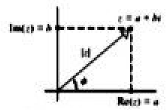
5.3 Complex Vector Spaces

Because the characteristic equation of any square matrix can have complex solutions, the notions of complex eigenvalues and eigenvectors arise naturally, even within the context of matrices with real entries. In this section we will discuss this idea and apply our results to study symmetric matrices in more detail. A review of the essentials of complex numbers appears in the back of this text.

Review of Complex Numbers

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

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



▲ Figure 5.3.1

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

Complex Eigenvalues

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

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

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

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

is

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

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

Vectors in C^n

A vector space in which scalars are allowed to be complex numbers is called a ***complex vector space***. In this section we will be concerned only with the following complex generalization of the real vector space R^n.

> **DEFINITION 1** If n is a positive integer, then a ***complex n-tuple*** is a sequence of n complex numbers $(v_1, v_2, \ldots, v_n)$. The set of all complex n-tuples is called ***complex n-space*** and is denoted by C^n. Scalars are complex numbers, and the operations of addition, subtraction, and scalar multiplication are performed componentwise.

The terminology used for n-tuples of real numbers applies to complex n-tuples without change. Thus, if $v_1, v_2, \ldots, v_n$ are complex numbers, then we call $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ a ***vector*** in C^n and $v_1, v_2, \ldots, v_n$ its ***components***. Some examples of vectors in C^3 are

$$\mathbf{u} = (1 + i, -4i, 3 + 2i), \quad \mathbf{v} = (0, i, 5), \quad \mathbf{w} = \left(6 - \sqrt{2}i, 9 + \tfrac{1}{2}i, \pi i\right)$$

Every vector

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) = (a_1 + b_1 i, a_2 + b_2 i, \ldots, a_n + b_n i)$$

in C^n can be split into ***real and imaginary parts*** as

$$\mathbf{v} = (a_1, a_2, \ldots, a_n) + i(b_1, b_2, \ldots, b_n)$$

which we also denote as

$$\mathbf{v} = \text{Re}(\mathbf{v}) + i\,\text{Im}(\mathbf{v})$$

where

$$\text{Re}(\mathbf{v}) = (a_1, a_2, \ldots, a_n) \quad \text{and} \quad \text{Im}(\mathbf{v}) = (b_1, b_2, \ldots, b_n)$$

The vector

$$\bar{\mathbf{v}} = (\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n) = (a_1 - b_1 i, a_2 - b_2 i, \ldots, a_n - b_n i)$$

is called the ***complex conjugate*** of $\mathbf{v}$ and can be expressed in terms of $\text{Re}(\mathbf{v})$ and $\text{Im}(\mathbf{v})$ as

$$\bar{\mathbf{v}} = (a_1, a_2, \ldots, a_n) - i(b_1, b_2, \ldots, b_n) = \text{Re}(\mathbf{v}) - i\,\text{Im}(\mathbf{v}) \tag{2}$$

It follows that the vectors in R^n can be viewed as those vectors in C^n whose imaginary part is zero; or stated another way, a vector $\mathbf{v}$ in C^n is in R^n if and only if $\bar{\mathbf{v}} = \mathbf{v}$.

In this section we will need to distinguish between matrices whose entries *must* be real numbers, called ***real matrices***, and matrices whose entries may be *either* real numbers or complex numbers, called ***complex matrices***. When convenient, you can think of a real matrix as a complex matrix each of whose entries has a zero imaginary part. The standard operations on real matrices carry over without change to complex matrices, and all of the familiar properties of matrices continue to hold.

If A is a complex matrix, then $\text{Re}(A)$ and $\text{Im}(A)$ are the matrices formed from the real and imaginary parts of the entries of A, and $\bar{A}$ is the matrix formed by taking the complex conjugate of each entry in A.

▶ EXAMPLE 1 Real and Imaginary Parts of Vectors and Matrices

> As you might expect, if A is a complex matrix, then A and $\bar{A}$ can be expressed in terms of $\text{Re}(A)$ and $\text{Im}(A)$ as
>
> $$A = \text{Re}(A) + i\,\text{Im}(A)$$
> $$\bar{A} = \text{Re}(A) - i\,\text{Im}(A)$$

Let

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

Then

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

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

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

Algebraic Properties of the Complex Conjugate

The next two theorems list some properties of complex vectors and matrices that we will need in this section. Some of the proofs are given as exercises.

where

$$\text{Re}(\mathbf{v}) = (a_1, a_2, \ldots, a_n) \quad \text{and} \quad \text{Im}(\mathbf{v}) = (b_1, b_2, \ldots, b_n)$$

The vector

$$\overline{\mathbf{v}} = (\overline{v}_1, \overline{v}_2, \ldots, \overline{v}_n) = (a_1 - b_1 i, a_2 - b_2 i, \ldots, a_n - b_n i)$$

is called the ***complex conjugate*** of $\mathbf{v}$ and can be expressed in terms of $\text{Re}(\mathbf{v})$ and $\text{Im}(\mathbf{v})$ as

$$\overline{\mathbf{v}} = (a_1, a_2, \ldots, a_n) - i(b_1, b_2, \ldots, b_n) = \text{Re}(\mathbf{v}) - i\,\text{Im}(\mathbf{v}) \tag{2}$$

It follows that the vectors in R^n can be viewed as those vectors in C^n whose imaginary part is zero; or stated another way, a vector $\mathbf{v}$ in C^n is in R^n if and only if $\overline{\mathbf{v}} = \mathbf{v}$.

In this section we will need to distinguish between matrices whose entries *must* be real numbers, called ***real matrices***, and matrices whose entries may be *either* real numbers or complex numbers, called ***complex matrices***. When convenient, you can think of a real matrix as a complex matrix each of whose entries has a zero imaginary part. The standard operations on real matrices carry over without change to complex matrices, and all of the familiar properties of matrices continue to hold.

If A is a complex matrix, then $\text{Re}(A)$ and $\text{Im}(A)$ are the matrices formed from the real and imaginary parts of the entries of A, and $\overline{A}$ is the matrix formed by taking the complex conjugate of each entry in A.

▶ EXAMPLE 1 Real and Imaginary Parts of Vectors and Matrices

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

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

Let

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

Then

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

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

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

Algebraic Properties of the Complex Conjugate

The next two theorems list some properties of complex vectors and matrices that we will need in this section. Some of the proofs are given as exercises.

THEOREM 5.3.1 *If $\mathbf{u}$ and $\mathbf{v}$ are vectors in C^n, and if k is a scalar, then:*

(a) $\overline{\overline{\mathbf{u}}} = \mathbf{u}$
(b) $\overline{k\mathbf{u}} = \overline{k}\,\overline{\mathbf{u}}$
(c) $\overline{\mathbf{u} + \mathbf{v}} = \overline{\mathbf{u}} + \overline{\mathbf{v}}$
(d) $\overline{\mathbf{u} - \mathbf{v}} = \overline{\mathbf{u}} - \overline{\mathbf{v}}$

THEOREM 5.3.2 *If A is an $m \times k$ complex matrix and B is a $k \times n$ complex matrix, then:*

(a) $\overline{\overline{A}} = A$
(b) $\overline{(A^T)} = (\overline{A})^T$
(c) $\overline{AB} = \overline{A}\,\overline{B}$

The Complex Euclidean Inner Product

The following definition extends the notions of dot product and norm to C^n.

DEFINITION 2 If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in C^n, then the ***complex Euclidean inner product*** of $\mathbf{u}$ and $\mathbf{v}$ (also called the ***complex dot product***) is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = u_1\overline{v}_1 + u_2\overline{v}_2 + \cdots + u_n\overline{v}_n \tag{3}$$

We also define the ***Euclidean norm*** on C^n to be

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{|v_1|^2 + |v_2|^2 + \cdots + |v_n|^2} \tag{4}$$

The complex conjugates in (3) ensure that $\|\mathbf{v}\|$ is a real number, for without them the quantity $\mathbf{v} \cdot \mathbf{v}$ in (4) might be imaginary.

As in the real case, we call $\mathbf{v}$ a ***unit vector*** in C^n if $\|\mathbf{v}\| = 1$, and we say two vectors $\mathbf{u}$ and $\mathbf{v}$ are ***orthogonal*** if $\mathbf{u} \cdot \mathbf{v} = 0$.

▶ EXAMPLE 2 Complex Euclidean Inner Product and Norm

Find $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{v} \cdot \mathbf{u}$, $\|\mathbf{u}\|$, and $\|\mathbf{v}\|$ for the vectors

$$\mathbf{u} = (1 + i, i, 3 - i) \quad \text{and} \quad \mathbf{v} = (1 + i, 2, 4i)$$

Solution

$$\mathbf{u} \cdot \mathbf{v} = (1+i)(\overline{1+i}) + i(\overline{2}) + (3-i)(\overline{4i}) = (1+i)(1-i) + 2i + (3-i)(-4i) = -2 - 10i$$

$$\mathbf{v} \cdot \mathbf{u} = (1+i)(\overline{1+i}) + 2(\overline{i}) + (4i)(\overline{3-i}) = (1+i)(1-i) - 2i + 4i(3+i) = -2 + 10i$$

$$\|\mathbf{u}\| = \sqrt{|1+i|^2 + |i|^2 + |3-i|^2} = \sqrt{2 + 1 + 10} = \sqrt{13}$$

$$\|\mathbf{v}\| = \sqrt{|1+i|^2 + |2|^2 + |4i|^2} = \sqrt{2 + 4 + 16} = \sqrt{22} \ ◀$$

Recall from Table 1 of Section 3.2 that if $\mathbf{u}$ and $\mathbf{v}$ are *column vectors* in R^n, then their dot product can be expressed as

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$$

The analogous formulas in C^n are (verify)

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\overline{\mathbf{v}} = \overline{\mathbf{v}}^T\mathbf{u} \tag{5}$$

Example 2 reveals a major difference between the dot product on R^n and the complex dot product on C^n. For the dot product on R^n we always have $\mathbf{v} \cdot \mathbf{u} = \mathbf{u} \cdot \mathbf{v}$ (the *symmetry property*), but for the complex dot product the corresponding relationship is given by $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$, which is called its ***antisymmetry*** property. The following theorem is an analog of Theorem 3.2.2.

THEOREM 5.3.3 *If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in C^n, and if k is a scalar, then the complex Euclidean inner product has the following properties:*

(a) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$ [Antisymmetry property]
(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ [Distributive property]
(c) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ [Homogeneity property]
(d) $\mathbf{u} \cdot k\mathbf{v} = \overline{k}(\mathbf{u} \cdot \mathbf{v})$ [Antihomogeneity property]
(e) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$. [Positivity property]

Parts (*c*) and (*d*) of this theorem state that a scalar multiplying a complex Euclidean inner product can be regrouped with the first vector, but to regroup it with the second vector you must first take its complex conjugate. We will prove part (*d*), and leave the others as exercises.

Proof (*d*)

$$k(\mathbf{u}\cdot\mathbf{v}) = k(\overline{\mathbf{v}\cdot\mathbf{u}}) = \overline{\overline{k}}\,(\overline{\mathbf{v}\cdot\mathbf{u}}) = \overline{\overline{k}\,(\mathbf{v}\cdot\mathbf{u})} = \overline{(\overline{k}\mathbf{v})\cdot\mathbf{u}} = \mathbf{u}\cdot(\overline{k}\mathbf{v})$$

To complete the proof, substitute $\overline{k}$ for k and use the fact that $\overline{\overline{k}} = k$. ◄

Vector Concepts in C^n

Is R^n a subspace of C^n? Explain.

Except for the use of complex scalars, the notions of linear combination, linear independence, subspace, spanning, basis, and dimension carry over without change to C^n.

Eigenvalues and eigenvectors are defined for complex matrices exactly as for real matrices. If A is an $n \times n$ matrix with complex entries, then the complex roots of the characteristic equation $\det(\lambda I - A) = 0$ are called ***complex eigenvalues*** of A. As in the real case, λ is a complex eigenvalue of A if and only if there exists a nonzero vector $\mathbf{x}$ in C^n such that $A\mathbf{x} = \lambda\mathbf{x}$. Each such $\mathbf{x}$ is called a ***complex eigenvector*** of A corresponding to λ. The complex eigenvectors of A corresponding to λ are the nonzero solutions of the linear system $(\lambda I - A)\mathbf{x} = \mathbf{0}$, and the set of all such solutions is a subspace of C^n, called the ***complex eigenspace*** of A corresponding to λ.

The following theorem states that if a *real matrix* has complex eigenvalues, then those eigenvalues and their corresponding eigenvectors occur in conjugate pairs.

THEOREM 5.3.4 *If λ is an eigenvalue of a real $n \times n$ matrix A, and if $\mathbf{x}$ is a corresponding eigenvector, then $\overline{\lambda}$ is also an eigenvalue of A, and $\overline{\mathbf{x}}$ is a corresponding eigenvector.*

Proof Since λ is an eigenvalue of A and $\mathbf{x}$ is a corresponding eigenvector, we have

$$\overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\,\overline{\mathbf{x}} \tag{6}$$

However, $\overline{A} = A$, since A has real entries, so it follows from part (*c*) of Theorem 5.3.2 that

$$\overline{A\mathbf{x}} = \overline{A}\,\overline{\mathbf{x}} = A\overline{\mathbf{x}} \tag{7}$$

Equations (6) and (7) together imply that

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

in which $\overline{\mathbf{x}} \neq \mathbf{0}$ (why?); this tells us that $\overline{\lambda}$ is an eigenvalue of A and $\overline{\mathbf{x}}$ is a corresponding eigenvector. ◄

▶ **EXAMPLE 3 Complex Eigenvalues and Eigenvectors**

Find the eigenvalues and bases for the eigenspaces of

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

Solution The characteristic polynomial of A is

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

332 / 802

so the eigenvalues of A are $\lambda = i$ and $\lambda = -i$. Note that these eigenvalues are complex conjugates, as guaranteed by Theorem 5.3.4. To find the eigenvectors we must solve the system

$$\begin{bmatrix} \lambda+2 & 1 \\ -5 & \lambda-2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

with $\lambda = i$ and then with $\lambda = -i$. With $\lambda = i$, this system becomes

$$\begin{bmatrix} i+2 & 1 \\ -5 & i-2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{8}$$

We could solve this system by reducing the augmented matrix

$$\begin{bmatrix} i+2 & 1 & 0 \\ -5 & i-2 & 0 \end{bmatrix} \tag{9}$$

to reduced row echelon form by Gauss–Jordan elimination, though the complex arithmetic is somewhat tedious. A simpler procedure here is first to observe that the reduced row echelon form of (9) must have a row of zeros because (8) has nontrivial solutions. This being the case, each row of (9) must be a scalar multiple of the other, and hence the first row can be made into a row of zeros by adding a suitable multiple of the second row to it. Accordingly, we can simply set the entries in the first row to zero, then interchange the rows, and then multiply the new first row by $-\frac{1}{5}$ to obtain the reduced row echelon form

$$\begin{bmatrix} 1 & \frac{2}{5} - \frac{1}{5}i & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Thus, a general solution of the system is

$$x_1 = \left(-\tfrac{2}{5} + \tfrac{1}{5}i\right)t, \quad x_2 = t$$

This tells us that the eigenspace corresponding to $\lambda = i$ is one-dimensional and consists of all complex scalar multiples of the basis vector

$$\mathbf{x} = \begin{bmatrix} -\frac{2}{5} + \frac{1}{5}i \\ 1 \end{bmatrix} \tag{10}$$

As a check, let us confirm that $A\mathbf{x} = i\mathbf{x}$. We obtain

$$A\mathbf{x} = \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} -\frac{2}{5} + \frac{1}{5}i \\ 1 \end{bmatrix} = \begin{bmatrix} -2\left(-\frac{2}{5} + \frac{1}{5}i\right) - 1 \\ 5\left(-\frac{2}{5} + \frac{1}{5}i\right) + 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} - \frac{2}{5}i \\ i \end{bmatrix} = i\mathbf{x}$$

We could find a basis for the eigenspace corresponding to $\lambda = -i$ in a similar way, but the work is unnecessary since Theorem 5.3.4 implies that

$$\overline{\mathbf{x}} = \begin{bmatrix} -\frac{2}{5} - \frac{1}{5}i \\ 1 \end{bmatrix} \tag{11}$$

must be a basis for this eigenspace. The following computations confirm that $\overline{\mathbf{x}}$ is an eigenvector of A corresponding to $\lambda = -i$:

$$\begin{aligned} A\overline{\mathbf{x}} &= \begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} -\frac{2}{5} - \frac{1}{5}i \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -2\left(-\frac{2}{5} - \frac{1}{5}i\right) - 1 \\ 5\left(-\frac{2}{5} - \frac{1}{5}i\right) + 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5} + \frac{2}{5}i \\ -i \end{bmatrix} = -i\overline{\mathbf{x}} \end{aligned}$$ ◄

Since a number of our subsequent examples will involve 2×2 matrices with real entries, it will be useful to discuss some general results about the eigenvalues of such matrices. Observe first that the characteristic polynomial of the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix} = (\lambda - a)(\lambda - d) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$$

We can express this in terms of the trace and determinant of A as

$$\det(\lambda I - A) = \lambda^2 - \operatorname{tr}(A)\lambda + \det(A) \tag{12}$$

from which it follows that the characteristic equation of A is

$$\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = 0 \tag{13}$$

Now recall from algebra that if $ax^2 + bx + c = 0$ is a quadratic equation with real coefficients, then the *discriminant* $b^2 - 4ac$ determines the nature of the roots:

$$b^2 - 4ac > 0 \quad \text{[Two distinct real roots]}$$
$$b^2 - 4ac = 0 \quad \text{[One repeated real root]}$$
$$b^2 - 4ac < 0 \quad \text{[Two conjugate imaginary roots]}$$

Applying this to (13) with $a = 1$, $b = -\operatorname{tr}(A)$, and $c = \det(A)$ yields the following theorem.

THEOREM 5.3.5 *If A is a 2×2 matrix with real entries, then the characteristic equation of A is $\lambda^2 - \operatorname{tr}(A)\lambda + \det(A) = 0$ and*

(a) A has two distinct real eigenvalues if $\operatorname{tr}(A)^2 - 4\det(A) > 0$;

(b) A has one repeated real eigenvalue if $\operatorname{tr}(A)^2 - 4\det(A) = 0$;

(c) A has two complex conjugate eigenvalues if $\operatorname{tr}(A)^2 - 4\det(A) < 0$.

▶ EXAMPLE 4 **Eigenvalues of a 2 × 2 Matrix**

In each part, use Formula (13) for the characteristic equation to find the eigenvalues of

$$\text{(a) } A = \begin{bmatrix} 2 & 2 \\ -1 & 5 \end{bmatrix} \quad \text{(b) } A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix} \quad \text{(c) } A = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$$

Olga Taussky-Todd (1906–1995)

Historical Note Olga Taussky-Todd was one of the pioneering women in matrix analysis and the first woman appointed to the faculty at the California Institute of Technology. She worked at the National Physical Laboratory in London during World War II, where she was assigned to study flutter in supersonic aircraft. While there, she realized that some results about the eigenvalues of a certain 6×6 complex matrix could be used to answer key questions about the flutter problem that would otherwise have required laborious calculation. After World War II Olga Taussky-Todd continued her work on matrix-related subjects and helped to draw many known but disparate results about matrices into the coherent subject that we now call matrix theory.

[Image: Courtesy of the Archives, California Institute of Technology]

334 / 802

Solution (a) We have $\operatorname{tr}(A) = 7$ and $\det(A) = 12$, so the characteristic equation of A is

$$\lambda^2 - 7\lambda + 12 = 0$$

Factoring yields $(\lambda - 4)(\lambda - 3) = 0$, so the eigenvalues of A are $\lambda = 4$ and $\lambda = 3$.

Solution (b) We have $\operatorname{tr}(A) = 2$ and $\det(A) = 1$, so the characteristic equation of A is

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

Factoring this equation yields $(\lambda - 1)^2 = 0$, so $\lambda = 1$ is the only eigenvalue of A; it has algebraic multiplicity 2.

Solution (c) We have $\operatorname{tr}(A) = 4$ and $\det(A) = 13$, so the characteristic equation of A is

$$\lambda^2 - 4\lambda + 13 = 0$$

Solving this equation by the quadratic formula yields

$$\lambda = \frac{4 \pm \sqrt{(-4)^2 - 4(13)}}{2} = \frac{4 \pm \sqrt{-36}}{2} = 2 \pm 3i$$

Thus, the eigenvalues of A are $\lambda = 2 + 3i$ and $\lambda = 2 - 3i$. ◀

Symmetric Matrices Have Real Eigenvalues

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

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

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

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

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

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

then we obtain

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

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

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

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

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

A Geometric Interpretation of Complex Eigenvalues

The following theorem is the key to understanding the geometric significance of complex eigenvalues of real 2×2 matrices.

▲ Figure 5.3.2

THEOREM 5.3.7 *The eigenvalues of the real matrix*

$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \tag{15}$$

are $\lambda = a \pm bi$. If a and b are not both zero, then this matrix can be factored as

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \tag{16}$$

where ϕ is the angle from the positive x-axis to the ray that joins the origin to the point (a, b) (Figure 5.3.2).

▲ Figure 5.3.3

Geometrically, this theorem states that multiplication by a matrix of form (15) can be viewed as a rotation through the angle ϕ followed by a scaling with factor $|\lambda|$ (Figure 5.3.3).

Proof The characteristic equation of C is $(\lambda - a)^2 + b^2 = 0$ (verify), from which it follows that the eigenvalues of C are $\lambda = a \pm bi$. Assuming that a and b are not both zero, let ϕ be the angle from the positive x-axis to the ray that joins the origin to the point (a, b). The angle ϕ is an argument of the eigenvalue $\lambda = a + bi$, so we see from Figure 5.3.2 that

$$a = |\lambda|\cos\phi \quad \text{and} \quad b = |\lambda|\sin\phi$$

It follows from this that the matrix in (15) can be written as

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \frac{a}{|\lambda|} & -\frac{b}{|\lambda|} \\ \frac{b}{|\lambda|} & \frac{a}{|\lambda|} \end{bmatrix} = \begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \blacktriangleleft$$

The following theorem, whose proof is considered in the exercises, shows that every real 2×2 matrix with complex eigenvalues is similar to a matrix of form (15).

THEOREM 5.3.8 *Let A be a real 2×2 matrix with complex eigenvalues $\lambda = a \pm bi$ (where $b \neq 0$). If $\mathbf{x}$ is an eigenvector of A corresponding to $\lambda = a - bi$, then the matrix $P = [\text{Re}(\mathbf{x}) \quad \text{Im}(\mathbf{x})]$ is invertible and*

$$A = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}P^{-1} \tag{17}$$

▶ **EXAMPLE 5 A Matrix Factorization Using Complex Eigenvalues**

Factor the matrix in Example 3 into form (17) using the eigenvalue $\lambda = -i$ and the corresponding eigenvector that was given in (11).

Solution For consistency with the notation in Theorem 5.3.8, let us denote the eigenvector in (11) that corresponds to $\lambda = -i$ by $\mathbf{x}$ (rather than $\bar{\mathbf{x}}$ as before). For this λ and $\mathbf{x}$ we have

$$a = 0, \quad b = 1, \quad \text{Re}(\mathbf{x}) = \begin{bmatrix} -\frac{2}{5} \\ 1 \end{bmatrix}, \quad \text{Im}(\mathbf{x}) = \begin{bmatrix} -\frac{1}{5} \end{bmatrix}$$

336 / 802

Thus,

$$P = [\text{Re}(\mathbf{x}) \quad \text{Im}(\mathbf{x})] = \begin{bmatrix} -\frac{2}{5} & -\frac{1}{5} \\ 1 & 0 \end{bmatrix}$$

so A can be factored in form (17) as

$$\begin{bmatrix} -2 & -1 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} & -\frac{1}{5} \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -5 & -2 \end{bmatrix}$$

You may want to confirm this by multiplying out the right side. ◀

A Geometric Interpretation of Theorem 5.3.8

To clarify what Theorem 5.3.8 says geometrically, let us denote the matrices on the right side of (16) by S and R_ϕ, respectively, and then use (16) to rewrite (17) as

$$A = PSR_\phi P^{-1} = P\begin{bmatrix} |\lambda| & 0 \\ 0 & |\lambda| \end{bmatrix}\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}P^{-1} \tag{18}$$

If we now view P as the transition matrix from the basis $B = \{\text{Re}(\mathbf{x}), \text{Im}(\mathbf{x})\}$ to the standard basis, then (18) tells us that computing a product $A\mathbf{x}_0$ can be broken down into a three-step process:

Interpreting Formula (18)

Step 1. Map $\mathbf{x}_0$ from standard coordinates into B-coordinates by forming the product $P^{-1}\mathbf{x}_0$.

Step 2. Rotate and scale the vector $P^{-1}\mathbf{x}_0$ by forming the product $SR_\phi P^{-1}\mathbf{x}_0$.

Step 3. Map the rotated and scaled vector back to standard coordinates to obtain $A\mathbf{x}_0 = PSR_\phi P^{-1}\mathbf{x}_0$.

Power Sequences

There are many problems in which one is interested in how successive applications of a matrix transformation affect a specific vector. For example, if A is the standard matrix for an operator on R^n and $\mathbf{x}_0$ is some fixed vector in R^n, then one might be interested in the behavior of the power sequence

$$\mathbf{x}_0, \ A\mathbf{x}_0, \ A^2\mathbf{x}_0, \ldots, \ A^k\mathbf{x}_0, \ldots$$

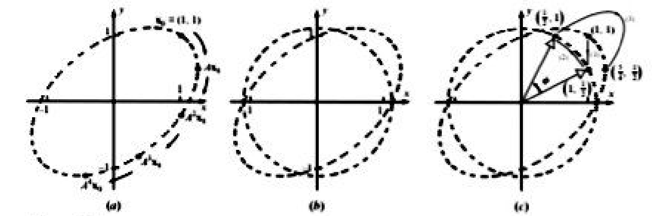
For example, if

$$A = \begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix} \quad \text{and} \quad \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

then with the help of a computer or calculator one can show that the first four terms in the power sequence are

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad A\mathbf{x}_0 = \begin{bmatrix} 1.25 \\ 0.5 \end{bmatrix}, \quad A^2\mathbf{x}_0 = \begin{bmatrix} 1.0 \\ -0.2 \end{bmatrix}, \quad A^3\mathbf{x}_0 = \begin{bmatrix} 0.35 \\ -0.82 \end{bmatrix}$$

With the help of MATLAB or a computer algebra system one can show that if the first 100 terms are plotted as ordered pairs (x, y), then the points move along the elliptical path shown in Figure 5.3.4*a*.



▲ Figure 5.3.4

To understand why the points move along an elliptical path, we will need to examine the eigenvalues and eigenvectors of A. We leave it for you to show that the eigenvalues of A are $\lambda = \frac{4}{5} \pm \frac{3}{5}i$ and that the corresponding eigenvectors are

$$\lambda_1 = \tfrac{4}{5} - \tfrac{3}{5}i: \quad \mathbf{v}_1 = \left(\tfrac{1}{2} + i, 1\right) \quad \text{and} \quad \lambda_2 = \tfrac{4}{5} + \tfrac{3}{5}i: \quad \mathbf{v}_2 = \left(\tfrac{1}{2} - i, 1\right)$$

If we take $\lambda = \lambda_1 = \frac{4}{5} - \frac{3}{5}i$ and $\mathbf{x} = \mathbf{v}_1 = \left(\frac{1}{2} + i, 1\right)$ in (17) and use the fact that $|\lambda| = 1$, then we obtain the factorization

$$\underset{A}{\begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix}} = \underset{P}{\begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix}} \underset{R_\phi}{\begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}} \underset{P^{-1}}{\begin{bmatrix} 0 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}} \tag{19}$$

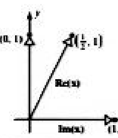
where R_ϕ is a rotation about the origin through the angle ϕ whose tangent is

$$\tan\phi = \frac{\sin\phi}{\cos\phi} = \frac{3/5}{4/5} = \frac{3}{4} \qquad \left(\phi = \tan^{-1}\tfrac{3}{4} \approx 36.9^\circ\right)$$

The matrix P in (19) is the transition matrix from the basis

$$B = \{\operatorname{Re}(\mathbf{x}), \operatorname{Im}(\mathbf{x})\} = \left\{\left(\tfrac{1}{2}, 1\right), (1, 0)\right\}$$

to the standard basis, and P^{-1} is the transition matrix from the standard basis to the basis B (Figure 5.3.5). Next, observe that if n is a positive integer, then (19) implies that



▲ Figure 5.3.5

$$A^n\mathbf{x}_0 = (PR_\phi P^{-1})^n\mathbf{x}_0 = PR_\phi^n P^{-1}\mathbf{x}_0$$

so the product $A^n\mathbf{x}_0$ can be computed by first mapping $\mathbf{x}_0$ into the point $P^{-1}\mathbf{x}_0$ in B-coordinates, then multiplying by R_ϕ^n to rotate this point about the origin through the angle $n\phi$, and then multiplying $R_\phi^n P^{-1}\mathbf{x}_0$ by P to map the resulting point back to standard coordinates. We can now see what is happening geometrically: In B-coordinates each successive multiplication by A causes the point $P^{-1}\mathbf{x}_0$ to advance through an angle ϕ, thereby tracing a circular orbit about the origin. However, the basis B is *skewed* (not orthogonal), so when the points on the circular orbit are transformed back to standard coordinates, the effect is to distort the circular orbit into the elliptical orbit traced by $A^n\mathbf{x}_0$ (Figure 5.3.4*b*). Here are the computations for the first step (successive steps are

illustrated in Figure 5.3.4*c*):

$$\begin{bmatrix} \frac{1}{2} & \frac{3}{4} \\ -\frac{3}{5} & \frac{11}{10} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}\begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} \qquad [\mathbf{x}_0 \text{ is mapped to } B\text{-coordinates.}]$$

$$= \begin{bmatrix} \frac{1}{2} & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix} \qquad [\text{The point } (1, \tfrac{1}{2}) \text{ is rotated through the angle } \phi.]$$

$$= \begin{bmatrix} \frac{5}{4} \\ \frac{1}{2} \end{bmatrix} \qquad [\text{The point } (\tfrac{1}{2}, 1) \text{ is mapped to standard coordinates.}]$$

Exercise Set 5.3

▶ In Exercises 1–2, find $\bar{\mathbf{u}}$, Re(**u**), Im(**u**), and $\|\mathbf{u}\|$. ◀

1. $\mathbf{u} = (2 - i, 4i, 1 + i)$

2. $\mathbf{u} = (6, 1 + 4i, 6 - 2i)$

▶ In Exercises 3–4, show that **u**, **v**, and k satisfy Theorem 5.3.1. ◀

3. $\mathbf{u} = (3 - 4i, 2 + i, -6i)$, $\mathbf{v} = (1 + i, 2 - i, 4)$, $k = i$

4. $\mathbf{u} = (6, 1 + 4i, 6 - 2i)$, $\mathbf{v} = (4, 3 + 2i, i - 3)$, $k = -i$

5. Solve the equation $i\mathbf{x} - 3\mathbf{v} = \bar{\mathbf{u}}$ for **x**, where **u** and **v** are the vectors in Exercise 3.

6. Solve the equation $(1 + i)\mathbf{x} + 2\mathbf{u} = \bar{\mathbf{v}}$ for **x**, where **u** and **v** are the vectors in Exercise 4.

▶ In Exercises 7–8, find $\bar{A}$, Re(A), Im(A), det(A), and tr(A). ◀

7. $A = \begin{bmatrix} -5i & 4 \\ 2 - i & 1 + 5i \end{bmatrix}$

8. $A = \begin{bmatrix} 4i & 2 - 3i \\ 2 + 3i & 1 \end{bmatrix}$

9. Let A be the matrix given in Exercise 7, and let B be the matrix

$$B = \begin{bmatrix} 1 - i \\ 2i \end{bmatrix}$$

Confirm that these matrices have the properties stated in Theorem 5.3.2.

10. Let A be the matrix given in Exercise 8, and let B be the matrix

$$B = \begin{bmatrix} 5i \\ 1 - 4i \end{bmatrix}$$

Confirm that these matrices have the properties stated in Theorem 5.3.2.

▶ In Exercises 11–12, compute $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{u} \cdot \mathbf{w}$, and $\mathbf{v} \cdot \mathbf{w}$, and show that the vectors satisfy Formula (5) and parts (*a*), (*b*), and (*c*) of Theorem 5.3.3. ◀

11. $\mathbf{u} = (i, 2i, 3)$, $\mathbf{v} = (4, -2i, 1 + i)$, $\mathbf{w} = (2 - i, 2i, 5 + 3i)$, $k = 2i$

12. $\mathbf{u} = (1 + i, 4, 3i)$, $\mathbf{v} = (3, -4i, 2 + 3i)$, $\mathbf{w} = (1 - i, 4i, 4 - 5i)$, $k = 1 + i$

13. Compute $\overline{(\mathbf{u} \cdot \bar{\mathbf{v}})} - \overline{\mathbf{w} \cdot \mathbf{u}}$ for the vectors **u**, **v**, and **w** in Exercise 11.

14. Compute $\overline{(i\mathbf{u} \cdot \mathbf{w}) + (\|\mathbf{u}\|\mathbf{v}) \cdot \mathbf{u}}$ for the vectors **u**, **v**, and **w** in Exercise 12.

▶ In Exercises 15–18, find the eigenvalues and bases for the eigenspaces of A. ◀

15. $A = \begin{bmatrix} 4 & -5 \\ 1 & 0 \end{bmatrix}$

16. $A = \begin{bmatrix} -1 & -5 \\ 4 & 7 \end{bmatrix}$

17. $A = \begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$

18. $A = \begin{bmatrix} 8 & 6 \\ -3 & 2 \end{bmatrix}$

▶ In Exercises 19–22, each matrix C has form (15). Theorem 5.3.7 implies that C is the product of a scaling matrix with factor $|\lambda|$ and a rotation matrix with angle ϕ. Find $|\lambda|$ and ϕ for which $-\pi < \phi \le \pi$. ◀

19. $C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

20. $C = \begin{bmatrix} 0 & 5 \\ -5 & 0 \end{bmatrix}$

21. $C = \begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$

22. $C = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{bmatrix}$

▶ In Exercises 23–26, find an invertible matrix P and a matrix C of form (15) such that $A = PCP^{-1}$. ◀

23. $A = \begin{bmatrix} -1 & -5 \\ 4 & 7 \end{bmatrix}$

24. $A = \begin{bmatrix} 4 & -5 \\ 1 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 8 & 6 \\ -3 & 2 \end{bmatrix}$

26. $A = \begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$