CHAPTER 9

Eigenvalues, Diagonalization, and Special Matrices

9.1 Eigenvalues and Eigenvectors

In this chapter, the term number refers to a real or complex number. Let **A** be an $n \times n$ matrix of numbers. A number λ is an *eigenvalue* of **A** if there is a nonzero $n \times 1$ matrix **E** such that

$$\mathbf{AE} = \lambda \mathbf{E}. \tag{9.1}$$

We call **E** an *eigenvector* associated with the eigenvalue λ.

We may think of an $n \times 1$ matrix of numbers as an n-vector, with real and/or complex components. If we consider **A** as a linear transformation mapping an n-vector **X** to an n-vector **AX**, then equation (9.1) holds when **A** moves **E** to a parallel vector $\lambda\mathbf{E}$. This is the geometric significance of an eigenvector.

If c is a nonzero number and $\mathbf{AE} = \lambda\mathbf{E}$, then

$$\mathbf{A}(c\mathbf{E}) = c\mathbf{AE} = c\lambda\mathbf{E} = \lambda(c\mathbf{E}).$$

This means that nonzero constant multiples of eigenvectors are eigenvectors (with the same eigenvalue).

EXAMPLE 9.1

Let

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Because

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0\begin{pmatrix} 0 \\ 4 \end{pmatrix},$$

then 0 is an eigenvalue of **A** with eigenvector

$$\mathbf{E} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

For any nonzero number α,

$$\begin{pmatrix} 0 \\ 4\alpha \end{pmatrix}$$

is also an eigenvector. Zero can be an eigenvalue, but an eigenvector must be a nonzero vector (at least one nonzero component). ♦

EXAMPLE 9.2

Let

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Then

$$\mathbf{A}\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

Therefore 1 is an eigenvalue with eigenvector

$$\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

or any nonzero constant times this matrix. Similarly,

$$\mathbf{A}\begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 4 \end{pmatrix} = (-1)\begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix}.$$

Therefore -1 is an eigenvalue with eigenvector

$$\begin{pmatrix} 1 \\ 2 \\ -4 \end{pmatrix},$$

or any nonzero multiple of this vector. ♦

We would like to be able to find all of the eigenvalues of a matrix. We will have $\mathbf{AE} = \lambda\mathbf{E}$, for some number λ and $n \times 1$ matrix **E**, exactly when

$$\lambda\mathbf{E} - \mathbf{AE} = \mathbf{O}.$$

This is equivalent to

$$(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{E} = \mathbf{O},$$

and this occurs exactly when the system

$$(\lambda\mathbf{I}_n - \mathbf{A})\mathbf{X} = \mathbf{O}$$

has a nontrivial solution **E**. The condition for this is that the coefficient matrix be singular (determinant zero), hence that

$$|\lambda \mathbf{I}_n - \mathbf{A}| = 0.$$

If expanded, the determinant on the left is a polynomial of degree n in the unknown λ, and is called the *characteristic polynomial* of **A**. Thus

$$p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_n - \mathbf{A}|.$$

This polynomial has n roots for λ (perhaps some repeated, perhaps some or all complex). These n numbers, counting multiplicities, are all of the eigenvalues of **A**. Corresponding to each eigenvalue λ, a nontrivial solution of

$$(\lambda \mathbf{I}_n - \mathbf{A})\mathbf{X} = \mathbf{O}$$

is an eigenvector.

We can summarize this discussion as follows.

THEOREM 9.1 *Eigenvalues and Eigenvectors of A*

Let **A** be an $n \times n$ matrix of numbers. Then

1. λ is an eigenvalue of **A** if and only if λ is a root of the characteristic polynomial of **A**. This occurs exactly when

$$p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_n - \mathbf{A}| = 0.$$

   Since $p_{\mathbf{A}}(\lambda)$ has degree n, **A** has n eigenvalues, counting each eigenvalue as many times as it appears as a root of $p_{\mathbf{A}}(\lambda)$.

2. If λ is an eigenvalue of **A**, then any nontrivial solution **E** of

$$(\lambda \mathbf{I}_n - \mathbf{A})\mathbf{X} = \mathbf{O}$$

   is an eigenvector of **A** associated with λ.

3. If **E** is an eigenvector associated with the eigenvalue λ, then so is $c\mathbf{E}$ for any nonzero number c. ♦

EXAMPLE 9.3

Let

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

as in Example 9.2. The characteristic polynomial is

$$p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_3 - \mathbf{A}| = \begin{vmatrix} \lambda - 1 & 1 & 0 \\ 0 & \lambda - 1 & -1 \\ 0 & 0 & \lambda + 1 \end{vmatrix} = (\lambda - 1)^2(\lambda + 1).$$

This polynomial has roots $1, 1, -1$ and these are the eigenvalues of **A**. The root 1 has multiplicity 2 and must be listed twice as an eigenvalue of **A**. **A** has three eigenvalues.

To find an eigenvector associated with the eigenvalue 1, put $\lambda = 1$ in (2) of the theorem and solve the system

$$((1)\mathbf{I}_3 - \mathbf{A})\mathbf{X} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This system of three equations in three unknowns has the general solution

$$\begin{pmatrix} \alpha \\ 0 \\ 0 \end{pmatrix}$$

and this is an eigenvector associated with 1 for any $\alpha \neq 0$.

For eigenvectors associated with -1, put $\lambda = -1$ in (2) of the theorem and solve

$$((-1)\mathbf{I}_3 - \mathbf{A})\mathbf{X} = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{X} = \mathbf{O}.$$

This system has the general solution

$$\begin{pmatrix} \beta \\ 2\beta \\ -4\beta \end{pmatrix}$$

and this is an eigenvector associated with -1 for any $\beta \neq 0$. ♦

EXAMPLE 9.4

Let

$$\mathbf{A} = \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix}.$$

The characteristic polynomial is

$$p_{\mathbf{A}}(\lambda) = |\lambda \mathbf{I}_2 - \mathbf{A}| = \left| \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} - \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix} \right| = \begin{vmatrix} \lambda - 1 & 2 \\ -2 & \lambda \end{vmatrix} = \lambda^2 - \lambda + 4,$$

with roots

$$\frac{1 + \sqrt{15}i}{2} \text{ and } \frac{1 - \sqrt{15}i}{2}$$

and these are the eigenvalues of **A**.

For an eigenvector corresponding to $(1 + \sqrt{15}i)/2$ solve $(((1 + \sqrt{15}i)/2)\mathbf{I}_2 - \mathbf{A})\mathbf{X} = \mathbf{O}$, which is

$$\left[\frac{1 + \sqrt{15}i}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & -2 \\ 2 & 0 \end{pmatrix} \right] \mathbf{X} = \mathbf{O}.$$

This is the system

$$\begin{pmatrix} \dfrac{-1 + \sqrt{15}i}{2} & 2 \\ -2 & \dfrac{1 + \sqrt{15}i}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This 2×2 system has general solution

$$\alpha \begin{pmatrix} 1 \\ (1 - \sqrt{15}i)/4 \end{pmatrix}.$$

This is an eigenvector associated with $(1 + \sqrt{15}i)/2$ for any $\alpha \neq 0$.

For eigenvectors associated with $(1-\sqrt{15}i)/2$, solve the 2×2 system

$$(((1-\sqrt{15}i)/2)\mathbf{I}_2 - \mathbf{A})\mathbf{X} = \mathbf{O}$$

to obtain

$$\beta \begin{pmatrix} 1 \\ (1+\sqrt{15}i)/4 \end{pmatrix}.$$

This is an eigenvector associated with $(1-\sqrt{15}i)/2$ for any $\beta \neq 0$. ♦

If $\mathbf{A}$ has real numbers as elements and $\lambda = \alpha + i\beta$ is an eigenvalue, then the conjugate $\overline{\lambda} = \alpha - i\beta$ is also an eigenvalue. This is because the characteristic polynomial of $\mathbf{A}$ has real coefficients in this case, so complex roots (eigenvalues of $\mathbf{A}$) occur in conjugate pairs. Furthermore, if $\mathbf{E}$ is an eigenvector corresponding to λ, then $\overline{\mathbf{E}}$ is an eigenvector corresponding to $\overline{\lambda}$, where we take the conjugate of a matrix by taking the conjugate of each of its elements. This can be seen by taking the conjugate of $\mathbf{AE} = \lambda\mathbf{E}$ to obtain

$$\overline{\mathbf{AE}} = \overline{\lambda}\,\overline{\mathbf{E}}.$$

Because $\mathbf{A}$ has real elements, $\overline{\mathbf{A}} = \mathbf{A}$ so

$$\mathbf{A}\overline{\mathbf{E}} = \overline{\lambda}\,\overline{\mathbf{E}}.$$

This observation can be seen in Example 9.4.

There is a general expression for the eigenvalues of a matrix that will be used soon to draw conclusions about eigenvalues of matrices having special properties.

LEMMA 9.1

Let $\mathbf{A}$ be an $n \times n$ matrix of numbers. Let λ be an eigenvalue of $\mathbf{A}$, with eigenvector $\mathbf{E}$. Then

$$\lambda = \frac{\overline{\mathbf{E}}^t \mathbf{AE}}{\overline{\mathbf{E}}^t \mathbf{E}}. \quad ♦ \tag{9.2}$$

Before giving the one line proof of this expression, examine what the right side means. Let

$$\mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

Then

$$\overline{\mathbf{E}}^t \mathbf{AE} = \begin{pmatrix} \overline{e_1} & \overline{e_2} & \cdots & \overline{e_n} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}.$$

This is a product of a $1 \times n$ matrix with an $n \times n$ matrix, then an $n \times 1$ matrix, hence is a 1×1 matrix, which we think of as a number. If we carry out this matrix product we obtain the number

$$\overline{\mathbf{E}}^t \mathbf{AE} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \overline{e_i} e_j.$$

For the denominator of equation (9.2) we have a $1 \times n$ matrix multiplied by an $n \times 1$ matrix, which is also a 1×1 matrix, or number. Specifically,

$$\overline{\mathbf{E}}^t\mathbf{E} = (\overline{e_1} \quad \overline{e_2} \quad \cdots \quad \overline{e_n})\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \sum_{j=1}^{n} \overline{e_j} e_j = \sum_{j=1}^{n} |e_j|^2.$$

Therefore the conclusion of Lemma 9.1 can be written

$$\lambda = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\overline{e_i}e_j}{\sum_{j=1}^{n}|e_j|^2}.$$

Proof of Lemma 9.1 Since $\mathbf{AE} = \lambda\mathbf{E}$, then

$$\overline{\mathbf{E}}^t\mathbf{AE} = \lambda\overline{\mathbf{E}}^t\mathbf{E},$$

yielding the conclusion of the lemma. ♦

When we discuss diagonalization, we will need to know if the eigenvectors of a matrix are linearly independent. The following theorem answers this question for the special case that the n eigenvalues of $\mathbf{A}$ are distinct (the characteristic polynomial has no repeated roots).

THEOREM 9.2

Suppose the $n \times n$ matrix $\mathbf{A}$ has n distinct eigenvalues. Then $\mathbf{A}$ has n linearly independent eigenvectors. ♦

To illustrate, in Example 9.4, $\mathbf{A}$ was 2×2 and had two distinct eigenvalues. The eigenvectors produced for each eigenvalue were linearly independent.

Proof We will show by induction that any k distinct eigenvalues have associated with them k linearly independent eigenvectors. For $k = 1$ there is nothing to show. Thus suppose $k \geq 2$ and the conclusion of the theorem is valid for any $k - 1$ distinct eigenvalues. This means that any $k - 1$ distinct eigenvalues have associated with them $k - 1$ distinct eigenvectors. Suppose $\mathbf{A}$ has k distinct eigenvalues $\lambda_1, \cdots, \lambda_k$ with corresponding eigenvectors $\mathbf{V}_1, \cdots, \mathbf{V}_k$. We want to show that these eigenvectors are linearly independent.

If they were linearly dependent, then there would be numbers $c_1, \cdots, c_k$, not all zero, such that

$$c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + \cdots + c_k\mathbf{V}_k = \mathbf{O}.$$

By relabeling if necessary, we may assume for convenience that $c_1 \neq 0$. Multiply this equation by $\lambda_1\mathbf{I}_n - \mathbf{A}$:

$$\begin{aligned}
\mathbf{O} &= (\lambda_1\mathbf{I}_n - \mathbf{A})(c_1\mathbf{V}_1 + c_2\mathbf{V}_2 + \cdots + c_k\mathbf{V}_k) \\
&= c_1(\lambda_1\mathbf{I}_n - \mathbf{A})\mathbf{V}_1 + c_2(\lambda_1\mathbf{I}_n - \mathbf{A})\mathbf{V}_2 \\
&\quad + \cdots + c_k(\lambda_1\mathbf{I}_n - \mathbf{A})\mathbf{V}_k \\
&= c_1(\lambda_1\mathbf{V}_1 - \lambda_1\mathbf{V}_1) + c_2(\lambda_1\mathbf{V}_2 - \lambda_2\mathbf{V}_2) \\
&\quad + \cdots + c_k(\lambda_1\mathbf{V}_k - \lambda_k\mathbf{V}_k) \\
&= c_2(\lambda_1 - \lambda_2)\mathbf{V}_1 + \cdots + c_k(\lambda_1 - \lambda_k)\mathbf{V}_k.
\end{aligned}$$

Now $\mathbf{V}_2, \cdots, \mathbf{V}_k$ are linearly independent by the inductive hypothesis, so these coefficients are all zero. But $\lambda_1 \neq \lambda_j$ for $j = 2, \cdots, k$ by the assumptions that the eigenvalues are distinct. Therefore

$$c_2 = \cdots = c_k = 0.$$

But then $c_1\mathbf{V}_1 = \mathbf{O}$. Since an eigenvalue cannot be the zero vector, this means that $c_1 = 0$ also. Therefore $\mathbf{V}_1, \dots, \mathbf{V}_k$ are linearly independent. By induction, this proves the theorem. ♦

In Example 9.3, the 3×3 matrix $\mathbf{A}$ had only two distinct eigenvalues, and only two linearly independent eigenvectors. However, the matrix of the next example has three linearly independent eigenvectors even though it has only two distinct eigenvalues. When eigenvalues are repeated, a matrix may or may not have n linearly independent eigenvectors.

EXAMPLE 9.5

Let

$$\mathbf{A} = \begin{pmatrix} 5 & -4 & 4 \\ 12 & -11 & 12 \\ 4 & -4 & 5 \end{pmatrix}.$$

The eigenvalues of $\mathbf{A}$ are $-3, 1, 1$, with 1 a repeated root of the characteristic polynomial. Corresponding to -3, we find an eigenvector

$$\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}.$$

Now look for an eigenvector corresponding to 1. We must solve the system

$$((1)\mathbf{I}_2 - \mathbf{A})\mathbf{X} = \begin{pmatrix} -4 & 4 & -4 \\ -12 & 12 & -12 \\ -4 & 4 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This system has the general solution

$$\alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

in which α and β are any numbers. With $\alpha = 1$ and $\beta = 0$, and then with $\alpha = 0$ and $\beta = 1$, we obtain two linearly independent eigenvectors associated with eigenvalue 1:

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

For this matrix $\mathbf{A}$, we can produce three linearly independent eigenvectors, even though the eigenvalues are not distinct. ♦

Eigenvalues and eigenvectors of special classes of matrices may exhibit special properties. Symmetric matrices form one such class. $\mathbf{A} = [a_{ij}]$ is *symmetric* if $a_{ij} = a_{ji}$ whenever $i \neq j$. This means that $\mathbf{A} = \mathbf{A}^t$, hence that each off-diagonal element is equal to its reflection across this main diagonal. For example,

$$\begin{pmatrix} -7 & -2-i & 1 & 14 \\ -2-i & 2 & -9 & 47i \\ 1 & -9 & -4 & \pi \\ 14 & 47i & \pi & 22 \end{pmatrix}$$

is symmetric.

It is a significant property of symmetric matrices that those with real elements have all real eigenvalues.

THEOREM 9.3 Eigenvalues of Real Symmetric Matrices

The eigenvalues of a real symmetric matrix are real. ♦

Proof By Lemma 9.1 (equation (9.2)), for any eigenvalue λ of $\mathbf{A}$, with eigenvector $\mathbf{E} = (e_1, \cdots, e_n)$,

$$\lambda = \frac{\overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}}{\overline{\mathbf{E}}^t \mathbf{E}}.$$

As noted previously, the denominator is

$$\overline{\mathbf{E}}^t \mathbf{E} = \sum_{j=1}^{n} |e_j|^2$$

and this is real. All we have to do is show that the numerator real, which we will do by showing that $\overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}$ equals its complex conjugate. First, because elements of $\mathbf{A}$ are real, each equals its own conjugate, so $\overline{\mathbf{A}} = \mathbf{A}$. Further, because $\mathbf{A}$ is symmetric, $\mathbf{A}^t = \mathbf{A}$. Therefore

$$\overline{\overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}} = \overline{\overline{\mathbf{E}}}^t \overline{\mathbf{A}}\overline{\mathbf{E}} = \overline{\overline{\mathbf{E}}}^t \mathbf{A}\overline{\mathbf{E}} = \mathbf{E}^t \mathbf{A}\overline{\mathbf{E}}.$$

But the last quantity is a 1×1 matrix, which equals its own transpose. Thus, continuing the last equation,

$$\mathbf{E}^t \mathbf{A}\overline{\mathbf{E}} = (\mathbf{E}^t \mathbf{A}\overline{\mathbf{E}})^t = (\overline{\mathbf{E}}^t)\mathbf{A}(\mathbf{E}^t)^t = \overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}.$$

The last two equations together show that $\overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}$ is its own conjugate, hence is real, proving the theorem. ♦

If the eigenvalues of a real matrix are all real, then associated eigenvectors will have real elements as well. In the case that $\mathbf{A}$ is also symmetric, we claim that eigenvectors associated with distinct eigenvalues must be orthogonal.

THEOREM 9.4 Orthogonality of Eigenvectors

Let $\mathbf{A}$ be a real symmetric matrix. Then eigenvectors associated with distinct eigenvalues are orthogonal.

Proof We can derive this result by a useful interplay between matrix and vector notation. Let λ and μ be distinct eigenvalues of $\mathbf{A}$, with eigenvectors, respectively,

$$\mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \text{ and } \mathbf{G} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{pmatrix}.$$

We have seen that

$$\mathbf{E} \cdot \mathbf{G} = e_1 g_1 + e_2 g_2 + \cdots + e_n g_n = \mathbf{E}^t \mathbf{G}.$$

Now use the facts that $\mathbf{A}\mathbf{E} = \lambda\mathbf{E}$, $\mathbf{A}\mathbf{G} = \mu\mathbf{G}$, and $\mathbf{A} = \mathbf{A}^t$ to write

$$\begin{aligned} \lambda\mathbf{E}^t\mathbf{G} &= (\mathbf{A}\mathbf{E})^t\mathbf{G} = (\mathbf{E}^t\mathbf{A}^t)\mathbf{G} \\ &= (\mathbf{E}^t\mathbf{A})\mathbf{G} = \mathbf{E}^t(\mathbf{A}\mathbf{G}) = \mathbf{E}^t(\mu\mathbf{G}) = \mu\mathbf{E}^t\mathbf{G}. \end{aligned}$$

But then

$$(\lambda - \mu)\mathbf{E}^t\mathbf{G} = 0.$$

Since $\lambda \neq \mu$, then $\mathbf{E}^t\mathbf{G} = \mathbf{E} \cdot \mathbf{G} = 0$. ♦

EXAMPLE 9.6

$$\mathbf{A} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 0 \end{pmatrix}$$

is a 3×3 symmetric matrix. The eigenvalues are 2, -1, and 4, with associated eigenvectors

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \text{ and } \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

These eigenvectors are mutually orthogonal. ♦

Finding eigenvalues of a matrix may be difficult because finding the roots of a polynomial can be difficult. In MAPLE, the command

```
eigenvals(A);
```

will list the eigenvalues of **A**, if n is not too large. The command

```
eigenvects(A);
```

will list each eigenvalue, its multiplicity, and, for each eigenvalue, as many linearly independent eigenvectors as are associated with that eigenvalue. We can also find the characteristic polynomial of **A** by

```
charpoly(A,t);
```

in which the variable of the polynomial is called t, but could be given any designation.

There is a method due to Gershgorin that enables us to place the eigenvalues inside disks in the complex plane. This is sometimes useful to get some idea of how the eigenvalues of a matrix are distributed.

THEOREM 9.5 *Gershgorin*

Let **A** be an $n \times n$ matrix of numbers. For $k = 1, 2, \cdots, n$ let

$$r_k = \sum_{j=1, j\neq k}^{n} |a_{kj}|.$$

Let C_k be the circle of radius r_k centered at (α_k, β_k), where $a_{kk} = \alpha_k + \beta_k i$. Then each eigenvalue of **A**, when plotted as a point in the complex plane, lies on or within one of the circles $C_1, \cdots, C_n$. ♦

C_k is the circle centered at the kth diagonal element a_{kk} of **A**, having radius equal to the sum of the magnitudes of the elements across row k, excluding the diagonal element occurring in that row.

EXAMPLE 9.7

Let

$$\mathbf{A} = \begin{pmatrix} 12i & 1 & 3 \\ 2 & -6 & 2+i \\ 3 & 1 & 5 \end{pmatrix}.$$

The characteristic polynomial of **A** is

$$p_{\mathbf{A}}(\lambda) = \lambda^3 + (1-12i)\lambda^2 - (43+13i)\lambda - 68 + 381i.$$

The Gershgorin circles have centers and radii:

$$C_1 : (0, 12), r_1 = 1 + 3 = 4,$$
$$C_2 : (-6, 0), r_2 = 2 + \sqrt{5}$$
$$C_3 : (5, 0), r_3 = 3 + 1 = 4.$$

Figure 9.1 shows these Gershgorin circles. The eigenvalues are in the disks determined by these circles. ♦

Gershgorin's theorem is not a way of approximating eigenvalues, since some of the disks may have large radii. However, sometimes important information that is revealed by these disks can be useful. For example, in studies of the stability of fluid flow it is important to know whether eigenvalues occur in the right half-plane.

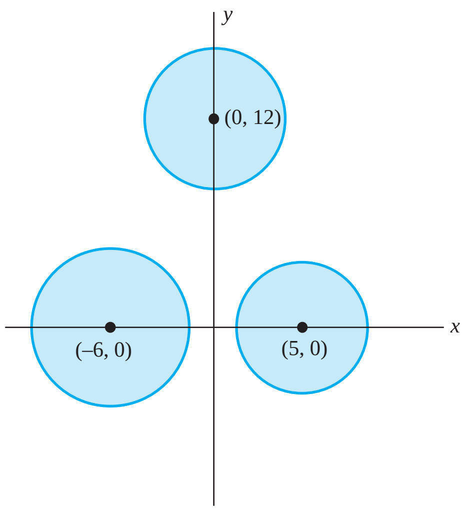

FIGURE 9.1 *Gerschgorin circles in Example 9.7.*

SECTION 9.1 PROBLEMS

In each of Problems 1 through 16, find the eigenvalues of the matrix. For each eigenvalue, find an eigenvector. Sketch the Gershgorin circles for the matrix and locate the eigenvalues as points in the plane.

1. $\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$

2. $\begin{pmatrix} -2 & 0 \\ 1 & 4 \end{pmatrix}$

3. $\begin{pmatrix} -5 & 0 \\ 1 & 2 \end{pmatrix}$

4. $\begin{pmatrix} 6 & -2 \\ -3 & 4 \end{pmatrix}$

5. $\begin{pmatrix} 1 & -6 \\ 2 & 2 \end{pmatrix}$

6. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

7. $\begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 3 \end{pmatrix}$

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

9. $\begin{pmatrix} -3 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

10. $\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \end{pmatrix}$

11. $\begin{pmatrix} -14 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

12. $\begin{pmatrix} 3 & 0 & 0 \\ 1 & -2 & -8 \\ 0 & -5 & 1 \end{pmatrix}$

13. $\begin{pmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ -5 & 0 & 7 \end{pmatrix}$

14. $\begin{pmatrix} -2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

15. $\begin{pmatrix} -4 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{pmatrix}$

16. $\begin{pmatrix} 5 & 1 & 0 & 9 \\ 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

In each of Problems 17 through 22, find the eigenvalues and associated eigenvectors of the matrix. Verify that eigenvectors associated with distinct eigenvalues are orthogonal.

17. $\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$

18. $\begin{pmatrix} -3 & 5 \\ 5 & 4 \end{pmatrix}$

19. $\begin{pmatrix} 6 & 1 \\ 1 & 4 \end{pmatrix}$

20. $\begin{pmatrix} -13 & 1 \\ 1 & 4 \end{pmatrix}$

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

22. $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}$

23. Suppose λ is an eigenvalue of $\mathbf{A}$ with eigenvector $\mathbf{E}$. Let k be a positive integer. Show that λ^k is an eigenvalue of $\mathbf{A}^k$ with eigenvector $\mathbf{E}$.

24. Let $\mathbf{A}$ be an $n \times n$ matrix of numbers. Show that the constant term in the characteristic polynomial of $\mathbf{A}$ is $(-1)^n|\mathbf{A}|$. Use this to show that any singular matrix must have 0 as an eigenvalue.

9.2 Diagonalization

Recall that the elements a_{ii} of a matrix make up its *main diagonal*. All other matrix elements are called *off-diagonal elements*.

A square matrix is called a *diagonal matrix* if all the off-diagonal elements are zero. A diagonal matrix has the appearance

$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & d_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & d_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & d_n \end{pmatrix}.$$

Diagonal matrices have many pleasant properties. Let **A** and **B** be $n \times n$ diagonal matrices with diagonal elements, respectively, a_{ii} and b_{ii}.

1. $\mathbf{A}+\mathbf{B}$ is diagonal with diagonal elements $a_{ii}+b_{ii}$.
2. **AB** is diagonal with diagonal elements $a_{ii}b_{ii}$.
3. $$|\mathbf{A}| = a_{11}a_{22}\cdots a_{nn},$$
   the product of the diagonal elements.
4. From (3), **A** is nonsingular exactly when each diagonal element is nonzero (so **A** has nonzero determinant). In this event, $\mathbf{A}^{-1}$ is the diagonal matrix having diagonal elements $1/a_{ii}$.
5. The eigenvalues of **A** are its diagonal elements.
6. $$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$
   with all zero elements except for 1 in the $i, 1$ place, is an eigenvector corresponding to the eigenvalue a_{ii}.

Most matrices are not diagonal. However, sometimes it is possible to transform a matrix to a diagonal one. This will enable us to transform some problems to simpler ones.

An $n \times n$ matrix **A** is *diagonalizable* if there is an $n \times n$ matrix **P** such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is a diagonal matrix. In this case we say that **P** *diagonalizes* **A**.

We will see that not every matrix is diagonalizable. The following result not only tells us exactly when **A** is diagonalizable, but also how to choose **P** to diagonalize **A**, and what $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ must look like.

THEOREM 9.6 *Diagonalization of a Matrix*

Let **A** be $n \times n$. Then **A** is diagonalizable if and only if **A** has n linearly independent eigenvectors. Furthermore, if **P** is the $n \times n$ matrix having these eigenvectors as columns, then $\mathbf{P}^{-1}\mathbf{AP}$ is the $n \times n$ diagonal matrix having the eigenvalues of **A** down its main diagonal, in the order in which the eigenvectors were chosen as columns of **P**.

In addition, if **Q** is any matrix that diagonalizes **A**, then necessarily the diagonal matrix $\mathbf{Q}^{-1}\mathbf{AQ}$ has the eigenvalues of **A** along its main diagonal, and the columns of **Q** must be eigenvectors of **A**, in the order in which the eigenvalues appear on the main diagonal of $\mathbf{Q}^{-1}\mathbf{AQ}$. ♦

We will prove the theorem after looking at three examples.

EXAMPLE 9.8

Let

$$\mathbf{A} = \begin{pmatrix} -1 & 4 \\ 0 & 3 \end{pmatrix}.$$

A has eigenvalues $-1, 3$ and corresponding linearly independent eigenvectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Form

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Determine

$$\mathbf{P}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

A simple computation shows that

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix},$$

a diagonal matrix with the eigenvalues of **A** on the main diagonal, in the order in which the eigenvectors were used to form the columns of

If we reverse the order of these eigenvectors as columns and define

$$\mathbf{Q} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

then

$$\mathbf{Q}^{-1}\mathbf{AQ} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$$

with the eigenvalues along the main diagonal, but now in the order reflecting the order of the eigenvectors used in forming the columns of **Q**. ♦

EXAMPLE 9.9

Let

$$\mathbf{A} = \begin{pmatrix} -1 & 1 & 3 \\ 2 & 1 & 4 \\ 1 & 0 & -2 \end{pmatrix}.$$

The eigenvalues are -1, $(-1+\sqrt{29})/2$ and $(-1-\sqrt{29})/2$, and corresponding eigenvectors are

$$\begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 3+\sqrt{29} \\ 10+2\sqrt{29} \\ 2 \end{pmatrix}, \begin{pmatrix} 3-\sqrt{29} \\ 10-2\sqrt{29} \\ 2 \end{pmatrix}.$$

These are linearly independent because the eigenvalues are distinct. Use these eigenvectors as columns of **P** to form

$$\mathbf{P} = \begin{pmatrix} 1 & 3+\sqrt{29} & 3-\sqrt{29} \\ -3 & 10+2\sqrt{29} & 10-2\sqrt{29} \\ 1 & 2 & 2 \end{pmatrix}.$$

We find that

$$\mathbf{P}^{-1} = \frac{\sqrt{29}}{812} \begin{pmatrix} 232/\sqrt{29} & -116/\sqrt{29} & 232/\sqrt{29} \\ 16-2\sqrt{29} & -1+\sqrt{29} & -19+5\sqrt{29} \\ -16-2\sqrt{29} & 1+\sqrt{29} & 19+5\sqrt{29} \end{pmatrix}$$

and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & (-1+\sqrt{29})/2 & 0 \\ 0 & 0 & (-1-\sqrt{29})/2 \end{pmatrix},$$

with the eigenvalues down the main diagonal in the order of the eigenvalues listed for columns of **P**.

In this example, $\mathbf{P}^{-1}$ is an unpleasant matrix. One of the values of Theorem 9.6 is that it tells us what $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ looks like, without actually having to determine $\mathbf{P}^{-1}$ and carry out this product. ♦

Although n distinct eigenvalues guarantee that **A** is diagonalizable, an $n \times n$ matrix with fewer than n distinct eigenvalues *may* still be diagonalizable. This will occur if we are able to find n linearly independent eigenvectors.

EXAMPLE 9.10

Let

$$\mathbf{A} = \begin{pmatrix} 5 & -4 & 4 \\ 12 & -11 & 12 \\ 4 & -4 & 5 \end{pmatrix}$$

as in Example 9.5. We found the eigenvalues $-3, 1, 1$, with a repeated eigenvalue. Nevertheless, we were able to find three linearly independent eigenvectors. Use these as columns to form

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 \\ 3 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

Then $\mathbf{P}$ diagonalizes $\mathbf{A}$:

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Again, we know this from Theorem 9.6, without explicitly computing the product $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. ♦

If $\mathbf{A}$ has fewer than n linearly independent eigenvectors, then $\mathbf{A}$ is not diagonalizable.

We will now prove Theorem 9.6.

Proof Let the eigenvalues of $\mathbf{A}$ be $\lambda_1, \lambda_2, \cdots, \lambda_n$ (not necessarily distinct). Suppose first that these eigenvalues have corresponding linearly independent eigenvectors $\mathbf{V}_1, \mathbf{V}_2, \cdots, \mathbf{V}_n$. These form the columns of $\mathbf{P}$, which we indicate by writing

$$\mathbf{P} = \begin{pmatrix} | & | & \cdots & | \\ \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_n \\ | & | & \cdots & | \end{pmatrix}.$$

$\mathbf{P}$ is nonsingular because its columns are linearly independent.

Let $\mathbf{D}$ be the $n \times n$ diagonal matrix having the eigenvalues of $\mathbf{A}$, in the given order, down the main diagonal. We want to prove that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}.$$

We will prove this by showing by direct computation that

$$\mathbf{A}\mathbf{P} = \mathbf{P}\mathbf{D}.$$

First, recall that the product $\mathbf{A}\mathbf{P}$ has as columns the product of $\mathbf{A}$ with the columns of $\mathbf{P}$. Thus

$$\begin{aligned} \text{column } j \text{ of } \mathbf{A}\mathbf{P} &= \mathbf{A}(\text{column } j \text{ of } \mathbf{P}) \\ &= \mathbf{A}(\mathbf{V}_j) = \lambda_j \mathbf{V}_j. \end{aligned}$$

Now compute $\mathbf{P}\mathbf{D}$. As a convenience in understanding the computation, write

$$\mathbf{V}_j = \begin{pmatrix} v_{1j} \\ v_{2j} \\ \cdots \\ v_{nj} \end{pmatrix}.$$

Then

$$\mathbf{P}\mathbf{D} = \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} & \cdots & \lambda_n v_{1n} \\ \lambda_1 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_n v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_1 v_{n1} & \lambda_2 v_{2n} & \cdots & \lambda_n v_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} | & | & \cdots & | \\ \lambda_1 \mathbf{V}_1 & \lambda_2 \mathbf{V}_2 & \cdots & \lambda_n \mathbf{V}_n \\ | & | & \cdots & | \end{pmatrix} = \mathbf{A}\mathbf{P},$$

since column j of this matrix is $\lambda_j \mathbf{V}_j$.

Thus far we have proved that, if **A** has n linearly independent eigenvectors, then **A** is diagonalizable and $\mathbf{P}^{-1}\mathbf{AP}$ is the diagonal matrix having the eigenvalues down the main diagonal, in the order in which the eigenvectors are seen as columns of **P**.

To prove the converse, now suppose that **A** is diagonalizable. We want to show that **A** has n linearly independent eigenvectors (regardless of whether the eigenvalues are distinct). Further, we want to show that, if $\mathbf{Q}^{-1}\mathbf{AQ}$ is a diagonal matrix, then the diagonal elements of this matrix are the eigenvalues of **A**, and the columns of **Q** are corresponding eigenvectors. Thus suppose that

$$\mathbf{Q}^{-1}\mathbf{AQ} = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} = \mathbf{D}.$$

Let $\mathbf{V}_j$ be column j of **Q**. These columns are linearly independent because **Q** is nonsingular. We will show that d_j is an eigenvalue of **A** with eigenvector $\mathbf{V}_j$.

From $\mathbf{Q}^{-1}\mathbf{AQ} = \mathbf{D}$, we have $\mathbf{AQ} = \mathbf{QD}$. Compute both sides of this equation separately. First, since the columns of **Q** are the $\mathbf{V}_j$'s, then

$$\mathbf{QD} = \begin{pmatrix} | & | & \cdots & | \\ \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_n \\ | & | & \cdots & | \end{pmatrix} \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

$$= \begin{pmatrix} | & | & \cdots & | \\ d_1\mathbf{V}_1 & d_2\mathbf{V}_2 & \cdots & d_n\mathbf{V}_n \\ | & | & \cdots & | \end{pmatrix},$$

which is a matrix having $d_j\mathbf{V}_j$ as column j. Now compute

$$\mathbf{AQ} = \mathbf{A}\begin{pmatrix} | & | & \cdots & | \\ \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_n \\ | & | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & | & \cdots & | \\ \mathbf{AV}_1 & \mathbf{AV}_2 & \cdots & \mathbf{AV}_n \\ | & | & \cdots & | \end{pmatrix},$$

which is a matrix having $\mathbf{AV}_j$ as column j. Since these matrices are equal, then

$$\mathbf{AV}_j = d_j\mathbf{V}_j$$

and this makes d_j an eigenvalue of **A** with eigenvector $\mathbf{V}_j$. ♦

Not every matrix is diagonalizable. We know from the theorem that a $n \times n$ matrix with fewer than n linearly independent eigenvectors is not diagonalizable.

EXAMPLE 9.11

Let

$$\mathbf{B} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

B has eigenvalues 1, 1, and all eigenvectors are constant multiples of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore **B** has as eigenvectors only nonzero multiples of one vector, and does not have two linearly independent eigenvectors. By the theorem, **B** is not diagonalizable.

Notice that, if **P** diagonalized **A**, then **P** would have to have eigenvectors of **B** as columns. Then **P** would have to have the form

$$\mathbf{P} = \begin{pmatrix} \alpha & \beta \\ 0 & 0 \end{pmatrix}$$

for some nonzero α and β. But this matrix is singular, with no inverse, because $|\mathbf{P}| = 0$. ♦

The key to diagonalizing **A** is the existence of n linearly independent eigenvectors. By Theorem 9.2, one circumstance in which this always happens is that **A** has n distinct eigenvalues.

COROLLARY 9.1

An $n \times n$ matrix with n distinct eigenvalues must be diagonalizable. ♦

EXAMPLE 9.12

Let

$$\mathbf{A} = \begin{pmatrix} -2 & 0 & 0 & 5 \\ 1 & 3 & 0 & 0 \\ 0 & 4 & 4 & 0 \\ 2 & 0 & 0 & -3 \end{pmatrix}.$$

A has eigenvalues $3, 4, (-5+\sqrt{41})/2$ and $(-5-\sqrt{41})/2$. Because these are distinct, **A** has 4 linearly independent eigenvectors and therefore is diagonalizable. There is a matrix **P**such that

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & (-5+\sqrt{41})/2 & 0 \\ 0 & 0 & 0 & (-5-\sqrt{41})/2 \end{pmatrix}.$$

We do not have to actually write down **P** (this would require finding eigenvectors) or compute $\mathbf{P}^{-1}$ to draw this conclusion. ♦

SECTION 9.2 PROBLEMS

In each of Problems 1 through 10, produce a matrix **P** that diagonalizes the given matrix, or show that the matrix is not diagonalizable. Determine $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$. *Hint*: Keep in mind that it is not necessary to compute **P** to know this product matrix.

1. $\begin{pmatrix} 0 & -1 \\ 4 & 3 \end{pmatrix}$

2. $\begin{pmatrix} 5 & 3 \\ 1 & 3 \end{pmatrix}$

3. $\begin{pmatrix} 1 & 0 \\ -4 & 1 \end{pmatrix}$

4. $\begin{pmatrix} -5 & 3 \\ 0 & 9 \end{pmatrix}$

5. $\begin{pmatrix} 5 & 0 & 0 \\ 1 & 0 & 3 \\ 0 & 0 & -2 \end{pmatrix}$

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

7. $\begin{pmatrix} -2 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$

8. $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}$

9. $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}$

10. $\begin{pmatrix} -2 & 0 & 0 & 0 \\ -4 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}$

11. Let $\mathbf{A}$ have eigenvalues $\lambda_1, \cdots, \lambda_n$, and suppose that $\mathbf{P}$ diagonalizes $\mathbf{A}$. Show that, for any positive integer k,

$$\mathbf{A}^k = \mathbf{P} \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n{}^k \end{pmatrix} \mathbf{P}^{-1}.$$

In each of Problems 12 through 15, use the idea of Problem 11 to compute the indicated power of the matrix.

12. $\mathbf{A} = \begin{pmatrix} -3 & -3 \\ -2 & 4 \end{pmatrix}$; $\mathbf{A}^{16}$

13. $\mathbf{A} = \begin{pmatrix} -1 & 0 \\ 1 & -5 \end{pmatrix}$; $\mathbf{A}^{18}$

14. $\mathbf{A} = \begin{pmatrix} -2 & 3 \\ 3 & -4 \end{pmatrix}$; $\mathbf{A}^{31}$

15. $\mathbf{A} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$; $\mathbf{A}^{43}$

16. Suppose $\mathbf{A}^2$ is diagonalizable. Prove that $\mathbf{A}$ is diagonalizable.

9.3 Some Special Types of Matrices

In this section, we will discuss several types of matrices having special properties.

9.3.1 Orthogonal Matrices

An $n \times n$ matrix is *orthogonal* if its transpose is its inverse:

$$\mathbf{A}^{-1} = \mathbf{A}^t.$$

In this event,

$$\mathbf{A}\mathbf{A}^t = \mathbf{A}^t\mathbf{A} = \mathbf{I}_n.$$

For example, it is routine to check that

$$\mathbf{A} = \begin{pmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & 2\sqrt{5} & -1\sqrt{5} \end{pmatrix}$$

is orthogonal. Just multiply this matrix by its transpose to obtain $\mathbf{I}_3$.

Because $(\mathbf{A}^t)^t = \mathbf{A}$, a matrix is orthogonal exactly when its transpose is orthogonal.

It is also easy to verify that an orthogonal matrix must have determinant 1 or -1.

THEOREM 9.7

If $\mathbf{A}$ is orthogonal, then $|\mathbf{A}| = \pm 1$. ♦

Proof Because a matrix and its transpose have the same determinant,

$$|\mathbf{I}_n| = 1 = |\mathbf{A}\mathbf{A}^{-1}| = |\mathbf{A}\mathbf{A}^t| = |\mathbf{A}||\mathbf{A}^t| = |\mathbf{A}|^2. \ ♦$$

The name "orthogonal matrix" derives from the following property.

THEOREM 9.8

Let **A** be an $n \times n$ matrix of real numbers. Then

1. **A** is orthogonal if and only the row vectors are mutually orthogonal unit vectors in R^n.
2. **A** is orthogonal if and only if the column vectors are mutually orthogonal unit vectors in R^n. ♦

We say that the row vectors of an orthogonal matrix form an *orthonormal set of vectors* in R^n. The column vectors also form an orthonormal set.

Proof The i, j element of $\mathbf{AA}^t$ is the dot product of row i of **A** with column j of $\mathbf{A}^t$, and this is the dot product of row i of **A** with row j of **A**.

If $i \neq j$, then this dot product is zero, because the $i, j-$ element of $\mathbf{I}_n$ is zero. And if $i = j$, then this dot product is 1 because the $i, i-$ element of $\mathbf{I}_n$ is 1. This proves that, if **A** is an orthogonal matrix, then its rows form an orthonormal set of vectors in R^n.

Conversely, suppose the rows are mutually orthogonal unit vectors in R^n. Then the i, j element of $\mathbf{AA}^t$ is 0 if $i \neq j$ and 1 if $i = j$, so $\mathbf{AA}^t = \mathbf{I}_n$.

By applying this argument to $\mathbf{A}^t$, this transpose is orthogonal if and only if its rows are orthogonal unit vectors, and these rows are the columns of **A**. ♦

We now know a lot about orthogonal matrices. We will use this information to determine all 2×2 real orthogonal matrices. Suppose

$$\mathbf{Q} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is orthogonal. What does this tell us about a, b, c and d? Because the row (column) vectors are mutually orthogonal unit vectors,

$$ac + bd = 0$$

$$ab + cd = 0$$

$$a^2 + b^2 = 1$$

$$c^2 + d^2 = 1.$$

Furthermore, $|\mathbf{Q}| = \pm 1$, so

$$ad - bc = 1 \text{ or } ad - bc = -1.$$

By analyzing these equations in all cases, we find that there must be some θ in $[0, 2\pi)$ such that $a = \cos(\theta)$ and $b = \sin(\theta)$, and **Q** must have one of the two forms:

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \text{ or } \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix},$$

depending on whether the determinant is 1 or -1. For example, with $\theta = \pi/6$, we obtain the orthogonal 2×2 matrices

$$\begin{pmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{pmatrix} \text{ or } \begin{pmatrix} \sqrt{3}/2 & 1/2 \\ 1/2 & -\sqrt{3}/2 \end{pmatrix}.$$

If we put Theorems 9.4 and 9.8 together, we obtain an interesting conclusion. Suppose **S** is a real, symmetric $n \times n$ matrix with n distinct eigenvalues. Then the associated eigenvectors are orthogonal. These may not be unit vectors. However, a scalar multiple of an eigenvector is still an eigenvector. Divide each eigenvector by its length and use these unit eigenvectors as columns of an orthogonal matrix **Q** that diagonalizes **S**. This proves the following.

THEOREM 9.9

An $n \times n$ real symmetric matrix with distinct eigenvalues can be diagonalized by an orthogonal matrix. ♦

EXAMPLE 9.13

Let

$$\mathbf{S} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 0 \end{pmatrix}.$$

This real, symmetric matrix has eigenvalues $2, -1, 4$, with corresponding eigenvectors

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \text{ and } \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

The matrix having these eigenvectors as columns will diagonalize **S**, but is not an orthogonal matrix because these eigenvectors do not all have length 1. Normalize the second and third eigenvectors by dividing them by their lengths, and then use these unit eigenvectors as columns of an orthogonal matrix **Q**:

$$\mathbf{Q} = \begin{pmatrix} 0 & 1/\sqrt{5} & 2/\sqrt{5} \\ 1 & 0 & 0 \\ 0 & 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}.$$

This orthogonal matrix also diagonalizes **S**. ♦

9.3.2 Unitary Matrices

We will use the following fact. If **W** is any matrix, then the operations of taking the transpose and the complex conjugate can be performed in either order:

$$\overline{(\mathbf{W}^t)} = (\overline{\mathbf{W}})^t.$$

This is verified by a routine calculation.

It is also straightforward to verify that the operations of taking a matrix inverse, and of taking its complex conjugate, can be performed in either order.

Now let **U** be an $n \times n$ matrix with complex elements.

We say that **U** is *unitary* if the inverse is the conjugate of the transpose (which is the same as the transpose of the conjugate):

$$\mathbf{U}^{-1} = \overline{\mathbf{U}}^t.$$

This means that

$$(\overline{\mathbf{U}})^t\mathbf{U} = \mathbf{U}(\overline{\mathbf{U}})^t = \mathbf{I}_n.$$

EXAMPLE 9.14

$$\mathbf{U} = \begin{pmatrix} i/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

It is routine to check that **U** is unitary. ♦

If $\mathbf{U}$ is a unitary matrix with real elements, then $\overline{\mathbf{U}} = \mathbf{U}$ and the condition of being unitary becomes $\mathbf{U}^{-1} = \mathbf{U}^t$. Therefore a real unitary matrix is orthogonal. In this sense unitary matrices are the extension of orthogonal matrices to allow complex matrix elements.

Since the rows (and columns) of an orthogonal matrix are mutually orthogonal unit vectors, we would expect a complex analogue of this condition for unitary matrices. If $(x_1, \cdots, x_n)$ and $(y_1, \cdots, y_n)$ are vectors in R^n, we can write

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

and obtain the dot product $\mathbf{X} \cdot \mathbf{Y}$ as the matrix product $\mathbf{X}^t\mathbf{Y}$, which is the 1×1 matrix (or number) $x_1y_1 + x_2y_2 + \cdots + x_ny_n$. In particular, the square of the length of $\mathbf{X}$ is

$$\mathbf{X}^t\mathbf{X} = x_1^2 + x_2^2 + \cdots + x_n^2.$$

To generalize this to the complex case, suppose we have complex $n-$ vectors $(z_1, z_2, \cdots, z_n)$ and $(w_1, w_2, \cdots, w_n)$. Let

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \text{ and } \mathbf{W} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

and define the dot product $\mathbf{Z} \cdot \mathbf{W}$ by

$$\mathbf{Z} \cdot \mathbf{W} = \overline{\mathbf{Z}}^t\mathbf{W}.$$

Then

$$\mathbf{Z} \cdot \mathbf{W} = \overline{z_1}w_1 + \overline{z_2}w_2 + \cdots + \overline{z_n}w_n.$$

In this way,

$$\mathbf{Z} \cdot \mathbf{Z} = \overline{z_1}z_1 + \overline{z_2}z_2 + \cdots + \overline{z_n}z_n = \sum_{j=1}^{n} |z_j|^2,$$

a real number, consistent with the interpretation of the dot product of a vector with itself as the square of the length. With this as background, we now define the complex analogue of an orthonormal set of vectors in R^n. We will say that complex $n-$ vectors $\mathbf{F}_1, \cdots, \mathbf{F}_r$ form a *unitary system* if $\mathbf{F}_j \cdot \mathbf{F}_k = 0$ if $j \neq k$, and each $\mathbf{F}_j$ has length 1 (that is, $\mathbf{F}_j \cdot \mathbf{F}_j = 1$).

A unitary system is an orthonormal set of vectors when each of the vectors has real components. With this background, we can state the unitary version of Theorem 9.8.

THEOREM 9.10

A complex matrix $\mathbf{U}$ is unitary if and only its row (column) vectors form a unitary system. ♦

We claim that the eigenvalues of a unitary matrix must have magnitude 1.

THEOREM 9.11

Let λ be an eigenvalue of a unitary matrix $\mathbf{U}$. Then $|\lambda| = 1$. ♦

This means that the eigenvalues of **U** lie on the unit circle about the origin in the complex plane. Since a real orthogonal matrix is also unitary, this also holds for real orthogonal matrices.

Proof Let λ be an eigenvalue of **U** with eigenvector **E**. We know that $\mathbf{UE} = \lambda\mathbf{E}$. Then $\overline{\mathbf{UE}} = \overline{\lambda}\overline{\mathbf{E}}$. Therefore,

$$(\overline{\mathbf{UE}})^t = \overline{\lambda}(\overline{\mathbf{E}})^t.$$

Then,

$$(\overline{\mathbf{E}})^t(\overline{\mathbf{U}})^t = \overline{\lambda}(\overline{\mathbf{E}})^t.$$

But **U** is unitary, so $\overline{\mathbf{U}}^t = \mathbf{U}^{-1}$. The last equation becomes

$$(\overline{\mathbf{E}})^t\mathbf{U}^{-1} = \overline{\lambda}(\overline{\mathbf{E}})^t.$$

Multiply both sides of this equation on the right by **UE**:

$$(\overline{\mathbf{E}})^t\mathbf{U}^{-1}\mathbf{UE} = \overline{\lambda}(\overline{\mathbf{E}})^t\mathbf{UE} = \overline{\lambda}(\overline{\mathbf{E}})^t\lambda\mathbf{E} = \overline{\lambda}\lambda\overline{\mathbf{E}}^t\mathbf{E}.$$

Now $\overline{\mathbf{E}}^t\mathbf{E}$ is the dot product of an eigenvector with itself, and so is a positive number. Dividing the last equation by $\overline{\mathbf{E}}^t\mathbf{E}$ yields the conclusion that $\overline{\lambda}\lambda = 1$. Then $|\lambda|^2 = 1$, proving the theorem. ♦

9.3.3 Hermitian and Skew-Hermitian Matrices

An $n \times n$ complex matrix **H** is *hermitian* if $\overline{\mathbf{H}} = \mathbf{H}^t$.

That is, a matrix is hermitian if its conjugate equals its transpose. If a hermitian matrix has real elements, then it must be symmetric, because then the matrix equals its conjugate, which equals its transpose.

An $n \times n$ complex matrix **S** is *skew-hermitian* if $\overline{\mathbf{S}} = -\mathbf{S}^t$.

Thus, **S** is skew-hermitian if its conjugate equals the negative of its transpose.

EXAMPLE 9.15

The matrix

$$\mathbf{H} = \begin{pmatrix} 15 & 8i & 6-2i \\ -8i & 0 & -4+i \\ 6+2i & -4-i & -3 \end{pmatrix}$$

is hermitian because

$$\overline{\mathbf{H}} = \begin{pmatrix} 15 & -8i & 6+2i \\ 8i & 0 & -4-i \\ 6-2i & -4+i & -3 \end{pmatrix} = \mathbf{H}^t.$$

The matrix

$$\mathbf{S} = \begin{pmatrix} 0 & 8i & 2i \\ 8i & 0 & 4i \\ 2i & 4i & 0 \end{pmatrix}$$

is skew-hermitian because

$$\overline{\mathbf{S}} = \begin{pmatrix} 0 & -8i & -2i \\ -8i & 0 & -4i \\ -2i & -4i & 0 \end{pmatrix} = -\mathbf{S}^t. \ \blacklozenge$$

We want to derive a result about eigenvalues of hermitian and skew-hermitian matrices. For this we need the following conclusions about the numerator of the general expression for eigenvalues in Lemma 9.1.

LEMMA 9.2

Let

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

be a complex $n \times 1$ matrix. Then

1. If $\mathbf{H}$ is $n \times n$ hermitian, then $\overline{\mathbf{Z}}^t\mathbf{HZ}$ is real.
2. If $\mathbf{S}$ is $n \times n$ skew-hermitian, then $\overline{\mathbf{Z}}^t\mathbf{HZ}$ is pure imaginary. ♦

Proof of Lemma 9.3 For condition (1), suppose $\mathbf{H}$ is hermitian, so that $\overline{\mathbf{H}}^t = \mathbf{H}$. Then

$$\overline{(\overline{\mathbf{Z}}^t\mathbf{HZ})} = (\overline{(\overline{\mathbf{Z}})^t})\overline{\mathbf{HZ}} = \mathbf{Z}^t\overline{\mathbf{HZ}}.$$

But $\overline{\mathbf{Z}}^t\mathbf{HZ}$ is a 1×1 matrix and so equals its own transpose. Continuing from the last equation, we have

$$\mathbf{Z}^t\overline{\mathbf{HZ}} = (\mathbf{Z}^t\overline{\mathbf{HZ}})^t = \overline{\mathbf{Z}}^t\overline{\mathbf{H}}^t(\mathbf{Z})^t = \overline{\mathbf{Z}}^t\mathbf{HZ}.$$

This shows that

$$\overline{(\overline{\mathbf{Z}}^t\mathbf{HZ})} = \overline{\mathbf{Z}}^t\mathbf{HZ}.$$

Since $\overline{\mathbf{Z}}^t\mathbf{HZ}$ equals its own conjugate, this quantity is real.

To prove condition (2), suppose $\mathbf{S}$ is skew-hermitian, so $\overline{\mathbf{S}}^t = -\mathbf{S}$. By an argument like that in the proof of condition (1), we find that

$$\overline{(\overline{\mathbf{Z}}^t\mathbf{SZ})} = -\overline{\mathbf{Z}}^t\mathbf{SZ}$$

If we write $\overline{\mathbf{Z}}^t\mathbf{SZ} = a + ib$, then the last equation means that

$$a - ib = -a - ib.$$

But then $a = -a$ so $a = 0$ and $\overline{\mathbf{Z}}^t\mathbf{SZ}$ is pure imaginary. This includes the possibility of a zero eigenvalue. ♦

This lemma absorbs most of the work we need for the following result, giving us information about eigenvalues.

THEOREM 9.12

1. The eigenvalues of a hermitian matrix are real.
2. The eigenvalues of a skew-hermitian are pure imaginary. ♦

Proof By Lemma 9.1, an eigenvalue λ of any $n \times n$ matrix $\mathbf{A}$, with corresponding eigenvector $\mathbf{E}$, satisfies

$$\lambda = \frac{\overline{\mathbf{E}}^t \mathbf{A}\mathbf{E}}{\overline{\mathbf{E}}^t \mathbf{E}}.$$

We know that the denominator of this quotient is a positive number. Now use Lemma 9.2. If $\mathbf{A}$ is hermitian, the numerator is real, so λ is real. If $\mathbf{A}$ is skew-hermitian then the numerator is pure imaginary, so λ is pure imaginary. ♦

Figure 9.2 shows a graphical representation of these conclusions about eigenvalues of matrices. When plotted as points in the complex plane, eigenvalues of a unitary (or orthogonal) matrix lie on the unit circle about the origin, eigenvalues of a hermitian matrix lie on the horizontal (real) axis, and eigenvalues of a skew-hermitian matrix are on the vertical (imaginary) axis.

9.3.4 Quadratic Forms

A *quadratic form* is an expression

$$\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}\overline{z_j}z_k$$

in which the a_{jk}'s and the z_j's are complex numbers. If these quantities are all real, we say that we have a *real quadratic form*.

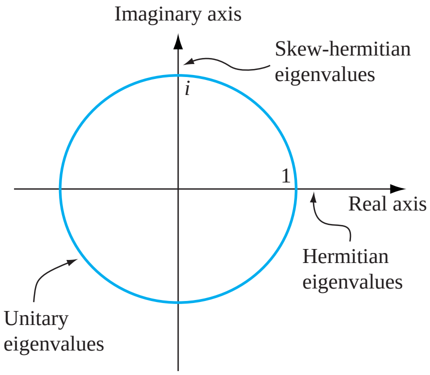

FIGURE 9.2 *Eigenvalues of unitary, hermitian, and skew-hermitian matrices.*

For $n = 2$, the quadratic form is

$$\sum_{j=1}^{2}\sum_{k=1}^{2} a_{jk}\overline{z_j}z_k = a_{11}\overline{z_1}z_1 + a_{12}\overline{z_1}z_2 + a_{21}z_1\overline{z_2} + a_{22}z_2\overline{z_2}.$$

The two middle terms are called *mixed product* terms, involving z_j and z_k with $j \neq k$.

If the quadratic form is real, then all of the numbers involved are real. In this case the conjugates play no role and this quadratic form can be written

$$\sum_{j=1}^{2}\sum_{k=1}^{2} a_{jk}x_j x_k = a_{11}x_1x_1 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2x_2$$

$$= a_1x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2.$$

As we have seen previously (in the discussion immediately preceding Lemma 9.1), we can let $\mathbf{A} = [a_{jk}]$ and write the complex quadratic form as $\overline{\mathbf{Z}}^t\mathbf{AZ}$, where

$$\mathbf{Z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}.$$

If all the quantities are real, we usually write this as $\mathbf{X}^t\mathbf{AX}$. In fact, any real quadratic form can be written in this way, with $\mathbf{A}$ a real symmetric matrix. We will illustrate this process.

EXAMPLE 9.16

Consider the real quadratic form

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 3x_1x_2 + 4x_2x_1 + 2x_2^2$$

$$= x_1^2 + 7x_1x_2 + 2x_2^2.$$

We can write the same quadratic form as

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix}\begin{pmatrix} 1 & 7/2 \\ 7/2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 7x_1x_2 + 2x_2^2$$

in which $\mathbf{A}$ is a symmetric matrix. ♦

This is important in developing a standard change of variables that is used to simplify quadratic forms by eliminating cross product terms.

THEOREM 9.13 Principal Axis Theorem

Let $\mathbf{A}$ be a real symmetric matrix with distinct eigenvalues $\lambda_1, \cdots \lambda_n$. Then there is an orthogonal matrix $\mathbf{Q}$ such that the change of variables $\mathbf{X} = \mathbf{QY}$ transforms the quadratic form $\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ij}x_ix_j$ to

$$\sum_{j=1}^{n} \lambda_j y_j^2.$$

Proof Let $\mathbf{Q}$ be an orthogonal matrix that diagonalizes $\mathbf{A}$. Then

$$\sum_{j=1}^{n}\sum_{k=1}^{n} a_{ij}x_i x_j = \mathbf{X}^t\mathbf{A}\mathbf{X}$$

$$= (\mathbf{Q}\mathbf{Y})^t\mathbf{A}\mathbf{Q}\mathbf{Y} = (\mathbf{Y}^t\mathbf{Q}^t)\mathbf{A}\mathbf{Q}\mathbf{Y}$$

$$= \mathbf{Y}^t(\mathbf{Q}^t\mathbf{A}\mathbf{Q})\mathbf{Y}$$

$$= \mathbf{Y}^t(\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q})\mathbf{Y}$$

$$= \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2. \blacklozenge$$

The expression $\sum_{j=1}^{n} \lambda_j y_j^2$ is called the *standard form* of $\mathbf{X}^t\mathbf{A}\mathbf{X}$.

EXAMPLE 9.17

Consider the quadratic form

$$x_1^2 - 7x_1x_2 + x_2^2.$$

This is $\mathbf{X}^t\mathbf{A}\mathbf{X}$, where

$$\begin{pmatrix} 1 & -7/2 \\ -7/2 & 1 \end{pmatrix}.$$

In general, the real quadratic form

$$ax_1^2 + bx_1x_2 + cx_2^2$$

can always be written as $\mathbf{X}^t\mathbf{A}\mathbf{X}$, with $\mathbf{A}$ the real symmetric matrix

$$\mathbf{A} = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}.$$

In this example, the eigenvalues of $\mathbf{A}$ are $-5/2, 9/2$ with corresponding eigenvectors

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Divide each eigenvector by its length to obtain columns of an orthogonal matrix $\mathbf{Q}$ that diagonalizes $\mathbf{A}$:

$$\mathbf{Q} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

The change of variables $\mathbf{X} = \mathbf{Q}\mathbf{Y}$ is equivalent to setting

$$x_1 = \frac{1}{\sqrt{2}}(y_1 - y_2)$$

$$x_2 = \frac{1}{\sqrt{2}}(y_1 + y_2).$$

This transforms the given quadratic form to its standard form

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 = -\frac{5}{2} y_1^2 + \frac{9}{2} y_2^2,$$

in which there are no cross product $y_1 y_2$ terms. ♦

SECTION 9.3 PROBLEMS

In each of Problems 1 through 12, find the eigenvalues and associated eigenvectors. Check that the eigenvectors associated with distinct eigenvalues are orthogonal. Find an orthogonal matrix that diagonalizes the matrix. Note Problems 17-22, Section 9.1.

1. $\begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix}$

2. $\begin{pmatrix} -3 & 5 \\ 5 & 4 \end{pmatrix}$

3. $\begin{pmatrix} 6 & 1 \\ 1 & 4 \end{pmatrix}$

4. $\begin{pmatrix} -13 & 1 \\ 1 & 4 \end{pmatrix}$

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

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

7. $\begin{pmatrix} 5 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}$

8. $\begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

9. $\begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}$

10. $\begin{pmatrix} 1 & 3 & 0 \\ 3 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

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

12. $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

In each of Problems 13 through 21, determine whether the matrix is unitary, hermitian, skew-hermitian, or none of these. Find the eigenvalues and associated eigenvectors. If the matrix is diagonalizable, write a matrix that diagonalizes it. In Problems 5 and 7, eigenvalues must be approximated, so only “approximate eigenvectors" can be found. It is instructive to try to diagonalize a matrix using approximate eigenvectors.

13. $\begin{pmatrix} 0 & 2i \\ 2i & 4 \end{pmatrix}$

14. $\begin{pmatrix} 3 & 4i \\ 4i & -5 \end{pmatrix}$

15. $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1-i \\ 0 & -1-i & 0 \end{pmatrix}$

16. $\begin{pmatrix} 1/\sqrt{2} & i/\sqrt{2} & 0 \\ -1/\sqrt{2} & i/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$

17. $\begin{pmatrix} 3 & 2 & 0 \\ 2 & 0 & i \\ 0 & -i & 0 \end{pmatrix}$

18. $\begin{pmatrix} -1 & 0 & 3-i \\ 0 & 1 & 0 \\ 3+i & 0 & 0 \end{pmatrix}$

19. $\begin{pmatrix} i & 1 & 0 \\ -1 & 0 & 2i \\ 0 & 2i & 0 \end{pmatrix}$

20. $\begin{pmatrix} 3i & 0 & 0 \\ -1 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$

21. $\begin{pmatrix} 8 & -1 & i \\ -1 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$

In each of Problems 22 through 28, determine a matrix **A** so that the quadratic form is $\mathbf{X}^t\mathbf{A}\mathbf{X}$, and find the standard form of the quadratic form.

22. $-5x_1^2 + 4x_1x_2 + 3x_2^2$
23. $4x_1^2 - 12x_1x_2 + x_2^2$
24. $-3x_1^2 + 4x_1x_2 + 7x_2^2$
25. $4x_1^2 - 4x_1x_2 + x_2^2$
26. $-6x_1x_2 + 4x_2^2$
27. $5x_1^2 + 4x_1x_2 + 2x_2^2$
28. $-2x_1x_2 + 2x_2^2$
29. Suppose **A** is hermitian. Show that

$$\overline{(\mathbf{A}\mathbf{A}^t)} = \overline{\mathbf{A}}\mathbf{A}.$$

30. Prove that the main diagonal elements of a hermitian matrix are real.
31. Prove that each main diagonal element of a skew-hermitian matrix is zero or pure imaginary.
32. Prove that the product of two unitary matrices is unitary.