Therefore the product XAX^t is a 1×1 matrix whose single entry is the dot product

$$\sum_{j=1}^{n} y_j x_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_i a_{ij} \right) x_j = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j .$$

Note: It is customary to identify the 1×1 matrix XAX^t with the sum in (5.9) and to call the product XAX^t a quadratic form. Equation (5.8) is written more simply as follows:

$$Q(x) = XAX^t .$$

EXAMPLE 1. Let $A = \begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix}$, $X = [x_1, x_2]$. Then we have

$$XA = [x_1, x_2] \begin{bmatrix} 1 & -1 \\ -3 & 5 \end{bmatrix} = [x_1 - 3x_2, -x_1 + 5x_2],$$

and hence

$$XAX^t = [x_1 - 3x_2, -x_1 + 5x_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 3x_2x_1 - x_1x_2 + 5x_2^2 .$$

EXAMPLE 2. Let $B = \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix}$, $X = [x_1, x_2]$. Then we have

$$XBX^t = [x_1, x_2] \begin{bmatrix} 1 & -2 \\ -2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 2x_2x_1 - 2x_1x_2 + 5x_2^2 .$$

In both Examples 1 and 2 the two mixed product terms add up to $-4x_1x_2$ so $XAX^t = XBX^t$. These examples show that different matrices can lead to the same quadratic form. Note that one of these matrices is symmetric. This illustrates the next theorem.

THEOREM 5.10. *For any $n \times n$ matrix A and any $1 \times n$ row matrix X we have $XAX^t = XBX^t$ where B is the symmetric matrix $B = \frac{1}{2}(A + A^t)$.*

Proof. Since XAX^t is a 1×1 matrix it is equal to its transpose, $XAX^t = (XAX^t)^t$. But the transpose of a product is the product of transposes in reversed order, so we have $(XAX^t)^t = XA^tX^t$. Therefore $XAX^t = \frac{1}{2}XAX^t + \frac{1}{2}XA^tX^t = XBX^t$.

5.13 Reduction of a real quadratic form to a diagonal form

A real symmetric matrix A is Hermitian. Therefore, by Theorem 5.7 it is similar to the diagonal matrix $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ of its eigenvalues. Moreover, we have $\Lambda = C^tAC$, where C is an orthogonal matrix. Now we show that C can be used to convert the quadratic form XAX^t to a diagonal form.

THEOREM 5.11. *Let XAX^t be the quadratic form associated with a real symmetric matrix A, and let C be an orthogonal matrix that converts A to a diagonal matrix $\Lambda = C^tAC$. Then we have*

$$XAX^t = Y\Lambda Y^t = \sum_{i=1}^{n} \lambda_i y_i^2,$$

where $Y = [y_1, \ldots, y_n]$ is the row matrix $Y = XC$, and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

Proof. Since C is orthogonal we have $C^{-1} = C^t$. Therefore the equation $Y = XC$ implies $X = YC^t$, and we obtain

$$XAX^t = (YC^t)A(YC^t)^t = Y(C^tAC)Y^t = Y\Lambda Y^t.$$

Note: Theorem 5.11 is described by saying that the linear transformation $Y = XC$ reduces the quadratic form XAX^t to a diagonal form $Y\Lambda Y^t$.

EXAMPLE 1. The quadratic form belonging to the identity matrix is

$$XIX^t = \sum_{i=1}^{n} x_i^2 = \|X\|^2,$$

the square of the length of the vector $X = (x_1, \ldots, x_n)$. A linear transformation $Y = XC$, where C is an orthogonal matrix, gives a new quadratic form $Y\Lambda Y^t$ with $\Lambda = CIC^t = CC^t = I$. Since $XIX^t = YIY^t$ we have $\|X\|^2 = \|Y\|^2$, so Y has the same length as X. A linear transformation which preserves the length of each vector is called an ***isometry.*** These transformations are discussed in more detail in Section 5.19.

EXAMPLE 2. Determine an orthogonal matrix C which reduces the quadratic form $Q(x) = 2x_1^2 + 4x_1x_2 + 5x_2^2$ to a diagonal form.

Solution. We write $Q(x) = XAX^t$, where $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$. This symmetric matrix was diagonalized in Example 1 following Theorem 5.7. It has the eigenvalues $\lambda_1 = 1$, $\lambda_2 = 6$, and an orthonormal set of eigenvectors u_1, u_2, where $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, $t = 1/\sqrt{5}$. An orthogonal diagonalizing matrix is $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. The corresponding diagonal form is

$$Y\Lambda Y^t = \lambda_1 y_1^2 + \lambda_2 y_2^2 = y_1^2 + 6y_2^2.$$

The result of Example 2 has a simple geometric interpretation, illustrated in Figure 5.1. The linear transformation $Y = XC$ can be regarded as a rotation which maps the basis $\boldsymbol{i}$, $\boldsymbol{j}$ onto the new basis u_1, u_2. A point with coordinates (x_1, x_2) relative to the first basis has new coordinates (y_1, y_2) relative to the second basis. Since $XAX^t = Y\Lambda Y^t$, the set of points (x_1, x_2) satisfying the equation $XAX^t = c$ for some c is identical with the set of points (y_1, y_2) satisfying $Y\Lambda Y^t = c$. The second equation, written as $y_1^2 + 6y_2^2 = c$, is

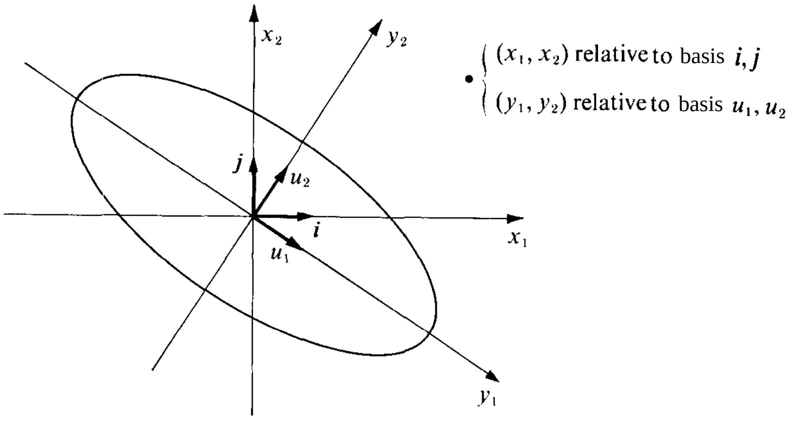


FIGURE 5.1 Rotation of axes by an orthogonal matrix. The ellipse has Cartesian equation $XAX^t = 9$ in the x_1x_2-system, and equation $Y\Lambda Y^t = 9$ in the y_1y_2-system.

the Cartesian equation of an ellipse if c > 0. Therefore the equation XAX^t = c, written as $2x_1^2 + 4x_1x_2 + 5x_2^2 = c$, represents the same ellipse in the original coordinate system. Figure 5.1 shows the ellipse corresponding to c = 9.

5.14 Applications to analytic geometry

The reduction of a quadratic form to a diagonal form can be used to identify the set of all points (x, y) in the plane which satisfy a Cartesian equation of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{5.10}$$

We shall find that this set is always a conic section, that is, an ellipse, hyperbola, parabola, or one of the degenerate cases (the empty set, a single point, or one or two straight lines). The type of conic is governed by the second-degree terms, that is, by the quadratic form $ax^2 + bxy + cy^2$. To conform with the notation used earlier, we write x_1 for x, x_2 for y, and express this quadratic form as a matrix product,

$$XAX^t = ax_1^2 + bx_1x_2 + cx_2^2,$$

where $X = [x_1, x_2]$ and $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. By a rotation Y = XC we reduce this form to a diagonal form $\lambda_1 y_1^2 + \lambda_2 y_2^2$, where λ_1, λ_2 are the eigenvalues of A. An orthonormal set of eigenvectors u_1, u_2 determines a new set of coordinate axes, relative to which the Cartesian equation (5.10) becomes

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + d'y_1 + e'y_2 + f = 0, \tag{5.11}$$

with new coefficients d' and e' in the linear terms. In this equation there is no mixed product term y_1y_2, so the type of conic is easily identified by examining the eigenvalues λ_1 and λ_2.

If the conic is not degenerate, Equation (5.11) represents an ellipse if λ_1, λ_2 have the same sign, a *hyperbola* if λ_1, λ_2 have opposite signs, and a *parabola* if either λ_1 or λ_2 is zero.. The three cases correspond to $\lambda_1\lambda_2 > 0$, $\lambda_1\lambda_2 < 0$, and $\lambda_1\lambda_2 = 0$. We illustrate with some specific examples.

EXAMPLE 1. $2x^2 + 4xy + 5y^2 + 4x + 13y - \frac{1}{4} = 0$. We rewrite this as

$$2x_1^2 + 4x_1x_2 + 5x_2^2 + 4x_1 + 13x_2 - \tfrac{1}{4} = 0. \tag{5.12}$$

The quadratic form $2x_1^2 + 4x_1x_2 + 5x_2^2$ is the one treated in Example 2 of the foregoing section. Its matrix has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 6$, and an orthonormal set of eigenvectors $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, where $t = 1/\sqrt{5}$. An orthogonal diagonalizing matrix is $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. This reduces the quadratic part of (5.12) to the form $y_1^2 + 6y_2^2$. To determine the effect on the linear part we write the equation of rotation $Y = XC$ in the form $X = YC^t$ and obtain

$$[x_1, x_2] = \frac{1}{\sqrt{5}}[y_1, y_2]\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}, \quad x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \quad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

Therefore the linear part $4x_1 + 13x_2$ is transformed to

$$\frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{13}{\sqrt{5}}(-y_1 + 2y_2) = -\sqrt{5}\,y_1 + 6\sqrt{5}\,y_2.$$

The transformed Cartesian equation becomes

$$y_1^2 + 6y_2^2 - \sqrt{5}\,y_1 + 6\sqrt{5}\,y_2 - \tfrac{1}{4} = 0.$$

By completing the squares in y_1 and y_2 we rewrite this as follows:

$$(y_1 - \tfrac{1}{2}\sqrt{5})^2 + 6(y_2 + \tfrac{1}{2}\sqrt{5})^2 = 9.$$

This is the equation of an ellipse with its center at the point $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ in the y_1y_2-system. The positive directions of the y_1 and y_2 axes are determined by the eigenvectors u_1 and u_2, as indicated in Figure 5.2.

We can simplify the equation further by writing

$$z_1 = y_1 - \tfrac{1}{2}\sqrt{5}, \qquad z_2 = y_2 + \tfrac{1}{2}\sqrt{5}.$$

Geometrically, this is the same as introducing a new system of coordinate axes parallel to the y_1y_2 axes but with the new origin at the center of the ellipse. In the z_1z_2-system the equation of the ellipse is simply

$$z_1^2 + 6z_2^2 = 9, \qquad \text{or} \qquad \frac{z_1^2}{9} + \frac{z_2^2}{3/2} = 1.$$

The ellipse and all three coordinate systems are shown in Figure 5.2.

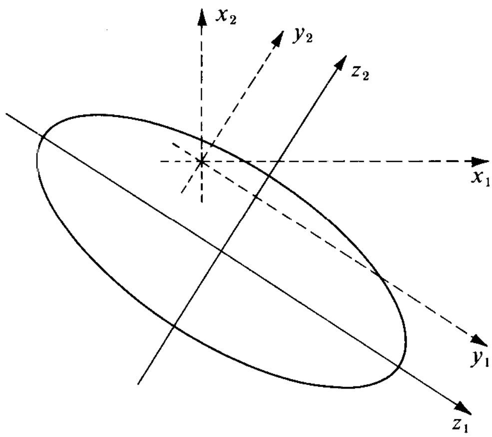


FIGURE 5.2 Rotation and translation of coordinate axes. The rotation $Y = XC$ is followed by the translation $z_1 = y_1 - \frac{1}{2}\sqrt{5}$, $z_2 = y_2 + \frac{1}{2}\sqrt{5}$.

EXAMPLE 2. $2x^2 - 4xy - y^2 - 4x + 10y - 13 = 0$. We rewrite this as

$$2x_1^2 - 4x_1x_2 - x_2^2 - 4x_1 + 10x_2 - 13 = 0.$$

The quadratic part is XAX^t, where $A = \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix}$. This matrix has the eigenvalues $\lambda_1 = 3$, $\lambda_2 = -2$. An orthonormal set of eigenvectors is $u_1 = t(2, -1)$, $u_2 = t(1, 2)$, where $t = 1/\sqrt{5}$. An orthogonal diagonalizing matrix is $C = t\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. The equation of rotation $X = YC^t$ gives us

$$x_1 = \frac{1}{\sqrt{5}}(2y_1 + y_2), \qquad x_2 = \frac{1}{\sqrt{5}}(-y_1 + 2y_2).$$

Therefore the transformed equation becomes

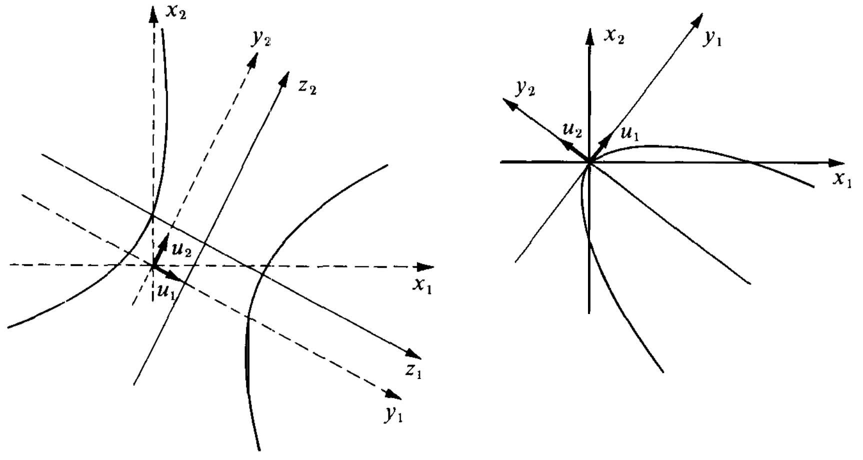
$$3y_1^2 - 2y_2^2 - \frac{4}{\sqrt{5}}(2y_1 + y_2) + \frac{10}{\sqrt{5}}(-y_1 + 2y_2) - 13 = 0,$$

or

$$3y_1^2 - 2y_2^2 - \frac{18}{\sqrt{5}}\, y_1 + \frac{16}{\sqrt{5}}\, y_2 - 13 = 0.$$

By completing the squares in y_1 and y_2 we obtain the equation

$$3(y_1 - \tfrac{3}{5}\sqrt{5})^2 - 2(y_2 - \tfrac{4}{5}\sqrt{5})^2 = 12,$$



(a) Hyperbola: $3z_1^2 - 2z_2^2 = 12$ (b) Parabola: $y_1^2 + y_2 = 0$

FIGURE 5.3 The curves in Examples 2 and 3.

which represents a hyperbola with its center at $(\frac{3}{5}\sqrt{5}, \frac{4}{5}\sqrt{5})$ in the y_1y_2-system. The translation $z_1 = y_1 - \frac{3}{5}\sqrt{5}$, $z_2 = y_2 - \frac{4}{5}\sqrt{5}$ simplifies this equation further to

$$3z_1^2 - 2z_2^2 = 12, \qquad \text{or} \qquad \frac{z_1^2}{4} - \frac{z_2^2}{6} = 1 .$$

The hyperbola is shown in Figure 5.3(a). The eigenvectors u_1 and u_2 determine the directions of the positive y_1 and y_2 axes.

EXAMPLE 3. $9x^2 + 24xy + 16y^2 - 20x + 15y = 0$. We rewrite this as

$$9x_1^2 + 24x_1x_2 + 16x_2^2 - 20x_1 + 15x_2 = 0 .$$

The symmetric matrix for the quadratic part is $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$. Its eigenvalues are $\lambda_1 = 25$, $\lambda_2 = 0$. An orthonormal set of eigenvectors is $u_1 = \frac{1}{5}(3, 4)$, $u_2 = \frac{1}{5}(-4, 3)$. An orthogonal diagonalizing matrix is $C = \frac{1}{5}\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$. The equation of rotation $X = YC^t$ gives us

$$x_1 = \tfrac{1}{5}(3y_1 - 4y_2), \qquad x_2 = \tfrac{1}{5}(4y_1 + 3y_2) .$$

Therefore the transformed Cartesian equation becomes

$$25y_1^2 - \tfrac{20}{5}(3y_1 - 4y_2) + \tfrac{15}{5}(4y_1 + 3y_2) = 0 .$$

This simplifies to $y_1^2 + y_2 = 0$, the equation of a parabola with its vertex at the origin. The parabola is shown in Figure 5.3(b).

EXAMPLE **4. *Degenerate cases.*** A knowledge of the eigenvalues alone does not reveal whether the Cartesian equation represents a degenerate conic section. For example, the three equations $x^2 + 2y^2 = 1$, $x^2 + 2y^2 = 0$, and $x^2 + 2y^2 = -1$ all have the same eigenvalues; the first represents a nondegenerate ellipse, the second is satisfied only by $(x, y) = (0, 0)$, and the third represents the empty set. The last two can be regarded as degenerate cases of the ellipse.

The graph of the equation $y^2 = 0$ is the x-axis. The equation $y^2 - 1 = 0$ represents the two parallel lines $y = 1$ a.nd $y = -1$. These can be regarded as degenerate cases of the parabola. The equation $x^2 - 4y^2 = 0$ represents two intersecting lines since it is satisfied if either $x - 2y = 0$ or $x + 2y = 0$. This can be regarded as a degenerate case of the hyperbola.

However, if the Cartesian equation $ax^2 + bxy + cy^2 + dx + ey + f = 0$ represents a nondegenerate conic section, then the ***type*** of conic can be determined quite easily. The characteristic polynomial of the matrix of the quadratic form $ax^2 + bxy + cy^2$ is

$$\det \begin{bmatrix} \lambda - a & -b/2 \\ -b/2 & \lambda - c \end{bmatrix} = \lambda^2 - (a + c)\lambda + (ac - \tfrac{1}{4}b^2) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

Therefore the product of the eigenvalues is

$$\lambda_1\lambda_2 = ac - \tfrac{1}{4}b^2 = \tfrac{1}{4}(4ac - b^2).$$

Since the type of conic is determined by the algebraic sign of the product $\lambda_1\lambda_2$, we see that the conic is an ***ellipse, hyperbola,*** or ***parabola,*** according as $4ac - b^2$ is ***positive, negative,*** or ***zero.*** The number $4ac - b^2$ is called the ***discriminant*** of the quadratic form $ax^2 + bxy + cy^2$. In Examples 1, 2 and 3 the discriminant has the values 34, -24, and 0, respectively.

5.15 Exercises

In each of Exercises 1 through 7, find (a) a symmetric matrix A for the quadratic form; (b) the eigenvalues of A; (c) an orthonormal set of eigenvectors; (d) an orthogonal diagonalizing matrix C.

1. 4x; $+ 4x_1x_2 + x_2^2$.
2. x_1x_2.
3. x_1^2 + 2x,x, $- x_2^2$.
4. 34x; $- 24x_1x_2 + 41x_2^2$.
5. $x_1^2 + x_1x_2 + x_1x_3 + x_2x_3$.
6. $2x_1^2 + 4x_1x_3 + x_2^2 - x_3^2$.
7. 3x; $+ 4x_1x_2 + 8x_1x_3 + 4x_2x_3$ + 3x;.

In each of Exercises 8 through 18, identify and make a sketch of the conic section represented by the Cartesian equation.

8. $y^2 - 2xy + 2x^2 - 5 = 0$.
9. $y^2 - 2xy + 5x = 0$.
10. $y^2 - 2xy + x^2 - 5x = 0$.
11. $5x^2 - 4xy + 2y^2 - 6 = 0$.
12. $19x^2 + 4xy + 16y^2 - 212x + 104y = 356$.
13. $9x^2 + 24xy + 16y^2 -$ 5:2x $+ 14y = 6$.
14. $5x^2 + 6xy + 5y^2$ -2 =O.
15. $x^2 + 2xy + y^2 - 2x + 2y + 3 = 0$.
16. $2x^2 + 4xy + 5y^2 - 2x - y - 4 = 0$.
17. $x^2 + 4xy - 2y^2 - 12 = 0$.
18. xy $+ y$ -2x -2 =O.
19. For what value (or values) of c will the graph of the Cartesian equation $2xy - 4x + 7y + c = 0$ be a pair of lines ?
20. If the equation $ax^2 + bxy + cy^2 = 1$ represents an ellipse, prove that the area of the region it bounds is $2\pi/\sqrt{4ac - b^2}$. This gives a geometric meaning to the discriminant $4ac - b^2$.

★ 5.16† Eigenvalues of a symmetric transformation obtained as values of its quadratic form

Now we drop the requirement that V be finite-dimensional and we find a relation between the eigenvalues of a symmetric operator and its quadratic form.

Suppose x is an eigenvector with norm 1 belonging to an eigenvalue λ. Then $T(x) = \lambda x$ so we have

$$Q(x) = (T(x), x) = (\lambda x, x) = \lambda(x, x) = \lambda, \tag{5.13}$$

since $(x, x) = 1$. The set of all x in V satisfying $(x, x) = 1$ is called the ***unit sphere*** in V. Equation (5.13) proves the following theorem.

THEOREM 5.12. *Let $T: V \to V$ be a symmetric transformation on a real Euclidean space V, and let $Q(x) = (T(x), x)$. Then the eigenvalues of T (if any exist) are to be found among the values that Q takes on the unit sphere in V.*

EXAMPLE. Let $V = V_2(R)$ with the usual basis (i, j) and the usual dot product as inner product. Let T be the symmetric transformation with matrix $A = \begin{bmatrix} 4 & 0 \\ 0 & 8 \end{bmatrix}$. Then the quadratic form of T is given by

$$Q(x) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij}x_i x_j = 4x_1^2 + 8x_2^2.$$

The eigenvalues of T are $\lambda_1 = 4$, $\lambda_2 = 8$. It is easy to see that these eigenvalues are, respectively, the minimum and maximum values which Q takes on the unit circle $x_1^2 + x_2^2 = 1$. In fact, on this circle we have

$$Q(x) = 4(x_1^2 + x_2^2) + 4x_2^2 = 4 + 4x_2^2, \qquad \text{where} \quad -1 \le x_2 \le 1.$$

This has its smallest value, 4, when $x_2 = 0$ and its largest value, 8, when $x_2 = \pm 1$.

Figure 5.4 shows the unit circle and two ellipses. The inner ellipse has the Cartesian equation $4x_1^2 + 8x_2^2 = 4$. It consists of all points $x = (x_1, x_2)$ in the plane satisfying $Q(x) = 4$. The outer ellipse has Cartesian equation $4x_1^2 + 8x_2^2 = 8$ and consists of all points satisfying $Q(x) = 8$. The points $(\pm 1, 0)$ where the inner ellipse touches the unit circle are eigenvectors belonging to the eigenvalue 4. The points $(0, \pm 1)$ on the outer ellipse are eigenvectors belonging to the eigenvalue 8.

The foregoing example illustrates extremal properties of eigenvalues which hold more generally. In the next section we will prove that the smallest and largest eigenvalues (if they exist) are always the minimum and maximum values which Q takes on the unit sphere. Our discussion of these extremal properties will make use of the following theorem on quadratic forms. It should be noted that this theorem does not require that V be finite dimensional.

† Starred sections can be omitted or postponed without loss in continuity.

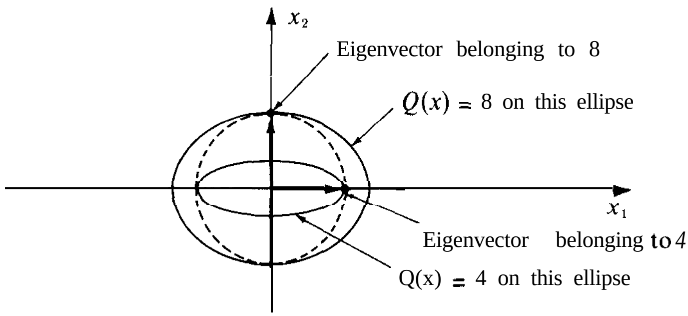


FIGURE 5.4 Geometric relation between the eigenvalues of T and the values of Q on the unit sphere, illustrated with a two-dimensional example.

THEOREM 5.13. *Let $T: V \to V$ be a symmetric transformation on a real Euclidean space V with quadratic form $Q(x) = (T(x), x)$. Assume that Q does not change sign on V. Then if $Q(x) = 0$ for some x in V we also have $T(x) = 0$. In other words, if Q does not change sign, then Q vanishes only on the null space of T.*

Proof. Assume $Q(x) = 0$ for some x in V and let y be any element in V. Choose any real t and consider $Q(x + ty)$. Using linearity of T, linearity of the inner product, and symmetry of T, we have

$$\begin{aligned} Q(x + ty) &= (T(x + ty), x + ty) = (T(x) + tT(y), x + ty) \\ &= (T(x), x) + t(T(x), y) + t(T(y), x) + t^2(T(y), y) \\ &= Q(x) + 2t(T(x), y) + t^2Q(y) = at + bt^2, \end{aligned}$$

where $a = 2(T(x), y)$ and $b = Q(y)$. If Q is nonnegative on V we have the inequality

$$at + bt^2 \geq 0 \qquad \text{for all real } t .$$

In other words, the quadratic polynomial $p(t) = at + bt^2$ has its minimum at $t = 0$. Hence $p'(0) = 0$. But $p'(0) = a = 2(T(x), y)$, so $(T(x), y) = 0$. Since y was arbitrary, we can in particular take $y = T(x)$, getting $(T(x), T(x)) = 0$. This proves that $T(x) = 0$.

If Q is nonpositive on V we get $p(t) = at + bt^2 \leq 0$ for all t, so p has its maximum at $t = 0$, and hence $p'(0) = 0$ as before.

★ 5.17 Extremal properties of eigenvalues of a symmetric transformation

Now we shall prove that the extreme values of a quadratic form on the unit sphere are eigenvalues'.

THEOREM 5.14. *Let $T: V \to V$ be a symmetric linear transformation on a real Euclidean space V, and let $Q(x) = (T(x), x)$. Among all values that Q takes on the unit sphere, assume* *there is an extremum*† *(maximum or minimum) at a point u with (u, u) = 1. Then u is an eigenvector for T; the corresponding eigenvalue is Q(u), the extreme value of Q on the unit sphere.*

Proof. Assume Q has a minimum at u. Then we have

$$(5.14) \qquad Q(x) \geq Q(u) \quad \text{for all } x \text{ with } (x, x) = 1.$$

Let $\lambda = Q(u)$. If $(x, x) = 1$ we have $Q(u) = \lambda(x, x) = (\lambda x, x)$ so inequality (5.14) can be written as

$$(5.15) \qquad (T(x), x) \geq (\lambda x, x)$$

provided $(x, x) = 1$. Now we prove that (5.15) is valid for *all* x in V. Suppose $\|x\| = a$. Then $x = ay$, where $\|y\| = 1$. Hence

$$(T(x), x) = (T(ay), ay) = a^2(T(y), y) \quad \text{and} \quad (\lambda x, x) = a^2(\lambda y, y).$$

But $(T(y), y) \geq (\lambda y, y)$ since $(y, y) = 1$. Multiplying both members of this inequality by a^2 we get (5.15) for $x = ay$.

Since $(T(x), x) - (\lambda x, x) = (T(x) - \lambda x, x)$, we can rewrite inequality (5.15) in the form $(T(x) - \lambda x, x) \geq 0$, or

$$(5.16) \qquad (S(x), x) \geq 0, \quad \text{where} \quad S = T - \lambda I.$$

When $x = u$ we have equality in (5.14) and hence also in (5.16). The linear transformation S is symmetric. Inequality (5.16) states that the quadratic form Q_1 given by $Q_1(x) = (S(x), x)$ is nonnegative on V. When $x = u$ we have $Q_1(u) = 0$. Therefore, by Theorem 5.13 we must have $S(u) = 0$. In other words, $T(u) = \lambda u$, so u is an eigenvector for T, and $\lambda = Q(u)$ is the corresponding eigenvalue. This completes the proof if Q has a minimum at u.

If there is a maximum at u all the inequalities in the foregoing proof are reversed and we apply Theorem 5.13 to the nonpositive quadratic form Q_1.

★ 5.18 The finite-dimensional case

Suppose now that $\dim V = n$. Then T has n real eigenvalues which can be arranged in increasing order, say

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n.$$

According to Theorem 5.14, the smallest eigenvalue λ_1 is the minimum of Q on the unit sphere, and the largest eigenvalue is the maximum of Q on the unit sphere. Now we shall show that the intermediate eigenvalues also occur as extreme values of Q, restricted to certain subsets of the unit sphere.

† If V is infinite-dimensional, the quadratic form Q need not have an extremum on the unit sphere. This will be the case when T has no eigenvalues. In the finite-dimensional case, Q always has a maximum and a minimum somewhere on the unit sphere. This follows as a consequence of a more general theorem on extreme values of continuous functions. For a special case of this theorem see Section 9.16.

Let u_1 be an eigenvector on the unit sphere which minimizes Q. Then $\lambda_1 = Q(u_1)$. If λ is an eigenvalue different from λ_1 any eigenvector belonging to λ must be orthogonal to u_1. Therefore it is natural to search for such an eigenvector on the orthogonal complement of the subspace spanned by u_1.

Let S be the subspace spanned by u_1. The orthogonal complement $S^\perp$ consists of all elements in V orthogonal to u_1. In particular, $S^\perp$ contains all eigenvectors belonging to eigenvalues $\lambda \neq \lambda_1$. It is easily verified that dim $S^\perp = n - 1$ and that T maps $S^\perp$ into itself.† Let S_{n-1} denote the unit sphere in the $(n - 1)$-dimensional subspace $S^\perp$. (The unit sphere S_{n-1} is a subset of the unit sphere in V.) Applying Theorem 5.14 to the subspace $S^\perp$ we find that $\lambda_2 = Q(u_2)$, where u_2 is a point which minimizes Q on S_{n-1}.

The next eigenvector λ_3 can be obtained in a similar way as the minimum value of Q on the unit sphere S_{n-2} in the $(n - 2)$-dimensional space consisting of those elements orthogonal to both u_1 and u_2. Continuing in this manner we find that each eigenvalue λ_k is the minimum value which Q takes on a unit sphere S_{n-k+1} in a subspace of dimension $n - k + 1$. The largest of these minima, λ_n, is also the ***maximum*** value which Q takes on each of the spheres S_{n-k+1}. The corresponding set of eigenvectors $u_1, \ldots, u_n$ form an orthonormal basis for V.

5.19 Unitary transformations

We conclude this chapter with a brief discussion of another important class of transformations known as unitary transformations. In the finite-dimensional case they are represented by unitary matrices.

DEFINITION. ***Let E be a Euclidean space and V a subspace of E. A linear transformation $T: V \to E$ is called unitary on V if we have***

$$(5.17) \qquad (T(x), T(y)) = (x, y) \quad \textit{for all } x \textit{ and } y \textit{ in } V.$$

When E is a real Euclidean space a unitary transformation is also called an orthogonal transformation.

Equation (5.17) is described by saying that T preserves inner products. Therefore it is natural to expect that T also preserves orthogonality and norms, since these are derived from the inner product.

THEOREM 5.15. ***If $T: V \to E$ is a unitary transformation on V, then for all x and y in V we have:***

(a) ***$(x, y) = 0$ implies $(T(x), T(y)) = 0$*** ***(T preserves orthogonality).***

(b) $\|T(x)\| = \|x\|$ ***(T preserves norms).***

(c) $\|T(x) - T(y)\| = \|x - y\|$ ***(T preserves distances).***

(d) ***T is invertible, and T^{-1} is unitary on $T(V)$.***

† This was done in the proof of Theorem 5.4, Section 5.6.

Proof. Part (a) follows at once from Equation (5.17). Part (b) follows by taking $x = y$ in (5.17). Part (c) follows from (b) because $T(x) - T(y) = T(x - y)$.

To prove (d) we use (b) which shows that $T(x) = 0$ implies $x = O$, so T is invertible. If $x \in T(V)$ and $y \in T(V)$ we can write $x = T(u)$, $y = T(v)$, so we have

$$(T^{-1}(x), T^{-1}(y)) = (u, v) = (T(u), T(v)) = (x, y).$$

Therefore T^{-1} is unitary on $T(V)$.

Regarding eigenvalues and eigenvectors we have the following theorem.

THEOREM 5.116. *Let $T: V \to E$ be a unitary transformation on V.*
(a) *If T has an eigenvalue λ, then $|\lambda| = 1$.*
(b) *If x and y are eigenvectors belonging to distinct eigenvalues λ and μ, then x and y are orthogonal.*
(c) *Zf $V = E$ and* dim $V = n$, *and if V is a complex space, then there exist eigenvectors $u_1, \ldots, u_n$ of T which form an orthonormal basis for V. The matrix of T relative to this basis is the diagonal matrix $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, where λ_k is the eigenvalue belonging to u_k.*

Proof. To prove (a), let x be an eigenvector belonging to il. Then $x \neq 0$ and $T(x) = \lambda x$. Taking $y = x$ in Equation (5.17) we get

$$(\lambda x, \lambda x) = (x, x) \qquad \text{or} \qquad \lambda\bar{\lambda}(x, x) = (x, x).$$

Since $(x, x) > 0$ and $\lambda\bar{\lambda} = |\lambda|^2$, this implies $|\lambda| = 1$.

To prove (b), write $T(x) = Lx$, $T(y) = \mu y$ and compute the inner product $(T(x), T(y))$ in two ways. We have

$$(T(x), T(y)) = (x, y)$$

since T is unitary. We also have

$$(T(x), T(y)) = (\lambda x, \mu y) = \lambda\bar{\mu}(x, y)$$

since x and y are eigenvectors. Therefore $\lambda\bar{\mu}(x, y) = (x, y)$, so $(x, y) = 0$ unless $\lambda\bar{\mu} = 1$. But $\lambda\bar{\lambda} = 1$ by (a), so if we had $\lambda\bar{\mu} = 1$ we would also have $\lambda\bar{\lambda} = \lambda\bar{\mu}$, $\bar{\lambda} = \bar{\mu}$, $\lambda = \mu$, which contradicts the assumption that λ and μ are distinct. Therefore $\lambda\bar{\mu} \neq 1$ and $(x, y) = 0$.

Part (c) is proved by induction on n in much the same way that we proved Theorem 5.4, the corresponding result for Hermitian operators. The only change required is in that part of the proof which shows that T maps $S^\perp$ into itself, where

$$S^\perp = \{x \mid x \in V, \quad (x, u_1) = 0\}.$$

Here u_1 is an eigenvector of T with eigenvalue λ_1. From the equation $T(u_1) = \lambda_1 u_1$ we find

$$u_1 = \lambda_1^{-1} T(u_1) = \bar{\lambda}_1 T(u_1)$$

since $\lambda_1\bar{\lambda}_1 = |\lambda_1|^2 = 1$. Now choose any x in $S^\perp$ and note that

$$(T(x), u_1) = (T(x), \bar{\lambda}_1 T(u_1)) = \lambda_1(T(x), T(u_1)) = \lambda_1(x, u_1) = 0.$$

Hence $T(x) \in S^\perp$ if $x \in S^\perp$, so T maps $S^\perp$ into itself. The rest of the proof is identical with that of Theorem 5.4, so we shall not repeat the details.

The next two theorems describe properties of unitary transformations on a finite-dimensional space. We give only a brief outline of the proofs.

THEOREM 5.17. *Assume* dim $V = n$ *and let* $E = (e_1, \ldots, e_n)$ *be a jixed basis for* V. *Then a linear transformation* $T: V \to V$ *is unitary if and only if*

(5.18) $$(T(e_i), T(e_j)) = (e_i, e_j) \quad \text{for all } i \text{ and } j.$$

In particular, if E *is orthonormal then* T *is unitary if and only if* T *maps* E *onto an orthonormal basis.*

Sketch of proof. Write $x = \sum x_i e_i$, $y = \sum y_j e_j$. Then we have

$$(x, y) = \left(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (e_i, e_j),$$

and

$$(T(x), T(y)) = \left(\sum_{i=1}^{n} x_i T(e_i), \sum_{j=1}^{n} y_j T(e_j)\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (T(e_i), T(e_j)).$$

Now compare (x, y) with $(T(x), T(y))$.

THEOREM 5.18. *Assume* dim $V = n$ *and let* $(e_1, \ldots, e_n)$ *be an orthonormal basis for* V. *Let* $A = (a_{ij})$ *be the matrix representation of a linear transformation* $T: V \to V$ *relative to this basis. Then* T *is unitary if and only if* A *is unitary, that is, if and only if*

(5.19) $$A^*A = I.$$

Sketch of proof. Since (e_i, e_j) is the ij-entry of the identity matrix, Equation (5.19) implies

(5.20) $$(e_i, e_j) = \sum_{k=1}^{n} \bar{a}_{ki} a_{kj} = \sum_{k=1}^{n} a_{ki} \bar{a}_{kj}.$$

Since A is the matrix of T we have $T(e_i) = \sum_{k=1}^{n} a_{ki} e_k$, $T(e_j) = \sum_{r=1}^{n} a_{rj} e_r$, so

$$(T(e_i), T(e_j)) = \left(\sum_{k=1}^{n} a_{ki} e_k, \sum_{r=1}^{n} a_{rj} e_r\right) = \sum_{k=1}^{n} \sum_{r=1}^{n} a_{ki} \bar{a}_{rj} (e_k, e_r) = \sum_{k=1}^{n} a_{ki} \bar{a}_{kj}.$$

Now compare this with (5.20) and use Theorem 5.17.