5.4 Orthogonal Diagonalization of Symmetric Matrices

We saw in Chapter 4 that a square matrix with real entries will not necessarily have real eigenvalues. Indeed, the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ has complex eigenvalues i and $-i$. We also discovered that not all square matrices are diagonalizable. The situation changes dramatically if we restrict our attention to real *symmetric* matrices. As we will show in this section, all of the eigenvalues of a real symmetric matrix are real, and such a matrix is always diagonalizable.

Recall that a symmetric matrix is one that equals its own transpose. Let's begin by studying the diagonalization process for a symmetric 2×2 matrix.

Example 5.16

If possible, diagonalize the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

Solution The characteristic polynomial of A is $\lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2)$, from which we see that A has eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 2$. Solving for the corresponding eigenvectors, we find

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

respectively. So A is diagonalizable, and if we set $P = [\mathbf{v}_1 \quad \mathbf{v}_2]$, then we know that $P^{-1}AP = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} = D$.

However, we can do better. Observe that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. So, if we normalize them to get the unit eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

and then take

$$Q = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

we have $Q^{-1}AQ = D$ also. But now Q is an *orthogonal* matrix, since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal set of vectors. Therefore, $Q^{-1} = Q^T$, and we have $Q^TAQ = D$. (Note that checking is easy, since computing Q^{-1} only involves taking a transpose!)

The situation in Example 5.16 is the one that interests us. It is important enough to warrant a new definition.

Definition A square matrix A is ***orthogonally diagonalizable*** if there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^TAQ = D$.

We are interested in finding conditions under which a matrix is orthogonally diagonalizable. Theorem 5.17 shows us where to look.

Theorem 5.17 If A is orthogonally diagonalizable, then A is symmetric.

Proof If A is orthogonally diagonalizable, then there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^TAQ = D$. Since $Q^{-1} = Q^T$, we have $Q^TQ = I = QQ^T$, so

$$QDQ^T = QQ^TAQQ^T = IAI = A$$

But then

$$A^T = (QDQ^T)^T = (Q^T)^TD^TQ^T = QDQ^T = A$$

since every diagonal matrix is symmetric. Hence, A is symmetric.

Remark Theorem 5.17 shows that the orthogonally diagonalizable matrices are all to be found *among* the symmetric matrices. It does *not* say that every symmetric matrix must be orthogonally diagonalizable. However, it is a remarkable fact that this indeed is true! Finding a proof for this amazing result will occupy us for much of the rest of this section.

a + bi

We next prove that we don't need to worry about *complex* eigenvalues when working with symmetric matrices with *real* entries.

Theorem 5.18 If A is a real symmetric matrix, then the eigenvalues of A are real.

Recall that the *complex conjugate* of a complex number $z = a + bi$ is the number $\bar{z} = a - bi$ (see Appendix C). To show that z is real, we need to show that $b = 0$. One way to do this is to show that $z = \bar{z}$, for then $bi = -bi$ (or $2bi = 0$), from which it follows that $b = 0$.

We can also extend the notion of complex conjugate to vectors and matrices by, for example, defining $\bar{A}$ to be the matrix whose entries are the complex conjugates of the entries of A; that is, if $A = [a_{ij}]$, then $\bar{A} = [\bar{a}_{ij}]$. The rules for complex conjugation extend easily to matrices; in particular, we have $\overline{AB} = \bar{A}\bar{B}$ for compatible matrices A and B.

Proof Suppose that λ is an eigenvalue of A with corresponding eigenvector $\mathbf{v}$. Then $A\mathbf{v} = \lambda\mathbf{v}$, and, taking complex conjugates, we have $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}}$. But then

$$A\bar{\mathbf{v}} = \bar{A}\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$$

since A is real. Taking transposes and using the fact that A is symmetric, we have

$$\bar{\mathbf{v}}^TA = \bar{\mathbf{v}}^TA^T = (A\bar{\mathbf{v}})^T = (\bar{\lambda}\bar{\mathbf{v}})^T = \bar{\lambda}\bar{\mathbf{v}}^T$$

Therefore,

$$\lambda(\bar{\mathbf{v}}^T\mathbf{v}) = \bar{\mathbf{v}}^T(\lambda\mathbf{v}) = \bar{\mathbf{v}}^T(A\mathbf{v}) = (\bar{\mathbf{v}}^TA)\mathbf{v} = (\bar{\lambda}\bar{\mathbf{v}}^T)\mathbf{v} = \bar{\lambda}(\bar{\mathbf{v}}^T\mathbf{v})$$

or $(\lambda - \bar{\lambda})(\bar{\mathbf{v}}^T\mathbf{v}) = 0$.

Now if $\mathbf{v} = \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix}$, then $\bar{\mathbf{v}} = \begin{bmatrix} a_1 - b_1 i \\ \vdots \\ a_n - b_n i \end{bmatrix}$, so

$$\bar{\mathbf{v}}^T\mathbf{v} = (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \neq 0$$

since $\mathbf{v} \neq \mathbf{0}$ (because it is an eigenvector). We conclude that $\lambda - \overline{\lambda} = 0$, or $\lambda = \overline{\lambda}$. Hence, λ is real.

Theorem 4.20 showed that, for any square matrix, eigenvectors corresponding to distinct eigenvalues are linearly independent. For symmetric matrices, something stronger is true: Such eigenvectors are *orthogonal*.

Theorem 5.19 If A is a symmetric matrix, then any two eigenvectors corresponding to distinct eigenvalues of A are orthogonal.

Proof Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors corresponding to the distinct eigenvalues $\lambda_1 \neq \lambda_2$ so that $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. Using $A^T = A$ and the fact that $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^T\mathbf{y}$ for any two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, we have

$$\begin{aligned}\lambda_1(\mathbf{v}_1\cdot\mathbf{v}_2) &= (\lambda_1\mathbf{v}_1)\cdot\mathbf{v}_2 = A\mathbf{v}_1\cdot\mathbf{v}_2 = (A\mathbf{v}_1)^T\mathbf{v}_2 \\ &= (\mathbf{v}_1^TA^T)\mathbf{v}_2 = (\mathbf{v}_1^TA)\mathbf{v}_2 = \mathbf{v}_1^T(A\mathbf{v}_2) \\ &= \mathbf{v}_1^T(\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1^T\mathbf{v}_2) = \lambda_2(\mathbf{v}_1\cdot\mathbf{v}_2)\end{aligned}$$

Hence, $(\lambda_1 - \lambda_2)(\mathbf{v}_1\cdot\mathbf{v}_2) = 0$. But $\lambda_1 - \lambda_2 \neq 0$, so $\mathbf{v}_1\cdot\mathbf{v}_2 = 0$, as we wished to show.

Example 5.17

Verify the result of Theorem 5.19 for

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

Solution The characteristic polynomial of A is $-\lambda^3 + 6\lambda^2 - 9\lambda + 4 = -(\lambda - 4)\cdot(\lambda - 1)^2$, from which it follows that the eigenvalues of A are $\lambda_1 = 4$ and $\lambda_2 = 1$. The corresponding eigenspaces are

$$E_4 = \operatorname{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \quad \text{and} \quad E_1 = \operatorname{span}\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right)$$

(Check this.) We easily verify that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = 0$$

from which it follows that every vector in E_4 is orthogonal to every vector in E_1. (Why?)

Remark Note that $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = 1$. Thus, eigenvectors corresponding to the *same* eigenvalue need not be orthogonal.

We can now prove the main result of this section. It is called the Spectral Theorem, since the set of eigenvalues of a matrix is sometimes called the ***spectrum*** of the matrix. (Technically, we should call Theorem 5.20 the *Real* Spectral Theorem, since there is a corresponding result for matrices with complex entries.)

Theorem 5.20 **The Spectral Theorem**

Let A be an $n \times n$ real matrix. Then A is symmetric if and only if it is orthogonally diagonalizable.

Spectrum is a Latin word meaning "image." When atoms vibrate, they emit light. And when light passes through a prism, it spreads out into a spectrum—a band of rainbow colors. Vibration frequencies correspond to the eigenvalues of a certain operator and are visible as bright lines in the spectrum of light that is emitted from a prism. Thus, we can literally see the eigenvalues of the atom in its spectrum, and for this reason, it is appropriate that the word *spectrum* has come to be applied to the set of all eigenvalues of a matrix (or operator).

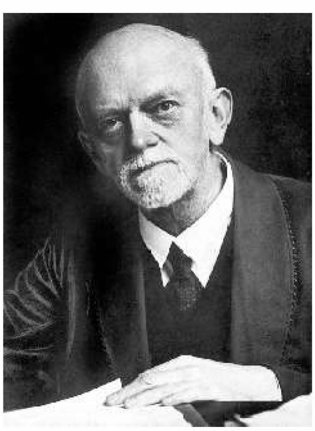
Proof We have already proved the "if" part as Theorem 5.17. To prove the "only if" implication, we proceed by induction on n. For $n = 1$, there is nothing to do, since a 1×1 matrix is already in diagonal form. Now assume that every $k \times k$ real symmetric matrix with real eigenvalues is orthogonally diagonalizable. Let $n = k + 1$ and let A be an $n \times n$ real symmetric matrix with real eigenvalues.

Let λ_1 be one of the eigenvalues of A and let $\mathbf{v}_1$ be a corresponding eigenvector. Then $\mathbf{v}_1$ is a *real* vector (why?) and we can assume that $\mathbf{v}_1$ is a unit vector, since otherwise we can normalize it and we will still have an eigenvector corresponding to λ_1. Using the Gram-Schmidt Process, we can extend $\mathbf{v}_1$ to an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$. Now we form the matrix

$$Q_1 = [\mathbf{v}_1 \quad \mathbf{v}_2 \cdots \mathbf{v}_n]$$

Then Q_1 is orthogonal, and

$$\begin{aligned} Q_1^T A Q_1 &= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} A [\mathbf{v}_1 \quad \mathbf{v}_2 \cdots \mathbf{v}_n] = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [A\mathbf{v}_1 \quad A\mathbf{v}_2 \cdots A\mathbf{v}_n] \\ &= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [\lambda_1\mathbf{v}_1 \quad A\mathbf{v}_2 \cdots A\mathbf{v}_n] \\ &= \left[\begin{array}{c|c} \lambda_1 & * \\ \hline \mathbf{0} & A_1 \end{array}\right] = B \end{aligned}$$


Baldwin H. Ward & Kathryn C. Ward/CORBIS

In a lecture he delivered at the University of Göttingen in 1905, the German mathematician David Hilbert (1862–1943) considered linear operators acting on certain infinite-dimensional vector spaces. Out of this lecture arose the notion of a quadratic form in infinitely many variables, and it was in this context that Hilbert first used the term *spectrum* to mean a complete set of eigenvalues. The spaces in question are now called *Hilbert spaces*.

Hilbert made major contributions to many areas of mathematics, among them integral equations, number theory, geometry, and the foundations of mathematics. In 1900, at the Second International Congress of Mathematicians in Paris, Hilbert gave an address entitled "The Problems of Mathematics." In it, he challenged mathematicians to solve 23 problems of fundamental importance during the coming century. Many of the problems have been solved—some were proved true, others false—and some may never be solved. Nevertheless, Hilbert's speech energized the mathematical community and is often regarded as the most influential speech ever given about mathematics.

since $\mathbf{v}_1^T(\lambda_1\mathbf{v}_1) = \lambda_1(\mathbf{v}_1^T\mathbf{v}_1) = \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = \lambda_1$ and $\mathbf{v}_i^T(\lambda_1\mathbf{v}_1) = \lambda_1(\mathbf{v}_i^T\mathbf{v}_1) = \lambda_1(\mathbf{v}_i \cdot \mathbf{v}_1) = 0$ for $i \neq 1$, because $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal set.

But

$$B^T = (Q_1^TAQ_1)^T = Q_1^TA^T(Q_1^T)^T = Q_1^TAQ_1 = B$$

so B is symmetric. Therefore, B has the block form

$$B = \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & A_1 \end{array}\right]$$

and A_1 is symmetric. Furthermore, B is similar to A (why?), so the characteristic polynomial of B is equal to the characteristic polynomial of A, by Theorem 4.22. By Exercise 39 in Section 4.3, the characteristic polynomial of A_1 divides the characteristic polynomial of A. It follows that the eigenvalues of A_1 are also eigenvalues of A and, hence, are real. We also see that A_1 has real entries. (Why?) Thus, A_1 is a $k \times k$ real symmetric matrix with real eigenvalues, so the induction hypothesis applies to it. Hence, there is an orthogonal matrix P_2 such that $P_2^TA_1P_2$ is a diagonal matrix—say, D_1. Now let

$$Q_2 = \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2 \end{array}\right]$$

Then Q_2 is an orthogonal $(k + 1)\times(k + 1)$ matrix, and therefore so is $Q = Q_1Q_2$. Consequently,

$$\begin{aligned} Q^TAQ &= (Q_1Q_2)^TA(Q_1Q_2) = (Q_2^TQ_1^T)A(Q_1Q_2) = Q_2^T(Q_1^TAQ_1)Q_2 = Q_2^TBQ_2 \\ &= \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2^T \end{array}\right]\left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & A_1 \end{array}\right]\left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2 \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & P_2^TA_1P_2 \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & D_1 \end{array}\right] \end{aligned}$$

which is a diagonal matrix. This completes the induction step, and we conclude that, for all $n \geq 1$, an $n \times n$ real symmetric matrix with real eigenvalues is orthogonally diagonalizable.

Example 5.18

Orthogonally diagonalize the matrix

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

Solution This is the matrix from Example 5.17. We have already found that the eigenspaces of A are

$$E_4 = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \quad \text{and} \quad E_1 = \text{span}\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right)$$

We need three orthonormal eigenvectors. First, we apply the Gram-Schmidt Process to

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

to obtain

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}$$

The new vector, which has been constructed to be orthogonal to $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, is still in E_1 (why?) and so is orthogonal to $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Thus, we have three mutually orthogonal vectors, and all we need to do is normalize them and construct a matrix Q with these vectors as its columns. We find that

$$Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix}$$

and it is straightforward to verify that

$$Q^TAQ = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The Spectral Theorem allows us to write a real symmetric matrix A in the form $A = QDQ^T$, where Q is orthogonal and D is diagonal. The diagonal entries of D are just the eigenvalues of A, and if the columns of Q are the orthonormal vectors $\mathbf{q}_1, \ldots, \mathbf{q}_n$, then, using the column-row representation of the product, we have

$$\begin{aligned} A = QDQ^T &= [\mathbf{q}_1 \ \cdots \ \mathbf{q}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \\ &= [\lambda_1\mathbf{q}_1 \ \cdots \ \lambda_n\mathbf{q}_n] \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \cdots + \lambda_n\mathbf{q}_n\mathbf{q}_n^T \end{aligned}$$

This is called the ***spectral decomposition*** of A. Each of the terms $\lambda_i\mathbf{q}_i\mathbf{q}_i^T$ is a rank 1 matrix, by Exercise 62 in Section 3.5, and $\mathbf{q}_i\mathbf{q}_i^T$ is actually the matrix of the projection onto the subspace spanned by $\mathbf{q}_i$. (See Exercise 25.) For this reason, the spectral decomposition

$$A = \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \cdots + \lambda_n\mathbf{q}_n\mathbf{q}_n^T$$

is sometimes referred to as the ***projection form of the Spectral Theorem.***

Example 5.19

Find the spectral decomposition of the matrix A from Example 5.18.

Solution From Example 5.18, we have:

$$\lambda_1 = 4, \qquad \lambda_2 = 1, \qquad \lambda_3 = 1$$

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

Therefore,

$$\mathbf{q}_1\mathbf{q}_1^T = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} [1/\sqrt{3} \quad 1/\sqrt{3} \quad 1/\sqrt{3}] = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

$$\mathbf{q}_2\mathbf{q}_2^T = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} [-1/\sqrt{2} \quad 0 \quad 1/\sqrt{2}] = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

$$\mathbf{q}_3\mathbf{q}_3^T = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} [-1/\sqrt{6} \quad 2/\sqrt{6} \quad -1/\sqrt{6}] = \begin{bmatrix} 1/6 & -1/3 & 1/6 \\ -1/3 & 2/3 & -1/3 \\ 1/6 & -1/3 & 1/6 \end{bmatrix}$$

so

$$\begin{aligned} A &= \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \lambda_3\mathbf{q}_3\mathbf{q}_3^T \\ &= 4\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix} \end{aligned}$$

which can be easily verified.

In this example, $\lambda_2 = \lambda_3$, so we could combine the last two terms $\lambda_2\mathbf{q}_2\mathbf{q}_2^T + \lambda_3\mathbf{q}_3\mathbf{q}_3^T$ to get

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

The rank 2 matrix $\mathbf{q}_2\mathbf{q}_2^T + \mathbf{q}_3\mathbf{q}_3^T$ is the matrix of a projection onto the two-dimensional subspace (i.e., the plane) spanned by $\mathbf{q}_2$ and $\mathbf{q}_3$. (See Exercise 26.)

Observe that the spectral decomposition expresses a symmetric matrix A explicitly in terms of its eigenvalues and eigenvectors. This gives us a way of constructing a matrix with given eigenvalues and (orthonormal) eigenvectors.

Example 5.20

Find a 2×2 matrix with eigenvalues $\lambda_1 = 3$ and $\lambda_2 = -2$ and corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

Solution We begin by normalizing the vectors to obtain an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2\}$, with

$$\mathbf{q}_1 = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \quad \text{and} \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

Now, we compute the matrix A whose spectral decomposition is

$$\begin{aligned} A &= \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \lambda_2 \mathbf{q}_2 \mathbf{q}_2^T \\ &= 3\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \end{bmatrix} - 2\begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}\begin{bmatrix} -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \\ &= 3\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix} - 2\begin{bmatrix} \frac{16}{25} & -\frac{12}{25} \\ -\frac{12}{25} & \frac{9}{25} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{5} & \frac{12}{5} \\ \frac{12}{5} & \frac{6}{5} \end{bmatrix} \end{aligned}$$

It is easy to check that A has the desired properties. (Do this.)

Exercises 5.4

Orthogonally diagonalize the matrices in Exercises 1–10 by finding an orthogonal matrix Q and a diagonal matrix D such that $Q^TAQ = D$.

1. $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

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

3. $A = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 9 & -2 \\ -2 & 6 \end{bmatrix}$

5. $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 2 & 3 & 0 \\ 3 & 2 & 4 \\ 0 & 4 & 2 \end{bmatrix}$

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

8. $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$

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

10. $A = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$

11. If $b \neq 0$, orthogonally diagonalize $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$.

12. If $b \neq 0$, orthogonally diagonalize $A = \begin{bmatrix} a & 0 & b \\ 0 & a & 0 \\ b & 0 & a \end{bmatrix}$.

13. Let A and B be orthogonally diagonalizable $n \times n$ matrices and let c be a scalar. Use the Spectral Theorem to prove that the following matrices are orthogonally diagonalizable:

(a) $A + B$ **(b)** cA **(c)** A^2

14. If A is an invertible matrix that is orthogonally diagonalizable, show that A^{-1} is orthogonally diagonalizable.

15. If A and B are orthogonally diagonalizable and $AB = BA$, show that AB is orthogonally diagonalizable.

16. If A is a symmetric matrix, show that every eigenvalue of A is nonnegative if and only if $A = B^2$ for some symmetric matrix B.

In Exercises 17–20, find a spectral decomposition of the matrix in the given exercise.

17. Exercise 1

18. Exercise 2

19. Exercise 5

20. Exercise 8

In Exercises 21 and 22, find a symmetric 2×2 matrix with eigenvalues λ_1 and λ_2 and corresponding orthogonal eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

21. $\lambda_1 = -1, \lambda_2 = 2, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

22. $\lambda_1 = 3, \lambda_2 = -3, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

In Exercises 23 and 24, find a symmetric 3×3 matrix with eigenvalues λ_1, λ_2, and λ_3 and corresponding orthogonal eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

23. $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},$

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$

24. $\lambda_1 = 1, \lambda_2 = -4, \lambda_3 = -4, \mathbf{v}_1 = \begin{bmatrix} 4 \\ 5 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix},$

$\mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$

25. Let $\mathbf{q}$ be a unit vector in $\mathbb{R}^n$ and let W be the subspace spanned by $\mathbf{q}$. Show that the orthogonal projection of a vector $\mathbf{v}$ onto W (as defined in Sections 1.2 and 5.2) is given by

$$\text{proj}_W(\mathbf{v}) = (\mathbf{q}\mathbf{q}^T)\mathbf{v}$$

and that the matrix of this projection is thus $\mathbf{q}\mathbf{q}^T$. [*Hint:* Remember that, for $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$.]

26. Let $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ be an orthonormal set of vectors in $\mathbb{R}^n$ and let W be the subspace spanned by this set.

(a) Show that the matrix of the orthogonal projection onto W is given by

$$P = \mathbf{q}_1\mathbf{q}_1^T + \cdots + \mathbf{q}_k\mathbf{q}_k^T$$

(b) Show that the projection matrix P in part (a) is symmetric and satisfies $P^2 = P$.

(c) Let $Q = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_k]$ be the $n \times k$ matrix whose columns are the orthonormal basis vectors of W. Show that $P = QQ^T$ and deduce that $\text{rank}(P) = k$.

27. Let A be an $n \times n$ real matrix, all of whose eigenvalues are real. Prove that there exist an orthogonal matrix Q and an upper triangular matrix T such that $Q^TAQ = T$. This very useful result is known as ***Schur's Triangularization Theorem.*** [*Hint:* Adapt the proof of the Spectral Theorem.]

28. Let A be a nilpotent matrix (see Exercise 56 in Section 4.2). Prove that there is an orthogonal matrix Q such that $Q^T AQ$ is upper triangular with zeros on its diagonal. [*Hint:* Use Exercise 27.]

5.5 Applications

Quadratic Forms

An expression of the form

$$ax^2 + by^2 + cxy$$

is called a ***quadratic form*** in x and y. Similarly,

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

is a quadratic form in x, y, and z. In words, a quadratic form is a sum of terms, each of which has total degree *two* in the variables. Therefore, $5x^2 - 3y^2 + 2xy$ is a quadratic form, but $x^2 + y^2 + x$ is not.

We can represent quadratic forms using matrices as follows:

$$ax^2 + by^2 + cxy = [x \quad y]\begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

and

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = [x \quad y \quad z]\begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(Verify these.) Each has the form $\mathbf{x}^T A\mathbf{x}$, where the matrix A is symmetric. This observation leads us to the following general definition.

Definition A ***quadratic form*** in n variables is a function $f: \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

where A is a symmetric $n \times n$ matrix and $\mathbf{x}$ is in $\mathbb{R}^n$. We refer to A as the ***matrix associated with f***.

Example 5.21

What is the quadratic form with associated matrix $A = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$?

Solution If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, then

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 5x_2^2 - 6x_1x_2$$

Observe that the *off-diagonal* entries $a_{12} = a_{21} = -3$ of A are *combined* to give the coefficient -6 of x_1x_2. This is true generally. We can expand a quadratic form in n variables $\mathbf{x}^T A\mathbf{x}$ as follows:

$$\mathbf{x}^T A\mathbf{x} = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + \sum_{i<j} 2a_{ij}x_ix_j$$

Thus, if $i \neq j$, the coefficient of x_ix_j is $2a_{ij}$.

Example 5.22

Find the matrix associated with the quadratic form

$$f(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 5x_3^2 + 6x_1x_2 - 3x_1x_3$$

Solution The coefficients of the squared terms x_i^2 go on the diagonal as a_{ii}, and the coefficients of the cross-product terms x_ix_j are split between a_{ij} and a_{ji}. This gives

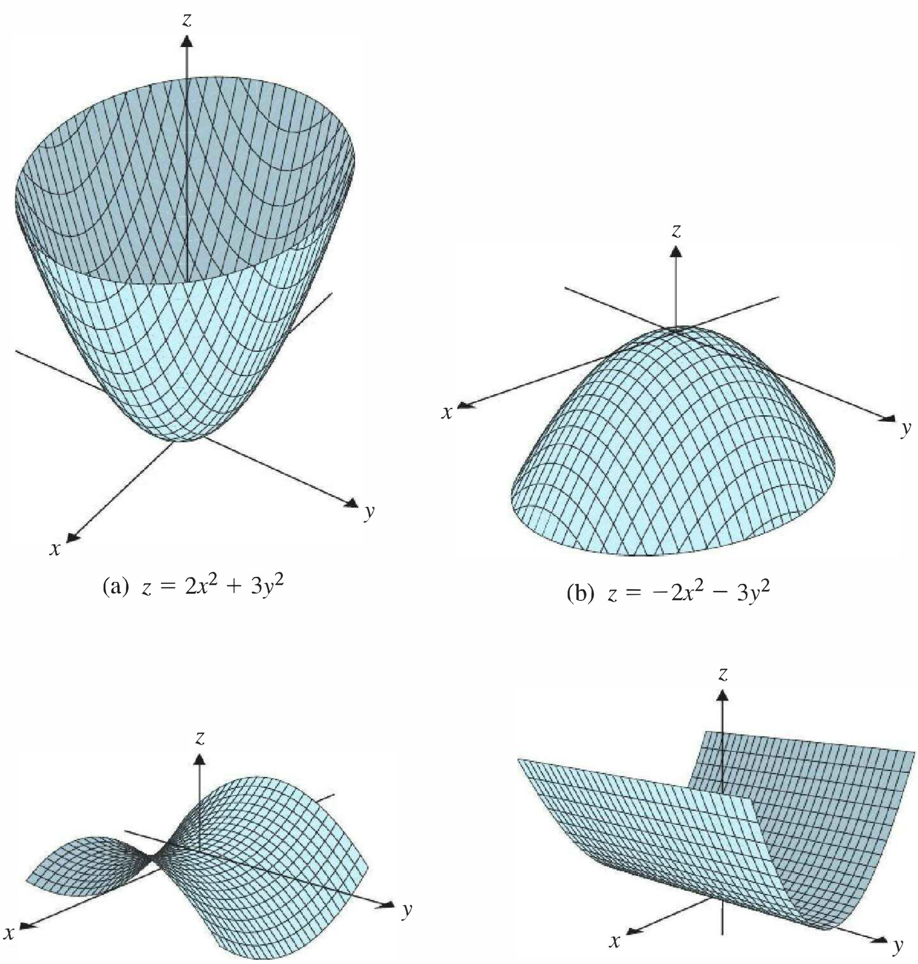
$$A = \begin{bmatrix} 2 & 3 & -\frac{3}{2} \\ 3 & -1 & 0 \\ -\frac{3}{2} & 0 & 5 \end{bmatrix}$$

so
$$f(x_1, x_2, x_3) = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 2 & 3 & -\frac{3}{2} \\ 3 & -1 & 0 \\ -\frac{3}{2} & 0 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

as you can easily check.

In the case of a quadratic form $f(x, y)$ in two variables, the graph of $z = f(x, y)$ is a surface in $\mathbb{R}^3$. Some examples are shown in Figure 5.12.

Observe that the effect of holding x or y constant is to take a cross section of the graph parallel to the yz or xz planes, respectively. For the graphs in Figure 5.12, all of these cross sections are easy to identify. For example, in Figure 5.12(a), the cross sections we get by holding x or y constant are all parabolas opening upward, so $f(x, y) \geq 0$ for all values of x and y. In Figure 5.12(c), holding x constant gives parabolas opening downward and holding y constant gives parabolas opening upward, producing a *saddle point*.



(a) $z = 2x^2 + 3y^2$

(b) $z = -2x^2 - 3y^2$

(c) $z = 2x^2 - 3y^2$

(d) $z = 2x^2$

Figure 5.12

Graphs of quadratic forms $f(x, y)$

What makes this type of analysis quite easy is the fact that these quadratic forms have no cross-product terms. The matrix associated with such a quadratic form is a diagonal matrix. For example,

$$2x^2 - 3y^2 = [x \quad y]\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

In general, the matrix of a quadratic form is a symmetric matrix, and we saw in Section 5.4 that such matrices can always be diagonalized. We will now use this fact to show that, for *every* quadratic form, we can eliminate the cross-product terms by means of a suitable change of variable.

Let $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in n variables, with A a symmetric $n \times n$ matrix. By the Spectral Theorem, there is an orthogonal matrix Q that diagonalizes A; that is, $Q^TAQ = D$, where D is a diagonal matrix displaying the eigenvalues of A. We now set

$$\mathbf{x} = Q\mathbf{y} \quad \text{or, equivalently,} \quad \mathbf{y} = Q^{-1}\mathbf{x} = Q^T\mathbf{x}$$

Substitution into the quadratic form yields

$$\begin{aligned} \mathbf{x}^TA\mathbf{x} &= (Q\mathbf{y})^TA(Q\mathbf{y}) \\ &= \mathbf{y}^TQ^TAQ\mathbf{y} \\ &= \mathbf{y}^TD\mathbf{y} \end{aligned}$$

which is a quadratic form without cross-product terms, since D is diagonal. Furthermore, if the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$, then Q can be chosen so that

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

If $\mathbf{y} = [y_1 \quad \cdots \quad y_n]^T$, then, with respect to these new variables, the quadratic form becomes

$$\mathbf{y}^TD\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

This process is called ***diagonalizing a quadratic form.*** We have just proved the following theorem, known as the ***Principal Axes Theorem.*** (The reason for this name will become clear in the next subsection.)

Theorem 5.21 **The Principal Axes Theorem**

Every quadratic form can be diagonalized. Specifically, if A is the $n \times n$ symmetric matrix associated with the quadratic form $\mathbf{x}^TA\mathbf{x}$ and if Q is an orthogonal matrix such that $Q^TAQ = D$ is a diagonal matrix, then the change of variable $\mathbf{x} = Q\mathbf{y}$ transforms the quadratic form $\mathbf{x}^TA\mathbf{x}$ into the quadratic form $\mathbf{y}^TD\mathbf{y}$, which has no cross-product terms. If the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$ and $\mathbf{y} = [y_1 \quad \cdots \quad y_n]^T$, then

$$\mathbf{x}^TA\mathbf{x} = \mathbf{y}^TD\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

Example 5.23

Find a change of variable that transforms the quadratic form

$$f(x_1, x_2) = 5x_1^2 + 4x_1x_2 + 2x_2^2$$

into one with no cross-product terms.

Solution The matrix of f is

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

with eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 1$. Corresponding unit eigenvectors are

$$\mathbf{q}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \quad \text{and} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

(Check this.) If we set

$$Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

then $Q^TAQ = D$. The change of variable $\mathbf{x} = Q\mathbf{y}$, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

converts f into

$$f(\mathbf{y}) = f(y_1, y_2) = [y_1 \quad y_2]\begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 6y_1^2 + y_2^2$$

The original quadratic form $\mathbf{x}^TA\mathbf{x}$ and the new one $\mathbf{y}^TD\mathbf{y}$ (referred to in the Principal Axes Theorem) are *equal* in the following sense. In Example 5.23, suppose we want to evaluate $f(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$ at $\mathbf{x} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$. We have

$$f(-1, 3) = 5(-1)^2 + 4(-1)(3) + 2(3)^2 = 11$$

In terms of the new variables,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -7/\sqrt{5} \end{bmatrix}$$

so

$$f(y_1, y_2) = 6y_1^2 + y_2^2 = 6(1/\sqrt{5})^2 + (-7/\sqrt{5})^2 = 55/5 = 11$$

exactly as before.

The Principal Axes Theorem has some interesting and important consequences. We will consider two of these. The first relates to the possible *values* that a quadratic form can take on.

Definition A quadratic form $f(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$ is classified as one of the following:

1. ***positive definite*** if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$
2. ***positive semidefinite*** if $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$
3. ***negative definite*** if $f(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$
4. ***negative semidefinite*** if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x}$
5. ***indefinite*** if $f(\mathbf{x})$ takes on both positive and negative values

A symmetric matrix A is called ***positive definite, positive semidefinite, negative definite, negative semidefinite,*** or ***indefinite*** if the associated quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ has the corresponding property.

The quadratic forms in parts (a), (b), (c), and (d) of Figure 5.12 are positive definite, negative definite, indefinite, and positive semidefinite, respectively. The Principal Axes Theorem makes it easy to tell if a quadratic form has one of these properties.

Theorem 5.22 Let A be an $n \times n$ symmetric matrix. The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is

a. positive definite if and only if all of the eigenvalues of A are positive.
b. positive semidefinite if and only if all of the eigenvalues of A are nonnegative.
c. negative definite if and only if all of the eigenvalues of A are negative.
d. negative semidefinite if and only if all of the eigenvalues of A are nonpositive.
e. indefinite if and only if A has both positive and negative eigenvalues.

You are asked to prove Theorem 5.22 in Exercise 27.

Example 5.24 Classify $f(x, y, z) = 3x^2 + 3y^2 + 3z^2 - 2xy - 2xz - 2yz$ as positive definite, negative definite, indefinite, or none of these.

Solution The matrix associated with f is

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

which has eigenvalues 1, 4, and 4. (Verify this.) Since all of these eigenvalues are positive, f is a positive definite quadratic form.

If a quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is positive definite, then, since $f(\mathbf{0}) = 0$, the *minimum* value of $f(\mathbf{x})$ is 0 and it occurs at the origin. Similarly, a negative definite quadratic form has a maximum at the origin. Thus, Theorem 5.22 allows us to solve certain types of maxima/minima problems easily, without resorting to calculus. A type of problem that falls into this category is the ***constrained optimization problem.***

It is often important to know the maximum or minimum values of a quadratic form subject to certain constraints. (Such problems arise not only in mathematics but also in statistics, physics, engineering, and economics.) We will be interested in finding the extreme values of $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ subject to the constraint that $\|\mathbf{x}\| = 1$. In the case of a quadratic form in two variables, we can visualize what the problem means. The graph of $z = f(x, y)$ is a surface in $\mathbb{R}^3$, and the constraint $\|\mathbf{x}\| = 1$ restricts the point (x, y) to the unit circle in the xy-plane. Thus, we are considering those points that lie simultaneously on the surface and on the unit cylinder perpendicular to the xy plane. These points form a curve lying on the surface, and we want the highest and lowest points on this curve. Figure 5.13 shows this situation for the quadratic form and corresponding surface in Figure 5.12(c).

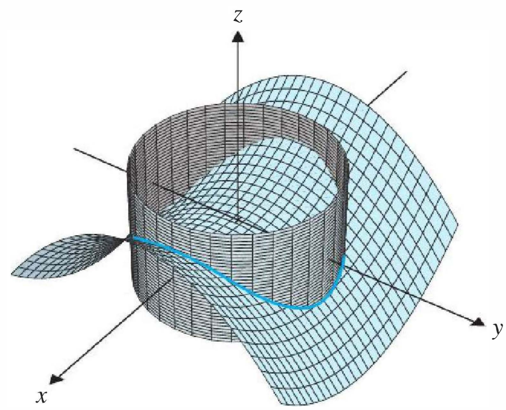


Figure 5.13
The intersection of $z = 2x^2 - 3y^2$ with the cylinder $x^2 + y^2 = 1$

In this case, the maximum and minimum values of $f(x, y) = 2x^2 - 3y^2$ (the highest and lowest points on the curve of intersection) are 2 and -3, respectively, which are just the eigenvalues of the associated matrix. Theorem 5.23 shows that this is always the case.

Theorem 5.23

Let $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be a quadratic form with associated $n \times n$ symmetric matrix A. Let the eigenvalues of A be $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$. Then the following are true, subject to the constraint $\|\mathbf{x}\| = 1$:

a. $\lambda_1 \ge f(\mathbf{x}) \ge \lambda_n$
b. The maximum value of $f(\mathbf{x})$ is λ_1, and it occurs when $\mathbf{x}$ is a unit eigenvector corresponding to λ_1.
c. The minimum value of $f(\mathbf{x})$ is λ_n, and it occurs when $\mathbf{x}$ is a unit eigenvector corresponding to λ_n.

Proof As usual, we begin by orthogonally diagonalizing A. Accordingly, let Q be an orthogonal matrix such that $Q^T AQ$ is the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Then, by the Principal Axes Theorem, the change of variable $\mathbf{x} = Q\mathbf{y}$ gives $\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y}$. Now note that $\mathbf{y} = Q^T\mathbf{x}$ implies that

$$\mathbf{y}^T\mathbf{y} = (Q^T\mathbf{x})^T(Q^T\mathbf{x}) = \mathbf{x}^T(Q^T)^T Q^T\mathbf{x} = \mathbf{x}^T QQ^T\mathbf{x} = \mathbf{x}^T\mathbf{x}$$

since $Q^T = Q^{-1}$. Hence, using $\mathbf{x} \cdot \mathbf{x} = \mathbf{x}^T\mathbf{x}$, we see that $\|\mathbf{y}\| = \sqrt{\mathbf{y}^T\mathbf{y}} = \sqrt{\mathbf{x}^T\mathbf{x}} = \|\mathbf{x}\| = 1$. Thus, if $\mathbf{x}$ is a unit vector, so is the corresponding $\mathbf{y}$, and the values of $\mathbf{x}^T A\mathbf{x}$ and $\mathbf{y}^T D\mathbf{y}$ are the same.

(a) To prove property (a), we observe that if $\mathbf{y} = [y_1 \quad \cdots \quad y_n]^T$, then

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\,\mathbf{y} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \\ &\le \lambda_1 y_1^2 + \lambda_1 y_2^2 + \cdots + \lambda_1 y_n^2 \\ &= \lambda_1 (y_1^2 + y_2^2 + \cdots + y_n^2) \\ &= \lambda_1 \|\mathbf{y}\|^2 \\ &= \lambda_1 \end{aligned}$$

Thus, $f(\mathbf{x}) \le \lambda_1$ for all $\mathbf{x}$ such that $\|\mathbf{x}\| = 1$. The proof that $f(\mathbf{x}) \ge \lambda_n$ is similar. (See Exercise 37.)

(b) If $\mathbf{q}_1$ is a unit eigenvector corresponding to λ_1, then $A\mathbf{q}_1 = \lambda_1\mathbf{q}_1$ and

$$f(\mathbf{q}_1) = \mathbf{q}_1^T A\mathbf{q}_1 = \mathbf{q}_1^T \lambda_1 \mathbf{q}_1 = \lambda_1(\mathbf{q}_1^T\mathbf{q}_1) = \lambda_1$$

This shows that the quadratic form actually takes on the value λ_1, and so, by property (a), it is the maximum value of $f(\mathbf{x})$ and it occurs when $\mathbf{x} = \mathbf{q}_1$.

(c) You are asked to prove this property in Exercise 38.

Example 5.25

Find the maximum and minimum values of the quadratic form $f(x_1, x_2) = 5x_1^2 + 4x_1x_2 + 2x_2^2$ subject to the constraint $x_1^2 + x_2^2 = 1$, and determine values of x_1 and x_2 for which each of these occurs.

Solution In Example 5.23, we found that f has the associated eigenvalues $\lambda_1 = 6$ and $\lambda_2 = 1$, with corresponding unit eigenvectors

$$\mathbf{q}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \quad \text{and} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

Therefore, the maximum value of f is 6 when $x_1 = 2/\sqrt{5}$ and $x_2 = 1/\sqrt{5}$. The minimum value of f is 1 when $x_1 = 1/\sqrt{5}$ and $x_2 = -2/\sqrt{5}$. (Observe that these extreme values occur twice—in opposite directions—since $-\mathbf{q}_1$ and $-\mathbf{q}_2$ are also unit eigenvectors for λ_1 and λ_2, respectively.)

Graphing Quadratic Equations

The general form of a quadratic equation in two variables x and y is

$$ax^2 + by^2 + cxy + dx + ey + f = 0$$

where at least one of a, b, and c is nonzero. The graphs of such quadratic equations are called ***conic sections*** (or ***conics***), since they can be obtained by taking cross sections of a (double) cone (i.e., slicing it with a plane). The most important of the conic sections are the ellipses (with circles as a special case), hyperbolas, and parabolas. These are called the ***nondegenerate*** conics. Figure 5.14 shows how they arise.

It is also possible for a cross section of a cone to result in a single point, a straight line, or a pair of lines. These are called ***degenerate*** conics. (See Exercises 59–64.)

The graph of a nondegenerate conic is said to be in ***standard position*** relative to the coordinate axes if its equation can be expressed in one of the forms in Figure 5.15.

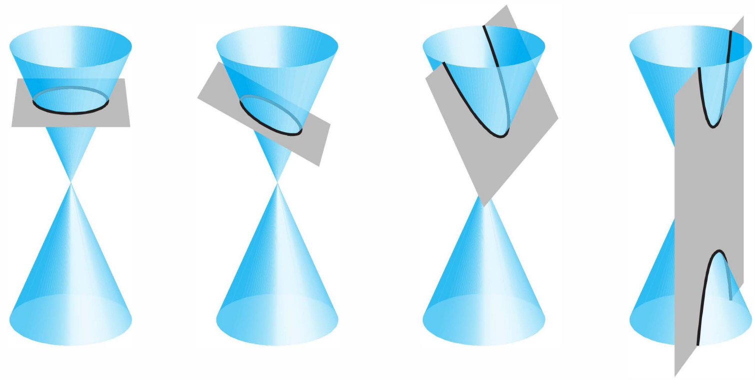


Figure 5.14
The nondegenerate conics

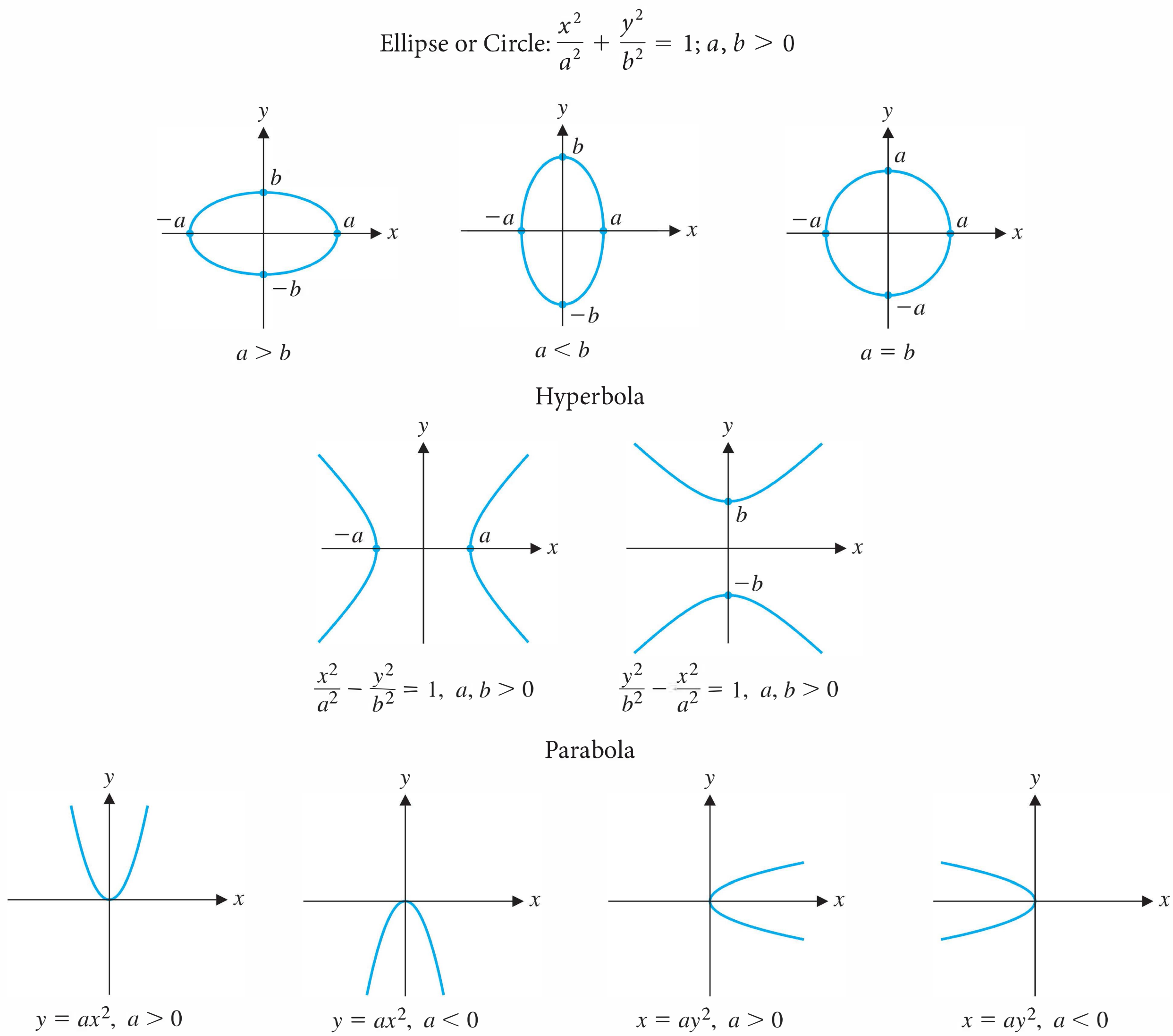


Figure 5.15
Nondegenerate conics in standard position

Example 5.26

If possible, write each of the following quadratic equations in the form of a conic in standard position and identify the resulting graph.

(a) $4x^2 + 9y^2 = 36$ (b) $4x^2 - 9y^2 + 1 = 0$ (c) $4x^2 - 9y = 0$

Solution (a) The equation $4x^2 + 9y^2 = 36$ can be written in the form

$$\frac{x^2}{9} + \frac{y^2}{4} = 1$$

so its graph is an ellipse intersecting the x-axis at $(\pm 3, 0)$ and the y-axis at $(0, \pm 2)$.

(b) The equation $4x^2 - 9y^2 + 1 = 0$ can be written in the form

$$\frac{y^2}{\frac{1}{9}} - \frac{x^2}{\frac{1}{4}} = 1$$

so its graph is a hyperbola, opening up and down, intersecting the y-axis at $(0, \pm\frac{1}{3})$.

(c) The equation $4x^2 - 9y = 0$ can be written in the form

$$y = \frac{4}{9}x^2$$

so its graph is a parabola opening upward.

If a quadratic equation contains too many terms to be written in one of the forms in Figure 5.15, then its graph is not in standard position. When there are additional terms but no xy term, the graph of the conic has been *translated* out of standard position.

Example 5.27

Identify and graph the conic whose equation is

$$x^2 + 2y^2 - 6x + 8y + 9 = 0$$

Solution We begin by grouping the x and y terms separately to get

$$(x^2 - 6x) + (2y^2 + 8y) = -9$$

or

$$(x^2 - 6x) + 2(y^2 + 4y) = -9$$

Next, we complete the squares on the two expressions in parentheses to obtain

$$(x^2 - 6x + 9) + 2(y^2 + 4y + 4) = -9 + 9 + 8$$

or

$$(x - 3)^2 + 2(y + 2)^2 = 8$$

We now make the substitutions $x' = x - 3$ and $y' = y + 2$, turning the above equation into

$$(x')^2 + 2(y')^2 = 8 \quad \text{or} \quad \frac{(x')^2}{8} + \frac{(y')^2}{4} = 1$$

This is the equation of an ellipse in standard position in the $x'y'$ coordinate system, intersecting the x'-axis at $(\pm 2\sqrt{2}, 0)$ and the y'-axis at $(0, \pm 2)$. The origin in the $x'y'$ coordinate system is at $x = 3$, $y = -2$, so the ellipse has been translated out of standard position 3 units to the right and 2 units down. Its graph is shown in Figure 5.16.

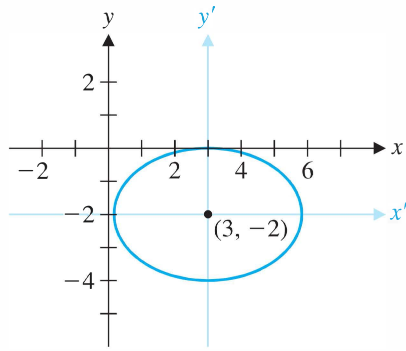


Figure 5.16
A translated ellipse

If a quadratic equation contains a cross-product term, then it represents a conic that has been *rotated*.

Example 5.28

Identify and graph the conic whose equation is

$$5x^2 + 4xy + 2y^2 = 6$$

Solution The left-hand side of the equation is a quadratic form, so we can write it in matrix form as $\mathbf{x}^T A\mathbf{x} = 6$, where

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

In Example 5.23, we found that the eigenvalues of A are 6 and 1, and a matrix Q that orthogonally diagonalizes A is

$$Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}$$

Observe that det $Q = -1$. In this example, we will interchange the columns of this matrix to make the determinant equal to $+1$. Then Q will be the matrix of a *rotation*, by Exercise 28 in Section 5.1. It is always possible to rearrange the columns of an orthogonal matrix Q to make its determinant equal to $+1$. (Why?) We set

$$Q = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

instead, so that

$$Q^T A Q = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = D$$

The change of variable $\mathbf{x} = Q\mathbf{x}'$ converts the given equation into the form $(\mathbf{x}')^T D\mathbf{x}' = 6$ by means of a rotation. If $\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, then this equation is just

$$(x')^2 + 6(y')^2 = 6 \quad \text{or} \quad \frac{(x')^2}{6} + (y')^2 = 1$$

which represents an ellipse in the $x'y'$ coordinate system.

To graph this ellipse, we need to know which vectors play the roles of $\mathbf{e}_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in the new coordinate system. (These two vectors locate the positions of the x' and y' axes.) But, from $\mathbf{x} = Q\mathbf{x}'$, we have

$$Q\mathbf{e}_1' = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

and

$$Q\mathbf{e}_2' = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

These are just the columns $\mathbf{q}_1$ and $\mathbf{q}_2$ of Q, which are the eigenvectors of A! The fact that these are orthonormal vectors agrees perfectly with the fact that the change of variable is just a rotation. The graph is shown in Figure 5.17.

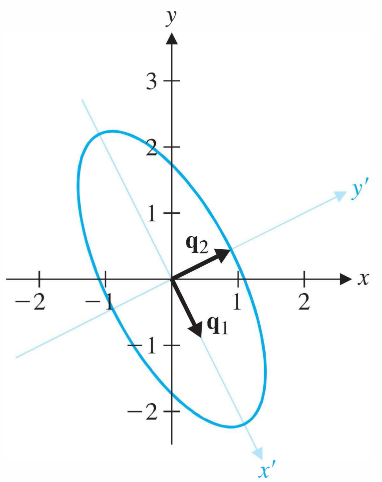


Figure 5.17
A rotated ellipse

You can now see why the Principal Axes Theorem is so named. If a real symmetric matrix A arises as the coefficient matrix of a quadratic equation, the eigenvectors of A give the directions of the principal axes of the corresponding graph.

It is possible for the graph of a conic to be both rotated and translated out of standard position, as illustrated in Example 5.29.

Example 5.29

Identify and graph the conic whose equation is

$$5x^2 + 4xy + 2y^2 - \frac{28}{\sqrt{5}}x - \frac{4}{\sqrt{5}}y + 4 = 0$$

Solution The strategy is to eliminate the cross-product term first. In matrix form, the equation is $\mathbf{x}^T A\mathbf{x} + B\mathbf{x} + 4 = 0$, where

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix}$$

The cross-product term comes from the quadratic form $\mathbf{x}^T A\mathbf{x}$, which we diagonalize as in Example 5.28 by setting $\mathbf{x} = Q\mathbf{x}'$, where

$$Q = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Then, as in Example 5.28,

$$\mathbf{x}^T A\mathbf{x} = (\mathbf{x}')^T D\mathbf{x}' = (x')^2 + 6(y')^2$$

But now we also have

$$B\mathbf{x} = BQ\mathbf{x}' = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = -4x' - 12y'$$

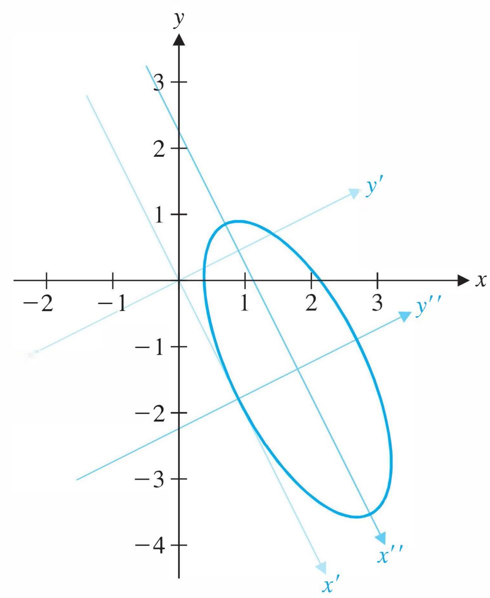


Figure 5.18

Thus, in terms of x' and y', the given equation becomes

$$(x')^2 + 6(y')^2 - 4x' - 12y' + 4 = 0$$

To bring the conic represented by this equation into standard position, we need to *translate* the $x'y'$ axes. We do so by completing the squares, as in Example 5.27. We have

$$((x')^2 - 4x' + 4) + 6((y')^2 - 2y' + 1) = -4 + 4 + 6 = 6$$

or
$$(x' - 2)^2 + 6(y' - 1)^2 = 6$$

This gives us the translation equations

$$x'' = x' - 2 \quad \text{and} \quad y'' = y' - 1$$

In the $x''y''$ coordinate system, the equation is simply

$$(x'')^2 + 6(y'')^2 = 6$$

which is the equation of an ellipse (as in Example 5.28). We can sketch this ellipse by first rotating and then translating. The resulting graph is shown in Figure 5.18.

The general form of a quadratic equation in three variables x, y, and z is

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$$

where at least one of $a, b, \ldots, f$ is nonzero. The graph of such a quadratic equation is called a ***quadric surface*** (or ***quadric***). Once again, to recognize a quadric we need

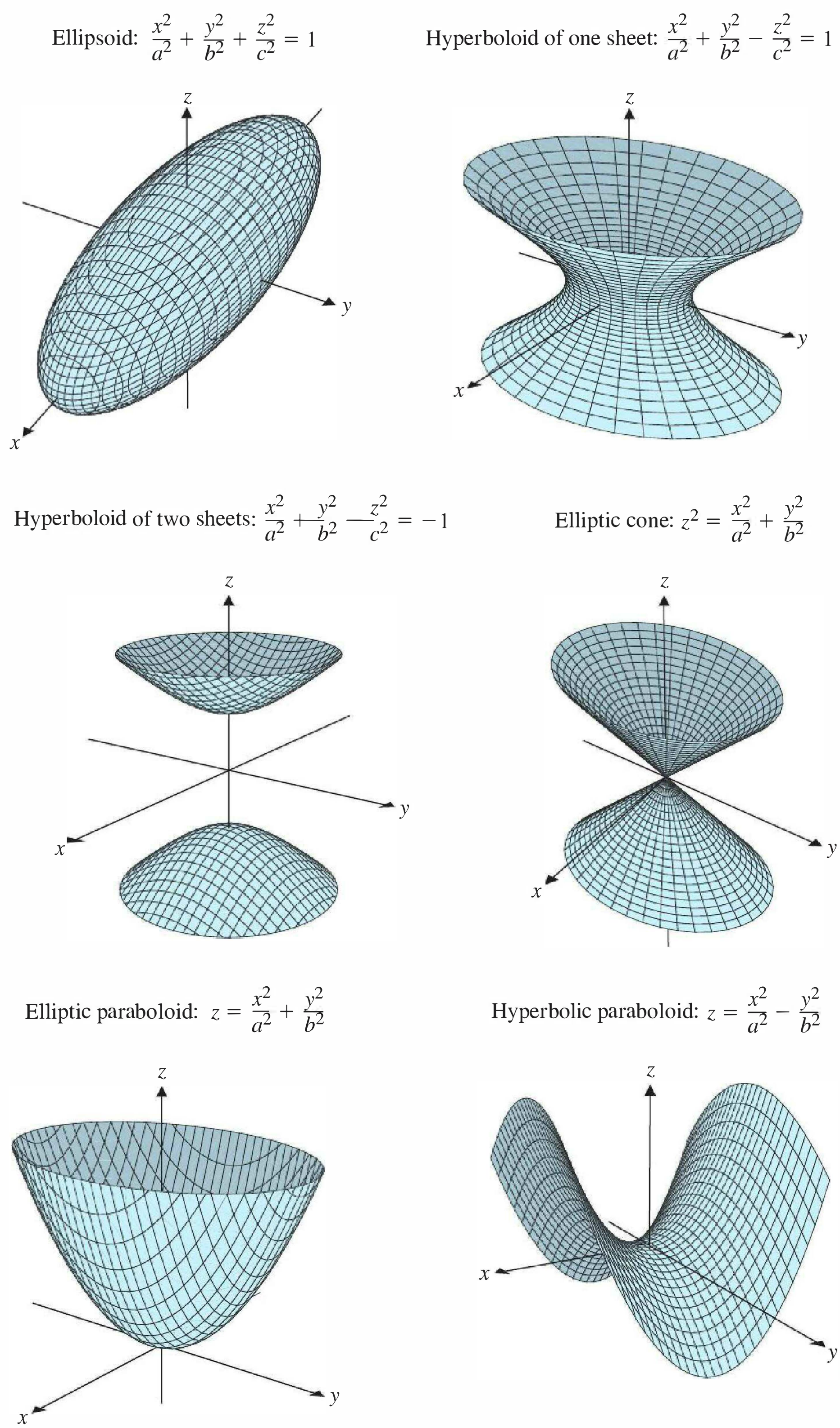


Figure 5.19
Quadric surfaces

to put it into standard position. Some quadrics in standard position are shown in Figure 5.19; others are obtained by permuting the variables.

Example 5.30

Identify the quadric surface whose equation is

$$5x^2 + 11y^2 + 2z^2 + 16xy + 20xz - 4yz = 36$$

Solution The equation can be written in matrix form as $\mathbf{x}^T A\mathbf{x} = 36$, where

$$A = \begin{bmatrix} 5 & 8 & 10 \\ 8 & 11 & -2 \\ 10 & -2 & 2 \end{bmatrix}$$

We find the eigenvalues of A to be 18, 9, and -9, with corresponding orthogonal eigenvectors

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

respectively. We normalize them to obtain

$$\mathbf{q}_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \text{and} \quad \mathbf{q}_3 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

and form the orthogonal matrix

$$Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

Note that in order for Q to be the matrix of a rotation, we require $\det Q = 1$, which is true in this case. (Otherwise, $\det Q = -1$, and swapping two columns changes the sign of the determinant.) Therefore,

$$Q^T AQ = D = \begin{bmatrix} 18 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -9 \end{bmatrix}$$

and, with the change of variable $\mathbf{x} = Q\mathbf{x}'$, we get $\mathbf{x}^T A\mathbf{x} = (\mathbf{x}')D\mathbf{x}' = 36$, so

$$18(x')^2 + 9(y')^2 - 9(z')^2 = 36 \quad \text{or} \quad \frac{(x')^2}{2} + \frac{(y')^2}{4} - \frac{(z')^2}{4} = 1$$

From Figure 5.19, we recognize this equation as the equation of a hyperboloid of one sheet. The x', y', and z' axes are in the directions of the eigenvectors $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$, respectively. The graph is shown in Figure 5.20.

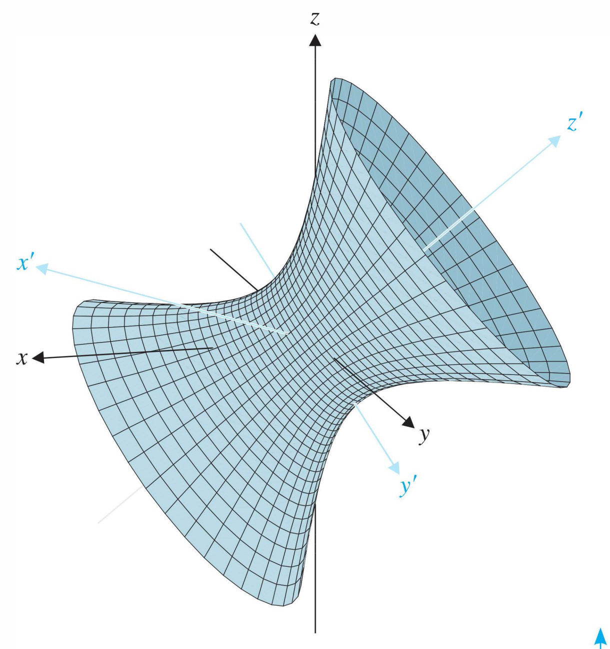


Figure 5.20
A hyperboloid of one sheet in nonstandard position

We can also identify and graph quadrics that have been translated out of standard position using the "complete-the-squares method" of Examples 5.27 and 5.29. You will be asked to do so in the exercises.

Exercises 5.5

Quadratic Forms

In Exercises 1–6, evaluate the quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ for the given A and **x**.

1. $A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$

2. $A = \begin{bmatrix} 5 & 1 \\ 1 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

3. $A = \begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 2 & 1 \\ -3 & 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

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

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

In Exercises 7–12, find the symmetric matrix A associated with the given quadratic form.

7. $x_1^2 + 2x_2^2 + 6x_1x_2$

8. x_1x_2

9. $3x^2 - 3xy - y^2$

10. $x_1^2 - x_3^2 + 8x_1x_2 - 6x_2x_3$

11. $5x_1^2 - x_2^2 + 2x_3^2 + 2x_1x_2 - 4x_1x_3 + 4x_2x_3$

12. $2x^2 - 3y^2 + z^2 - 4xz$

Diagonalize the quadratic forms in Exercises 13–18 by finding an orthogonal matrix Q such that the change of variable $\mathbf{x} = Q\mathbf{y}$ *transforms the given form into one with no cross-product terms. Give Q and the new quadratic form.*

13. $2x_1^2 + 5x_2^2 - 4x_1x_2$

14. $x^2 + 8xy + y^2$

15. $7x_1^2 + x_2^2 + x_3^2 + 8x_1x_2 + 8x_1x_3 - 16x_2x_3$

16. $x_1^2 + x_2^2 + 3x_3^2 - 4x_1x_2$

17. $x^2 + z^2 - 2xy + 2yz$

18. $2xy + 2xz + 2yz$

Classify each of the quadratic forms in Exercises 19–26 as positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite.

19. $x_1^2 + 2x_2^2$ **20.** $x_1^2 + x_2^2 - 2x_1x_2$

21. $-2x^2 - 2y^2 + 2xy$ **22.** $x^2 + y^2 + 4xy$

23. $2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$

24. $x_1^2 + x_2^2 + x_3^2 + 2x_1x_3$ **25.** $x_1^2 + x_2^2 - x_3^2 + 4x_1x_2$

26. $-x^2 - y^2 - z^2 - 2xy - 2xz - 2yz$

27. Prove Theorem 5.22.

28. Let $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ be a symmetric 2×2 matrix. Prove that A is positive definite if and only if $a > 0$ and $\det A > 0$. [*Hint:* $ax^2 + 2bxy + dy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(d - \frac{b^2}{a}\right)y^2$.]

29. Let B be an invertible matrix. Show that $A = B^TB$ is positive definite.

30. Let A be a positive definite symmetric matrix. Show that there exists an invertible matrix B such that $A = B^TB$. [*Hint:* Use the Spectral Theorem to write $A = QDQ^T$. Then show that D can be factored as C^TC for some invertible matrix C.]

31. Let A and B be positive definite symmetric $n \times n$ matrices and let c be a positive scalar. Show that the following matrices are positive definite.

(a) cA **(b)** A^2 **(c)** $A + B$
(d) A^{-1} (First show that A is necessarily invertible.)

32. Let A be a positive definite symmetric matrix. Show that there is a positive definite symmetric matrix B such that $A = B^2$. (Such a matrix B is called a ***square root*** of A.)

In Exercises 33–36, find the maximum and minimum values of the quadratic form $f(\mathbf{x})$ in the given exercise, subject to the constraint $\|\mathbf{x}\| = 1$, and determine the values of $\mathbf{x}$ for which these occur.

33. Exercise 20 **34.** Exercise 22

35. Exercise 23 **36.** Exercise 24

37. Finish the proof of Theorem 5.23(a).

38. Prove Theorem 5.23(c).

Graphing Quadratic Equations

In Exercises 39–44, identify the graph of the given equation.

39. $x^2 + 5y^2 = 25$ **40.** $x^2 - y^2 - 4 = 0$

41. $x^2 - y - 1 = 0$ **42.** $2x^2 + y^2 - 8 = 0$

43. $3x^2 = y^2 - 1$ **44.** $x = -2y^2$

In Exercises 45–50, use a translation of axes to put the conic in standard position. Identify the graph, give its equation in the translated coordinate system, and sketch the curve.

45. $x^2 + y^2 - 4x - 4y + 4 = 0$

46. $4x^2 + 2y^2 - 8x + 12y + 6 = 0$

47. $9x^2 - 4y^2 - 4y = 37$ **48.** $x^2 + 10x - 3y = -13$

49. $2y^2 + 4x + 8y = 0$

50. $2y^2 - 3x^2 - 18x - 20y + 11 = 0$

In Exercises 51–54, use a rotation of axes to put the conic in standard position. Identify the graph, give its equation in the rotated coordinate system, and sketch the curve.

51. $x^2 + xy + y^2 = 6$ **52.** $4x^2 + 10xy + 4y^2 = 9$

53. $4x^2 + 6xy - 4y^2 = 5$ **54.** $3x^2 - 2xy + 3y^2 = 8$

In Exercises 55–58, identify the conic with the given equation and give its equation in standard form.

55. $3x^2 - 4xy + 3y^2 - 28\sqrt{2}x + 22\sqrt{2}y + 84 = 0$

56. $6x^2 - 4xy + 9y^2 - 20x - 10y - 5 = 0$

57. $2xy + 2\sqrt{2}x - 1 = 0$

58. $x^2 - 2xy + y^2 + 4\sqrt{2}x - 4 = 0$

Sometimes the graph of a quadratic equation is a straight line, a pair of straight lines, or a single point. We refer to such a graph as a ***degenerate conic.*** *It is also possible that the equation is not satisfied for any values of the variables, in which case there is no graph at all and we refer to the conic as an* ***imaginary conic.*** *In Exercises 59–64, identify the conic with the given equation as either degenerate or imaginary and, where possible, sketch the graph.*

59. $x^2 - y^2 = 0$ **60.** $x^2 + 2y^2 + 2 = 0$

61. $3x^2 + y^2 = 0$ **62.** $x^2 + 2xy + y^2 = 0$

63. $x^2 - 2xy + y^2 + 2\sqrt{2}x - 2\sqrt{2}y = 0$

64. $2x^2 + 2xy + 2y^2 + 2\sqrt{2}x - 2\sqrt{2}y + 6 = 0$

65. Let A be a symmetric 2×2 matrix and let k be a scalar. Prove that the graph of the quadratic equation $\mathbf{x}^TA\mathbf{x} = k$ is

(a) a hyperbola if $k \neq 0$ and $\det A < 0$
(b) an ellipse, circle, or imaginary conic if $k \neq 0$ and $\det A > 0$
(c) a pair of straight lines or an imaginary conic if $k \neq 0$ and $\det A = 0$
(d) a pair of straight lines or a single point if $k = 0$ and $\det A \neq 0$
(e) a straight line if $k = 0$ and $\det A = 0$
[*Hint:* Use the Principal Axes Theorem.]

In Exercises 66–73, identify the quadric with the given equation and give its equation in standard form.

66. $4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz = 8$

67. $x^2 + y^2 + z^2 - 4yz = 1$

68. $-x^2 - y^2 - z^2 + 4xy + 4xz + 4yz = 12$

69. $2xy + z = 0$

70. $16x^2 + 100y^2 + 9z^2 - 24xz - 60x - 80z = 0$

71. $x^2 + y^2 - 2z^2 + 4xy - 2xz + 2yz - x + y + z = 0$

72. $10x^2 + 25y^2 + 10z^2 - 40xz + 20\sqrt{2}x + 50y + 20\sqrt{2}z = 15$

73. $11x^2 + 11y^2 + 14z^2 + 2xy + 8xz - 8yz - 12x + 12y + 12z = 6$

74. Let A be a real 2×2 matrix with complex eigenvalues $\lambda = a \pm bi$ such that $b \neq 0$ and $|\lambda| = 1$. Prove that every trajectory of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ lies on an ellipse. [*Hint:* Theorem 4.43 shows that if $\mathbf{v}$ is an eigenvector corresponding to $\lambda = a - bi$, then the matrix $P = [\text{Re}\,\mathbf{v} \quad \text{Im}\,\mathbf{v}]$ is invertible and $A = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}P^{-1}$. Set $B = (PP^T)^{-1}$. Show that the quadratic $\mathbf{x}^TB\mathbf{x} = k$ defines an ellipse for all $k > 0$, and prove that if $\mathbf{x}$ lies on this ellipse, so does $A\mathbf{x}$.]

Chapter Review

Key Definitions and Concepts

fundamental subspaces of a matrix, 380
Gram-Schmidt Process, 389
orthogonal basis, 370
orthogonal complement of a subspace, 378
orthogonal matrix, 374
orthogonal projection, 382
orthogonal set of vectors, 369
Orthogonal Decomposition Theorem, 384
orthogonally diagonalizable matrix, 400
orthonormal basis, 372
orthonormal set of vectors, 372
properties of orthogonal matrices, 374–376
QR factorization, 393
Rank Theorem, 386
spectral decomposition, 405
Spectral Theorem, 403

Review Questions

1. Mark each of the following statements true or false:

(a) Every orthonormal set of vectors is linearly independent.

(b) Every nonzero subspace of $\mathbb{R}^n$ has an orthogonal basis.

(c) If A is a square matrix with orthonormal rows, then A is an orthogonal matrix.

(d) Every orthogonal matrix is invertible.

(e) If A is a matrix with $\det A = 1$, then A is an orthogonal matrix.

(f) If A is an $m \times n$ matrix such that $(\text{row}(A))^{\perp} = \mathbb{R}^n$, then A must be the zero matrix.

(g) If W is a subspace of $\mathbb{R}^n$ and $\mathbf{v}$ is a vector in $\mathbb{R}^n$ such that $\text{proj}_W(\mathbf{v}) = \mathbf{0}$, then $\mathbf{v}$ must be the zero vector.

(h) If A is a symmetric, orthogonal matrix, then $A^2 = I$.

(i) Every orthogonally diagonalizable matrix is invertible.

(j) Given any n real numbers $\lambda_1, \ldots, \lambda_n$, there exists a symmetric $n \times n$ matrix with $\lambda_1, \ldots, \lambda_n$ as its eigenvalues.

2. Find all values of a and b such that

$$\left\{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} a \\ b \\ 3 \end{bmatrix}\right\} \text{ is an orthogonal set of vectors.}$$

3. Find the coordinate vector $[\mathbf{v}]_{\mathcal{B}}$ of $\mathbf{v} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$ with respect to the orthogonal basis

$$\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\} \text{ of } \mathbb{R}^3$$

4. The coordinate vector of a vector $\mathbf{v}$ with respect to an orthonormal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{R}^2$ is $[\mathbf{v}]_\mathcal{B} = \begin{bmatrix} -3 \\ 1/2 \end{bmatrix}$. If $\mathbf{v}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$, find all possible vectors $\mathbf{v}$.

5. Show that $\begin{bmatrix} 6/7 & 2/7 & 3/7 \\ -1/\sqrt{5} & 0 & 2/\sqrt{5} \\ 4/7\sqrt{5} & -15/7\sqrt{5} & 2/7\sqrt{5} \end{bmatrix}$ is an orthogonal matrix.

6. If $\begin{bmatrix} 1/2 & a \\ b & c \end{bmatrix}$ is an orthogonal matrix, find all possible values of a, b, and c.

7. If Q is an orthogonal $n \times n$ matrix and $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is an orthonormal set in $\mathbb{R}^n$, prove that $\{Q\mathbf{v}_1, \ldots, Q\mathbf{v}_k\}$ is an orthonormal set.

8. If Q is an $n \times n$ matrix such that the angles $\angle(Q\mathbf{x}, Q\mathbf{y})$ and $\angle(\mathbf{x}, \mathbf{y})$ are equal for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, prove that Q is an orthogonal matrix.

In Questions 9–12, find a basis for $W^\perp$.

9. W is the line in $\mathbb{R}^2$ with general equation
$2x - 5y = 0$

10. W is the line in $\mathbb{R}^3$ with parametric equations
$x = t$
$y = 2t$
$z = -t$

11. $W = \text{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} \right\}$

12. $W = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 2 \end{bmatrix} \right\}$

13. Find bases for each of the four fundamental subspaces of

$$A = \begin{bmatrix} 1 & -1 & 2 & 1 & 3 \\ -1 & 2 & -2 & 1 & -2 \\ 2 & 1 & 4 & 8 & 9 \\ 3 & -5 & 6 & -1 & 7 \end{bmatrix}$$

14. Find the orthogonal decomposition of

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

with respect to

$$W = \text{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

15. (a) Apply the Gram-Schmidt Process to

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

to find an orthogonal basis for $W = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$.

(b) Use the result of part (a) to find a QR factorization of $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

16. Find an orthogonal basis for $\mathbb{R}^4$ that contains the vectors $\begin{bmatrix} 1 \\ 0 \\ 2 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}$.

17. Find an orthogonal basis for the subspace

$$W = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} : x_1 + x_2 + x_3 + x_4 = 0 \right\} \text{ of } \mathbb{R}^4$$

18. Let $A = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$.

(a) Orthogonally diagonalize A.
(b) Give the spectral decomposition of A.

19. Find a symmetric matrix with eigenvalues $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = -2$ and eigenspaces

$$E_1 = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right), E_{-2} = \text{span}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \right)$$

20. If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$ and

$$A = c_1\mathbf{v}_1\mathbf{v}_1^T + c_2\mathbf{v}_2\mathbf{v}_2^T + \cdots + c_n\mathbf{v}_n\mathbf{v}_n^T$$

prove that A is a symmetric matrix with eigenvalues $c_1, c_2, \ldots, c_n$ and corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.