26. The transition matrix in Example 5 has the property that both its rows and its columns add up to 1. In general, a matrix A is said to be *doubly stochastic* if both A and A^T are stochastic. Let A be an $n \times n$ doubly stochastic matrix whose eigenvalues satisfy

$$\lambda_1 = 1 \quad \text{and} \quad |\lambda_j| < 1 \quad \text{for } j = 2, 3, \ldots, n$$

Show that if $\mathbf{e}$ is the vector in $\mathbb{R}^n$ whose entries are all equal to 1, then the Markov chain will converge to the steady-state vector $\mathbf{x} = \frac{1}{n}\mathbf{e}$ for any starting vector $\mathbf{x}_0$. Thus, for a doubly stochastic transition matrix, the steady-state vector will assign equal probabilities to all possible outcomes.

27. Let A be the PageRank transition matrix and let $\mathbf{x}_k$ be a vector in the Markov chain with starting probability vector $\mathbf{x}_0$. Since n is very large, the direct multiplication $\mathbf{x}_{k+1} = A\mathbf{x}_k$ is computationally intensive. However, the computation can be simplified dramatically if we take advantage of the structured components of A given in equation (5). Because M is sparse, the multiplication $\mathbf{w}_k = M\mathbf{x}_k$ is computationally much simpler. Show that if we set

$$\mathbf{b} = \frac{1-p}{n}\mathbf{e}$$

then

$$E\mathbf{x}_k = \mathbf{e} \quad \text{and} \quad \mathbf{x}_{k+1} = p\mathbf{w}_k + \mathbf{b}$$

where M, E, $\mathbf{e}$, and p are as defined in equation (5).

28. Use the definition of the matrix exponential to compute e^A for each of the following matrices:

(a) $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$ **(b)** $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

29. Compute e^A for each of the following matrices:

(a) $A = \begin{pmatrix} -2 & -1 \\ 6 & 3 \end{pmatrix}$ **(b)** $A = \begin{pmatrix} 3 & 4 \\ -2 & -3 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$

30. In each of the following, solve the initial value problem $\mathbf{Y}' = A\mathbf{Y}$, $\mathbf{Y}(0) = \mathbf{Y}_0$, by computing $e^{tA}\mathbf{Y}_0$:

(a) $A = \begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 2 & 3 \\ -1 & -2 \end{pmatrix}, \quad \mathbf{Y}_0 = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$

(c) $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}, \quad \mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

(d) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & -1 \end{pmatrix}, \quad \mathbf{Y}_0 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

31. Let λ be an eigenvalue of an $n \times n$ matrix A and let $\mathbf{x}$ be an eigenvector belonging to λ. Show that e^λ is an eigenvalue of e^A and $\mathbf{x}$ is an eigenvector of e^A belonging to e^λ.

32. Show that e^A is nonsingular for any diagonalizable matrix A.

33. Let A be a diagonalizable matrix with characteristic polynomial

$$p(\lambda) = a_1\lambda^n + a_2\lambda^{n-1} + \cdots + a_{n+1}$$

(a) Show that if D is a diagonal matrix whose diagonal entries are the eigenvalues of A, then

$$p(D) = a_1 D^n + a_2 D^{n-1} + \cdots + a_{n+1} I = O$$

(b) Show that $p(A) = O$.

(c) Show that if $a_{n+1} \neq 0$, then A is nonsingular and $A^{-1} = q(A)$ for some polynomial q of degree less than n.

6.4 Hermitian Matrices

Let $\mathbb{C}^n$ denote the vector space of all n-tuples of complex numbers. The set $\mathbb{C}$ of all complex numbers will be taken as our field of scalars. We have already seen that a matrix A with real entries may have complex eigenvalues and eigenvectors. In this section, we study matrices with complex entries and look at the complex analogues of symmetric and orthogonal matrices.

Complex Inner Products

If $\alpha = a + bi$ is a complex scalar, the length of α is given by

$$|\alpha| = \sqrt{\bar{\alpha}\alpha} = \sqrt{a^2 + b^2}$$

The length of a vector $\mathbf{z} = (z_1, z_2, \ldots, z_n)^T$ in $\mathbb{C}^n$ is given by

$$\begin{aligned}
\|\mathbf{z}\| &= \left(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2\right)^{1/2} \\
&= (\bar{z}_1 z_1 + \bar{z}_2 z_2 + \cdots + \bar{z}_n z_n)^{1/2} \\
&= \left(\bar{\mathbf{z}}^T \mathbf{z}\right)^{1/2}
\end{aligned}$$

As a notational convenience, we write $\mathbf{z}^H$ for the transpose of $\bar{\mathbf{z}}$. Thus

$$\bar{\mathbf{z}}^T = \mathbf{z}^H \quad \text{and} \quad \|\mathbf{z}\| = (\mathbf{z}^H \mathbf{z})^{1/2}$$

Definition

Let V be a vector space over the complex numbers. An **inner product** on V is an operation that assigns, to each pair of vectors $\mathbf{z}$ and $\mathbf{w}$ in V, a complex number $\langle \mathbf{z}, \mathbf{w} \rangle$ satisfying the following conditions:

I. $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$, with equality if and only if $\mathbf{z} = \mathbf{0}$.
II. $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ for all $\mathbf{z}$ and $\mathbf{w}$ in V.
III. $\langle \alpha\mathbf{z} + \beta\mathbf{w}, \mathbf{u} \rangle = \alpha\langle \mathbf{z}, \mathbf{u} \rangle + \beta\langle \mathbf{w}, \mathbf{u} \rangle$.

Note that for a complex inner product space, $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$, rather than $\langle \mathbf{w}, \mathbf{z} \rangle$. If we make the proper modifications to allow for this difference, the theorems on real inner product spaces in Chapter 5, Section 5, will all be valid for complex inner product spaces. In particular, let us recall Theorem 5.5.2: If $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for a real inner product space V and

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{u}_i$$

then

$$c_i = \langle \mathbf{u}_i, \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{u}_i \rangle \quad \text{and} \quad \|\mathbf{x}\|^2 = \sum_{i=1}^{n} c_i^2$$

In the case of a complex inner product space, if $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is an orthonormal basis and

$$\mathbf{z} = \sum_{i=1}^{n} c_i \mathbf{w}_i$$

then

$$c_i = \langle \mathbf{z}, \mathbf{w}_i \rangle, \bar{c}_i = \langle \mathbf{w}_i, \mathbf{z} \rangle \quad \text{and} \quad \|\mathbf{z}\|^2 = \sum_{i=1}^{n} c_i \bar{c}_i$$

We can define an inner product on $\mathbb{C}^n$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z} \tag{1}$$

for all $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^n$. We leave it to the reader to verify that (1) actually does define an inner product on $\mathbb{C}^n$. The complex inner product space $\mathbb{C}^n$ is similar to the real inner product space $\mathbb{R}^n$. The main difference is that in the complex case it is necessary to conjugate before transposing when taking an inner product.

$\mathbb{R}^n$	$\mathbb{C}^n$
$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^T\mathbf{x}$	$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H\mathbf{z}$
$\mathbf{x}^T\mathbf{y} = \mathbf{y}^T\mathbf{x}$	$\mathbf{z}^H\mathbf{w} = \overline{\mathbf{w}^H\mathbf{z}}$
$\|\mathbf{x}\|^2 = \mathbf{x}^T\mathbf{x}$	$\|\mathbf{z}\|^2 = \mathbf{z}^H\mathbf{z}$

EXAMPLE 1 If

$$\mathbf{z} = \begin{Bmatrix} 5+i \\ 1-3i \end{Bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{Bmatrix} 2+i \\ -2+3i \end{Bmatrix}$$

then

$$\begin{aligned} \mathbf{w}^H\mathbf{z} &= (2-i,\ -2-3i)\begin{Bmatrix} 5+i \\ 1-3i \end{Bmatrix} = (11-3i) + (-11+3i) = 0 \\ \mathbf{z}^H\mathbf{z} &= |5+i|^2 + |1-3i|^2 = 36 \\ \mathbf{w}^H\mathbf{w} &= |2+i|^2 + |-2+3i|^2 = 18 \end{aligned}$$

It follows that $\mathbf{z}$ and $\mathbf{w}$ are orthogonal and

$$\|\mathbf{z}\| = 6, \qquad \|\mathbf{w}\| = 3\sqrt{2}$$

Hermitian Matrices

Let $M = (m_{ij})$ be an $m \times n$ matrix with $m_{ij} = a_{ij} + ib_{ij}$ for each i and j. We may write M in the form

$$M = A + iB$$

where $A = (a_{ij})$ and $B = (b_{ij})$ have real entries. We define the conjugate of M by

$$\overline{M} = A - iB$$

Thus, $\overline{M}$ is the matrix formed by conjugating each of the entries of M. The transpose of $\overline{M}$ will be denoted by M^H. The vector space of all $m \times n$ matrices with complex entries is denoted by $\mathbb{C}^{m\times n}$. If A and B are elements of $\mathbb{C}^{m\times n}$ and $C \in \mathbb{C}^{n\times r}$, then the following rules are easily verified (see Exercise 9):

I. $(A^H)^H = A$
II. $(\alpha A + \beta B)^H = \overline{\alpha}A^H + \overline{\beta}B^H$
III. $(AC)^H = C^H A^H$

Definition A matrix M is said to be **Hermitian** if $M = M^H$.

EXAMPLE 2 The matrix

$$M = \begin{Bmatrix} 3 & 2-i \\ 2+i & 4 \end{Bmatrix}$$

is Hermitian, since

$$M^H = \begin{Bmatrix} \overline{3} & \overline{2-i} \\ \overline{2+i} & \overline{4} \end{Bmatrix}^T = \begin{Bmatrix} 3 & 2-i \\ 2+i & 4 \end{Bmatrix} = M$$

If M is a matrix with real entries, then $M^H = M^T$. In particular, if M is a real symmetric matrix, then M is Hermitian. Thus, we may view Hermitian matrices as the complex analogue of real symmetric matrices. Hermitian matrices have many nice properties, as we shall see in the next theorem.

Theorem 6.4.1 The eigenvalues of a Hermitian matrix are all real. Furthermore, eigenvectors belonging to distinct eigenvalues are orthogonal.

Proof Let A be a Hermitian matrix. Let λ be an eigenvalue of A and let $\mathbf{x}$ be an eigenvector belonging to λ. If $\alpha = \mathbf{x}^H A\mathbf{x}$, then

$$\overline{\alpha} = \alpha^H = (\mathbf{x}^H A\mathbf{x})^H = \mathbf{x}^H A\mathbf{x} = \alpha$$

Thus, α is real. It follows that

$$\alpha = \mathbf{x}^H A\mathbf{x} = \mathbf{x}^H \lambda\mathbf{x} = \lambda\|\mathbf{x}\|^2$$

and hence

$$\lambda = \frac{\alpha}{\|\mathbf{x}\|^2}$$

is real. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors belonging to distinct eigenvalues λ_1 and λ_2, respectively, then

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = \mathbf{x}_1^H A^H\mathbf{x}_2 = \mathbf{x}_1^H A\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

and

$$(A\mathbf{x}_1)^H\mathbf{x}_2 = (\mathbf{x}_2^H A\mathbf{x}_1)^H = (\lambda_1\mathbf{x}_2^H\mathbf{x}_1)^H = \lambda_1\mathbf{x}_1^H\mathbf{x}_2$$

Consequently,

$$\lambda_1\mathbf{x}_1^H\mathbf{x}_2 = \lambda_2\mathbf{x}_1^H\mathbf{x}_2$$

and since $\lambda_1 \neq \lambda_2$, it follows that

$$\langle\mathbf{x}_2, \mathbf{x}_1\rangle = \mathbf{x}_1^H\mathbf{x}_2 = 0$$

Definition An $n \times n$ matrix U is said to be **unitary** if its column vectors form an orthonormal set in $\mathbb{C}^n$.

Thus, U is unitary if and only if $U^H U = I$. If U is unitary, then, since the column vectors are orthonormal, U must have rank n. It follows that

$$U^{-1} = IU^{-1} = U^H U U^{-1} = U^H$$

A real unitary matrix is an orthogonal matrix.

Corollary 6.4.2 If the eigenvalues of a Hermitian matrix A are distinct, then there exists a unitary matrix U that diagonalizes A.

Proof Let $\mathbf{x}_i$ be an eigenvector belonging to λ_i for each eigenvalue λ_i of A. Let $\mathbf{u}_i = (1/\|\mathbf{x}_i\|)\mathbf{x}_i$. Thus, $\mathbf{u}_i$ is a unit eigenvector belonging to λ_i for each i. It follows from Theorem 6.4.1 that $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthonormal set in $\mathbb{C}^n$. Let U be the matrix whose ith column vector is $\mathbf{u}_i$ for each i; then U is unitary and U diagonalizes A. ■

EXAMPLE 3 Let

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

Find a unitary matrix U that diagonalizes A.

Solution

The eigenvalues of A are $\lambda_1 = 3$ and $\lambda_2 = 0$, with corresponding eigenvectors $\mathbf{x}_1 = (1-i, 1)^T$ and $\mathbf{x}_2 = (-1, 1+i)^T$. Let

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1 = \frac{1}{\sqrt{3}}(1-i, 1)^T$$

and

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2 = \frac{1}{\sqrt{3}}(-1, 1+i)^T$$

Thus

$$U = \frac{1}{\sqrt{3}}\begin{pmatrix} 1-i & -1 \\ 1 & 1+i \end{pmatrix}$$

and

$$\begin{aligned} U^H AU &= \frac{1}{3}\begin{pmatrix} 1+i & 1 \\ -1 & 1-i \end{pmatrix}\begin{pmatrix} 2 & 1-i \\ 1+i & 1 \end{pmatrix}\begin{pmatrix} 1-i & -1 \\ 1 & 1+i \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$ ■

Actually, Corollary 6.4.2 is valid even if the eigenvalues of A are not distinct. To show this, we will first prove the following theorem:

Theorem 6.4.3 Schur's Theorem

For each $n \times n$ matrix A, there exists a unitary matrix U such that $U^H AU$ is upper triangular.

Proof The proof is by induction on n. The result is obvious if $n = 1$. Assume that the hypothesis holds for $k \times k$ matrices, and let A be a $(k+1) \times (k+1)$ matrix. Let λ_1 be an eigenvalue of A, and let $\mathbf{w}_1$ be a unit eigenvector belonging to λ_1. Using the Gram–Schmidt process, construct $\mathbf{w}_2, \ldots, \mathbf{w}_{k+1}$ such that $\{\mathbf{w}_1, \ldots, \mathbf{w}_{k+1}\}$ is an orthonormal basis for $\mathbb{C}^{k+1}$. Let W be the matrix whose ith column vector is $\mathbf{w}_i$ for $i = 1, \ldots, k+1$. Then, by construction, W is unitary. The first column of $W^H AW$ will be $W^H A\mathbf{w}_1$.

$$W^H A\mathbf{w}_1 = \lambda_1 W^H \mathbf{w}_1 = \lambda_1 \mathbf{e}_1$$

Thus, $W^H AW$ is a matrix of the form

$$\left[\begin{array}{c|cccc} \lambda_1 & \times & \times & \cdots & \times \\ \hline 0 & & & & \\ \vdots & & & M & \\ 0 & & & & \end{array}\right]$$

where M is a $k \times k$ matrix. By the induction hypothesis, there exists a $k \times k$ unitary matrix V_1 such that $V_1^H M V_1 = T_1$, where T_1 is upper triangular. Let

$$V = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & V_1 & \\ 0 & & & \end{array}\right]$$

$$V^H W^H AWV = \left[\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & V_1^H M V_1 & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array}\right] = T$$

Let $U = WV$. The matrix U is unitary, since

$$U^H U = (WV)^H WV = V^H W^H WV = I$$

and $U^H AU = T$. ■

The factorization $A = UTU^H$ is often referred to as the *Schur decomposition* of A. In the case that A is Hermitian, the matrix T will be diagonal.

Theorem 6.4.4 Spectral Theorem

If A is Hermitian, then there exists a unitary matrix U that diagonalizes A.

Proof By Theorem 6.4.3, there is a unitary matrix U such that $U^H AU = T$, where T is upper triangular. Furthermore,

$$T^H = (U^H AU)^H = U^H A^H U = U^H AU = T$$

Therefore, T is Hermitian and consequently must be diagonal. ■

In the case of a real symmetric matrix, the diagonalizing matrix U will be an orthogonal matrix. The following example shows how to determine the matrix U. Later in this section we give a formal proof that all real symmetric matrices have orthogonal diagonalizing matrices.

EXAMPLE 4 Given

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

find an orthogonal matrix U that diagonalizes A.

Solution

The characteristic polynomial

$$p(\lambda) = -\lambda^3 + 3\lambda^2 + 9\lambda + 5 = (1+\lambda)^2(5-\lambda)$$

has roots $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 5$. Computing eigenvectors in the usual way, we see that $\mathbf{x}_1 = (1, 0, 1)^T$ and $\mathbf{x}_2 = (-2, 1, 0)^T$ form a basis for the eigenspace $N(A + I)$. We can apply the Gram–Schmidt process to obtain an orthonormal basis for the eigenspace corresponding to $\lambda_1 = \lambda_2 = -1$:

$$\begin{aligned} \mathbf{u}_1 &= \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1 = \frac{1}{\sqrt{2}}(1, 0, 1)^T \\ \mathbf{p} &= \left(\mathbf{x}_2^T\mathbf{u}_1\right)\mathbf{u}_1 = -\sqrt{2}\mathbf{u}_1 = (-1, 0, 1)^T \\ \mathbf{x}_2 - \mathbf{p} &= (-1, 1, 1)^T \\ \mathbf{u}_2 &= \frac{1}{\|\mathbf{x}_2 - \mathbf{p}\|}(\mathbf{x}_2 - \mathbf{p}) = \frac{1}{\sqrt{3}}(-1, 1, 1)^T \end{aligned}$$

The eigenspace corresponding to $\lambda_3 = 5$ is spanned by $\mathbf{x}_3 = (-1, -2, 1)^T$. Since $\mathbf{x}_3$ must be orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$ (Theorem 6.4.1), we need only normalize

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3\|}\mathbf{x}_3 = \frac{1}{\sqrt{6}}(-1, -2, 1)^T$$

Thus, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal set and

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

diagonalizes A. ■

It follows from Theorem 6.4.4 that each Hermitian matrix A can be factored into a product UDU^H, where U is unitary and D is diagonal. Since U diagonalizes A, it follows that the diagonal elements of D are the eigenvalues of A and the column vectors of U are eigenvectors of A. Thus, A cannot be defective. It has a complete set of eigenvectors that form an orthonormal basis for $\mathbb{C}^n$. This is, in a sense, the ideal situation. We have seen how to express a vector as a linear combination of orthonormal basis elements (Theorem 5.5.2), and the action of A on any linear combination of eigenvectors can easily be determined. Thus, if A has an orthonormal set of eigenvectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ and $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$, then

$$A\mathbf{x} = c_1\lambda_1\mathbf{u}_1 + \cdots + c_n\lambda_n\mathbf{u}_n$$

Furthermore,

$$c_i = \langle \mathbf{x}, \mathbf{u}_i \rangle = \mathbf{u}_i^H\mathbf{x}$$

or, equivalently, $\mathbf{c} = U^H\mathbf{x}$. Hence,

$$A\mathbf{x} = \lambda_1(\mathbf{u}_1^H\mathbf{x})\mathbf{u}_1 + \cdots + \lambda_n(\mathbf{u}_n^H\mathbf{x})\mathbf{u}_n$$

The Real Schur Decomposition

If A is a real $n \times n$ matrix, then it is possible to obtain a factorization that resembles the Schur decomposition of A, but involves only real matrices. In this case, $A = QTQ^T$, where Q is an orthogonal matrix and T is a real matrix of the form

$$T = \begin{Bmatrix} B_1 & \times & \cdots & \times \\ & B_2 & & \times \\ & O & \ddots & \\ & & & B_j \end{Bmatrix} \tag{2}$$

where the B_i's are either 1×1 or 2×2 matrices. Each 2×2 block will correspond to a pair of complex conjugate eigenvalues of A. The matrix T is referred to as the *real Schur form* of A. The proof that every real $n \times n$ matrix A has such a factorization depends on the property that, for each pair of complex conjugate eigenvalues of A, there is a two-dimensional subspace of $\mathbb{R}^n$ that is invariant under A.

Definition A subspace S of $\mathbb{R}^n$ is said to be **invariant** under a matrix A if, for each $\mathbf{x} \in S$, $A\mathbf{x} \in S$.

Lemma 6.4.5 Let A be a real $n \times n$ matrix with eigenvalue $\lambda_1 = a + bi$ (where a and b are real and $b \neq 0$), and let $\mathbf{z}_1 = \mathbf{x} + i\mathbf{y}$ (where $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$) be an eigenvector belonging to λ_1. If $S = \text{Span}(\mathbf{x}, \mathbf{y})$, then $\dim S = 2$ and S is invariant under A.

Proof Since λ is complex, $\mathbf{y}$ must be nonzero; otherwise we would have $A\mathbf{z} = A\mathbf{x}$ (a real vector) equal to $\lambda\mathbf{z} = \lambda\mathbf{x}$ (a complex vector). Since A is real, $\lambda_2 = a - bi$ is also an eigenvalue of A and $\mathbf{z}_2 = \mathbf{x} - i\mathbf{y}$ is an eigenvector belonging to λ_2. If there were a scalar c such that $\mathbf{x} = c\mathbf{y}$, then $\mathbf{z}_1$ and $\mathbf{z}_2$ would both be multiples of $\mathbf{y}$ and could not be independent. However, $\mathbf{z}_1$ and $\mathbf{z}_2$ belong to distinct eigenvalues, so they must be linearly independent. Therefore, $\mathbf{x}$ cannot be a multiple of $\mathbf{y}$ and hence $S = \text{Span}(\mathbf{x}, \mathbf{y})$ has dimension 2.

To show the invariance of S, note that since $A\mathbf{z}_1 = \lambda_1 \mathbf{z}_1$, the real and imaginary parts of both sides must agree. Thus,

$$\begin{aligned} A\mathbf{z}_1 &= A\mathbf{x} + iA\mathbf{y} \\ \lambda_1 \mathbf{z}_1 &= (a + bi)(\mathbf{x} + i\mathbf{y}) = (a\mathbf{x} - b\mathbf{y}) + i(b\mathbf{x} + a\mathbf{y}) \end{aligned}$$

and it follows that

$$A\mathbf{x} = a\mathbf{x} - b\mathbf{y} \quad \text{and} \quad A\mathbf{y} = b\mathbf{x} + a\mathbf{y}$$

If $\mathbf{w} = c_1\mathbf{x} + c_2\mathbf{y}$ is any vector in S, then

$$A\mathbf{w} = c_1 A\mathbf{x} + c_2 A\mathbf{y} = c_1(a\mathbf{x} - b\mathbf{y}) + c_2(b\mathbf{x} + a\mathbf{y}) = (c_1 a + c_2 b)\mathbf{x} + (c_2 a - c_1 b)\mathbf{y}$$

So $A\mathbf{w}$ is in S, and hence S is invariant under A. ■

Using this lemma, we can a prove version of Schur's theorem for matrices with real entries. As before, the proof will be by induction.

Theorem 6.4.6 The Real Schur Decomposition

If A is an $n \times n$ matrix with real entries, then A can be factored into a product QTQ^T, where Q is an orthogonal matrix and T is in Schur form (2).

Proof In the case that $n = 2$, if the eigenvalues of A are real, we can take $\mathbf{q}_1$ to be a unit eigenvector belonging to the first eigenvalue λ_1 and let $\mathbf{q}_2$ be any unit vector that is orthogonal to $\mathbf{q}_1$. If we set $Q = (\mathbf{q}_1, \mathbf{q}_2)$, then Q is an orthogonal matrix. If we set $T = Q^T A Q$, then the first column of T is

$$Q^T A \mathbf{q}_1 = \lambda_1 Q^T \mathbf{q}_1 = \lambda_1 \mathbf{e}_1$$

So T is upper triangular and $A = QTQ^T$. If the eigenvalues of A are complex, then we simply set $T = A$ and $Q = I$. So every 2×2 real matrix has a real Schur decomposition.

Now let A be a $k \times k$ matrix where $k \geq 3$ and assume that, for $2 \leq m < k$, every $m \times m$ real matrix has a Schur decomposition of the form (2). Let λ_1 be an eigenvalue of A. If λ_1 is real, let $\mathbf{q}_1$ be a unit eigenvector belonging to λ_1 and choose $\mathbf{q}_2, \mathbf{q}_3, \ldots, \mathbf{q}_n$ so that $Q_1 = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n)$ is an orthogonal matrix. As in the proof of Schur's theorem, it follows that the first column of $Q_1^T A Q_1$ will be $\lambda_1 \mathbf{e}_1$. In the case that λ_1 is complex, let $\mathbf{z} = \mathbf{x} + i\mathbf{y}$ (where $\mathbf{x}$ and $\mathbf{y}$ are real) be an eigenvector belonging to λ_1 and let $S = \text{Span}(\mathbf{x}, \mathbf{y})$. By Lemma 6.4.5, $\dim S = 2$ and S is invariant under A. Let $\{\mathbf{q}_1, \mathbf{q}_2\}$ be an orthonormal basis for S. Choose $\mathbf{q}_3, \mathbf{q}_4, \ldots, \mathbf{q}_n$ so that $Q_1 = (\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n)$ is an orthogonal matrix. Since S is invariant under A, it follows that

$$A\mathbf{q}_1 = b_{11}\mathbf{q}_1 + b_{21}\mathbf{q}_2 \quad \text{and} \quad A\mathbf{q}_2 = b_{12}\mathbf{q}_1 + b_{22}\mathbf{q}_2$$

for some scalars $b_{11}, b_{21}, b_{12}, b_{22}$ and hence the first two columns of $Q_1^T A Q_1$ will be

$$(Q_1^T A \mathbf{q}_1, Q_1^T A \mathbf{q}_2) = (b_{11}\mathbf{e}_1 + b_{21}\mathbf{e}_2, b_{12}\mathbf{e}_1 + b_{22}\mathbf{e}_2)$$

So, in general, $Q_1^T A Q_1$ will be a matrix of block form

$$Q_1^T A Q_1 = \begin{pmatrix} B_1 & X \\ O & A_1 \end{pmatrix}$$

where

$$\begin{array}{ll} B_1 = (\lambda_1) \text{ and } A_1 \text{ is } (k-1)\times(k-1) & \text{if } \lambda_1 \text{ is real} \\ B_1 \text{ is } 2\times 2 \text{ and } A_1 \text{ is } (k-2)\times(k-2) & \text{if } \lambda_1 \text{ is complex.} \end{array}$$

In either case, we can apply our induction hypothesis to A_1 and obtain a Schur decomposition $A_1 = UT_1U^T$. Let us assume that the Schur form T_1 has $j-1$ diagonal blocks $B_2, B_3, \ldots, B_j$. If we set

$$Q_2 = \begin{pmatrix} I & O \\ O & Q_1 \end{pmatrix} \quad \text{and} \quad Q = Q_1 Q_2$$

then both Q_2 and Q are $k \times k$ orthogonal matrices. If we then set $T = Q^T A Q$, we will obtain a matrix in the Schur form (2), and it follows that A will have Schur decomposition QTQ^T. ■

In the case that all of the eigenvalues of A are real, the real Schur form T will be upper triangular. In the case that A is real and symmetric, then, since all of the eigenvalues of A are real, T must be upper triangular; however, in this case T must also be symmetric. So we end up with a diagonalization of A. Thus, for real symmetric matrices, we have the following version of the Spectral Theorem:

Corollary 6.4.7 Spectral Theorem—Real Symmetric Matrices

If A is a real symmetric matrix, then there is an orthogonal matrix Q that diagonalizes A; that is, $Q^T A Q = D$, where D is diagonal.

Normal Matrices

There are non-Hermitian matrices that possess complete orthonormal sets of eigenvectors. For example, skew-symmetric and skew-Hermitian matrices have this property. (A is *skew Hermitian* if $A^H = -A$.) If A is any matrix with a complete orthonormal set of eigenvectors, then $A = UDU^H$, where U is unitary and D is a diagonal matrix (whose diagonal elements may be complex). In general, $D^H \neq D$ and, consequently,

$$A^H = UD^HU^H \neq A$$

However,

$$AA^H = UDU^HUD^HU^H = UDD^HU^H$$

and

$$A^HA = UD^HU^HUDU^H = UD^HDU^H$$

Since

$$D^H D = DD^H = \begin{bmatrix} |\lambda_1|^2 & & & \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ & & & |\lambda_n|^2 \end{bmatrix}$$

it follows that

$$AA^H = A^H A$$

Definition A matrix A is said to be **normal** if $AA^H = A^H A$.

We have shown that if a matrix has a complete orthonormal set of eigenvectors, then it is normal. The converse is also true.

Theorem 6.4.8 A matrix A is normal if and only if A possesses a complete orthonormal set of eigenvectors.

Proof In view of the preceding remarks, we need only show that a normal matrix A has a complete orthonormal set of eigenvectors. By Theorem 6.4.3, there exists a unitary matrix U and a triangular matrix T such that $T = U^H AU$. We claim that T is also normal. To see this, note that

$$T^H T = U^H A^H UU^H AU = U^H A^H AU$$

and

$$TT^H = U^H AUU^H A^H U = U^H AA^H U$$

Since $A^H A = AA^H$, it follows that $T^H T = TT^H$. Comparing the diagonal elements of TT^H and $T^H T$, we see that

$$\begin{aligned} |t_{11}|^2 + |t_{12}|^2 + |t_{13}|^2 + \cdots + |t_{1n}|^2 &= |t_{11}|^2 \\ |t_{22}|^2 + |t_{23}|^2 + \cdots + |t_{2n}|^2 &= |t_{12}|^2 + |t_{22}|^2 \\ &\vdots \\ |t_{nn}|^2 &= |t_{1n}|^2 + |t_{2n}|^2 + |t_{3n}|^2 + \cdots + |t_{nn}|^2 \end{aligned}$$

It follows that $t_{ij} = 0$ whenever $i \neq j$. Thus, U diagonalizes A and the column vectors of U are eigenvectors of A. ■

SECTION 6.4 EXERCISES

1. For each of the following pairs of vectors $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^2$, compute (i) $\|\mathbf{z}\|$, (ii) $\|\mathbf{w}\|$, (iii) $\langle \mathbf{z}, \mathbf{w}\rangle$, and (iv) $\langle \mathbf{w}, \mathbf{z}\rangle$:

(a) $\mathbf{z} = \begin{Bmatrix} 4+2i \\ 4i \end{Bmatrix}, \quad \mathbf{w} = \begin{Bmatrix} -2 \\ 2+i \end{Bmatrix}$

(b) $\mathbf{z} = \begin{Bmatrix} 1+i \\ 2i \\ 3-i \end{Bmatrix}, \quad \mathbf{w} = \begin{Bmatrix} 2-4i \\ 5 \\ 2i \end{Bmatrix}$

2. Let

$$\mathbf{z}_1 = \begin{Bmatrix} \dfrac{1+i}{2} \\ \dfrac{1-i}{2} \end{Bmatrix} \quad \text{and} \quad \mathbf{z}_2 = \begin{Bmatrix} \dfrac{i}{\sqrt{2}} \\ -\dfrac{1}{\sqrt{2}} \end{Bmatrix}$$

(a) Show that $\{\mathbf{z}_1, \mathbf{z}_2\}$ is an orthonormal set in $\mathbb{C}^2$.

(b) Write the vector $\mathbf{z} = \begin{Bmatrix} 2+4i \\ -2i \end{Bmatrix}$ as a linear combination of $\mathbf{z}_1$ and $\mathbf{z}_2$.

3. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be an orthonormal basis for $\mathbb{C}^2$, and let $\mathbf{z} = (4+2i)\mathbf{u}_1 + (6-5i)\mathbf{u}_2$.

(a) What are the values of $\mathbf{u}_1^H\mathbf{z}$, $\mathbf{z}^H\mathbf{u}_1$, $\mathbf{u}_2^H\mathbf{z}$, and $\mathbf{z}^H\mathbf{u}_2$?

(b) Determine the value of $\|\mathbf{z}\|$.

4. Which of the matrices that follow are Hermitian? Normal?

(a) $\begin{Bmatrix} 1-i & 2 \\ 2 & 3 \end{Bmatrix}$ **(b)** $\begin{Bmatrix} 1 & 2-i \\ 2+i & -1 \end{Bmatrix}$

(c) $\begin{Bmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{Bmatrix}$

(d) $\begin{Bmatrix} \dfrac{1}{\sqrt{2}}i & \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}}i \end{Bmatrix}$

(e) $\begin{Bmatrix} 0 & i & 1 \\ i & 0 & -2+i \\ -1 & 2+i & 0 \end{Bmatrix}$

(f) $\begin{Bmatrix} 3 & 1+i & i \\ 1-i & 1 & 3 \\ -i & 3 & 1 \end{Bmatrix}$

5. Find an orthogonal or unitary diagonalizing matrix for each of the following:

(a) $\begin{Bmatrix} 2 & 1 \\ 1 & 2 \end{Bmatrix}$ **(b)** $\begin{Bmatrix} 1 & 3+i \\ 3-i & 4 \end{Bmatrix}$

(c) $\begin{Bmatrix} 2 & i & 0 \\ -i & 2 & 0 \\ 0 & 0 & 2 \end{Bmatrix}$ **(d)** $\begin{Bmatrix} 2 & 1 & 1 \\ 1 & 3 & -2 \\ 1 & -2 & 3 \end{Bmatrix}$

(e) $\begin{Bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{Bmatrix}$ **(f)** $\begin{Bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{Bmatrix}$

(g) $\begin{Bmatrix} 4 & 2 & -2 \\ 2 & 1 & -1 \\ -2 & -1 & 1 \end{Bmatrix}$

6. Show that the diagonal entries of a Hermitian matrix must be real.

7. Let A be a Hermitian matrix and let $\mathbf{x}$ be a vector in $\mathbb{C}^n$. Show that if $c = \mathbf{x}A\mathbf{x}^H$, then c is real.

8. Let A be a Hermitian matrix and let $B = iA$. Show that B is skew Hermitian.

9. Let A and C be matrices in $\mathbb{C}^{m\times n}$ and let $B \in \mathbb{C}^{n\times r}$. Prove each of the following rules:

(a) $(A^H)^H = A$

(b) $(\alpha A + \beta C)^H = \overline{\alpha}A^H + \overline{\beta}C^H$

(c) $(AB)^H = B^H A^H$

10. Let A and B be Hermitian matrices. Answer *true* or *false* for each of the statements that follow. In each case, explain or prove your answer.

(a) The eigenvalues of AB are all real.

(b) The eigenvalues of ABA are all real.

11. Show that

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H\mathbf{z}$$

defines an inner product on $\mathbb{C}^n$.

12. Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ be vectors in $\mathbb{C}^n$ and let α and β be complex scalars. Show that

$$\langle \mathbf{z}, \alpha\mathbf{x} + \beta\mathbf{y} \rangle = \overline{\alpha}\langle \mathbf{z}, \mathbf{x} \rangle + \overline{\beta}\langle \mathbf{z}, \mathbf{y} \rangle$$

13. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be an orthonormal basis for a complex inner product space V, and let

$$\mathbf{z} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n$$
$$\mathbf{w} = b_1\mathbf{u}_1 + b_2\mathbf{u}_2 + \cdots + b_n\mathbf{u}_n$$

Show that

$$\langle \mathbf{z}, \mathbf{w} \rangle = \sum_{i=1}^{n} \overline{b_i} a_i$$

14. Given that

$$A = \begin{Bmatrix} 4 & 0 & 0 \\ 0 & 1 & i \\ 0 & -i & 1 \end{Bmatrix}$$

find a matrix B such that $B^H B = A$.

15. Let U be a unitary matrix. Prove that

(a) U is normal.

(b) $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{C}^n$.

(c) if λ is an eigenvalue of U, then $|\lambda| = 1$.

16. Let $\mathbf{u}$ be a unit vector in $\mathbb{C}^n$ and define $U = I - 2\mathbf{u}\mathbf{u}^H$. Show that U is both unitary and Hermitian and, consequently, is its own inverse.

17. Show that if a matrix U is both unitary and Hermitian, then any eigenvalue of U must equal either 1 or -1.

18. Let A be a 2×2 matrix with Schur decomposition UTU^H and suppose that $t_{12} \neq 0$. Show that

(a) the eigenvalues of A are $\lambda_1 = t_{11}$ and $\lambda_2 = t_{22}$.
(b) $\mathbf{u}_1$ is an eigenvector of A belonging to $\lambda_1 = t_{11}$.
(c) $\mathbf{u}_2$ is not an eigenvector of A belonging to $\lambda_2 = t_{22}$.

19. Let A be a 5×5 matrix with real entries. Let $A = QTQ^T$ be the real Schur decomposition of A, where T is a block matrix of the form given in equation (2). What are the possible block structures for T in each of the following cases?

(a) All of the eigenvalues of A are real.
(b) A has three real eigenvalues and two complex eigenvalues.
(c) A has one real eigenvalue and four complex eigenvalues.

20. Let A be a $n \times n$ matrix with Schur decomposition UTU^H. Show that if the diagonal entries of T are all distinct, then there is an upper triangular matrix R such that $X = UR$ diagonalizes A.

21. Show that $M = A + iB$ (where A and B real matrices) is skew Hermitian if and only if A is skew symmetric and B is symmetric.

22. Show that if A is skew Hermitian and λ is an eigenvalue of A, then λ is purely imaginary (i.e., $\lambda = bi$, where b is real).

23. Show that if A is a normal matrix, then each of the following matrices must also be normal:

(a) A^H **(b)** $I + A$ **(c)** A^2

24. Let A be a real 2×2 matrix with the property that $a_{21}a_{12} > 0$, and let

$$r = \sqrt{a_{21}/a_{12}} \quad \text{and} \quad S = \begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}$$

Compute $B = SAS^{-1}$. What can you conclude about the eigenvalues and eigenvectors of B? What can you conclude about the eigenvalues and eigenvectors of A? Explain.

25. Let $p(x) = -x^3 + cx^2 + (c+3)x + 1$, where c is a real number. Let

$$C = \begin{pmatrix} c & c+3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

and let

$$A = \begin{pmatrix} -1 & 2 & -c-3 \\ 1 & -1 & c+2 \\ -1 & 1 & -c-1 \end{pmatrix}$$

(a) Compute $A^{-1}CA$.
(b) Show that C is the companion matrix of $p(x)$ and use the result from part (a) to prove that $p(x)$ will have only real roots, regardless of the value of c.

26. Let A be a Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ and orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. Show that

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

27. Let

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

Write A as a sum $\lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T$, where λ_1 and λ_2 are eigenvalues and $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthonormal eigenvectors.

28. Let A be a Hermitian matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$. For any nonzero vector $\mathbf{x}$ in $\mathbb{R}^n$, the *Rayleigh quotient* $\rho(\mathbf{x})$ is defined by

$$\rho(\mathbf{x}) = \frac{\langle A\mathbf{x}, \mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle} = \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}}$$

(a) If $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$, show that

$$\rho(\mathbf{x}) = \frac{|c_1|^2\lambda_1 + |c_2|^2\lambda_2 + \cdots + |c_n|^2\lambda_n}{\|\mathbf{c}\|^2}$$

(b) Show that

$$\lambda_n \leq \rho(\mathbf{x}) \leq \lambda_1$$

(c) Show that

$$\max_{\mathbf{x} \neq \mathbf{0}} \rho(\mathbf{x}) = \lambda_1 \quad \text{and} \quad \min_{\mathbf{x} \neq \mathbf{0}} \rho(\mathbf{x}) = \lambda_n$$

29. Given $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times n}$, the equation

$$AX - XB = C \tag{3}$$

is known as *Sylvester's equation*. An $m \times n$ matrix X is said to be a solution if it satisfies (3).

(a) Show that if B has Schur decomposition $B = UTU^H$, then Sylvester's equation can be transformed into an equation of the form $AY - YT = G$, where $Y = XU$ and $G = CU$.
(b) Show that

$$(A - t_{11}I)\mathbf{y}_1 = \mathbf{g}_1$$

$$(A - t_{jj}I)\mathbf{y}_j = \mathbf{g}_j + \sum_{i=1}^{j-1} t_{ij}\mathbf{y}_i, \; j = 2, \ldots, n$$

(c) Show that if A and B have no common eigenvalues, then Sylvester's equation has a solution.