8.4 Eigenbases. Diagonalization. Quadratic Forms

So far we have emphasized properties of eigen*values*. We now turn to general properties of eigen*vectors*. Eigenvectors of an $n \times n$ matrix $\mathbf{A}$ may (or may not!) form a basis for R^n. If we are interested in a transformation $\mathbf{y} = \mathbf{Ax}$, such an "**eigenbasis**" (basis of eigenvectors)—if it exists—is of great advantage because then we can represent any $\mathbf{x}$ in R^n uniquely as a linear combination of the eigenvectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$, say,

$$\mathbf{x} = c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_n\mathbf{x}_n.$$

And, denoting the corresponding (not necessarily distinct) eigenvalues of the matrix $\mathbf{A}$ by $\lambda_1, \cdots, \lambda_n$, we have $\mathbf{Ax}_j = \lambda_j\mathbf{x}_j$, so that we simply obtain

$$\begin{aligned} \mathbf{y} = \mathbf{Ax} &= \mathbf{A}(c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_n) \\ &= c_1\mathbf{Ax}_1 + \cdots + c_n\mathbf{Ax}_n \\ &= c_1\lambda_1\mathbf{x}_1 + \cdots + c_n\lambda_n\mathbf{x}_n. \end{aligned} \tag{1}$$

This shows that we have decomposed the complicated action of $\mathbf{A}$ on an arbitrary vector $\mathbf{x}$ into a sum of simple actions (multiplication by scalars) on the eigenvectors of $\mathbf{A}$. This is the point of an eigenbasis.

Now if the n eigenvalues are all different, we do obtain a basis:

THEOREM 1

Basis of Eigenvectors

*If an $n \times n$ matrix $\mathbf{A}$ has n **distinct** eigenvalues, then $\mathbf{A}$ has a basis of eigenvectors $\mathbf{x}_1, \cdots, \mathbf{x}_n$ for R^n.*

PROOF All we have to show is that $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are linearly independent. Suppose they are not. Let r be the largest integer such that $\{\mathbf{x}_1, \cdots, \mathbf{x}_r\}$ is a linearly independent set. Then $r < n$ and the set $\{\mathbf{x}_1, \cdots, \mathbf{x}_r, \mathbf{x}_{r+1}\}$ is linearly dependent. Thus there are scalars $c_1, \cdots, c_{r+1}$, not all zero, such that

$$c_1\mathbf{x}_1 + \cdots + c_{r+1}\mathbf{x}_{r+1} = \mathbf{0} \tag{2}$$

(see Sec. 7.4). Multiplying both sides by $\mathbf{A}$ and using $\mathbf{Ax}_j = \lambda_j\mathbf{x}_j$, we obtain

$$\mathbf{A}(c_1\mathbf{x}_1 + \cdots + c_{r+1}\mathbf{x}_{r+1}) = c_1\lambda_1\mathbf{x}_1 + \cdots + c_{r+1}\lambda_{r+1}\mathbf{x}_{r+1} = \mathbf{A0} = \mathbf{0}. \tag{3}$$

To get rid of the last term, we subtract λ_{r+1} times (2) from this, obtaining

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r = \mathbf{0}.$$

Here $c_1(\lambda_1 - \lambda_{r+1}) = 0, \cdots, c_r(\lambda_r - \lambda_{r+1}) = 0$ since $\{x_1, \cdots, x_r\}$ is linearly independent. Hence $c_1 = \cdots = c_r = 0$, since all the eigenvalues are distinct. But with this, (2) reduces to $c_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$, hence $c_{r+1} = 0$, since $\mathbf{x}_{r+1} \neq \mathbf{0}$ (an eigenvector!). This contradicts the fact that not all scalars in (2) are zero. Hence the conclusion of the theorem must hold. ■

EXAMPLE 1 **Eigenbasis. Nondistinct Eigenvalues. Nonexistence**

The matrix $\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ has a basis of eigenvectors $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ corresponding to the eigenvalues $\lambda_1 = 8$, $\lambda_2 = 2$. (See Example 1 in Sec. 8.2.)

Even if not all n eigenvalues are different, a matrix $\mathbf{A}$ may still provide an eigenbasis for R^n. See Example 2 in Sec. 8.1, where $n = 3$.

On the other hand, $\mathbf{A}$ ***may not have enough linearly independent eigenvectors to make up a basis***. For instance, $\mathbf{A}$ in Example 3 of Sec. 8.1 is

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \text{and has only one eigenvector} \qquad \begin{bmatrix} k \\ 0 \end{bmatrix} \qquad (k \neq 0, \text{ arbitrary}).$$

Actually, eigenbases exist under much more general conditions than those in Theorem 1. An important case is the following.

THEOREM 2 **Symmetric Matrices**

A symmetric matrix has an orthonormal basis of eigenvectors for R^n.

For a proof (which is involved) see Ref. [B3], vol. 1, pp. 270–272.

EXAMPLE 2 **Orthonormal Basis of Eigenvectors**

The first matrix in Example 1 is symmetric, and an orthonormal basis of eigenvectors is $[1/\sqrt{2} \quad 1/\sqrt{2}]^{\mathsf{T}}$, $[1/\sqrt{2} \quad -1/\sqrt{2}]^{\mathsf{T}}$.

Similarity of Matrices. Diagonalization

Eigenbases also play a role in reducing a matrix $\mathbf{A}$ to a diagonal matrix whose entries are the eigenvalues of $\mathbf{A}$. This is done by a "similarity transformation," which is defined as follows (and will have various applications in numerics in Chap. 20).

DEFINITION **Similar Matrices. Similarity Transformation**

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular!) $n \times n$ matrix $\mathbf{P}$. This transformation, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

The key property of this transformation is that it preserves the eigenvalues of $\mathbf{A}$:

THEOREM 3 **Eigenvalues and Eigenvectors of Similar Matrices**

If $\hat{\mathbf{A}}$ is similar to $\mathbf{A}$, then $\hat{\mathbf{A}}$ has the same eigenvalues as $\mathbf{A}$.

Furthermore, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue.

similar matrix -->
same eigenvalue
eigenvector y=P^(-1)x

PROOF From $\mathbf{Ax} = \lambda\mathbf{x}$ (λ an eigenvalue, $\mathbf{x} \neq \mathbf{0}$) we get $\boxed{\mathbf{P}^{-1}\mathbf{Ax} = \lambda\mathbf{P}^{-1}\mathbf{x}}$. Now $\mathbf{I} = \mathbf{PP}^{-1}$. By this ***identity trick*** the equation $\mathbf{P}^{-1}\mathbf{Ax} = \lambda\mathbf{P}^{-1}\mathbf{x}$ gives

$$\boxed{\mathbf{P}^{-1}\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{AIx} = \mathbf{P}^{-1}\mathbf{APP}^{-1}\mathbf{x} = (\mathbf{P}^{-1}\mathbf{AP})\mathbf{P}^{-1}\mathbf{x} = \hat{\mathbf{A}}(\mathbf{P}^{-1}\mathbf{x}) = \lambda\mathbf{P}^{-1}\mathbf{x}.}$$

Hence λ is an eigenvalue of $\hat{\mathbf{A}}$ and $\mathbf{P}^{-1}\mathbf{x}$ a corresponding eigenvector. Indeed, $\mathbf{P}^{-1}\mathbf{x} \neq \mathbf{0}$ because $\mathbf{P}^{-1}\mathbf{x} = \mathbf{0}$ would give $\mathbf{x} = \mathbf{Ix} = \mathbf{PP}^{-1}\mathbf{x} = \mathbf{P0} = \mathbf{0}$, contradicting $\mathbf{x} \neq \mathbf{0}$. ■

EXAMPLE 3 **Eigenvalues and Vectors of Similar Matrices**

Let,
$$\mathbf{A} = \begin{bmatrix} 6 & -3 \\ 4 & -1 \end{bmatrix} \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}.$$

Then
$$\hat{\mathbf{A}} = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 6 & -3 \\ 4 & -1 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}.$$

1. 先求出eigenvalue
2. 再求出eigenvector
   x1, x2
3. X=[x1 x2]
4. D=X^(-1) A X
   =[lambda_1 0
   0 lambda_2]
5. 可求出 A^m

Here $\mathbf{P}^{-1}$ was obtained from (4*) in Sec. 7.8 with det $\mathbf{P} = 1$. We see that $\hat{\mathbf{A}}$ has the eigenvalues $\lambda_1 = 3$, $\lambda_2 = 2$. The characteristic equation of $\mathbf{A}$ is $(6 - \lambda)(-1 - \lambda) + 12 = \lambda^2 - 5\lambda + 6 = 0$. It has the roots (the eigenvalues of $\mathbf{A}$) $\lambda_1 = 3$, $\lambda_2 = 2$, confirming the first part of Theorem 3.

We confirm the second part. From the first component of $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ we have $(6 - \lambda)x_1 - 3x_2 = 0$. For $\lambda = 3$ this gives $3x_1 - 3x_2 = 0$, say, $\mathbf{x}_1 = [1 \quad 1]^\mathsf{T}$. For $\lambda = 2$ it gives $4x_1 - 3x_2 = 0$, say, $\mathbf{x}_2 = [3 \quad 4]^\mathsf{T}$. In Theorem 3 we thus have

$$\mathbf{y}_1 = \mathbf{P}^{-1}\mathbf{x}_1 = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{y}_2 = \mathbf{P}^{-1}\mathbf{x}_2 = \begin{bmatrix} 4 & -3 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Indeed, these are eigenvectors of the diagonal matrix $\hat{\mathbf{A}}$.

Perhaps we see that $\mathbf{x}_1$ and $\mathbf{x}_2$ are the columns of $\mathbf{P}$. This suggests the general method of transforming a matrix $\mathbf{A}$ to diagonal form $\mathbf{D}$ by using $\mathbf{P} = \mathbf{X}$, the matrix with eigenvectors as columns. ■

By a suitable similarity transformation we can now transform a matrix $\mathbf{A}$ to a diagonal matrix $\mathbf{D}$ whose diagonal entries are the eigenvalues of $\mathbf{A}$:

THEOREM 4 **Diagonalization of a Matrix**

If an $n \times n$ matrix $\mathbf{A}$ has a basis of eigenvectors, then

$$\boxed{\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX}} \tag{5}$$

is diagonal, with the eigenvalues of $\mathbf{A}$ as the entries on the main diagonal. Here $\mathbf{X}$ is the matrix with these eigenvectors as column vectors. Also,

$$\boxed{\mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X}} \qquad (m = 2, 3, \cdots). \tag{5*}$$

期末考會考二階矩陣 A, 求 A^(100), A^(100) = X D^(100) X^(-1)

PROOF Let $\mathbf{x}_1, \cdots, \mathbf{x}_n$ be a basis of eigenvectors of $\mathbf{A}$ for R^n. Let the corresponding eigenvalues of $\mathbf{A}$ be $\lambda_1, \cdots, \lambda_n$, respectively, so that $\mathbf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1, \cdots, \mathbf{A}\mathbf{x}_n = \lambda_n\mathbf{x}_n$. Then $\boxed{\mathbf{X} = [\mathbf{x}_1 \cdots \mathbf{x}_n]}$ has rank n, by Theorem 3 in Sec. 7.4. Hence $\mathbf{X}^{-1}$ exists by Theorem 1 in Sec. 7.8. We claim that

$$\boxed{\mathbf{A}\mathbf{X} = \mathbf{A}[\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n] = [\mathbf{A}\mathbf{x}_1 \quad \cdots \quad \mathbf{A}\mathbf{x}_n] = [\lambda_1\mathbf{x}_1 \quad \cdots \quad \lambda_n\mathbf{x}_n] = \mathbf{X}\mathbf{D}} \tag{6}$$

where $\mathbf{D}$ is the diagonal matrix as in (5). The fourth equality in (6) follows by direct calculation. (Try it for $n = 2$ and then for general n.) The third equality uses $\mathbf{A}\mathbf{x}_k = \lambda_k\mathbf{x}_k$. The second equality results if we note that the first column of $\mathbf{AX}$ is $\mathbf{A}$ times the first column of $\mathbf{X}$, which is $\mathbf{x}_1$, and so on. For instance, when $n = 2$ and we write $\mathbf{x}_1 = [x_{11} \quad x_{21}]$, $\mathbf{x}_2 = [x_{12} \quad x_{22}]$, we have

$$\begin{aligned}\mathbf{AX} = \mathbf{A}[\mathbf{x}_1 \quad \mathbf{x}_2] &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \\ &= \begin{array}{cc} \left[\begin{matrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{22}x_{22} \end{matrix}\right] \\ \text{Column 1} \qquad\qquad \text{Column 2} \end{array} = [\mathbf{A}\mathbf{x}_1 \quad \mathbf{A}\mathbf{x}_2].\end{aligned}$$

If we multiply (6) by $\mathbf{X}^{-1}$ from the left, we obtain (5). Since (5) is a similarity transformation, Theorem 3 implies that $\mathbf{D}$ has the same eigenvalues as $\mathbf{A}$. Equation (5*) follows if we note that

$$\boxed{\mathbf{D}^2 = \mathbf{DD} = (\mathbf{X}^{-1}\mathbf{AX})(\mathbf{X}^{-1}\mathbf{AX}) = \mathbf{X}^{-1}\mathbf{A}(\mathbf{X}\mathbf{X}^{-1})\mathbf{AX} = \mathbf{X}^{-1}\mathbf{AAX} = \mathbf{X}^{-1}\mathbf{A}^2\mathbf{X},} \quad \text{etc.} \ \blacksquare$$

EXAMPLE 4 **Diagonalization**

Diagonalize

$$\mathbf{A} = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}.$$

Solution. The characteristic determinant gives the characteristic equation $-\lambda^3 - \lambda^2 + 12\lambda = 0$. The roots (eigenvalues of $\mathbf{A}$) are $\lambda_1 = 3$, $\lambda_2 = -4$, $\lambda_3 = 0$. By the Gauss elimination applied to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1, \lambda_2, \lambda_3$ we find eigenvectors and then $\mathbf{X}^{-1}$ by the Gauss–Jordan elimination (Sec. 7.8, Example 1). The results are

$$\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}, \qquad \mathbf{X}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

Calculating $\mathbf{AX}$ and multiplying by $\mathbf{X}^{-1}$ from the left, we thus obtain

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{AX} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}\begin{bmatrix} -3 & -4 & 0 \\ 9 & 4 & 0 \\ -3 & -12 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad \blacksquare$$

Quadratic Forms. Transformation to Principal Axes

By definition, a **quadratic form** Q in the components $x_1, \cdots, x_n$ of a vector $\mathbf{x}$ is a sum of n^2 terms, namely,

$$
\boxed{\begin{aligned}
Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} &= \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_jx_k \\
&= \quad a_{11}x_1^2 \quad + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\
&\quad + a_{21}x_2x_1 + a_{22}x_2^2 \quad + \cdots + a_{2n}x_2x_n \\
&\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2.
\end{aligned}} \tag{7}
$$

$\mathbf{A} = [a_{jk}]$ is called the **coefficient matrix** of the form. We may assume that $\mathbf{A}$ is ***symmetric***, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms; see the following example.

EXAMPLE 5 Quadratic Form. Symmetric Coefficient Matrix

Let

$$
\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 4x_1x_2 + 6x_2x_1 + 2x_2^2 = 3x_1^2 + 10x_1x_2 + 2x_2^2.
$$

Here $4 + 6 = 10 = 5 + 5$. From the corresponding ***symmetric*** matrix $\mathbf{C} = [c_{jk}]$, where $c_{jk} = \frac{1}{2}(a_{jk} + a_{kj})$, thus $c_{11} = 3$, $c_{12} = c_{21} = 5$, $c_{22} = 2$, we get the same result; indeed,

$$
\mathbf{x}^{\mathsf{T}}\mathbf{C}\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 5x_1x_2 + 5x_2x_1 + 2x_2^2 = 3x_1^2 + 10x_1x_2 + 2x_2^2. \qquad \blacksquare
$$

Quadratic forms occur in physics and geometry, for instance, in connection with conic sections (ellipses $x_1^2/a^2 + x_2^2/b^2 = 1$, etc.) and quadratic surfaces (cones, etc.). Their transformation to principal axes is an important practical task related to the diagonalization of matrices, as follows.

By Theorem 2, the *symmetric* coefficient matrix $\mathbf{A}$ of (7) has an orthonormal basis of eigenvectors. Hence if we take these as column vectors, we obtain a matrix $\mathbf{X}$ that is orthogonal, so that $\mathbf{X}^{-1} = \mathbf{X}^{\mathsf{T}}$. From (5) we thus have $\mathbf{A} = \mathbf{X}\mathbf{D}\mathbf{X}^{-1} = \mathbf{X}\mathbf{D}\mathbf{X}^{\mathsf{T}}$. Substitution into (7) gives

$$
Q = \mathbf{x}^{\mathsf{T}}\mathbf{X}\mathbf{D}\mathbf{X}^{\mathsf{T}}\mathbf{x}. \tag{8}
$$

If we set $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, then, since $\mathbf{X}^{\mathsf{T}} = \mathbf{X}^{-1}$, we have $\mathbf{X}^{-1}\mathbf{x} = \mathbf{y}$ and thus obtain

$$
\mathbf{x} = \mathbf{X}\mathbf{y}. \tag{9}
$$

Furthermore, in (8) we have $\mathbf{x}^{\mathsf{T}}\mathbf{X} = (\mathbf{X}^{\mathsf{T}}\mathbf{x})^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}$ and $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, so that Q becomes simply

$$
Q = \mathbf{y}^{\mathsf{T}}\mathbf{D}\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2. \tag{10}
$$

This proves the following basic theorem.

THEOREM 5

Principal Axes Theorem

The substitution (9) *transforms a quadratic form*

$$Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_jx_k \qquad (a_{kj} = a_{jk})$$

to the principal axes form or **canonical form** (10), *where* $\lambda_1, \cdots, \lambda_n$ *are the* (*not necessarily distinct*) *eigenvalues of the* (*symmetric!*) *matrix* $\mathbf{A}$, *and* $\mathbf{X}$ *is an orthogonal matrix with corresponding eigenvectors* $\mathbf{x}_1, \cdots, \mathbf{x}_n$, *respectively, as column vectors.*

EXAMPLE 6 **Transformation to Principal Axes. Conic Sections**

Find out what type of conic section the following quadratic form represents and transform it to principal axes:

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128.$$

Solution. We have $Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$, where

A選擇為symmetric matrix

$$\boxed{\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix},} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This gives the characteristic equation $(17 - \lambda)^2 - 15^2 = 0$. It has the roots $\boxed{\lambda_1 = 2, \lambda_2 = 32.}$ Hence (10) becomes

可以轉換成 Q = lambda_1 y_1^2 + lambda_2 y_2^2

$$\boxed{Q = 2y_1^2 + 32y_2^2.}$$

We see that $Q = 128$ represents the ellipse $2y_1^2 + 32y_2^2 = 128$, that is,

$$\frac{y_1^2}{8^2} + \frac{y_2^2}{2^2} = 1.$$

If we want to know the direction of the principal axes in the x_1x_2-coordinates, we have to determine normalized eigenvectors from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1 = 2$ and $\lambda = \lambda_2 = 32$ and then use (9). We get

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

hence

行列式值 |X|=1

$$\mathbf{x} = \mathbf{X}\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad \begin{aligned} x_1 &= y_1/\sqrt{2} - y_2/\sqrt{2} \\ x_2 &= y_1/\sqrt{2} + y_2/\sqrt{2}. \end{aligned}$$

This is a 45° rotation. Our results agree with those in Sec. 8.2, Example 1, except for the notations. See also Fig. 160 in that example. ■

PROBLEM SET 8.4

1–5 SIMILAR MATRICES HAVE EQUAL EIGENVALUES

Verify this for $\mathbf{A}$ and $\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{AP}$. If $\mathbf{y}$ is an eigenvector of $\mathbf{P}$, show that $\mathbf{x} = \mathbf{Py}$ are eigenvectors of $\mathbf{A}$. Show the details of your work.

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

2. $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 7 & -5 \\ 10 & -7 \end{bmatrix}$

3. $\mathbf{A} = \begin{bmatrix} 8 & -4 \\ 2 & 2 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0.28 & 0.96 \\ -0.96 & 0.28 \end{bmatrix}$

4. $\mathbf{A} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 3 & 2 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 3 & 0 & 5 \end{bmatrix},$
   $\lambda_1 = 3$

5. $\mathbf{A} = \begin{bmatrix} -5 & 0 & 15 \\ 3 & 4 & -9 \\ -5 & 0 & 15 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. **PROJECT. Similarity of Matrices.** Similarity is basic, for instance, in designing numeric methods.

   (a) Trace. By definition, the **trace** of an $n \times n$ matrix $\mathbf{A} = [a_{jk}]$ is the sum of the diagonal entries,

   $$\text{trace } \mathbf{A} = a_{11} + a_{22} + \cdots + a_{nn}.$$

   Show that the trace equals the sum of the eigenvalues, each counted as often as its algebraic multiplicity indicates. Illustrate this with the matrices $\mathbf{A}$ in Probs. 1, 3, and 5.

   (b) Trace of product. Let $\mathbf{B} = [b_{jk}]$ be $n \times n$. Show that similar matrices have equal traces, by first proving

   $$\text{trace } \mathbf{AB} = \sum_{i=1}^{n} \sum_{l=1}^{n} a_{il}b_{li} = \text{trace } \mathbf{BA}.$$

   (c) Find a relationship between $\hat{\mathbf{A}}$ in (4) and $\hat{\hat{\mathbf{A}}} = \mathbf{PAP}^{-1}$.

   (d) Diagonalization. What can you do in (5) if you want to change the order of the eigenvalues in $\mathbf{D}$, for instance, interchange $d_{11} = \lambda_1$ and $d_{22} = \lambda_2$?

7. **No basis.** Find further 2×2 and 3×3 matrices without eigenbasis.

8. **Orthonormal basis.** Illustrate Theorem 2 with further examples.

9–16 DIAGONALIZATION OF MATRICES

Find an eigenbasis (a basis of eigenvectors) and diagonalize. Show the details.

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

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

11. $\begin{bmatrix} -19 & 7 \\ -42 & 16 \end{bmatrix}$

12. $\begin{bmatrix} -4.3 & 7.7 \\ 1.3 & 9.3 \end{bmatrix}$

13. $\begin{bmatrix} 4 & 0 & 0 \\ 12 & -2 & 0 \\ 21 & -6 & 1 \end{bmatrix}$

14. $\begin{bmatrix} -5 & -6 & 6 \\ -9 & -8 & 12 \\ -12 & -12 & 16 \end{bmatrix}, \quad \lambda_1 = -2$

15. $\begin{bmatrix} 4 & 3 & 3 \\ 3 & 6 & 1 \\ 3 & 1 & 6 \end{bmatrix}, \quad \lambda_1 = 10$

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

17–23 PRINCIPAL AXES. CONIC SECTIONS

What kind of conic section (or pair of straight lines) is given by the quadratic form? Transform it to principal axes. Express $\mathbf{x}^\mathsf{T} = [x_1 \quad x_2]$ in terms of the new coordinate vector $\mathbf{y}^\mathsf{T} = [y_1 \quad y_2]$, as in Example 6.

17. $7x_1^2 + 6x_1x_2 + 7x_2^2 = 200$
18. $3x_1^2 + 8x_1x_2 - 3x_2^2 = 10$
19. $3x_1^2 + 22x_1x_2 + 3x_2^2 = 0$
20. $9x_1^2 + 6x_1x_2 + x_2^2 = 10$
21. $x_1^2 - 12x_1x_2 + x_2^2 = 70$
22. $4x_1^2 + 12x_1x_2 + 13x_2^2 = 16$
23. $-11x_1^2 + 84x_1x_2 + 24x_2^2 = 156$

24. Definiteness. A quadratic form $Q(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ and its (symmetric!) matrix $\mathbf{A}$ are called **(a) positive definite** if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, **(b) negative definite** if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq \mathbf{0}$, **(c) indefinite** if $Q(\mathbf{x})$ takes both positive and negative values. (See Fig. 162.) [$Q(\mathbf{x})$ and $\mathbf{A}$ are called *positive semidefinite* (*negative semidefinite*) if $Q(\mathbf{x}) \geqq 0$ ($Q(\mathbf{x}) \leqq 0$) for all $\mathbf{x}$.] Show that a necessary and sufficient condition for (a), (b), and (c) is that the eigenvalues of $\mathbf{A}$ are (a) all positive, (b) all negative, and (c) both positive and negative.

Hint. Use Theorem 5.

25. Definiteness. A necessary and sufficient condition for positive definiteness of a quadratic form $Q(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ with *symmetric* matrix $\mathbf{A}$ is that all the **principal minors** are positive (see Ref. [B3], vol. 1, p. 306), that is,

$$a_{11} > 0, \qquad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} > 0, \qquad \cdots, \qquad \det \mathbf{A} > 0.$$

Show that the form in Prob. 22 is positive definite, whereas that in Prob. 23 is indefinite.

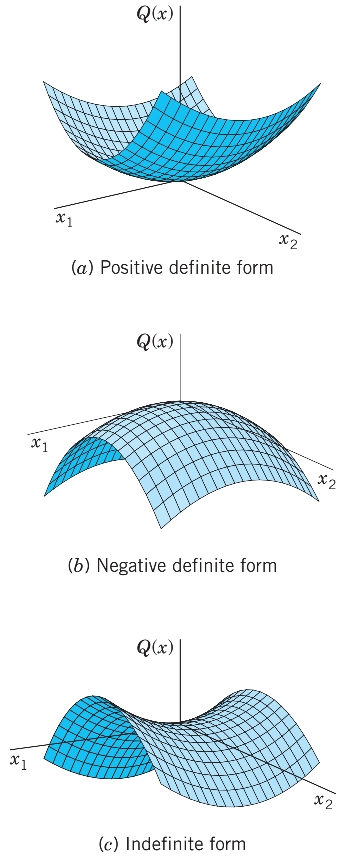

Fig. 162. Quadratic forms in two variables (Problem 24)

8.5 Complex Matrices and Forms. *Optional*

The three classes of matrices in Sec. 8.3 have complex counterparts which are of practical interest in certain applications, for instance, in quantum mechanics. This is mainly because of their spectra as shown in Theorem 1 in this section. The second topic is about extending quadratic forms of Sec. 8.4 to complex numbers. (The reader who wants to brush up on complex numbers may want to consult Sec. 13.1.)

Notations

$\overline{\mathbf{A}} = [\overline{a}_{jk}]$ is obtained from $\mathbf{A} = [a_{jk}]$ by replacing each entry $a_{jk} = \alpha + i\beta$ (α, β real) with its complex conjugate $\overline{a}_{jk} = \alpha - i\beta$. Also, $\overline{\mathbf{A}}^\mathsf{T} = [\overline{a}_{kj}]$ is the transpose of $\overline{\mathbf{A}}$, hence the conjugate transpose of $\mathbf{A}$.

EXAMPLE 1 Notations

If $\mathbf{A} = \begin{bmatrix} 3+4i & 1-i \\ 6 & 2-5i \end{bmatrix}$, then $\overline{\mathbf{A}} = \begin{bmatrix} 3-4i & 1+i \\ 6 & 2+5i \end{bmatrix}$ and $\overline{\mathbf{A}}^\mathsf{T} = \begin{bmatrix} 3-4i & 6 \\ 1+i & 2+5i \end{bmatrix}$. ■