Chapter 8
Unitary and Hermitian Matrices

8.1 Unitary Matrices

A complex square matrix $U \in \mathbf{C}^{n\times n}$ that satisfies

$$U^h U = U U^h = I$$

is called **unitary**. If U is a real unitary matrix then

$$U^t U = U U^t = I$$

and is U called **orthogonal**. Equivalently, a complex matrix U is unitary if $U^{-1} = U^h$, and a real matrix is orthogonal if $U^{-1} = U^t$. Note that the columns of an $n \times n$ unitary (orthogonal) matrix form an orthonormal basis for $\mathbf{C}^{n\times 1}$ ($\mathbf{R}^{n\times 1}$).

If U is unitary, then

$$\langle U\mathbf{x} \,|\, U\mathbf{y} \rangle = \mathbf{x}^h U^h U \mathbf{y} = \mathbf{x}^h \mathbf{y} = \langle \mathbf{x} \,|\, \mathbf{y} \rangle \quad \text{for all } \ \mathbf{x}, \mathbf{y} \in \mathbf{C}^{n\times 1}$$

Consequently, $\| U\mathbf{x} \| = \| \mathbf{x} \|$ for all $\mathbf{x} \in \mathbf{C}^{n\times 1}$, and if $\{ \mathbf{x}_1, \ldots, \mathbf{x}_k \}$ is an orthonormal set, then so is $\{ U\mathbf{x}_1, \ldots, U\mathbf{x}_k \}$. Also,

$$\theta(U\mathbf{x}, U\mathbf{y}) = \theta(\mathbf{x}, \mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbf{C}^{n\times 1}$. In other words, a mapping by a unitary transformation preserves norms and angles.

Example 8.1

It can easily be verified that

$$R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

is an orthogonal matrix.

Let

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{y} = R\mathbf{x} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}$$

Then

$$\| \mathbf{y} \|^2 = \frac{(x_1 - x_2)^2 + (x_1 + x_2)^2}{2} = x_1^2 + x_2^2 = \| \mathbf{x} \|^2$$

Expressing R as

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad \theta = \frac{\pi}{4}$$

we observe that the transformation $\mathbf{y} = R\mathbf{x}$ corresponds to a counterclockwise rotation in the plane by an angle of $\theta = \pi/4$ (see Example 5.3).

If

$$R_1 = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \quad \text{and} \quad R_2 = \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}$$

are two such rotation matrices corresponding to conterclockwise rotations by θ_1 and θ_2, then we expect that $R = R_2R_1$ should also be a rotation matrix corresponding to a conterclockwise rotation by $\theta = \theta_1 + \theta_2$. Indeed, simple trigonometric identities give

$$\begin{aligned} R &= \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 & -\cos\theta_2\sin\theta_1 - \sin\theta_2\cos\theta_1 \\ \cos\theta_2\sin\theta_1 + \sin\theta_2\cos\theta_1 & \cos\theta_2\cos\theta_1 - \sin\theta_2\sin\theta_1 \end{bmatrix} \\ &= \begin{bmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{bmatrix} \end{aligned}$$

Note that R is also an orthogonal matrix.

Rotation matrices in the 3-space can be defined similarly (see Exercise 8.3).

If U is a unitary matrix, then

$$1 = \det(U^hU) = (\det U^h)(\det U) = (\det U)^*(\det U) = |\det U|^2$$

so that $|\det U| = 1$. If U is orthogonal then $\det U$ is real, and therefore

$$\det U = \mp 1$$

As a simple example, the reader can verify that $\det U = 1$ for the rotation matrix in Example 8.1.

Structure of unitary matrices is characterized by the following theorem.

Theorem 8.1 *Let $U \in \mathbf{C}^{n\times n}$ be a unitary matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then*

a) $|\lambda_i| = 1, i = 1, \ldots, n$

b) there exists a unitary matrix $P \in \mathbf{C}^{n\times n}$ such that

$$P^hUP = D = diag[\lambda_1, \ldots, \lambda_n]$$

Proof We use induction on n.

For $n = 1$, $U = u$ (a scalar) with $\lambda = u$. Then $U^hU = u^*u = |u|^2 = 1$, and the result is trivially true with $P = 1$ and $D = u$.

Suppose that (a) and (b) are true for all unitary matrices of order $n-1$, and consider a unitary matrix $U = U_1$ of order n. Let λ_1 be an eigenvalue of U, and let $\mathbf{v}_1$ be a unit eigenvector (scaled to have unity norm) associated with λ_1. Choose V_1 such that

$$P_1 = [\,\mathbf{v}_1 \;\; V_1\,]$$

is a unitary matrix. (Columns of V_1 complete $\{\mathbf{v}_1\}$ to an orthonormal basis for $\mathbf{C}^{n\times 1}$.) Then

$$\mathbf{v}_1^h V_1 = \mathbf{0}_{1\times(n-1)} \quad \text{and} \quad V_1^h \mathbf{v}_1 = \mathbf{0}_{(n-1)\times 1}$$

from which we obtain

$$\begin{aligned} P_1^h U_1 P_1 &= \begin{bmatrix} \mathbf{v}_1^h \\ V_1^h \end{bmatrix} U_1 [\,\mathbf{v}_1 \;\; V_1\,] &&= \begin{bmatrix} \mathbf{v}_1^h U_1 \mathbf{v}_1 & \mathbf{v}_1^h U_1 V_1 \\ V_1^h U_1 \mathbf{v}_1 & V_1^h U_1 V_1 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \mathbf{v}_1^h \mathbf{v}_1 & \lambda_1^* \mathbf{v}_1^h V_1 \\ \lambda_1 V_1^h \mathbf{v}_1 & V_1^h U_1 V_1 \end{bmatrix} &&= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & U_2 \end{bmatrix} \end{aligned}$$

Since

$$(P_1^h U_1 P_1)^h (P_1^h U_1 P_1) = P_1^h U_1^h P_1 P_1^h U_1 P_1 = I$$

we have

$$\begin{bmatrix} \lambda_1^* \lambda_1 & \mathbf{0} \\ \mathbf{0} & U_2^h U_2 \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}$$

which implies that $\lambda_1^* \lambda_1 = 1$, that is, $|\,\lambda_1\,|^2 = 1$, and that U_2 is unitary. Let U_2 have the eigenvalues $\lambda_2, \ldots, \lambda_n$. Then they are also eigenvalues of $U = U_1$. Since U_2 is of order $n-1$, by induction hypothesis $|\,\lambda_2\,|^2 = \cdots = |\,\lambda_n\,|^2 = 1$ and there exists a unitary matrix P_2 such that

$$P_2^h U_2 P_2 = D_2 = \operatorname{diag}[\,\lambda_2, \,\ldots\,, \lambda_n\,]$$

Let

$$P = P_1 \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix}$$

Then

$$\begin{aligned} P^h U P &= \begin{bmatrix} 1 & \\ & P_2^h \end{bmatrix} P_1^h U_1 P_1 \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & \\ & P_2^h \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & U_2 \end{bmatrix} \begin{bmatrix} 1 & \\ & P_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & \mathbf{0} \\ \mathbf{0} & D_2 \end{bmatrix} \;=\; D_1 \end{aligned}$$

Theorem 8.1 simply states that eigenvalues of a unitary (orthogonal) matrix are located on the unit circle in the complex plane, that such a matrix can always be diagonalized (even if it has multiple eigenvalues), and that a modal matrix can be chosen to be unitary (orthogonal).

Example 8.2

The matrix

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$$

is unitary as

$$U^h U = \frac{1}{2}\begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Its characteristic equation

$$s^2 - \sqrt{2}s + 1 = 0$$

gives the eigenvalues $\lambda_{1,2} = (1 \mp i)/\sqrt{2}$. We observe that $|\,\lambda_{1,2}\,| = 1$ as expected. An eigenvector associated with λ_1 is found by solving

$$(U - \lambda_1 I)\mathbf{v} = \frac{1}{\sqrt{2}}\begin{bmatrix} -i & i \\ i & -i \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

as

$$\mathbf{v}_1 = \mathrm{col}\,[\,1, 1\,]$$

Similarly, an eigenvector associated with λ_2 is found by solving

$$(U - \lambda_2 I)\mathbf{v} = \frac{1}{\sqrt{2}}\begin{bmatrix} i & i \\ i & i \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

as

$$\mathbf{v}_1 = \mathrm{col}\,[\,-1, 1\,]$$

Note that we need not look specifically for an eigenvector $\mathbf{v}_2$ that is orthogonal to $\mathbf{v}_1$; eigenvectors of a unitary matrix associated with distinct eigenvalues are orthogonal (see Exercise 8.11).

Normalizing the eigenvectors, we obtain a unitary modal matrix

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

The reader can easily verify that

$$P^h U P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1+i & \\ & 1-i \end{bmatrix}$$

8.2 Hermitian Matrices

Recall that a matrix $A \in \mathbf{C}^{n\times n}$ is called Hermitian if $A^h = A$, and that a real Hermitian matrix is symmetric.

The following theorem characterizes structure of Hermitian matrices.

Theorem 8.2 *Let $A \in \mathbf{C}^{n\times n}$ be a Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then*

a) $\lambda_i^* = \lambda_i, i = 1, \ldots, n$, that is, eigenvalues of A are real

b) there exists a unitary matrix $P \in \mathbf{C}^{n\times n}$ such that

$$P^h A P = D = diag\,[\,\lambda_1,\, \ldots,\, \lambda_n\,]$$

Proof If $\mathbf{v}$ is a unit eigenvector of A associated with an eigenvalue λ, then

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

and

$$\mathbf{v}^h A = \mathbf{v}^h A^h = (A\mathbf{v})^h = (\lambda\mathbf{v})^h = \lambda^* \mathbf{v}^h$$

Premultiplying both sides of the first equality by $\mathbf{v}^h$, postmultiplying both sides of the second equality by $\mathbf{v}$, and noting that $\mathbf{v}^h\mathbf{v} = \|\,\mathbf{v}\,\|^2 = 1$, we get

$$\mathbf{v}^h A\mathbf{v} = \lambda = \lambda^*$$

Hence all eigenvalues of A are real.

The existence of a unitary modal matrix P that diagonalizes A can be shown by following almost the same lines as in the proof of Theorem 8.1, and is left to the reader as an exercise.

Hence, like unitary matrices, Hermitian (symmetric) matrices can always be diagonalized by means of a unitary (orthogonal) modal matrix.

Example 8.3

The real symmetrix matrix

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

has the characteristic polynomial $d(s) = (s-1)^2(s-7)$. We observe that the eigenvalues are real.

Two linearly independent eigenvectors associated with the multiple eigenvalue $\lambda_1 = 1$ can be found by solving

$$(A - \lambda_1 I)\mathbf{v} = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} \mathbf{v} = \mathbf{0}$$

as

$$\mathbf{v}_{11} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_{22} = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

Applying the Gram-Schmidt process to $\{\mathbf{v}_{11}, \mathbf{v}_{12}\}$, and normalizing the orthogonal eigenvector generated by the process, we obtain two orthonormal eigenvectors associated with $\lambda_1 = 1$ as

$$\mathbf{u}_{11} = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{u}_{12} = \frac{1}{\sqrt{30}} \begin{bmatrix} -2 \\ -1 \\ 5 \end{bmatrix}$$

An eigenvector associated with $\lambda_2 = 7$ is found as

$$\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Like the eigenvectors of a unitary matrix, eigenvectors of a Hermitian matrix associated with distinct eigenvalues are also orthogonal (see Exercise 8.11). Therefore, we need not specifically look for an eigenvector $\mathbf{v}_2$ that is orthogonal to $\mathbf{v}_{11}$ and $\mathbf{v}_{12}$. After normalizing $\mathbf{v}_2$, we obtain a unit eigenvector associated with $\lambda_2 = 7$ as

$$\mathbf{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

The reader can verify that the modal matrix

$$P = [\, \mathbf{u}_{11} \;\; \mathbf{u}_{12} \;\; \mathbf{u}_2 \,] = \begin{bmatrix} -\frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{30}} & \frac{2}{\sqrt{6}} \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{5}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

is orthogonal and that

$$P^t A P = \mathrm{diag}\,[\, 1, 1, 7 \,]$$

8.3 Quadratic Forms

Let $S \in \mathbf{R}^{n \times n}$ be a symmetric matrix and let $\mathbf{x} \in \mathbf{R}^{n \times 1}$. An expression of the form

$$q(\mathbf{x}) = \mathbf{x}^t S \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} x_i x_j \tag{8.1}$$

is called a **quadratic form** in $\mathbf{x}$. Note that $q(\mathbf{x})$ is a scalar for every $\mathbf{x} \in \mathbf{R}^{n \times 1}$.

Clearly $q(\mathbf{0}) = 0$. If $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$, then $q(\mathbf{x})$ is said to be **positive definite**. If $q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$, and $q(\mathbf{y}) = 0$ for at least one $\mathbf{y} \neq \mathbf{0}$, then $q(\mathbf{x})$ is said to be **positive semi-definite**. $q(\mathbf{x})$ is said to be negative definite (negative semi-definite) if $-q(\mathbf{x})$ is positive definite (positive semi-definite), and indefinite if it is neither positive definite nor negative definite.

A real symmetric matrix S is said to be positive definite (positive semi-definite, negative definite, negative semi-definite, indefinite) if the associated quadratic form $q(\mathbf{x}) = \mathbf{x}^t S \mathbf{x}$ is positive definite (positive semi-definite, negative definite, negative semi-definite, indefinite).

Example 8.4

The quadratic form

$$q_1(x_1, x_2) = x_1^2 + 2x_1x_2 + 4x_2^2$$

involving the real variables x_1 and x_2 can be written as

$$q_1(\mathbf{x}) = [\, x_1 \;\; x_2 \,] \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^t S_1 \mathbf{x}$$

Note that the diagonal elements of S are the coefficients of the square terms x_1^2 and x_2^2, and the symmetrically located off-diagonal elements are one half of the coefficient of the cross-product term x_1x_2. Since

$$q_1(x_1, x_2) = (x_1 + x_2)^2 + 3x_2^2$$

q_1 is positive definite ($q_1 \geq 0$, and $q_1 = 0$ implies $x_1 = x_2 = 0$). Therefore, the symmetric matrix

$$S_1 = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

is also positive definite.

The quadratic form

$$q_2(x_1, x_2) = x_1^2 + 4x_1x_2 + 4x_2^2 = (x_1 + 2x_2)^2$$

is positive semi-definite, because $q_2 \geq 0$ and $q_2 = 0$ for any $x_1 = -2x_2 \neq 0$. Hence the matrix of q_2

$$S_2 = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

is also positive semi-definite.

The quadratic form

$$q_3(x_1, x_2) = x_1^2 + 6x_1x_2 + 4x_2^2$$

is indefinite, because $q_3(1, 0) = 1 > 0$ and $q_3(1, -1) = -1 < 0$. Thus its matrix

$$S_3 = \begin{bmatrix} 1 & 3 \\ 3 & 4 \end{bmatrix}$$

is indefinite.

In Example 8.4 we established positive definiteness of a quadratic form by expressing it as a linear combination of square terms, which is not always as easy as it was in this example. A systematic way of testing sign properties of a quadratic form is based on Theorem 8.2, and is described below.

Since $P^tSP = D$ for some orthogonal matrix P, a change of the variables as

$$\mathbf{x} = P\tilde{\mathbf{x}} \tag{8.2}$$

transforms the quadratic form in (8.1) into

$$q(\mathbf{x}) = \tilde{\mathbf{x}}^t P^t S P \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^t D \tilde{\mathbf{x}} = \tilde{q}(\tilde{\mathbf{x}}) \tag{8.3}$$

Since P is nonsingular, it represents a change of basis in $\mathbf{R}^{n\times 1}$. Therefore, q and $\tilde{q}$ are equivalent and thus have the same sign property. Also, since

$$\tilde{q}(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}^t D \tilde{\mathbf{x}} = \sum_{i=1}^{n} \lambda_i \tilde{x}_i^2$$

sign of $\tilde{q}$ is completely determined by the eigenvalues λ_i of S. We conclude that a symmetric matrix (whose eigenvalues are real) is positive (negative) definite if and only if all eigenvalues are positive (negative), positive (negative) semi-definite if and only if all eigenvalues are nonnegative (nonpositive) and at least one eigenvalue is zero, and indefinite if and only if it has both positive and negative eigenvalues.

Example 8.5

The matrix S_1 in Example 8.4 has the eigenvalues

$$\lambda_1 = (5+\sqrt{13})/2 \approx 4.3028\,, \quad \lambda_2 = (5-\sqrt{13})/2 \approx 0.6972$$

Since both eigenvalues are positive, S_1 and hence q_1 are positive definite.

S_2 has the eigenvalues

$$\lambda_1 = 4\,, \quad \lambda_2 = 0$$

and therefore it is positive semi-definite.

The indefinite matrix S_3 has a positive and a negative eigenvalue:

$$\lambda_1 = (5+\sqrt{45})/2 \approx 5.8541\,, \quad \lambda_2 = (5-\sqrt{45})/2 \approx -0.8541$$

8.3.1 Bounds of Quadratic Forms

Let $S \in \mathbf{R}^{n\times n}$ be a symmetric matrix with real eigenvalues $\lambda_1, \ldots, \lambda_n$, and associated orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ that form a basis for $\mathbf{R}^{n\times 1}$. For an arbitrary $\mathbf{x} \in \mathbf{R}^{n\times 1}$ expressed as

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

we have

$$\|\,\mathbf{x}\,\|^2 = \sum_{i=1}^{n} \|\,\alpha_i \mathbf{v}_i\,\|^2 = \sum_{i=1}^{n} |\,\alpha_i\,|^2$$

Considering

$$\mathbf{x}^t S \mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \mathbf{v}_i^t S \mathbf{v}_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j \lambda_i \mathbf{v}_i^t \mathbf{v}_j = \sum_{i=1}^{n} \lambda_i\,|\,\alpha_i\,|^2$$

we find that

$$\lambda_{\min} \sum_{i=1}^{n} |\,\alpha_i\,|^2 \le \mathbf{x}^t S \mathbf{x} \le \lambda_{\max} \sum_{i=1}^{n} |\,\alpha_i\,|^2$$

or equivalently

$$\lambda_{\min}\,\|\,\mathbf{x}\,\|^2 \le \mathbf{x}^t S \mathbf{x} \le \lambda_{\max}\,\|\,\mathbf{x}\,\|^2 \tag{8.4}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and and maximum eigenvalues of S. Clearly, equality on either side holds if $\mathbf{x}$ equals the corresponding eigenvector. (8.4) establishes bounds on a quadratic form $q(\mathbf{x}) = \mathbf{x}^t S \mathbf{x}$. It also provides an alternative explanation to the relation between sign-definiteness of a quadratic form and the eigenvalues of its symmetric matrix.

8.3.2 Quadratic Forms in Complex Variables

A quadratic form can also be formed by a complex vector $\mathbf{z} \in \mathbf{C}^{n\times 1}$ just by replacing the real symmetric matrix S in (8.1) by a complex Hermitian matrix H. Thus a quadratic form in $\mathbf{z} \in \mathbf{C}^{n\times 1}$ is

$$q(\mathbf{z}) = \mathbf{z}^h H \mathbf{z} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij} z_i^* z_j \tag{8.5}$$

where $H \in \mathbf{C}^{n\times n}$ is Hermitian. Although $\mathbf{z}$ is complex, since

$$q^*(\mathbf{z}) = (\mathbf{z}^h H \mathbf{z})^h = \mathbf{z}^h H^h \mathbf{z} = \mathbf{z}^h H \mathbf{z} = q(\mathbf{z})$$

the quadratic form in (8.5) is real. This allows us to extend the definitions of definite and semi-definite quadratic forms in real variables and real symmetric matrices to quadratic forms in complex variables and complex Hermitian matrices.

Example 8.6

Consider the quadratic form

$$q(z_1, z_2) = \mathbf{z}^h H \mathbf{z} = [\, z_1^* \;\; z_2^* \,] \begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = z_1^* z_1 + i z_1^* z_2 - i z_1 z_2^* + 2 z_2^* z_2$$

in the complex variables z_1 and z_2.

Rewriting the quadratic form as

$$q(z_1, z_2) = (z_1 + i z_2)^* (z_1 + i z_2) + z_2^* z_2 = |\, z_1 + i z_2 \,|^2 + |\, z_2 \,|^2$$

we observe that q is positive definite. Thus the Hermitian matrix

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

of the quadratic form is also positive definite.

Note that by letting $z_1 = x_1 + i y_1$ and $z_2 = x_2 + i y_2$, we can express q as

$$\begin{aligned} q(z_1, z_2) &= |\, (x_1 - y_2) + i(x_2 + y_1) \,|^2 + |\, z_2 \,|^2 \\ &= (x_1 - y_2)^2 + (x_2 + y_1)^2 + x_2^2 + y_2^2 \\ &= x_1^2 + y_1^2 - 2x_1 y_2 + 2x_2 y_1 + 2x_2^2 + 2y_2^2 \\ &= Q(x_1, x_2, y_1, y_2) \end{aligned}$$

which is a quadratic form in real variables.

Expressions similar to (8.3) and (8.4) can easily be derived for a quadratic form in complex variables. With $\mathbf{z} = P\tilde{\mathbf{z}}$, where P is a unitary modal matrix of H, the quadratic form $q(\mathbf{z}) = \mathbf{z}^h H \mathbf{z}$ is transformed into

$$q(\mathbf{z}) = \tilde{\mathbf{z}}^h P^h H P \tilde{\mathbf{z}} = \tilde{\mathbf{z}}^h D \tilde{\mathbf{z}} = \tilde{q}(\tilde{\mathbf{z}}) = \sum_{i=1}^{n} \lambda_i |\, \tilde{z}_i \,|^2 \tag{8.6}$$

Again, the sign properties of $q(\mathbf{z})$ can be deduced from the eigenvalues of H. For example, the Hermitian matrix of the quadratic form in Example 8.6 has the real eigenvalues

$$\lambda_1 = (3 + \sqrt{5})/2 \approx 2.6180\,, \quad \lambda_1 = (3 - \sqrt{5})/2 \approx 0.3280$$

Since both eigenvalues are positive we conclude that the quadratic form (and hence its matrix) are positive definite.

Similarly, for $q(\mathbf{z}) = \mathbf{z}^h H \mathbf{z}$ (8.4) becomes

$$\lambda_{\min} \| \mathbf{z} \|^2 \leq \mathbf{z}^h H \mathbf{z} \leq \lambda_{\max} \| \mathbf{z} \|^2 \tag{8.7}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and and maximum eigenvalues of H.

Finally, we observe from Example 8.6 that a quadratic form in complex variables is equivalent to a quadratic form in real variables (which are the real and imaginary parts of the complex variables). To prove this statement in general, let

$$\mathbf{z} = \mathbf{x} + i\mathbf{y}\,, \quad H = S + iK$$

in (8.5), where $\mathbf{x}$ and $\mathbf{y}$ are real, S is a real symmetric matrix, and K is a real skew-symmetric matrix. Then, noting that

$$\mathbf{x}^t K \mathbf{x} = \mathbf{y}^t K \mathbf{y} = 0\,, \quad \mathbf{x}^t S \mathbf{y} = \mathbf{y}^t S \mathbf{x}$$

$q(\mathbf{z})$ can be expressed as

$$\begin{aligned} q(\mathbf{z}) &= (\mathbf{x}^t - i\mathbf{y}^t)(S + iK)(\mathbf{x} + i\mathbf{y}) \\ &= \mathbf{x}^t S \mathbf{x} - \mathbf{x}^t K \mathbf{y} + \mathbf{y}^t K \mathbf{x} + \mathbf{y}^t S \mathbf{y} \\ &= [\,\mathbf{x}^t \;\; \mathbf{y}^t\,] \begin{bmatrix} S & -K \\ K & S \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \;=\; Q(\mathbf{x}, \mathbf{y}) \end{aligned} \tag{8.8}$$

involving real quantities only.

Since the quadratic forms $q(\mathbf{x})$ and $Q(\mathbf{y}, \mathbf{z})$ in (8.8) are equivalent, the eigenvalues of their matrices $H = S + iK$ and

$$\tilde{H} = \begin{bmatrix} S & -K \\ K & S \end{bmatrix} \tag{8.9}$$

must be related. This relation is studied in Exercise 8.12.

8.3.3 Conic Sections and Quadric Surfaces

Recall from analytic geometry that an equation of the form

$$s_{11}x_1^2 + 2s_{12}x_1x_2 + s_{22}x_2^2 = 1$$

where not all coefficients are zero, defines a central conic in the x_1x_2 plane. A suitable way to investigate the properties of such a conic is to rewrite the defining equation in compact form as

$$\mathbf{x}^t S \mathbf{x} = 1 \tag{8.10}$$

with the obvious definitions of $\mathbf{x}$ and S.

Let S have the eigenvalues $\lambda_1 \geq \lambda_2$ and an orthogonal modal matrix P such that

$$P^t S P = D = \operatorname{diag}[\,\lambda_1, \lambda_2\,]$$

Then a change of coordinate system as $\mathbf{x} = P\tilde{\mathbf{x}}$ transforms the equation of the conic into

$$\tilde{\mathbf{x}}^t D \tilde{\mathbf{x}} = \lambda_1 \tilde{x}_1^2 + \lambda_2 \tilde{x}_2^2 = 1$$

Depending on the signs of the (real) eigenvalues λ_1 and λ_2, we consider the following distinct cases:

a) $\lambda_1 \geq \lambda_2 > 0$ (S positive definite): Letting $a_1 = 1/\sqrt{\lambda_1},\ a_2 = 1/\sqrt{\lambda_2}$, the equation of the conic takes the form

$$\frac{\tilde{x}_1^2}{a_1^2} + \frac{\tilde{x}_2^2}{a_2^2} = 1$$

which represents an ellipse in the $\tilde{x}_1\tilde{x}_2$ plane with axes of length a_1 and a_2.

b) $\lambda_1 > \lambda_2 = 0$ (S positive semi-definite): Again, letting $a_1 = 1/\sqrt{\lambda_1}$, the equation becomes

$$\tilde{x}_1^2 = a_1^2$$

which represents two straight lines $\tilde{x}_1 = \mp a_1$.

c) $\lambda_1 > 0 > \lambda_2$ (S indefinite): With $a_1 = 1/\sqrt{\lambda_1},\ a_2 = 1/\sqrt{-\lambda_2}$, we have

$$\frac{\tilde{x}_1^2}{a_1^2} - \frac{\tilde{x}_2^2}{a_2^2} = 1$$

which represents a hyperbola.

d) $0 \geq \lambda_1 \geq \lambda_2$ (S negative definite or negative semi-definite): In these cases no point in the $\tilde{x}_1\tilde{x}_2$ plane (and therefore, no point in the x_1x_2 plane) satisfies the equation.

Note that the parabola, which is another conic, does not occur in any of the cases considered above, because it is not a central conic. A more general equation, which also includes the parabola, is considered in Exercise 8.28.

Example 8.7

Consider the equation

$$(10-c)x_1^2 + 2(6+c)x_1x_2 + (10-c)x_2^2 = \mathbf{x}^t S\mathbf{x} = 1$$

where c is a real parameter, and

$$S = \begin{bmatrix} 10-c & 6+c \\ 6+c & 10-c \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The matrix S has the eigenvalues

$$\lambda_1 = 16\,, \quad \lambda_2 = 4 - 2c$$

and an orthogonal modal matrix

$$P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

which is independent of the parameter c. Thus the change of basis as in (8.2) transforms the equation of the conic into

$$16\tilde{x}_1^2 + (4-2c)\tilde{x}_2^2 = 1$$

For $c = 0, 2, 4$, the equation becomes and represents

$$\begin{array}{lrcl} c=0: & 16\tilde{x}_1^2 + 4\tilde{x}_2^2 = 1 & : & \text{an elipse} \\ c=2: & 16\tilde{x}_1^2 = 1 & : & \text{two parallel lines} \\ c=4: & 16\tilde{x}_1^2 - 4\tilde{x}_2^2 = 1 & : & \text{a hyporbala} \end{array}$$

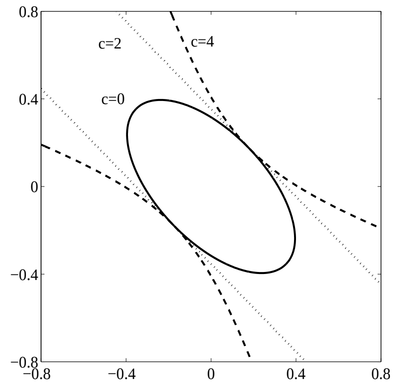


Figure 8.1: Conics in Example 8.7

The loci of the points that satisfy the given equation for each case are shown in Figure 8.1. Note that the transformation in (8.2) corresponds to a counter-clock-wise rotation of the coordinate axes by 45°.

A three-dimensional version of (8.10) defines a **quadric surface** in the $x_1x_2x_3$ space. As in the two-dimensional case, a change of coordinate system as $\mathbf{x} = P\tilde{\mathbf{x}}$ transforms the equation of the quadric into

$$\tilde{\mathbf{x}}^t D\tilde{\mathbf{x}} = \lambda_1\tilde{x}_1^2 + \lambda_2\tilde{x}_2^2 + \lambda_3\tilde{x}_3^2 = 1$$

where, without loss of generality, we may assume $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Then we have the following distinct cases:

a) $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ (S positive definite): The quadric equation can be expressed as

$$\frac{\tilde{x}_1^2}{a_1^2} + \frac{\tilde{x}_2^2}{a_2^2} + \frac{\tilde{x}_3^2}{a_3^2} = 1$$

which represents an ellipsoid with axes of length a_1, a_2 and a_3.

b) $\lambda_1 \geq \lambda_2 > \lambda_3 = 0$ (S positive semi-definite): Now the equation becomes

$$\frac{\tilde{x}_1^2}{a_1^2} + \frac{\tilde{x}_2^2}{a_2^2} = 1$$

which represents an elliptic cylinder.

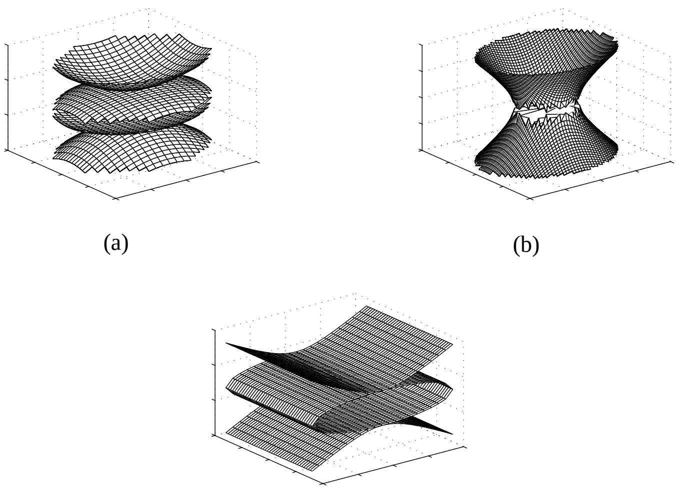
c) $\lambda_1 \geq \lambda_2 > 0 > \lambda_3$ (S indefinite): We have a hyperboloid of one sheet described as

$$\frac{\tilde{x}_1^2}{a_1^2} + \frac{\tilde{x}_2^2}{a_2^2} - \frac{\tilde{x}_3^2}{a_3^2} = 1$$

d) $\lambda_1 > \lambda_2 = \lambda_3 = 0$ (S positive semi-definite): The equation becomes

$$\frac{\tilde{x}_1^2}{a_1^2} = 1$$

which represents two parallel planes.



(a) (b) (c)

Figure 8.2: Central quadric surfaces

e) $\lambda_1 > \lambda_2 = 0 > \lambda_3$ (S indefinite): We have a hyperbolic cylinder described as

$$\frac{\tilde{x}_1^2}{a_1^2} - \frac{\tilde{x}_3^2}{a_3^2} = 1$$

f) $\lambda_1 > 0 > \lambda_2 \geq \lambda_3$ (S indefinite): The equation

$$\frac{\tilde{x}_1^2}{a_1^2} - \frac{\tilde{x}_2^2}{a_2^2} - \frac{\tilde{x}_3^2}{a_3^2} = 1$$

represents a hyperboloid of two sheets.

g) $0 \geq \lambda_1 \geq \lambda_2 \geq \lambda_3$ (S negative definite or negative semi-definite): The equation is never satisfied.

The quadric surfaces corresponding to the above cases are illustrated in Figure 8.2, where (a) contains an ellipsoid and a hyperboloid of two sheets, (b) a hyperboloid of one sheet, and (c) an elliptic and a hyperbolic cylinder.

8.4 The Singular Value Decomposition

Many applications of linear algebra require knowledge of the rank of a matrix, construction of bases for its row and column spaces or their orthogonal complements, or computation of projections onto these subspaces. Such applications usually involve matrices whose elements are given only approximately (e.g., as a result of some measurement). In such cases, approximate answers, together with a measure of approximation, make more sense than exact answers. For example, determining whether a given matrix is close (according to a specified measure) to a matrix of defective rank may be more significant than computing the rank of the given matrix itself. In theory, the rank of a matrix can easily be computed using the Gaussian elimination. However, as we noted in Example 4.4, it is not reliable when the matrix has nearly linearly dependent rows (or columns). This may pose serious problems in practical situations.

Singular value decomposition (SVD) is a computationally reliable method of transforming a given matrix into a simple form, from which its rank, bases for its column and row spaces and their orthogonal complements and projections onto these subspaces can be computed easily. In the following we shall first prove the SVD Theorem and then study its uses.

8.4.1 Singular Value Decomposition Theorem

Theorem 8.3 (The Singular Value Decomposition) *Let $A \in \mathbf{C}^{m\times n}$. There exist unitary matrices $U \in \mathbf{C}^{m\times m}$ and $V \in \mathbf{C}^{n\times n}$ such that*

$$U^h AV = \Sigma = \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} \tag{8.11}$$

where

$$\Sigma_1 = diag\,[\,\sigma_1,\ \dots\ ,\sigma_k\,]$$

with $\sigma_1 \geq \cdots \geq \sigma_k > 0$ for some $k \leq m, n$.

Proof $A^hA \in \mathbf{C}^{n\times n}$ is Hermitian, and since

$$\mathbf{x}^h A^h A\mathbf{x} = (A\mathbf{x})^h(A\mathbf{x}) = \|\, A\mathbf{x}\,\|^2 \geq 0 \quad \text{for all } \mathbf{x}$$

it is at least positive semi-definite. Let the eigenvalues of A^hA be

$$\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_k^2 > 0 = \sigma_{k+1}^2 = \cdots = \sigma_n^2$$

and let

$$V = [\,V_1\ \ V_2\,]$$

be a unitary modal matrix of A^hA, where V_1 and V_2 consist of the orthonormal eigenvectors associated with the nonzero and zero eigenvalues. (If A^hA is positive definite then $k = n$, and $V = V_1$.) Then

$$A^hAV_1 = V_1\Sigma_1^2 \implies V_1^hA^hAV_1 = \Sigma_1^2$$

and

$$A^hAV_2 = O \implies V_2^hA^hAV_2 = O \implies AV_2 = O$$

Let

$$U_1 = AV_1\Sigma_1^{-1}$$

Since

$$U_1^h U_1 = \Sigma_1^{-1} V_1^h A^h A V_1 \Sigma_1^{-1} = \Sigma_1^{-1} \Sigma_1^2 \Sigma_1^{-1} = I$$

columns of U_1 are orthonormal. Choose U_2 such that

$$U = [\, U_1 \;\; U_2 \,]$$

is unitary. (Columns of U_2 complete the columns of U_1 to an orthonormal basis for $\mathbf{C}^{m\times 1}$.) Then

$$\begin{aligned} U^h A V &= \begin{bmatrix} U_1^h A V_1 & U_1^h A V_2 \\ U_2^h A V_1 & U_2^h A V_2 \end{bmatrix} = \begin{bmatrix} \Sigma_1^{-1} V_1^h A^h A V_1 & O \\ U_2^h U_1 \Sigma_1 & O \end{bmatrix} \\ &= \begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix} = \Sigma \end{aligned}$$

The non-negative scalars $\sigma_i, i = 1, \ldots, n$, are called the **singular values** of A, and the columns of V and U are called the right and left **singular vectors** of A, respectively. Although the proof of Theorem 8.3 provides a constructive method, in practice, the singular value decomposition of A is obtained by means of a different algorithm which involves unitary transformations that do not require computation of the eigenvalues and eigenvectors of $A^h A$. The MATLAB command `[U,S,V]=svd(A)` that computes the singular value decomposition of A uses such an algoritm.

The following results are immediate consequences of Theorem 8.3:

a) $k = r$, that is, the number of nonzero singular values is the rank of A.

b) $A \in \mathbf{C}^{n\times n}$ is nonsingular if and only if all its singular values are positive.

c) The right singular vectors of A are the eigenvectors of $A^h A$.

d) The left singular vectors of A are the eigenvectors of AA^h.

e) $A^h A$ and AA^h have the same nonzero eigenvalues, $\sigma_1, \ldots, \sigma_k$.

f) If $A = U\Sigma V^h$ then $A^h = V\Sigma^h U^h$. Thus A and A^h have the same nonzero singular values.

g) If $A \in \mathbf{C}^{n\times n}$ is Hermitian with (real) eigenvalues $\lambda_i, i = 1, \ldots, n$, then its singular values are $\sigma_i = |\,\lambda_i\,|, i = 1, \ldots, n$.

(a) follows from (8.11) on noting that U and V are nonsingular, and (b) is a direct consequence of (a). (c) is the definition used in the proof of Theorem 8.3. (d) follows from

$$AA^h U = U\Sigma V^h V \Sigma^h U^h U = U\Sigma\Sigma^h$$

(e) and (f) are obvious. Finally, (g) is a result of the fact that if A is Hermitian with eigenvalues λ_i then $A^h A = A^2$ has the eigenvalues λ_i^2, so that singular values of A are $\sigma_i = \sqrt{\lambda_i^2} = |\,\lambda_i\,|$.

The following corollary of Theorem 8.3 characterizes various subspaces associated with a matrix. The proof is left to the reader (see Exercise 8.20).

Corollary 8.3.1 *Let $A \in \mathbf{C}^{m\times n}$ have the singular value decomposition*

$$A = U\Sigma V^h = [\,U_1 \;\; U_2\,]\begin{bmatrix} \Sigma_1 & O \\ O & O \end{bmatrix}\begin{bmatrix} V_1^h \\ V_2^h \end{bmatrix} = U_1\Sigma_1 V_1^h$$

Then

a) $im\,(U_1) = im\,(A)$

b) $im\,(U_2) = ker\,(A^h)$

c) $im\,(V_1) = im\,(A^h)$

d) $im\,(V_2) = ker\,(A)$

Thus $U_1U_1^h$ is the orthogonal projection matrix on $\mathrm{im}\,(A)$, $U_2U_2^h$ is the orthogonal projection matrix on $\ker(A^h)$, etc.

As a consequence of Corollary 8.3.1 we also have

$$\mathbf{C}^{m\times 1} = \mathrm{im}\,(U_1) \overset{\perp}{\oplus} \mathrm{im}\,(U_2) = \mathrm{im}\,(A) \overset{\perp}{\oplus} \ker\,(A^h)$$

and the dual relation

$$\mathbf{C}^{n\times 1} = \mathrm{im}\,(V_1) \overset{\perp}{\oplus} \mathrm{im}\,(V_2) = \mathrm{im}\,(A^h) \overset{\perp}{\oplus} \ker\,(A)$$

where the symbol $\overset{\perp}{\oplus}$ denotes a direct sum decomposition into orthogonal subspaces.

Example 8.8

Let us find the singular value decomposition of

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

Eigenvalues of

$$A^tA = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$$

are $\lambda_1 = 6$ and $\lambda_2 = 4$. Hence

$$\sigma_1 = \sqrt{6}\,, \quad \sigma_2 = 2$$

An orthogonal modal matrix of A^tA can be found as

$$V = V_1 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

Letting

$$U_1 = AV_1\Sigma_1^{-1} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \end{bmatrix}$$

and completing the columns of U_1 to an orthonormal basis for $\mathbf{R}^{3\times 1}$, the singular value decomposition of A is obtained as

$$\begin{aligned} A &= U\Sigma V^t \\ &= \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \end{aligned}$$

The reader can verify that

$$AA^t = \begin{bmatrix} 4 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 4 \end{bmatrix}$$

has the eigenvalues $\lambda_1 = 6, \lambda_2 = 4, \lambda_3 = 0$, and that columns of U are orthonormal eigenvectors of AA^t.

Finally, from

$$A \xrightarrow{\text{e.c.o.}} \begin{bmatrix} 2 & -2 \\ 2 & 0 \\ 2 & 2 \end{bmatrix} \xrightarrow{\text{e.c.o.}} U_1$$

we observe that $\text{im}\,(A) = \text{im}\,(U_1)$, verifying the result of Corollary 8.3.1(a). Other results of the corollary can be verified similarly.

8.4.2 The Least-Squares Problem and The Pseudoinverse

Corollary 8.1 is especially useful in solving least-squares problems. Recall that the linear system

$$A\mathbf{x} = \mathbf{y} \tag{8.12}$$

has a solution if and only if $\mathbf{y} \in \text{im}\,(A)$. If not, we look for solution(s) of the consistent system

$$A\mathbf{x} = \mathbf{y}_A \tag{8.13}$$

where $\mathbf{y}_A$ is the orthogonal projection of $\mathbf{y}$ on $\text{im}\,(A)$. Such solutions are called least-square solutions as they minimize the squared error $\| A\mathbf{x} - \mathbf{y} \|_2^2$.

Let A have the singular value decomposition

$$A = U\Sigma V^h = U_1\Sigma_1 V_1^h$$

Then the orthogonal projection of $\mathbf{y}$ on $\text{im}\,(A)$ is

$$\mathbf{y}_A = U_1 U_1^h \mathbf{y}$$

and (8.13) becomes

$$U_1\Sigma_1 V_1^h \mathbf{x} = U_1 U_1^h \mathbf{y} \tag{8.14}$$

It is left to the reader (Exercise 8.22) to show that the above $m\times n$ system is equivalent to the $r \times n$ system

$$V_1^h \mathbf{x} = \Sigma_1^{-1} U_1^h \mathbf{y} \tag{8.15}$$

In other words, the least-squares solutions of (8.12) are precisely the solutions of (8.15). In particular,

$$\mathbf{x} = \boldsymbol{\phi}_{LS} = V_1 \Sigma_1^{-1} U_1^h \mathbf{y} = A^\dagger \mathbf{y} \tag{8.16}$$

is a least-squares solution with minimum 2-norm (Exercise 8.22). Note that $\boldsymbol{\phi}_{LS}$ can be written in open form as

$$\boldsymbol{\phi}_{LS} = \frac{\mathbf{u}_1^h \mathbf{y}}{\sigma_1} \mathbf{v}_1 + \cdots + \frac{\mathbf{u}_r^h \mathbf{y}}{\sigma_r} \mathbf{v}_r \tag{8.17}$$

Consider the matrix

$$A^\dagger = V_1 \Sigma_1^{-1} U_1^h \tag{8.18}$$

that appears in (8.16). Since

$$\begin{aligned} AA^\dagger A &= U_1 \Sigma_1 V_1^h V_1 \Sigma_1^{-1} U_1^h U_1 \Sigma_1 V_1^h &= U_1 \Sigma_1 V_1^h &= A \\ A^\dagger A A^\dagger &= V_1 \Sigma_1^{-1} U_1^h U_1 \Sigma_1 V_1^h V_1 \Sigma_1^{-1} U_1^h &= V_1 \Sigma_1^{-1} U_1^h &= A^\dagger \end{aligned}$$

it follows that $A^\dagger$ is a generalized inverse of A, called the **Moore-Penrose generalized inverse** or the **pseudoinverse** of A. An interesting property of $A^\dagger$ is that since

$$\begin{aligned} A^\dagger A &= V_1 \Sigma_1^{-1} U_1^h U_1 \Sigma_1 V_1^h &= V_1 V_1^h \\ AA^\dagger &= U_1 \Sigma_1 V_1^h V_1 \Sigma_1^{-1} U_1^h &= U_1 U_1^h \end{aligned}$$

$A^\dagger A$ and $AA^\dagger$ are both Hermitian.[1] Moreover, $A^\dagger$ reduces to a left inverse when $r(A) = n$ (in which case $V_1 V_1^h = I_n$), to a right inverse when $r(A) = m$ (in which case $U_1 U_1^h = I_m$), and to A^{-1} when A is square and nonsingular.

Example 8.9

Consider a linear system with

$$A = \begin{bmatrix} 5 & 0 & 5 \\ 1 & 1 & 2 \\ 0 & 5 & 5 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 4 \\ -15 \\ 2 \end{bmatrix}$$

Following the procedure of Section 7.4.2, the orthogonal projection of $\mathbf{y}$ on $\mathrm{im}\,(A)$ is computed as

$$R = \begin{bmatrix} 5 & 0 \\ 1 & 1 \\ 0 & 5 \end{bmatrix}, \quad \mathbf{y}_A = R(R^t R)^{-1} R^t = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

The general solution of $A\mathbf{x} = \mathbf{y}_A$ is then obtained as

$$\mathbf{x} = \begin{bmatrix} 0.2 \\ -0.2 \\ 0.0 \end{bmatrix} + c \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

[1] In fact, $\hat{A}_G = A^\dagger$ is the unique generalized inverse such that $\hat{A}_G A$ and $A\hat{A}_G$ are both Hermitian. We will not prove this fact.

which characterizes all least-squares solutions. Among these least-squares solutions

$$\mathbf{x} = \boldsymbol{\phi}_{LS} = \begin{bmatrix} 0.2 \\ -0.2 \\ 0.0 \end{bmatrix}$$

(corresponding to $c = 0$) has the minimum 2-norm.

The singular value decomposition of A produces

$$\Sigma = \mathrm{diag}\,[\,9, 5, 0\,]$$

$$U = \frac{1}{3\sqrt{6}} \begin{bmatrix} 5 & 3\sqrt{3} & \sqrt{2} \\ 2 & 0 & -5\sqrt{2} \\ 5 & -3\sqrt{3} & \sqrt{2} \end{bmatrix}, \quad V = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & \sqrt{3} & \sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \\ 2 & 0 & -\sqrt{2} \end{bmatrix}$$

Then the pseudoinverse of A is obtained as

$$\begin{aligned} A^{\dagger} &= \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & \sqrt{3} \\ 1 & -\sqrt{3} \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 5 \end{bmatrix}^{-1} \frac{1}{3\sqrt{6}} \begin{bmatrix} 5 & 2 & 5 \\ 3\sqrt{3} & 0 & -3\sqrt{3} \end{bmatrix} \\ &= \frac{1}{810} \begin{bmatrix} 106 & 10 & -56 \\ -56 & 10 & 106 \\ 50 & 20 & 50 \end{bmatrix} \end{aligned}$$

and the expression in (8.16) gives the same minimum norm least-squares solution as obtained above. (Alternatively, (8.17) produces the same solution.)

The results can be checked using the MATLAB commands `[U,S,V]=svd(A)`, which produces the singular value decomposition of A, and `API=pinv(A)`, which computes the pseudoinverse of A. The reader is also urged to verify that $A^{\dagger}A$ and $AA^{\dagger}$ are both symmetric.

8.4.3 The SVD and Matrix Norms

Recall that

$$\|\,A\,\|_p = \max_{\|\,\mathbf{x}\,\|_p=1} \{\,\|\,A\mathbf{x}\,\|_p\,\}$$

is the matrix norm subordinate to the p-vector norm. Since

$$\|\,A\mathbf{x}\,\|_2^2 = \mathbf{x}^h A^h A\mathbf{x} \le \lambda_{\max}(A^h A)\,\|\,\mathbf{x}\,\|^2$$

with equality holding for the eigenvector corresponding to $\lambda_{\max}(A^h A)$, we find that

$$\|\,A\,\|_2 = \sqrt{\lambda_{\max}(A^h A)} = \sigma_1 \tag{8.19}$$

This is a significant result, which states that the matrix norm subordinate to the Euclidean vector norm is its largest singular value.

An equally significant result, which is left to the reader to prove (Exercise 8.25), is that

$$\|\,A\,\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2} \tag{8.20}$$

To appreciate the significance of (8.20), let

$$A' = U\Sigma' V^h$$

where

$$\Sigma' = \begin{bmatrix} \Sigma_1' & O \\ O & O \end{bmatrix}$$

with $\Sigma_1' = \operatorname{diag}[\,\sigma_1, \ldots, \sigma_q\,]$ and $q < r$. Then $r(A') = q$ and

$$\|\, A - A' \,\|_F = \sqrt{\sigma_{q+1}^2 + \cdots + \sigma_r^2} \tag{8.21}$$

Moreover, if B is any other matrix with $r(B) = q$ then

$$\|\, A - B \,\|_F \geq \|\, A - A' \,\|_F$$

Thus A' is the best rank-q approximation to A in Frobenius norm.[2]

Example 8.10

Consider the following matrix generated by the MATLAB command `A=rand(3,3)`.

$$A = \begin{bmatrix} 0.4103 & 0.3529 & 0.1389 \\ 0.8936 & 0.8132 & 0.2028 \\ 0.0579 & 0.0099 & 0.1987 \end{bmatrix}$$

The command `[U,S,V]=svd(A)` produces

$$S = \begin{bmatrix} 1.3485 & 0 & 0 \\ 0 & 0.1941 & 0 \\ 0 & 0 & 0.0063 \end{bmatrix}$$

in addition to U and V that are not shown here. Since all singular values of A are nonzero, we conclude that A is nonsingular, that is, it has rank $r(A) = 3$. However, the third singular value is quite small compared to the others, which suggests that A is close to a rank-2 (i.e., singular) matrix.

Let

$$Q = \begin{bmatrix} 1.3485 & 0 & 0 \\ 0 & 0.1941 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$B = UQV^t = \begin{bmatrix} 0.4065 & 0.3568 & 0.1398 \\ 0.8953 & 0.8114 & 0.2023 \\ 0.0589 & 0.0088 & 0.1985 \end{bmatrix}$$

Then B has rank $r(B) = 2$ (which can be verified by MATLAB), and is the closest rank-2 matrix to A in Frobenius norm. The command `norm(A-B,'fro')` computes the Frobenius norm of the difference of A and B as

$$\|\, A - B \,\|_F = 0.0063$$

as expected.

[2] The proof of this result is beyond the scope of this book.

Let A be a nonsingular matrix of order n having a singular value decomposition

$$A = U\Sigma V^h$$

where $\Sigma = \mathrm{diag}\,[\,\sigma_1, \ldots, \sigma_n\,]$ with $\sigma_1 \geq \cdots \geq \sigma_n > 0$. Then

$$A^{-1} = V\Sigma^{-1}U^h \tag{8.22}$$

which shows that A^{-1} has the singular values $\sigma_n^{-1} \geq \cdots \geq \sigma_1^{-1} > 0$.[3] Thus

$$\|\, A^{-1} \,\|_2 = \sigma_n^{-1}$$

The ratio

$$\mu = \frac{\sigma_1}{\sigma_n} = \|\, A \,\|_2 \,\|\, A^{-1} \,\|_2 \geq 1$$

is called the **condition number** of A, and is a measure of linear independence of the columns of A. The larger the condition number of a matrix, the closer it is to being singular.

Example 8.11

The matrix

$$A = \begin{bmatrix} 0.2676 & 0.5111 & 0.7627 \\ 0.6467 & 0.6931 & 0.5241 \\ 0.7371 & 0.6137 & 0.1690 \end{bmatrix}$$

has the singular values

$$\sigma_1 = 1.6564\,, \quad \sigma_2 = 0.5411\,, \quad \sigma_3 = 0.0001$$

as calculated and displayed up to the fourth decimal digit by MATLAB.

The wide separation of the singular values indicate that A is nearly singular. Indeed, the condition number of A, computed by the MATLAB command `cond(A)` as [4]

$$\mu = 1.6305e + 004$$

implies that A is badly conditioned, and that A^{-1} has large element values of the order of 10^4. The reader can verify this observation by computing A^{-1} using MATLAB.

8.5 Exercises

1. (a) Verify that the following matrices are unitary.
   (b) Find a unitary modal matrix for each of these matrices and diagonalize them.
   (c) Use MATLAB command `eig` to verify your results in part (b).

$$A = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

[3]The expression in (8.22) is similar to the singular value decomposition of A^{-1} except that its singular values are in ascending order.

[4]The difference between μ and σ_1/σ_3 is due to the limited number of digits used to display the values.

2. Prove that the product of two unitary matrices is unitary. Verify this result for the matrices A an B in Exercise 8.1.

3. In the xyz space a counterclockwise rotation about the z axis by an angle of θ_z is represented by the matrix

$$R_z = \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   (a) Determine the structure of the rotation matrices R_y and R_x about the y and x axes.

   (b) Show that R_x, R_y and R_z are orthogonal.

   (c) Characterize an invariant subspace for each of R_x, R_y and R_z.

4. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}_{n\times n}$$

   (a) Show that A is orthogonal.

   (b) Find the eigenvalues and eigenvectors of A.

   (c) Find a unitary modal matrix and the Jordan form of A.

5. Use MATLAB command `eig` to verify your answer to Exercise 8.4 for $n = 2, 3, 4$.

6. Let $\mathbf{b}$ be a unit vector. Show that $A = I - \mathbf{b}\mathbf{b}^t$ is a symmetric projection matrix.

7. Prove the **Schur's theorem**: *Let A be an $n\times n$ complex matrix. There exists a unitary matrix U such that U^hAU is an upper triangular matrix with diagonal elements being the eigenvalues of A.* Hint: Refer to Exercise 5.13.

8. $A \in \mathcal{C}^{n\times n}$ is said to be an involution if $A^2 = I$. Show that any two of the following imply the third.

   (a) A is Hermitian

   (b) A is unitary

   (c) A is an involution.

9. Verify the result of Exercise 8.8 for the matrices B and C in Exercise 8.1.

10. Prove that a Hermitian matrix A has a unitary modal matrix, and thus complete the proof of Theorem 8.2. Hint: Let $\mathbf{v}_1$ be a unit eigenvector of $A = A_1$ associated with some eigenvalue λ_1. Choose V_1 such that $P_1 = [\,\mathbf{v}_1 \;\; V_1\,]$ is unitary, and consider $P_1^hAP_1$.

11. (a) Show that eigenvectors of a unitary matrix associated with distinct eigenvalues are orthogonal.

    (b) Show that eigenvectors of a Hermitian matrix associated with distinct eigenvalues are orthogonal.

12. Show that if λ is an eigenvalue of the Hermitian matrix $H = S + iK$ then it is also an eigenvalue of $\tilde{H}$ in (8.9) and vice versa. Hint: Let $\mathbf{v} = \mathbf{u} + i\mathbf{w}$ be an eigenvector of H associated with λ and consider the real and imaginary parts of the expression

$$(\lambda I - H)\mathbf{v} = \mathbf{0}$$

13. Let S be an $n \times n$ real symmetric matrix with an orthogonal modal matrix Q and the diagonal Jordan form D. Find a modal matrix and the Jordan form of

$$H = \begin{bmatrix} 0 & jS \\ -jS & 0 \end{bmatrix}$$

in terms of Q and D. Hint: Let

$$\tilde{H} = \begin{bmatrix} Q^t & O \\ O & Q^t \end{bmatrix} \begin{bmatrix} 0 & jS \\ -jS & 0 \end{bmatrix} \begin{bmatrix} Q & O \\ O & Q \end{bmatrix} = \begin{bmatrix} 0 & jD \\ -jD & 0 \end{bmatrix}$$

and find eigenvalues and eigenvectors of $\tilde{H}$.

14. An $n \times n$ complex matrix A is said to be **normal** if it satisfies $A^h A = AA^h$. Clearly, unitary and Hermitian matrices are normal.

    (a) Show that a normal triangular matrix must be diagonal.

    (b) Prove that A can be diagonalized by a unitary similarity transformation if and only if it is normal. Hint: To prove sufficiency, use the Schur's theorem and the result of part (a).

15. Verify the result of Exercise 8.14 for the matrix

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

which is neither unitary nor Hermitian.

16. Investigate the sign properties of the following quadratic forms.

    (a) $q(x, y) = 2x^2 + 8xy + 2y^2$

    (b) $q(x_1, x_2, x_3) = 2x_1^2 + x_2^2 + 7x_3^2 - 2x_1x_2 + 2x_1x_3 - 4x_2x_3$

    (c) $q(z_1, z_2) = 2|z_1|^2 + |z_2|^2 + 2\,\mathrm{Im}\,(z_1 z_2^*)$

    (d) $q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$

17. Let $A \in \mathbf{C}^{m \times n}$. Prove the following.

    (a) $A^h A$ and AA^h are non-negative-definite.

    (b) $A^h A$ is positive definite if and only if $r(A) = n$.

    (c) AA^h is positive definite if and only if $r(A) = m$.

18. (a) Show that

$$Q = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

    is positive definite.

    (b) Show that $\langle \mathbf{x} \,|\, \mathbf{y} \rangle_Q = \mathbf{x}^t Q \mathbf{y}$ is an inner product in $\mathbf{R}^{3 \times 1}$.

    (c) Apply GSOP to $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ to generate an orthogonal basis for $\mathbf{R}^{3 \times 1}$, where orthogonality is defined with respect to $\langle \cdot \,|\, \cdot \rangle_Q$.

(d) Find the orthogonal projection of $\mathbf{e}_3$ on span $(\mathbf{e}_1, \mathbf{e}_2)$.

19. Obtain singular value decompositions of

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

(a) by hand,

(b) by using MATLAB.

20. Prove Corollary 8.1. Hint: $\mathbf{C}^{m\times 1} = \operatorname{im}(U_1) \overset{\perp}{\oplus} \operatorname{im}(U_2)$ and $\mathbf{C}^{n\times 1} = \operatorname{im}(V_1) \overset{\perp}{\oplus} \operatorname{im}(V_2)$.

21. Let

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

Obtain the singular value decompositions of A and A^{100}.

22. (a) Show that the systems in (8.14) and (8.15) have the same solution(s).

(b) Let $\mathbf{x} = \boldsymbol{\phi}$ be any solution of a consistent linear system $A\mathbf{x} = \mathbf{y}$, and let $\boldsymbol{\phi}_0$ be the orthogonal projection of $\boldsymbol{\phi}$ on $\ker(A)$. Prove that $\mathbf{x} = \boldsymbol{\phi} - \boldsymbol{\phi}_0$ is the unique minimum 2-norm solution of $A\mathbf{x} = \mathbf{y}$.

(c) Prove that $\mathbf{x} = \boldsymbol{\phi}_{LS} = V_1\Sigma_1^{-1}U_1^h\mathbf{y}$ is the minimum 2-norm least-squares solution of $A\mathbf{x} = \mathbf{y}$.

23. (a) Find all least-squares solutions of $A\mathbf{x} = \mathbf{y}$, where

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & 3 & 2 \\ 0 & 1 & 1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 9 \\ -9 \\ 9 \\ -9 \end{bmatrix}$$

(b) Among all least-squares solutions find the one with minimum Euclidean norm.

(c) Use the MATLAB command `x=pinv(C)*y` to verify your answer in part (b).

24. Show that the minimum 2-norm least squares solution to $C\mathbf{x} = \mathbf{y}$, where C is as in Exercise 8.19 and $\mathbf{y} = \operatorname{col}[\,y_1, y_2, y_3, y_4\,]$ is given by

$$\mathbf{x} = \boldsymbol{\phi}_{LS} = \frac{1}{6}\begin{bmatrix} 2y_1 - y_2 + 2y_3 - y_4 \\ y_1 + y_2 + y_3 + y_4 \\ -y_1 + 2y_2 - y_3 + 2y_4 \end{bmatrix}$$

Verify your result by computing the minimum norm least-squares solution of $C\mathbf{x} = \mathbf{y}$ by using the MATLAB command `x=pinv(C)*y` for several randomly generated $\mathbf{y}$ vectors.

25. (a) Prove (8.20). Hint: From

$$A = U_1\Sigma_1V_1^h = \sum_{k=1}^{n} \sigma_k \mathbf{u}_k \mathbf{v}_k^h$$

it follows that

$$a_{ij} = \sum_{k=1}^{r} \sigma_k u_{ik} v_{jk}^*$$

where

$$\mathbf{u}_k = \begin{bmatrix} u_{1k} \\ \vdots \\ u_{mk} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_k = \begin{bmatrix} v_{1k} \\ \vdots \\ v_{nk} \end{bmatrix}$$

are the kth left and right singular vectors of A. Manipulate the expression

$$\| A \|_F = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} a_{ij}^* = \sum_{i=1}^{m} \sum_{j=1}^{n} (\sum_{k=1}^{r} \sigma_k u_{ik} v_{jk}^*)(\sum_{l=1}^{r} \sigma_l u_{il} v_{jl}^*)$$

(b) Prove (8.21).

26. Generate a 4×3 random matrix D using the MATLAB command `rand`, and let $E = C + 0.001D$, where C is the matrix in Exercise 8.19. Using MATLAB

    (a) obtain the singular value decomposition of E,

    (b) compute a rank-2 matrix F which is closest to E in Frobenius norm,

    (c) compute $\| E - F \|_F$ and $\| E - C \|_F$.

27. Let

$$A = \begin{bmatrix} 0.5055 & 0.6412 & 0.8035 \\ 0.1693 & 0.0162 & 0.6978 \\ 0.5247 & 0.8369 & 0.4619 \end{bmatrix}$$

    Using MATLAB

    (a) Compute the condition number and the inverse of A.

    (b) Find the best rank-1 and rank-2 approximations A_1 and A_2 of A in Frobenius norm. Compute also $\| A - A_1 \|_F$ and $\| A - A_2 \|_F$, and comment on the results.

28. (Application) The most general expression of a conic in the $x_1 x_2$ plane is

$$s_{11} x_1^2 + 2 s_{12} x_1 x_2 + s_{22} x_2^2 + 2 r_1 x_1 + 2 r_2 x_2 = 1$$

    Let the equation be expressed in compact form as

$$\mathbf{x}^t S \mathbf{x} + 2 \mathbf{r}^t \mathbf{x} = 1$$

    If $S = O$ and $\mathbf{r} = \mathbf{0}$, then the solution set is empty. If $S = O$ and $\mathbf{r} \neq \mathbf{0}$, then the conic degenerates into a straight line.

    Suppose $S \neq O$. Let S have the eigenvalues $\lambda_1 \geq \lambda_2$ and an orthogonal modal matrix P such that $P^t S P = D = \operatorname{diag}[\, \lambda_1, \lambda_2 \,]$. Then a change of the coordinate system as

$$\mathbf{x} = P \tilde{\mathbf{x}} - \mathbf{x}_o \,, \quad \tilde{\mathbf{x}} = P^t (\mathbf{x} + \mathbf{x}_o) \tag{8.23}$$

    transforms the equation of the conic into

$$\tilde{\mathbf{x}}^t D \tilde{\mathbf{x}} + 2(\mathbf{r} - S\mathbf{x}_o)^t P \tilde{\mathbf{x}} = 1 + 2\mathbf{r}^t \mathbf{x}_o - \mathbf{x}_o^t S \mathbf{x}_o$$

    The transformation in (8.23) corresponds to a rotation of the coordinate axes, followed by a shift of the origin as illustrated in Figure 8.3. The purpose of shifting the origin of the coordinate system after the rotation is to eliminate, if posible, the linear term $2(\mathbf{r} - S\mathbf{x}_o)^t P \tilde{\mathbf{x}}$ so that the equation takes the form of that of a central conic.

    The following cases need to be considered:

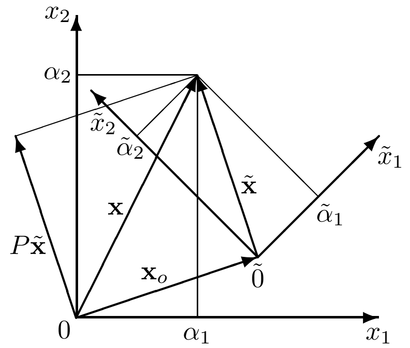


Figure 8.3: The coordinate transformation in (8.23)

(a) $\lambda_1 \geq \lambda_2 > 0$. In this case D is nonsingular. Show that, a choice of $\mathbf{x}_o = S^{-1}\mathbf{r}$, reduces the equation of the conic into

$$\frac{\tilde{x}_1^2}{a_1^2} + \frac{\tilde{x}_2^2}{a_2^2} = 1$$

which represents an ellipse in the $\tilde{x}_1\tilde{x}_2$ plane.

(b) $\lambda_1 > 0 = \lambda_2$. This case is more difficult to analyze than the previous one, because S is singular now. Show that a suitable choice of $\mathbf{x}_o$ reduces the equation of the conic into either

$$\tilde{x}_1^2 = c^2$$

which represents two parallel lines, or into

$$c\tilde{x}_1^2 + \tilde{x}_2 = 0$$

which represents a parabola.

(c) $\lambda_1 > 0 > \lambda_2$. As in case (a), choose $\mathbf{x}_o = S^{-1}\mathbf{r}$. Work out the details to show that, depending on the value of $1 + \mathbf{r}^t S^{-1}\mathbf{r}$, the equation either represents two intersecting straight lines or a hyperbola.

(d) $\lambda_1 = 0 > \lambda_2$. This case is similar to case (b). Show that the solution set may be empty, consist of a single point, two parallel straight lines, or a parabola.

(e) $0 > \lambda_1 \geq \lambda_2$. This case is similar to case (a). Show that in addition to being an ellipse, the solution set may also be empty or consist of a single point.

29. (a) Use MATLAB to plot the graph of the conic described by

$$x_1^2 + px_1x_2 + 2x_2^2 - 3x_2 - 6 = 0$$

in a rectangular region $-5 \leq x_1, x_2 \leq 5$ of the x_1x_2 plane for each of the values $p = -1$, $p = 2\sqrt{2}$ and $p = 4$.

(b) Transform the equation of the conic into one of the standard forms in Exercise 8.28 for each of the given values of the parameter p.