



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 $\mathbf{A}$** .

From this rather innocent looking vector equation flows an amazing amount of relevant theory and an incredible richness of applications. Indeed, eigenvalue problems come up all the time in engineering, physics, geometry, numerics, theoretical mathematics, biology, environmental science, urban planning, economics, psychology, and other areas. Thus, in your career you are likely to encounter eigenvalue problems.

We start with a basic and thorough introduction to eigenvalue problems in Sec. 8.1 and explain (1) with several simple matrices. This is followed by a section devoted entirely to applications ranging from mass–spring systems of physics to population control models of environmental science. We show you these diverse examples to train your skills in modeling and solving eigenvalue problems. Eigenvalue problems for real symmetric, skew-symmetric, and orthogonal matrices are discussed in Sec. 8.3 and their complex counterparts (which are important in modern physics) in Sec. 8.5. In Sec. 8.4 we show how by diagonalizing a matrix, we obtain its eigenvalues.

COMMENT. *Numerics for eigenvalues (Secs. 20.6–20.9) can be studied immediately after this chapter.*

Prerequisite: Chap. 7.

Sections that may be omitted in a shorter course: 8.4, 8.5.

References and Answers to Problems: App. 1 Part B, App. 2.

The following chart identifies where different types of eigenvalue problems appear in the book.

Topic	Where to find it
Matrix Eigenvalue Problem (algebraic eigenvalue problem)	Chap. 8
Eigenvalue Problems in Numerics	Secs. 20.6–20.9
Eigenvalue Problem for ODEs (Sturm–Liouville problems)	Secs. 11.5, 11.6
Eigenvalue Problems for Systems of ODEs	Chap. 4
Eigenvalue Problems for PDEs	Secs. 12.3–12.11

8.1 The Matrix Eigenvalue Problem. Determining Eigenvalues and Eigenvectors

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

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

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 $[30 \ 40]^T = 10 [3 \ 4]^T$, 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.

4.0教過求
eigenvalue, eigenvector

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.

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. (This was all demonstrated in our intuitive opening example. Can you see that the second equation in that example satisfies (1) with $\lambda = 10$ and $\mathbf{x} = [3 \ 4]^T$, and $\mathbf{A}$ the given 2×2 matrix? Write it out.) Now why do we require $\mathbf{x}$ to be nonzero? The reason is that $\mathbf{x} = \mathbf{0}$ is always a solution of (1) for any value of λ , because $\mathbf{A}\mathbf{0} = \mathbf{0}$. This is of no interest.

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

Now, with the new terminology for (1), we can just say that the problem of determining the eigenvalues and eigenvectors of a matrix is called an eigenvalue problem. (However, more precisely, we are considering an algebraic eigenvalue problem, as opposed to an eigenvalue problem involving an ODE or PDE, as considered in Secs. 11.5 and 12.3, or an integral equation.)

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

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

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

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

This can be written in matrix notation

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

because (1) is $\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{A}\mathbf{x} - \lambda\mathbf{I}\mathbf{x} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, 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,

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

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 $\mathbf{A}$.

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

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

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

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

(b₂) *Eigenvector of $\mathbf{A}$ corresponding to λ_2 .* For $\lambda = \lambda_2 = -6$, equation (2*) becomes

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

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 $\mathbf{A}$ corresponding to $\lambda_2 = -6$ is

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

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 $[3 \ 4]^T$ and $[-1 \ 1]^T$, respectively. The reader may want to verify this. ■

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

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

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

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

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:

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

$\mathbf{A} - \lambda \mathbf{I}$ 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 $\mathbf{A}$.

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.

For larger n , the actual computation of eigenvalues will, in general, require the use of Newton's method (Sec. 19.2) or another numeric approximation method in Secs. 20.7–20.9.

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. This is what we did in Example 1 and shall do again in the examples below. (To prevent misunderstandings: *numeric approximation methods*, such as in Sec. 20.8, may determine *eigenvectors first*.)

Eigenvectors have the following properties.

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 $\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$ and $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ imply $\mathbf{A}(\mathbf{w} + \mathbf{x}) = \mathbf{A}\mathbf{w} + \mathbf{A}\mathbf{x} = \lambda\mathbf{w} + \lambda\mathbf{x} = \lambda(\mathbf{w} + \mathbf{x})$ and $\mathbf{A}(k\mathbf{w}) = k(\mathbf{A}\mathbf{w}) = k(\lambda\mathbf{w}) = \lambda(k\mathbf{w})$; 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, $\mathbf{x}_1 = [1 \ 2]^T$ in Example 1 has the length $\|\mathbf{x}_1\| = \sqrt{1^2 + 2^2} = \sqrt{5}$; hence $[1/\sqrt{5} \ 2/\sqrt{5}]^T$ is a normalized eigenvector (a unit eigenvector).

Examples 2 and 3 will illustrate that an $n \times n$ matrix may have n linearly independent eigenvectors, or it may have fewer than n . In Example 4 we shall see that a *real* matrix may have *complex* eigenvalues and eigenvectors.

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

For $\lambda = -3$ the characteristic matrix

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

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

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

名詞，不算太重要。
可由下面例子看出其意義。

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 $m_\lambda = M_\lambda = 2$. In general, $m_\lambda \leq M_\lambda$, 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:

EXAMPLE 3 Algebraic Multiplicity, Geometric Multiplicity. Positive Defect

The characteristic equation of the matrix

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

Hence $\lambda = 0$ is an eigenvalue of algebraic multiplicity $M_0 = 2$. But its geometric multiplicity is only $m_0 = 1$, since eigenvectors result from $-0x_1 + x_2 = 0$, hence $x_2 = 0$, in the form $[x_1 \ 0]^T$. Hence for $\lambda = 0$ the defect is $\Delta_0 = 1$.

Similarly, the characteristic equation of the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{is} \quad \det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} 3 - \lambda & 2 \\ 0 & 3 - \lambda \end{vmatrix} = (3 - \lambda)^2 = 0.$$

Hence $\lambda = 3$ is an eigenvalue of algebraic multiplicity $M_3 = 2$, but its geometric multiplicity is only $m_3 = 1$, since eigenvectors result from $0x_1 + 2x_2 = 0$ in the form $[x_1 \ 0]^T$. ■

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

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

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

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

In the next section we shall need the following simple theorem.

THEOREM 3

Eigenvalues of the Transpose

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

PROOF Transposition does not change the value of the characteristic determinant, as follows from Theorem 2d in Sec. 7.7. ■

Having gained a first impression of matrix eigenvalue problems, we shall illustrate their importance with some typical applications in Sec. 8.2.

PROBLEM SET 8.1

1–16 EIGENVALUES, EIGENVECTORS

Find the eigenvalues. Find the corresponding eigenvectors. Use the given λ or factor in Probs. 11 and 15.

1. $\begin{bmatrix} 3.0 & 0 \\ 0 & -0.6 \end{bmatrix}$

2. $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 5 & -2 \\ 9 & -6 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

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

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

7. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

8. $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

9. $\begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$

10. $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$

11. $\begin{bmatrix} 6 & 2 & -2 \\ 2 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix}, \lambda = 3$

12. $\begin{bmatrix} 3 & 5 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 13 & 5 & 2 \\ 2 & 7 & -8 \\ 5 & 4 & 7 \end{bmatrix}$

14. $\begin{bmatrix} 2 & 0 & -1 \\ 0 & \frac{1}{2} & 0 \\ 1 & 0 & 4 \end{bmatrix}$

15. $\begin{bmatrix} -1 & 0 & 12 & 0 \\ 0 & -1 & 0 & 12 \\ 0 & 0 & -1 & -4 \\ 0 & 0 & -4 & -1 \end{bmatrix}, (\lambda + 1)^2$

16. $\begin{bmatrix} -3 & 0 & 4 & 2 \\ 0 & 1 & -2 & 4 \\ 2 & 4 & -1 & -2 \\ 0 & 2 & -2 & 3 \end{bmatrix}$

17–20 LINEAR TRANSFORMATIONS AND EIGENVALUES

Find the matrix $\mathbf{A}$ in the linear transformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{x} = [x_1 \ x_2]^T$ ($\mathbf{x} = [x_1 \ x_2 \ x_3]^T$) are Cartesian coordinates. Find the eigenvalues and eigenvectors and explain their geometric meaning.

17. Counterclockwise rotation through the angle $\pi/2$ about the origin in R^2 .

18. Reflection about the x_1 -axis in R^2 .

19. Orthogonal projection (perpendicular projection) of R^2 onto the x_2 -axis.

20. Orthogonal projection of R^3 onto the plane $x_2 = x_1$.

21–25 GENERAL PROBLEMS

21. **Nonzero defect.** Find further 2×2 and 3×3 matrices with positive defect. See Example 3.

22. **Multiple eigenvalues.** Find further 2×2 and 3×3 matrices with multiple eigenvalues. See Example 2.

23. **Complex eigenvalues.** Show that the eigenvalues of a real matrix are real or complex conjugate in pairs.

24. **Inverse matrix.** Show that $\mathbf{A}^{-1}$ exists if and only if the eigenvalues $\lambda_1, \dots, \lambda_n$ are all nonzero, and then $\mathbf{A}^{-1}$ has the eigenvalues $1/\lambda_1, \dots, 1/\lambda_n$.

25. **Transpose.** Illustrate Theorem 3 with examples of your own.

8.2 Some Applications of Eigenvalue Problems

We have selected some typical examples from the wide range of applications of matrix eigenvalue problems. The last example, that is, Example 4, shows an application involving vibrating springs and ODEs. It falls into the domain of Chapter 4, which covers matrix eigenvalue problems related to ODE's modeling mechanical systems and electrical

networks. Example 4 is included to keep our discussion independent of Chapter 4. (However, the reader not interested in ODEs may want to skip Example 4 without loss of continuity.)

EXAMPLE 1

Stretching of an Elastic Membrane

An elastic membrane in the x_1x_2 -plane with boundary circle $x_1^2 + x_2^2 = 1$ (Fig. 160) is stretched so that a point $P: (x_1, x_2)$ goes over into the point $Q: (y_1, y_2)$ given by

$$(1) \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{A}\mathbf{x} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \quad \text{in components,} \quad \begin{aligned} y_1 &= 5x_1 + 3x_2 \\ y_2 &= 3x_1 + 5x_2. \end{aligned}$$

Find the **principal directions**, that is, the directions of the position vector $\mathbf{x}$ of P for which the direction of the position vector $\mathbf{y}$ of Q is the same or exactly opposite. What shape does the boundary circle take under this deformation?

Solution. We are looking for vectors $\mathbf{x}$ such that $\mathbf{y} = \lambda\mathbf{x}$. Since $\mathbf{y} = \mathbf{A}\mathbf{x}$, this gives $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, the equation of an eigenvalue problem. In components, $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ is

$$(2) \quad \begin{aligned} 5x_1 + 3x_2 &= \lambda x_1 & \text{or} & & (5 - \lambda)x_1 + 3x_2 &= 0 \\ 3x_1 + 5x_2 &= \lambda x_2 & & & 3x_1 + (5 - \lambda)x_2 &= 0. \end{aligned}$$

The characteristic equation is

$$(3) \quad \begin{vmatrix} 5 - \lambda & 3 \\ 3 & 5 - \lambda \end{vmatrix} = (5 - \lambda)^2 - 9 = 0.$$

Its solutions are $\lambda_1 = 8$ and $\lambda_2 = 2$. These are the eigenvalues of our problem. For $\lambda = \lambda_1 = 8$, our system (2) becomes

$$\begin{aligned} -3x_1 + 3x_2 &= 0, & \text{Solution } x_2 &= x_1, \text{ } x_1 \text{ arbitrary,} \\ 3x_1 - 3x_2 &= 0. & \text{for instance, } x_1 &= x_2 = 1. \end{aligned}$$

For $\lambda_2 = 2$, our system (2) becomes

$$\begin{aligned} 3x_1 + 3x_2 &= 0, & \text{Solution } x_2 &= -x_1, \text{ } x_1 \text{ arbitrary,} \\ 3x_1 + 3x_2 &= 0. & \text{for instance, } x_1 &= 1, x_2 = -1. \end{aligned}$$

We thus obtain as eigenvectors of $\mathbf{A}$, for instance, $[1 \ 1]^T$ corresponding to λ_1 and $[1 \ -1]^T$ corresponding to λ_2 (or a nonzero scalar multiple of these). These vectors make 45° and 135° angles with the positive x_1 -direction. They give the principal directions, the answer to our problem. The eigenvalues show that in the principal directions the membrane is stretched by factors 8 and 2, respectively; see Fig. 160.

Accordingly, if we choose the principal directions as directions of a new Cartesian u_1u_2 -coordinate system, say, with the positive u_1 -semi-axis in the first quadrant and the positive u_2 -semi-axis in the second quadrant of the x_1x_2 -system, and if we set $u_1 = r \cos \phi$, $u_2 = r \sin \phi$, then a boundary point of the unstretched circular membrane has coordinates $\cos \phi$, $\sin \phi$. Hence, after the stretch we have

$$z_1 = 8 \cos \phi, \quad z_2 = 2 \sin \phi.$$

Since $\cos^2 \phi + \sin^2 \phi = 1$, this shows that the deformed boundary is an ellipse (Fig. 160)

$$(4) \quad \frac{z_1^2}{8^2} + \frac{z_2^2}{2^2} = 1.$$

這一題重要，但在
p.344第6題還會再看
一次。

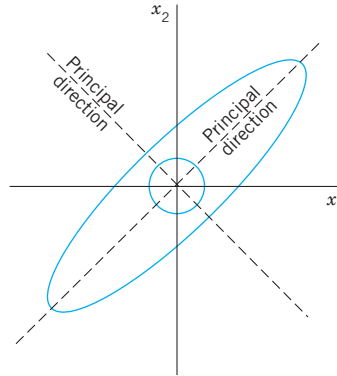


Fig. 160. Undeformed and deformed membrane in Example 1

EXAMPLE 2 Eigenvalue Problems Arising from Markov Processes

例2~4不必看

Markov processes as considered in Example 13 of Sec. 7.2 lead to eigenvalue problems if we ask for the limit state of the process in which the state vector $\mathbf{x}$ is reproduced under the multiplication by the stochastic matrix $\mathbf{A}$ governing the process, that is, $\mathbf{A}\mathbf{x} = \mathbf{x}$. Hence $\mathbf{A}$ should have the eigenvalue 1, and $\mathbf{x}$ should be a corresponding eigenvector. This is of practical interest because it shows the long-term tendency of the development modeled by the process.

In that example,

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.1 & 0 \\ 0.2 & 0.9 & 0.2 \\ 0.1 & 0 & 0.8 \end{bmatrix}. \quad \text{For the transpose,} \quad \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.9 & 0 \\ 0 & 0.2 & 0.8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Hence $\mathbf{A}^T$ has the eigenvalue 1, and the same is true for $\mathbf{A}$ by Theorem 3 in Sec. 8.1. An eigenvector $\mathbf{x}$ of $\mathbf{A}$ for $\lambda = 1$ is obtained from

$$\mathbf{A} - \mathbf{I} = \begin{bmatrix} -0.3 & 0.1 & 0 \\ 0.2 & -0.1 & 0.2 \\ 0.1 & 0 & -0.2 \end{bmatrix}, \quad \text{row-reduced to} \quad \begin{bmatrix} -\frac{3}{10} & \frac{1}{10} & 0 \\ 0 & -\frac{1}{30} & \frac{1}{5} \\ 0 & 0 & 0 \end{bmatrix}.$$

Taking $x_3 = 1$, we get $x_2 = 6$ from $-x_2/30 + x_3/5 = 0$ and then $x_1 = 2$ from $-3x_1/10 + x_2/10 = 0$. This gives $\mathbf{x} = [2 \ 6 \ 1]^T$. It means that in the long run, the ratio Commercial:Industrial:Residential will approach 2:6:1, provided that the probabilities given by $\mathbf{A}$ remain (about) the same. (We switched to ordinary fractions to avoid rounding errors.)

EXAMPLE 3 Eigenvalue Problems Arising from Population Models. Leslie Model

The Leslie model describes age-specified population growth, as follows. Let the oldest age attained by the females in some animal population be 9 years. Divide the population into three age classes of 3 years each. Let the “Leslie matrix” be

$$(5) \quad \mathbf{L} = [l_{jk}] = \begin{bmatrix} 0 & 2.3 & 0.4 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}$$

where l_{1k} is the average number of daughters born to a single female during the time she is in age class k , and $l_{j,j-1}$ ($j = 2, 3$) is the fraction of females in age class $j - 1$ that will survive and pass into class j . (a) What is the number of females in each class after 3, 6, 9 years if each class initially consists of 400 females? (b) For what initial distribution will the number of females in each class change by the same proportion? What is this rate of change?

Solution. (a) Initially, $\mathbf{x}_{(0)}^T = [400 \ 400 \ 400]$. After 3 years,

$$\mathbf{x}_{(3)} = \mathbf{L}\mathbf{x}_{(0)} = \begin{bmatrix} 0 & 2.3 & 0.4 \\ 0.6 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix} \begin{bmatrix} 400 \\ 400 \\ 400 \end{bmatrix} = \begin{bmatrix} 1080 \\ 240 \\ 120 \end{bmatrix}.$$

Similarly, after 6 years the number of females in each class is given by $\mathbf{x}_{(6)}^T = (\mathbf{L}\mathbf{x}_{(3)})^T = [600 \ 648 \ 72]$, and after 9 years we have $\mathbf{x}_{(9)}^T = (\mathbf{L}\mathbf{x}_{(6)})^T = [1519.2 \ 360 \ 194.4]$.

(b) Proportional change means that we are looking for a distribution vector $\mathbf{x}$ such that $\mathbf{L}\mathbf{x} = \lambda\mathbf{x}$, where λ is the rate of change (growth if $\lambda > 1$, decrease if $\lambda < 1$). The characteristic equation is (develop the characteristic determinant by the first column)

$$\det(\mathbf{L} - \lambda\mathbf{I}) = -\lambda^3 - 0.6(-2.3\lambda - 0.3 \cdot 0.4) = -\lambda^3 + 1.38\lambda + 0.072 = 0.$$

A positive root is found to be (for instance, by Newton's method, Sec. 19.2) $\lambda = 1.2$. A corresponding eigenvector $\mathbf{x}$ can be determined from the characteristic matrix

$$\mathbf{A} - 1.2\mathbf{I} = \begin{bmatrix} -1.2 & 2.3 & 0.4 \\ 0.6 & -1.2 & 0 \\ 0 & 0.3 & -1.2 \end{bmatrix}, \quad \text{say,} \quad \mathbf{x} = \begin{bmatrix} 1 \\ 0.5 \\ 0.125 \end{bmatrix}$$

where $x_3 = 0.125$ is chosen, $x_2 = 0.5$ then follows from $0.3x_2 - 1.2x_3 = 0$, and $x_1 = 1$ from $-1.2x_1 + 2.3x_2 + 0.4x_3 = 0$. To get an initial population of 1200 as before, we multiply $\mathbf{x}$ by $1200/(1 + 0.5 + 0.125) = 738$. *Answer:* Proportional growth of the numbers of females in the three classes will occur if the initial values are 738, 369, 92 in classes 1, 2, 3, respectively. The growth rate will be 1.2 per 3 years. ■

EXAMPLE 4 Vibrating System of Two Masses on Two Springs (Fig. 161)

Mass-spring systems involving several masses and springs can be treated as eigenvalue problems. For instance, the mechanical system in Fig. 161 is governed by the system of ODEs

$$\begin{aligned} (6) \quad y_1'' &= -3y_1 - 2(y_1 - y_2) = -5y_1 + 2y_2 \\ y_2'' &= -2(y_2 - y_1) = 2y_1 - 2y_2 \end{aligned}$$

where y_1 and y_2 are the displacements of the masses from rest, as shown in the figure, and primes denote derivatives with respect to time t . In vector form, this becomes

$$(7) \quad \mathbf{y}'' = \begin{bmatrix} y_1'' \\ y_2'' \end{bmatrix} = \mathbf{A}\mathbf{y} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

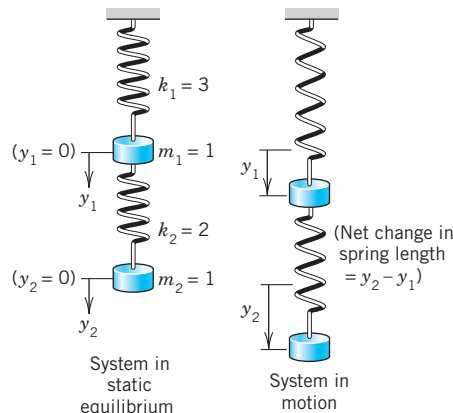


Fig. 161. Masses on springs in Example 4

We try a vector solution of the form

$$(8) \quad \mathbf{y} = \mathbf{x}e^{\omega t}.$$

This is suggested by a mechanical system of a single mass on a spring (Sec. 2.4), whose motion is given by exponential functions (and sines and cosines). Substitution into (7) gives

$$\omega^2 \mathbf{x}e^{\omega t} = \mathbf{A}\mathbf{x}e^{\omega t}.$$

Dividing by $e^{\omega t}$ and writing $\omega^2 = \lambda$, we see that our mechanical system leads to the eigenvalue problem

$$(9) \quad \mathbf{A}\mathbf{x} = \lambda\mathbf{x} \quad \text{where } \lambda = \omega^2.$$

From Example 1 in Sec. 8.1 we see that $\mathbf{A}$ has the eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -6$. Consequently, $\omega = \pm\sqrt{-1} = \pm i$ and $\sqrt{-6} = \pm i\sqrt{6}$, respectively. Corresponding eigenvectors are

$$(10) \quad \mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

From (8) we thus obtain the four complex solutions [see (10), Sec. 2.2]

$$\begin{aligned} \mathbf{x}_1 e^{\pm it} &= \mathbf{x}_1 (\cos t \pm i \sin t), \\ \mathbf{x}_2 e^{\pm i\sqrt{6}t} &= \mathbf{x}_2 (\cos \sqrt{6}t \pm i \sin \sqrt{6}t). \end{aligned}$$

By addition and subtraction (see Sec. 2.2) we get the four real solutions

$$\mathbf{x}_1 \cos t, \quad \mathbf{x}_1 \sin t, \quad \mathbf{x}_2 \cos \sqrt{6}t, \quad \mathbf{x}_2 \sin \sqrt{6}t.$$

A general solution is obtained by taking a linear combination of these,

$$\mathbf{y} = \mathbf{x}_1(a_1 \cos t + b_1 \sin t) + \mathbf{x}_2(a_2 \cos \sqrt{6}t + b_2 \sin \sqrt{6}t)$$

with arbitrary constants a_1, b_1, a_2, b_2 (to which values can be assigned by prescribing initial displacement and initial velocity of each of the two masses). By (10), the components of $\mathbf{y}$ are

$$\begin{aligned} y_1 &= a_1 \cos t + b_1 \sin t + 2a_2 \cos \sqrt{6}t + 2b_2 \sin \sqrt{6}t \\ y_2 &= 2a_1 \cos t + 2b_1 \sin t - a_2 \cos \sqrt{6}t - b_2 \sin \sqrt{6}t. \end{aligned}$$

These functions describe harmonic oscillations of the two masses. Physically, this had to be expected because we have neglected damping. ■

PROBLEM SET 8.2

1–6 ELASTIC DEFORMATIONS

Given $\mathbf{A}$ in a deformation $\mathbf{y} = \mathbf{A}\mathbf{x}$, find the principal directions and corresponding factors of extension or contraction. Show the details.

1. $\begin{bmatrix} 3.0 & 1.5 \\ 1.5 & 3.0 \end{bmatrix}$

2. $\begin{bmatrix} 2.0 & 0.4 \\ 0.4 & 2.0 \end{bmatrix}$

3. $\begin{bmatrix} 7 & \sqrt{6} \\ \sqrt{6} & 2 \end{bmatrix}$

4. $\begin{bmatrix} 5 & 2 \\ 2 & 13 \end{bmatrix}$

5. $\begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$

6. $\begin{bmatrix} 1.25 & 0.75 \\ 0.75 & 1.25 \end{bmatrix}$

7–9 MARKOV PROCESSES

Find the limit state of the Markov process modeled by the given matrix. Show the details.

7. $\begin{bmatrix} 0.2 & 0.5 \\ 0.8 & 0.5 \end{bmatrix}$

8. $\begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.1 & 0.6 \end{bmatrix}$

9. $\begin{bmatrix} 0.6 & 0.1 & 0.2 \\ 0.4 & 0.1 & 0.4 \\ 0 & 0.8 & 0.4 \end{bmatrix}$

10–12 AGE-SPECIFIC POPULATION

Find the growth rate in the Leslie model (see Example 3) with the matrix as given. Show the details.

$$10. \begin{bmatrix} 0 & 9.0 & 5.0 \\ 0.4 & 0 & 0 \\ 0 & 0.4 & 0 \end{bmatrix} \quad 11. \begin{bmatrix} 0 & 3.45 & 0.60 \\ 0.90 & 0 & 0 \\ 0 & 0.45 & 0 \end{bmatrix}$$

$$12. \begin{bmatrix} 0 & 3.0 & 2.0 & 2.0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix}$$

13–15 LEONTIEF MODELS¹

- 13. Leontief input–output model.** Suppose that three industries are interrelated so that their outputs are used as inputs by themselves, according to the 3×3 consumption matrix

$$\mathbf{A} = [a_{jk}] = \begin{bmatrix} 0.1 & 0.5 & 0 \\ 0.8 & 0 & 0.4 \\ 0.1 & 0.5 & 0.6 \end{bmatrix}$$

where a_{jk} is the fraction of the output of industry k consumed (purchased) by industry j . Let p_j be the price charged by industry j for its total output. A problem is to find prices so that for each industry, total expenditures equal total income. Show that this leads to $\mathbf{A}\mathbf{p} = \mathbf{p}$, where $\mathbf{p} = [p_1 \ p_2 \ p_3]^T$, and find a solution $\mathbf{p}$ with nonnegative p_1, p_2, p_3 .

- 14.** Show that a consumption matrix as considered in Prob. 13 must have column sums 1 and always has the eigenvalue 1.
- 15. Open Leontief input–output model.** If not the whole output but only a portion of it is consumed by the

industries themselves, then instead of $\mathbf{A}\mathbf{x} = \mathbf{x}$ (as in Prob. 13), we have $\mathbf{x} - \mathbf{A}\mathbf{x} = \mathbf{y}$, where $\mathbf{x} = [x_1 \ x_2 \ x_3]^T$ is produced, $\mathbf{A}\mathbf{x}$ is consumed by the industries, and, thus, $\mathbf{y}$ is the net production available for other consumers. Find for what production $\mathbf{x}$ a given demand vector $\mathbf{y} = [0.1 \ 0.3 \ 0.1]^T$ can be achieved if the consumption matrix is

$$\mathbf{A} = \begin{bmatrix} 0.1 & 0.4 & 0.2 \\ 0.5 & 0 & 0.1 \\ 0.1 & 0.4 & 0.4 \end{bmatrix}.$$

16–20 GENERAL PROPERTIES OF EIGENVALUE PROBLEMS

Let $\mathbf{A} = [a_{jk}]$ be an $n \times n$ matrix with (not necessarily distinct) eigenvalues $\lambda_1, \dots, \lambda_n$. Show.

- 16. Trace.** The sum of the main diagonal entries, called the *trace* of $\mathbf{A}$, equals the sum of the eigenvalues of $\mathbf{A}$.
- 17. “Spectral shift.”** $\mathbf{A} - k\mathbf{I}$ has the eigenvalues $\lambda_1 - k, \dots, \lambda_n - k$ and the same eigenvectors as $\mathbf{A}$.
- 18. Scalar multiples, powers.** $k\mathbf{A}$ has the eigenvalues $k\lambda_1, \dots, k\lambda_n$. $\mathbf{A}^m$ ($m = 1, 2, \dots$) has the eigenvalues $\lambda_1^m, \dots, \lambda_n^m$. The eigenvectors are those of $\mathbf{A}$.
- 19. Spectral mapping theorem.** The “polynomial matrix”

$$p(\mathbf{A}) = k_m \mathbf{A}^m + k_{m-1} \mathbf{A}^{m-1} + \dots + k_1 \mathbf{A} + k_0 \mathbf{I}$$

has the eigenvalues

$$p(\lambda_j) = k_m \lambda_j^m + k_{m-1} \lambda_j^{m-1} + \dots + k_1 \lambda_j + k_0$$

where $j = 1, \dots, n$, and the same eigenvectors as $\mathbf{A}$.

- 20. Perron’s theorem.** A Leslie matrix $\mathbf{L}$ with positive $l_{12}, l_{13}, l_{21}, l_{32}$ has a positive eigenvalue. (This is a special case of the Perron–Frobenius theorem in Sec. 20.7, which is difficult to prove in its general form.)

8.3 Symmetric, Skew-Symmetric, and Orthogonal Matrices

We consider three classes of real square matrices that, because of their remarkable properties, occur quite frequently in applications. The first two matrices have already been mentioned in Sec. 7.2. The goal of Sec. 8.3 is to show their remarkable properties.

¹WASSILY LEONTIEF (1906–1999). American economist at New York University. For his input–output analysis he was awarded the Nobel Prize in 1973.

DEFINITIONS

這一小節，了解名詞及其意義就可。

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 $\mathbf{A}$,

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

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. Every skew-symmetric matrix has all main diagonal entries zero. (Can you prove this?) ■

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

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

EXAMPLE 2 Illustration of Formula (4)

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

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.

This basic theorem (and an extension of it) will be proved in Sec. 8.5.

EXAMPLE 3 Eigenvalues of Symmetric and Skew-Symmetric Matrices

The matrices in (1) and (7) of Sec. 8.2 are symmetric and have real eigenvalues. The skew-symmetric matrix in Example 1 has the eigenvalues 0, $-25i$, and $25i$. (Verify this.) The following matrix has the real eigenvalues 1 and 5 but is not symmetric. Does this contradict Theorem 1?

$$\begin{bmatrix} 3 & 4 \\ 1 & 3 \end{bmatrix}$$

Orthogonal Transformations and Orthogonal Matrices

Orthogonal transformations are transformations

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

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. It can be shown that any orthogonal transformation in the plane or in three-dimensional space is a **rotation** (possibly combined with a reflection in a straight line or a plane, respectively).

The main reason for the importance of orthogonal matrices is as follows.

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*

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

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*

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

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

$$(9) \quad \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = (\mathbf{A}\mathbf{a})^T \mathbf{A}\mathbf{b} = \mathbf{a}^T \mathbf{A}^T \mathbf{A}\mathbf{b} = \mathbf{a}^T \mathbf{I}\mathbf{b} = \mathbf{a}^T \mathbf{b} = \mathbf{a} \cdot \mathbf{b}.$$

From this the invariance of $\|\mathbf{a}\|$ follows if we set $\mathbf{b} = \mathbf{a}$.

Orthogonal matrices have further interesting properties as follows.

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,

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

PROOF (a) Let $\mathbf{A}$ be orthogonal. Then $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^T\mathbf{A} = \mathbf{I}$. In terms of column vectors $\mathbf{a}_1, \dots, \mathbf{a}_n$,

$$(11) \quad \mathbf{I} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{A}^T\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} [\mathbf{a}_1 \cdots \mathbf{a}_n] = \begin{bmatrix} \mathbf{a}_1^T \mathbf{a}_1 & \mathbf{a}_1^T \mathbf{a}_2 & \cdots & \mathbf{a}_1^T \mathbf{a}_n \\ \cdot & \cdot & \cdots & \cdot \\ \mathbf{a}_n^T \mathbf{a}_1 & \mathbf{a}_n^T \mathbf{a}_2 & \cdots & \mathbf{a}_n^T \mathbf{a}_n \end{bmatrix}.$$

The last equality implies (10), by the definition of the $n \times n$ unit matrix $\mathbf{I}$. From (3) it follows that the inverse of an orthogonal matrix is orthogonal (see CAS Experiment 12). Now the column vectors of $\mathbf{A}^{-1} (= \mathbf{A}^T)$ are the row vectors of $\mathbf{A}$. Hence the row vectors of $\mathbf{A}$ also form an orthonormal system.

(b) Conversely, if the column vectors of $\mathbf{A}$ satisfy (10), the off-diagonal entries in (11) must be 0 and the diagonal entries 1. Hence $\mathbf{A}^T\mathbf{A} = \mathbf{I}$, as (11) shows. Similarly, $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. This implies $\mathbf{A}^T = \mathbf{A}^{-1}$ because also $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ and the inverse is unique. Hence $\mathbf{A}$ is orthogonal. Similarly when the row vectors of $\mathbf{A}$ form an orthonormal system, by what has been said at the end of part (a). ■

THEOREM 4**Determinant of an Orthogonal Matrix**

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

PROOF From $\det \mathbf{AB} = \det \mathbf{A} \det \mathbf{B}$ (Sec. 7.8, Theorem 4) and $\det \mathbf{A}^T = \det \mathbf{A}$ (Sec. 7.7, Theorem 2d), we get for an orthogonal matrix

$$1 = \det \mathbf{I} = \det(\mathbf{A}\mathbf{A}^{-1}) = \det(\mathbf{A}\mathbf{A}^T) = \det \mathbf{A} \det \mathbf{A}^T = (\det \mathbf{A})^2. \quad \blacksquare$$

EXAMPLE 4**Illustration of Theorems 3 and 4**

The last matrix in Example 1 and the matrix in (6) illustrate Theorems 3 and 4 because their determinants are -1 and $+1$, as you should verify. ■

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.

PROOF The first part of the statement holds for any real matrix $\mathbf{A}$ because its characteristic polynomial has real coefficients, so that its zeros (the eigenvalues of $\mathbf{A}$) must be as indicated. The claim that $|\lambda| = 1$ will be proved in Sec. 8.5. ■

EXAMPLE 5 Eigenvalues of an Orthogonal Matrix

The orthogonal matrix in Example 1 has the characteristic equation

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

Now one of the eigenvalues must be real (why?), hence $+1$ or -1 . Trying, we find -1 . Division by $\lambda + 1$ gives $-(\lambda^2 - 5\lambda/3 + 1) = 0$ and the two eigenvalues $(5 + i\sqrt{11})/6$ and $(5 - i\sqrt{11})/6$, which have absolute value 1. Verify all of this. ■

Looking back at this section, you will find that the numerous basic results it contains have relatively short, straightforward proofs. This is typical of large portions of matrix eigenvalue theory.

PROBLEM SET 8.3

1–10 SPECTRUM

Are the following matrices symmetric, skew-symmetric, or orthogonal? Find the spectrum of each, thereby illustrating Theorems 1 and 5. Show your work in detail.

1. $\begin{bmatrix} 0.8 & 0.6 \\ -0.6 & 0.8 \end{bmatrix}$

2. $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$

3. $\begin{bmatrix} 2 & 8 \\ -8 & 2 \end{bmatrix}$

4. $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$

5. $\begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & -2 \\ 0 & -2 & 5 \end{bmatrix}$

6. $\begin{bmatrix} a & k & k \\ k & a & k \\ k & k & a \end{bmatrix}$

7. $\begin{bmatrix} 0 & 9 & -12 \\ -9 & 0 & 20 \\ 12 & -20 & 0 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$

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

10. $\begin{bmatrix} \frac{4}{9} & \frac{8}{9} & \frac{1}{9} \\ -\frac{7}{9} & \frac{4}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{1}{9} & \frac{8}{9} \end{bmatrix}$

11. **WRITING PROJECT. Section Summary.** Summarize the main concepts and facts in this section, giving illustrative examples of your own.

12. **CAS EXPERIMENT. Orthogonal Matrices.**

(a) **Products. Inverse.** Prove that the product of two orthogonal matrices is orthogonal, and so is the inverse of an orthogonal matrix. What does this mean in terms of rotations?

(b) **Rotation.** Show that (6) is an orthogonal transformation. Verify that it satisfies Theorem 3. Find the inverse transformation.

(c) **Powers.** Write a program for computing powers $\mathbf{A}^m$ ($m = 1, 2, \dots$) of a 2×2 matrix $\mathbf{A}$ and their spectra. Apply it to the matrix in Prob. 1 (call it $\mathbf{A}$). To what rotation does $\mathbf{A}$ correspond? Do the eigenvalues of $\mathbf{A}^m$ have a limit as $m \rightarrow \infty$?

(d) Compute the eigenvalues of $(0.9\mathbf{A})^m$, where $\mathbf{A}$ is the matrix in Prob. 1. Plot them as points. What is their limit? Along what kind of curve do these points approach the limit?

(e) Find $\mathbf{A}$ such that $\mathbf{y} = \mathbf{A}\mathbf{x}$ is a counterclockwise rotation through 30° in the plane.

13–20 GENERAL PROPERTIES

13. **Verification.** Verify the statements in Example 1.

14. Verify the statements in Examples 3 and 4.

15. **Sum.** Are the eigenvalues of $\mathbf{A} + \mathbf{B}$ sums of the eigenvalues of $\mathbf{A}$ and of $\mathbf{B}$?

16. **Orthogonality.** Prove that eigenvectors of a symmetric matrix corresponding to different eigenvalues are orthogonal. Give examples.

17. **Skew-symmetric matrix.** Show that the inverse of a skew-symmetric matrix is skew-symmetric.

18. Do there exist nonsingular skew-symmetric $n \times n$ matrices with odd n ?

19. **Orthogonal matrix.** Do there exist skew-symmetric orthogonal 3×3 matrices?

20. **Symmetric matrix.** Do there exist nondiagonal symmetric 3×3 matrices that are orthogonal?