrix} i \\ 2 \end{bmatrix} \qquad \text{and} \qquad \mathbf{Ab} = \begin{bmatrix} -0.8 + 3.2i \\ -2.6 + 0.6i \end{bmatrix},$$

as one can readily verify. This gives $(\overline{\mathbf{A}\mathbf{a}})^{\mathsf{T}}\mathbf{Ab} = -2 + 2i$, illustrating Theorem 2. The matrix is unitary. Its columns form a unitary system,

$$\bar{\mathbf{a}}_1^{\mathsf{T}}\mathbf{a}_1 = -0.8i \cdot 0.8i + 0.6^2 = 1, \qquad \bar{\mathbf{a}}_1^{\mathsf{T}}\mathbf{a}_2 = -0.8i \cdot 0.6 + 0.6 \cdot 0.8i = 0,$$

$$\bar{\mathbf{a}}_2^{\mathsf{T}}\mathbf{a}_2 = 0.6^2 + (-0.8i)0.8i = 1$$

and so do its rows. Also, $\det \mathbf{A} = -1$. The eigenvalues are $0.6 + 0.8i$ and $-0.6 + 0.8i$, with eigenvectors $[1 \quad 1]^{\mathsf{T}}$ and $[1 \quad -1]^{\mathsf{T}}$, respectively. ■