
1 Complex Matrices

Nearly everything we have studied in this book would remain true if the phrase *real number* were replaced by *complex number* wherever it occurs. Then we would deal with matrices with complex entries, systems of linear equations with complex coefficients (and complex solutions), determinants of complex matrices, and vector spaces with scalar multiplication by any complex number allowed. Moreover, the proofs of most theorems about (the real version of) these concepts extend easily to the complex case. It is not our intention here to give a full treatment of complex linear algebra. However, we will carry the theory far enough to give another proof of the Real Spectral Theorem (??).

The set of complex numbers is denoted $\mathbb{C}$. We will use only the most basic properties of these numbers (mainly conjugation and absolute values), and the reader can find this material in [Complex Numbers](#).

If $n \geq 1$, we denote the set of all n -tuples of complex numbers by $\mathbb{C}^n$. As with $\mathbb{R}^n$, these n -tuples will be written either as row or column matrices and will be referred to as *vectors*. We define vector operations on $\mathbb{C}^n$ as follows:

$$\begin{aligned} [v_1, v_2, \dots, v_n] + [w_1, w_2, \dots, w_n] &= [v_1 + w_1, v_2 + w_2, \dots, v_n + w_n] \\ u[v_1, v_2, \dots, v_n] &= [uv_1, uv_2, \dots, uv_n] \quad \text{for } u \text{ in } \mathbb{C} \end{aligned}$$

With these definitions, $\mathbb{C}^n$ satisfies the axioms for a vector space (with complex scalars) given in [Abstract Vector Spaces](#). Thus we can speak of spanning sets for $\mathbb{C}^n$, of linearly independent subsets, and of bases. In all cases, the definitions are identical to the real case, except that the scalars are allowed to be complex numbers. In particular, the standard basis of $\mathbb{R}^n$ remains a basis of $\mathbb{C}^n$, called the *standard basis* of $\mathbb{C}^n$.

A Generalization of the Dot Product for Complex Vectors

There is a generalization to $\mathbb{C}^n$ of the dot product in $\mathbb{R}^n$.

Definition 1. Given $\mathbf{z} = [z_1, z_2, \dots, z_n]$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]$ in $\mathbb{C}^n$, define their *standard inner product* $\langle \mathbf{z}, \mathbf{w} \rangle$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1 \overline{w_1} + z_2 \overline{w_2} + \dots + z_n \overline{w_n}$$

where $\overline{w}$ is the conjugate of the complex number w .

Clearly, if $\mathbf{z}$ and $\mathbf{w}$ actually lie in $\mathbb{R}^n$, then $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z} \cdot \mathbf{w}$ is the usual dot product.

Example 1. If $\mathbf{z} = [2, 1 - i, 2i, 3 - i]$ and $\mathbf{w} = [1 - i, -1, -i, 3 + 2i]$, then

$$\begin{aligned} \langle \mathbf{z}, \mathbf{w} \rangle &= 2(1 + i) + (1 - i)(-1) + (2i)(i) + (3 - i)(3 - 2i) = 6 - 6i \\ \langle \mathbf{z}, \mathbf{z} \rangle &= 2 \cdot 2 + (1 - i)(1 + i) + (2i)(-2i) + (3 - i)(3 + i) = 20 \end{aligned}$$

Note that $\langle \mathbf{z}, \mathbf{w} \rangle$ is a complex number in general, as opposed to requiring an inner product to be real as we do in [Inner Product Spaces](#). However, if $\mathbf{w} = \mathbf{z} = [z_1, z_2, \dots, z_n]$, the definition gives $\langle \mathbf{z}, \mathbf{z} \rangle = |z_1|^2 + \dots + |z_n|^2$ which is a nonnegative real number, equal to 0 if and only if $\mathbf{z} = \mathbf{0}$. This explains the conjugation in the definition of $\langle \mathbf{z}, \mathbf{w} \rangle$, and it gives ?? of the following theorem.

Theorem 1. Let $\mathbf{z}, \mathbf{z}_1, \mathbf{w}$, and $\mathbf{w}_1$ denote vectors in $\mathbb{C}^n$, and let λ denote a complex number.

1. $\langle \mathbf{z} + \mathbf{z}_1, \mathbf{w} \rangle = \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{z}_1, \mathbf{w} \rangle$ and $\langle \mathbf{z}, \mathbf{w} + \mathbf{w}_1 \rangle = \langle \mathbf{z}, \mathbf{w} \rangle + \langle \mathbf{z}, \mathbf{w}_1 \rangle$.
2. $\langle \lambda \mathbf{z}, \mathbf{w} \rangle = \lambda \langle \mathbf{z}, \mathbf{w} \rangle$ and $\langle \mathbf{z}, \lambda \mathbf{w} \rangle = \overline{\lambda} \langle \mathbf{z}, \mathbf{w} \rangle$.

3. $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$.

4. $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$, and $\langle \mathbf{z}, \mathbf{z} \rangle = 0$ if and only if $\mathbf{z} = \mathbf{0}$.

Proof We leave ?? and ?? to the reader (Practice Problem ??), and ?? has already been proved. To prove ??, write $\mathbf{z} = [z_1, z_2, \dots, z_n]$ and $\mathbf{w} = [w_1, w_2, \dots, w_n]$. Then

$$\begin{aligned} \overline{\langle \mathbf{w}, \mathbf{z} \rangle} &= \overline{(w_1 \bar{z}_1 + \dots + w_n \bar{z}_n)} = \bar{w}_1 \bar{\bar{z}}_1 + \dots + \bar{w}_n \bar{\bar{z}}_n \\ &= z_1 \bar{w}_1 + \dots + z_n \bar{w}_n = \langle \mathbf{z}, \mathbf{w} \rangle \end{aligned}$$

■

Definition 2. As for the dot product on $\mathbb{R}^n$, property ?? enables us to define the *norm* or *length* $\|\mathbf{z}\|$ of a vector $\mathbf{z} = (z_1, z_2, \dots, z_n)$ in $\mathbb{C}^n$:

$$\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \dots + |z_n|^2}$$

The only properties of the norm function we will need are the following (the proofs are left to the reader):

Theorem 2. If $\mathbf{z}$ is any vector in $\mathbb{C}^n$, then

1. $\|\mathbf{z}\| \geq 0$ and $\|\mathbf{z}\| = 0$ if and only if $\mathbf{z} = \mathbf{0}$.
2. $\|\lambda \mathbf{z}\| = |\lambda| \|\mathbf{z}\|$ for all complex numbers λ .

A vector $\mathbf{u}$ in $\mathbb{C}^n$ is called a *unit vector* if $\|\mathbf{u}\| = 1$. Property ?? in Theorem ?? then shows that if $\mathbf{z} \neq \mathbf{0}$ is any nonzero vector in $\mathbb{C}^n$, then $\mathbf{u} = \frac{1}{\|\mathbf{z}\|} \mathbf{z}$ is a unit vector.

Example 2. In $\mathbb{C}^4$, find a unit vector $\mathbf{u}$ that is a positive real multiple of $\mathbf{z} = \begin{bmatrix} 1 - i \\ i \\ 2 \\ 3 + 4i \end{bmatrix}$.

Explanation. $\|\mathbf{z}\| = \sqrt{2 + 1 + 4 + 25} = \sqrt{32} = 4\sqrt{2}$, so take $\mathbf{u} = \frac{1}{4\sqrt{2}} \mathbf{z}$.

A matrix $A = [a_{ij}]$ is called a *complex matrix* if every entry a_{ij} is a complex number. The notion of conjugation for complex numbers extends to matrices as follows: Define the *conjugate* of $A = [a_{ij}]$ to be the matrix

$$\bar{A} = [\bar{a}_{ij}]$$

obtained from A by conjugating every entry. Then (using Appendix ??)

$$\overline{A + B} = \bar{A} + \bar{B} \quad \text{and} \quad \overline{AB} = \bar{A} \bar{B}$$

holds for all (complex) matrices of appropriate size.

Transposition of complex matrices is defined just as in the real case, and the following notion is fundamental.

Definition 3. The *conjugate transpose* A^H of a complex matrix A is defined by

$$A^H = (\bar{A})^T = \overline{(A^T)}$$

Observe that $A^H = A^T$ when A is real.

Remark 1. Other notations for A^H are A^* and $A^\dagger$.

Example 3.

$$\begin{bmatrix} 3 & 1-i & 2+i \\ 2i & 5+2i & -i \end{bmatrix}^H = \begin{bmatrix} 3 & -2i \\ 1+i & 5-2i \\ 2-i & i \end{bmatrix}$$

The following properties of A^H follow easily from the rules for transposition of real matrices and extend these rules to complex matrices. Note the conjugate in property ??.

Theorem 3. Let A and B denote complex matrices, and let λ be a complex number.

1. $(A^H)^H = A$.
2. $(A + B)^H = A^H + B^H$.
3. $(\lambda A)^H = \bar{\lambda} A^H$.
4. $(AB)^H = B^H A^H$.

Hermitian and Unitary Matrices

If A is a real symmetric matrix, it is clear that $A^H = A$. The complex matrices that satisfy this condition turn out to be the most natural generalization of the real symmetric matrices:

Definition 4. A square complex matrix A is called *Hermitian* if $A^H = A$, equivalently $\bar{A} = A^T$.

Remark 2. The name Hermitian honours Charles Hermite (1822–1901), a French mathematician who worked primarily in analysis and is remembered as the first to show that the number e from calculus is transcendental—that is, e is not a root of any polynomial with integer coefficients.

Hermitian matrices are easy to recognize because the entries on the main diagonal must be real, and the “reflection” of each off-diagonal entry in the main diagonal must be the conjugate of that entry.

Example 4. $\begin{bmatrix} 3 & i & 2+i \\ -i & -2 & -7 \\ 2-i & -7 & 1 \end{bmatrix}$ is Hermitian, whereas $\begin{bmatrix} 1 & i \\ i & -2 \end{bmatrix}$ and $\begin{bmatrix} 1 & i \\ -i & i \end{bmatrix}$ are not.

The following theorem extends Theorem ??, and gives a very useful characterization of Hermitian matrices in terms of the standard inner product in $\mathbb{C}^n$.

Theorem 4. An $n \times n$ complex matrix A is Hermitian if and only if

$$\langle Az, w \rangle = \langle z, Aw \rangle$$

for all n -tuples z and w in $\mathbb{C}^n$.

Proof If A is Hermitian, we have $A^T = \bar{A}$. If $\mathbf{z}$ and $\mathbf{w}$ are columns in $\mathbb{C}^n$, then $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z}^T \bar{\mathbf{w}}$, so

$$\langle A\mathbf{z}, \mathbf{w} \rangle = (A\mathbf{z})^T \bar{\mathbf{w}} = \mathbf{z}^T A^T \bar{\mathbf{w}} = \mathbf{z}^T \bar{A} \bar{\mathbf{w}} = \mathbf{z}^T (\overline{A\mathbf{w}}) = \langle \mathbf{z}, A\mathbf{w} \rangle$$

To prove the converse, let $\mathbf{e}_j$ denote column j of the identity matrix. If $A = [a_{ij}]$, the condition gives

$$\bar{a}_{ij} = \langle \mathbf{e}_i, A\mathbf{e}_j \rangle = \langle A\mathbf{e}_i, \mathbf{e}_j \rangle = a_{ji}$$

Hence $\bar{A} = A^T$, so A is Hermitian. ■

Let A be an $n \times n$ complex matrix. As in the real case, a complex number λ is called an *eigenvalue* of A if $A\mathbf{x} = \lambda\mathbf{x}$ holds for some column $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{C}^n$. In this case $\mathbf{x}$ is called an *eigenvector* of A corresponding to λ .

Definition 5. Let A be an $n \times n$ complex matrix. Then

$$c_A(z) = \det(zI - A)$$

is a degree n polynomial (in z) called the *characteristic polynomial* of A .

Remark 3. Some books define the characteristic polynomial to be $\det(A - zI)$ (the left side of the characteristic equation ??). Our definition ensures that the characteristic polynomial is a monic polynomial (i.e., its leading coefficient is 1). The two polynomials will have the same roots, which are the eigenvalues of the matrix A .

If A is an $n \times n$ matrix, the characteristic polynomial $c_A(z)$ is a polynomial of degree n and the eigenvalues of A are just the roots of $c_A(z)$. In most of our examples these roots have been *real* numbers (in fact, the examples have been carefully chosen so this will be the case!); but it need not happen, even when the characteristic polynomial has real coefficients. For example, if $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ then $c_A(z) = z^2 + 1$ has roots i and $-i$, where i is a complex number satisfying $i^2 = -1$. Therefore, we have to deal with the possibility that the eigenvalues of a (real) square matrix might be complex numbers.

For a complex matrix, $c_A(z)$ has complex coefficients (possibly nonreal). However, an argument like that given in Exploration ?? of [The Characteristic Equation](#) still works to show that the eigenvalues of A are the roots (possibly complex) of $c_A(z)$.

It is at this point that the advantage of working with complex numbers becomes apparent. The real numbers are incomplete in the sense that the characteristic polynomial of a real matrix may fail to have all its roots real. However, this difficulty does not occur for the complex numbers. The so-called fundamental theorem of algebra ensures that *every* polynomial of positive degree with complex coefficients has a complex root. Hence every square complex matrix A has a (complex) eigenvalue. Indeed (see ??), $c_A(z)$ factors completely as follows:

$$c_A(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n)$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A (with possible repetitions due to multiple roots).

The next result extends Theorem ??, which asserts that eigenvectors of a symmetric real matrix corresponding to distinct eigenvalues are orthogonal. In the complex context, two n -tuples $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^n$ are said to be *orthogonal* if $\langle \mathbf{z}, \mathbf{w} \rangle = 0$.

Theorem 5. Suppose A is a Hermitian matrix.

1. The eigenvalues of A are real.

2. Eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof Let λ and μ be eigenvalues of A with (nonzero) eigenvectors $\mathbf{z}$ and $\mathbf{w}$. Then $A\mathbf{z} = \lambda\mathbf{z}$ and $A\mathbf{w} = \mu\mathbf{w}$, so Theorem ?? gives

$$\lambda\langle\mathbf{z}, \mathbf{w}\rangle = \langle\lambda\mathbf{z}, \mathbf{w}\rangle = \langle A\mathbf{z}, \mathbf{w}\rangle = \langle\mathbf{z}, A\mathbf{w}\rangle = \langle\mathbf{z}, \mu\mathbf{w}\rangle = \bar{\mu}\langle\mathbf{z}, \mathbf{w}\rangle \quad (1)$$

If $\mu = \lambda$ and $\mathbf{w} = \mathbf{z}$, this becomes $\lambda\langle\mathbf{z}, \mathbf{z}\rangle = \bar{\lambda}\langle\mathbf{z}, \mathbf{z}\rangle$. Because $\langle\mathbf{z}, \mathbf{z}\rangle = \|\mathbf{z}\|^2 \neq 0$, this implies $\lambda = \bar{\lambda}$. Thus λ is real, proving (1). Similarly, μ is real, so equation (??) gives $\lambda\langle\mathbf{z}, \mathbf{w}\rangle = \mu\langle\mathbf{z}, \mathbf{w}\rangle$. If $\lambda \neq \mu$, this implies $\langle\mathbf{z}, \mathbf{w}\rangle = 0$, proving (2). ■

Corollary 1. Let A be a real symmetric matrix. Then the eigenvalues of A are real.

Proof Symmetric real matrices are Hermitian, and so the result follows immediately from Theorem ??. ■

The Real Spectral Theorem (??) asserts that every real symmetric matrix A is orthogonally diagonalizable—that is $U^T A U$ is diagonal where U is an orthogonal matrix ($U^{-1} = U^T$). The next theorem identifies the complex analogs of these orthogonal real matrices.

Definition 6. As in the real case, a set of nonzero vectors $\{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_m\}$ in $\mathbb{C}^n$ is called *orthogonal* if $\langle\mathbf{z}_i, \mathbf{z}_j\rangle = 0$ whenever $i \neq j$, and it is *orthonormal* if, in addition, $\|\mathbf{z}_i\| = 1$ for each i .

Theorem 6. The following are equivalent for an $n \times n$ complex matrix A .

1. A is invertible and $A^{-1} = A^H$.
2. The rows of A are an orthonormal set in $\mathbb{C}^n$.
3. The columns of A are an orthonormal set in $\mathbb{C}^n$.

Proof If $A = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_n \\ | & | & & | \end{bmatrix}$ is a complex matrix with j th column $\mathbf{c}_j$, then $A^T \bar{A} = [\langle\mathbf{c}_i, \mathbf{c}_j\rangle]$, as in Theorem ??. Now ?? $\Leftrightarrow$?? follows, and ?? $\Leftrightarrow$?? is proved in the same way. ■

Definition 7. A square complex matrix U is called *unitary* if $U^{-1} = U^H$.

Thus a real matrix is unitary if and only if it is orthogonal.

Example 5. The matrix $A = \begin{bmatrix} 1+i & 1 \\ 1-i & i \end{bmatrix}$ has orthogonal columns, but the rows are not orthogonal. Normalizing the columns gives the unitary matrix $\frac{1}{2} \begin{bmatrix} 1+i & \sqrt{2} \\ 1-i & \sqrt{2}i \end{bmatrix}$.

Given a real symmetric matrix A , we saw in [Orthogonal Matrices and Symmetric Matrices](#) a procedure for finding an orthogonal matrix Q such that $Q^T A Q$ is diagonal (see Example ??). The following example illustrates Theorem ?? and shows that the technique works for complex matrices.

Example 6. Consider the Hermitian matrix $A = \begin{bmatrix} 3 & 2+i \\ 2-i & 7 \end{bmatrix}$. Find the eigenvalues of A , find two orthonormal eigenvectors, and so find a unitary matrix U such that $U^H A U$ is diagonal.

Explanation. The characteristic polynomial of A is

$$c_A(z) = \det(zI - A) = \det \begin{bmatrix} x-3 & -2-i \\ -2+i & x-7 \end{bmatrix} = (z-2)(z-8)$$

Hence the eigenvalues are 2 and 8 (both real as expected), and corresponding eigenvectors are $\begin{bmatrix} 2+i \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2-i \end{bmatrix}$ (orthogonal as expected). Each has length $\sqrt{6}$, so let $U = \frac{1}{\sqrt{6}} \begin{bmatrix} 2+i & 1 \\ -1 & 2-i \end{bmatrix}$ be the unitary matrix with the normalized eigenvectors as columns.

Then $U^H A U = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$ is diagonal.

Unitary Diagonalization

An $n \times n$ complex matrix A is called *unitarily diagonalizable* if $U^H A U$ is diagonal for some unitary matrix U . As Example ?? suggests, we are going to prove that every Hermitian matrix is unitarily diagonalizable. However, with only a little extra effort, we can get a very important theorem that has this result as an easy consequence.

A complex matrix is called *upper triangular* if every entry below the main diagonal is zero. We owe the following theorem to Issai Schur.

Remark 4. Issai Schur (1875–1941) was a German mathematician who did fundamental work in the theory of representations of groups as matrices.

Theorem 7 (Schur's Theorem). If A is any $n \times n$ complex matrix, there exists a unitary matrix U such that

$$U^H A U = T$$

is upper triangular. Moreover, the entries on the main diagonal of T are the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A (including multiplicities).

Proof We use induction on n , mirroring the form of the proof of ??. If $n = 1$, A is already upper triangular. If $n > 1$, assume the theorem is valid for $(n-1) \times (n-1)$ complex matrices. Let λ_1 be an eigenvalue of A , and let $\mathbf{y}_1$ be an eigenvector with $\|\mathbf{y}_1\| = 1$. Next, the (complex analog of the) Gram-Schmidt process provides $\mathbf{y}_2, \dots, \mathbf{y}_n$ such that

$\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n\}$ is an orthonormal basis of $\mathbb{C}^n$. If $U_1 = \begin{bmatrix} | & | & & | \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n \\ | & | & & | \end{bmatrix}$ is the matrix with these vectors as its columns, then

$$U_1^H A U_1 = \begin{bmatrix} \lambda_1 & X_1 \\ 0 & A_1 \end{bmatrix}$$

in block form. Now apply induction to find a unitary $(n-1) \times (n-1)$ matrix W_1 such that $W_1^H A_1 W_1 = T_1$ is upper

triangular. Then $U_2 = \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix}$ is a unitary $n \times n$ matrix. Hence $U = U_1 U_2$ is unitary (using Theorem ??), and

$$\begin{aligned} U^H A U &= U_2^H (U_1^H A U_1) U_2 \\ &= \begin{bmatrix} 1 & 0 \\ 0 & W_1^H \end{bmatrix} \begin{bmatrix} \lambda_1 & X_1 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & X_1 W_1 \\ 0 & T_1 \end{bmatrix} \end{aligned}$$

is upper triangular. Finally, A and $U^H A U = T$ have the same eigenvalues by (the complex version of) Theorem ??, and they are the diagonal entries of T because T is upper triangular. ■

The fact that similar matrices have the same traces and determinants gives the following consequence of Schur's theorem.

Corollary 2. Let A be an $n \times n$ complex matrix, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ denote the eigenvalues of A , including multiplicities. Then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n \quad \text{and} \quad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Schur's theorem asserts that every complex matrix can be “unitarily triangularized.” However, we cannot substitute “unitarily diagonalized” here. In fact, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, there is no invertible complex matrix U at all such that $U^{-1} A U$ is diagonal. However, the situation is much better for Hermitian matrices.

Theorem 8. If A is Hermitian, there is a unitary matrix U such that $U^H A U$ is diagonal.

Proof By Schur's theorem, let $U^H A U = T$ be upper triangular where U is unitary. Since A is Hermitian, this gives

$$T^H = (U^H A U)^H = U^H A^H U^{HH} = U^H A U = T$$

This means that T is both upper and lower triangular. Hence T is actually diagonal. ■

The Real Spectral Theorem asserts that a real matrix A is symmetric if and only if it is orthogonally diagonalizable (that is, $P^T A P$ is diagonal for some real orthogonal matrix P). Theorem ?? is the complex analog of half of this result. However, the converse is false for complex matrices: There exist unitarily diagonalizable matrices that are not Hermitian.

Example 7. Show that the non-Hermitian matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is unitarily diagonalizable.

Explanation. The characteristic polynomial is $c_A(z) = z^2 + 1$. Hence the eigenvalues are i and $-i$, and it is easy to verify that $\begin{bmatrix} i \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ i \end{bmatrix}$ are corresponding eigenvectors. Moreover, these eigenvectors are orthogonal and both have length $\sqrt{2}$, so $U = \frac{1}{\sqrt{2}} \begin{bmatrix} i & -1 \\ -1 & i \end{bmatrix}$ is a unitary matrix such that $U^H A U = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$ is diagonal.

There is a very simple way to characterize those complex matrices that are unitarily diagonalizable. To this end, an $n \times n$ complex matrix N is called *normal* if $NN^H = N^H N$. It is clear that every Hermitian or unitary matrix is normal, as is the matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ in Example ??. In fact we have the following result.

Theorem 9. An $n \times n$ complex matrix A is unitarily diagonalizable if and only if A is normal.

Proof Assume first that $U^H A U = D$, where U is unitary and D is diagonal. Then $DD^H = D^H D$ as is easily verified. Because $DD^H = U^H(AA^H)U$ and $D^H D = U^H(A^H A)U$, it follows by cancellation that $AA^H = A^H A$.

Conversely, assume A is normal—that is, $AA^H = A^H A$. By Schur's theorem, let $U^H A U = T$, where T is upper triangular and U is unitary. Then T is normal too:

$$TT^H = U^H(AA^H)U = U^H(A^H A)U = T^H T$$

Hence it suffices to show that a normal $n \times n$ upper triangular matrix T must be diagonal. We induct on n ; it is clear if $n = 1$. If $n > 1$ and $T = [t_{ij}]$, then equating $(1, 1)$ -entries in TT^H and $T^H T$ gives

$$|t_{11}|^2 + |t_{12}|^2 + \dots + |t_{1n}|^2 = |t_{11}|^2$$

This implies $t_{12} = t_{13} = \dots = t_{1n} = 0$, so $T = \begin{bmatrix} t_{11} & 0 \\ 0 & T_1 \end{bmatrix}$ in block form. Hence $T = \begin{bmatrix} \bar{t}_{11} & 0 \\ 0 & T_1^H \end{bmatrix}$ so $TT^H = T^H T$ implies $T_1 T_1^H = T_1^H T_1$. Thus T_1 is diagonal by induction, and the proof is complete. ■

We conclude this section by using Schur's theorem (Theorem ??) to prove a famous theorem about matrices. Recall that the characteristic polynomial of a square matrix A is defined by $c_A(z) = \det(zI - A)$, and that the eigenvalues of A are just the roots of $c_A(z)$.

Theorem 10 (Cayley-Hamilton Theorem). If A is an $n \times n$ complex matrix, then $c_A(A) = 0$; that is, A is a root of its characteristic polynomial.

Remark 5. Named after the English mathematician Arthur Cayley (1821–1895) and William Rowan Hamilton (1805–1865), an Irish mathematician famous for his work on physical dynamics.

Proof If $p(z)$ is any polynomial with complex coefficients, then $p(P^{-1}AP) = P^{-1}p(A)P$ for any invertible complex matrix P . Hence, by Schur's theorem, we may assume that A is upper triangular. Then the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A appear along the main diagonal, so

$$c_A(z) = (z - \lambda_1)(z - \lambda_2)(z - \lambda_3) \cdots (z - \lambda_n)$$

Thus

$$c_A(A) = (A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)$$

Note that each matrix $A - \lambda_i I$ is upper triangular. Now observe:

1. $A - \lambda_1 I$ has zero first column because column 1 of A is $(\lambda_1, 0, 0, \dots, 0)^T$.
2. Then $(A - \lambda_1 I)(A - \lambda_2 I)$ has the first two columns zero because the second column of $(A - \lambda_2 I)$ is $(b, 0, 0, \dots, 0)^T$ for some constant b .
3. Next $(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I)$ has the first three columns zero because column 3 of $(A - \lambda_3 I)$ is $(c, d, 0, \dots, 0)^T$ for some constants c and d .

Continuing in this way we see that $(A - \lambda_1 I)(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I)$ has all n columns zero; that is, $c_A(A) = 0$. ■

Practice Problems

Problem 1 In each case, compute the norm of the complex vector.

1. $(1, 1 - i, -2, i)$

... ..

2. $(1 - i, 1 + i, 1, -1)$

3. $(2 + i, 1 - i, 2, 0, -i)$

... ..

4. $(-2, -i, 1 + i, 1 - i, 2i)$

Problem 2 In each case, determine whether the two vectors are orthogonal.

1. $\begin{bmatrix} 4 \\ -3i \\ 2 + i \end{bmatrix}, \begin{bmatrix} i \\ 2 \\ 2 - 4i \end{bmatrix}$ (Yes/ No)

2. $\begin{bmatrix} i \\ -i \\ 2 + i \end{bmatrix}, \begin{bmatrix} i \\ i \\ 2 - i \end{bmatrix}$ (Yes/ No)

3. $\begin{bmatrix} 1 \\ 1 \\ i \\ i \end{bmatrix}, \begin{bmatrix} 1 \\ i \\ -i \\ 1 \end{bmatrix}$ (Yes/ No)

4. $\begin{bmatrix} 4 + 4i \\ 2 + i \\ 2i \end{bmatrix}, \begin{bmatrix} -1 + i \\ 2 \\ 3 - 2i \end{bmatrix}$ (Yes/ No)

Problem 3 A subset U of $\mathbb{C}^n$ is called a complex subspace of $\mathbb{C}^n$ if it contains 0 and if, given $\mathbf{v}$ and $\mathbf{w}$ in U , both $\mathbf{v} + \mathbf{w}$ and $z\mathbf{v}$ lie in U (z any complex number). In each case, determine whether U is a complex subspace of $\mathbb{C}^3$.

1. $U = \{(w, \bar{w}, 0) \mid w \text{ in } \mathbb{C}\}$

2. $U = \{(w, 2w, a) \mid w \text{ in } \mathbb{C}, a \text{ in } \mathbb{R}\}$

3. $U = \mathbb{R}^3$ (Subspace/ Not a Subspace)

4. $U = \{(v + w, v - 2w, v) \mid v, w \text{ in } \mathbb{C}\}$

Problem 4 In each case, find a basis over $\mathbb{C}$, and determine the dimension of the complex subspace U of $\mathbb{C}^3$ (see the previous exercise).

-
1. $U = \{(w, v + w, v - iw) \mid v, w \text{ in } \mathbb{C}\}$
 2. $U = \{(iv + w, 0, 2v - w) \mid v, w \text{ in } \mathbb{C}\}$
 3. $U = \{(u, v, w) \mid iu - 3v + (1 - i)w = 0; u, v, w \text{ in } \mathbb{C}\}$
 4. $U = \{(u, v, w) \mid 2u + (1 + i)v - iw = 0; u, v, w \text{ in } \mathbb{C}\}$

Problem 5 In each case, determine whether the given matrix is Hermitian, unitary, or normal.

1. $\begin{bmatrix} 1 & -i \\ i & i \end{bmatrix}$
2. $\begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$
3. $\begin{bmatrix} 1 & i \\ -i & 2 \end{bmatrix}$
4. $\begin{bmatrix} 1 & -i \\ i & -1 \end{bmatrix}$
5. $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$
6. $\begin{bmatrix} 1 & 1+i \\ 1+i & i \end{bmatrix}$
7. $\begin{bmatrix} 1+i & 1 \\ -i & -1+i \end{bmatrix}$
8. $\frac{1}{\sqrt{2}|z|} \begin{bmatrix} z & z \\ \bar{z} & -\bar{z} \end{bmatrix}, z \neq 0$

Problem 6 Show that a matrix N is normal if and only if $\overline{N}N^T = N^T\overline{N}$.

Problem 7 Let $A = \begin{bmatrix} z & \bar{v} \\ v & w \end{bmatrix}$ where v , w , and z are complex numbers. Characterize in terms of v , w , and z when A is

1. Hermitian
2. unitary

3. normal.

Problem 8 In each case, find a unitary matrix U such that $U^H A U$ is diagonal.

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

2. $A = \begin{bmatrix} 4 & 3-i \\ 3+i & 1 \end{bmatrix}$

3. $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}; a, b, \text{ real}$

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

5. $A = \begin{bmatrix} 1 & 0 & 1+i \\ 0 & 2 & 0 \\ 1-i & 0 & 0 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1+i \\ 0 & 1-i & 2 \end{bmatrix}$

Problem 9 Show that $\langle Ax, y \rangle = \langle x, A^H y \rangle$ holds for all $n \times n$ matrices A and for all n -tuples x and y in $\mathbb{C}^n$.

Problem 10 1. Prove ?? and ?? of Theorem ??.

2. Prove Theorem ??.

3. Prove Theorem ??.

Problem 11 a. Show that A is Hermitian if and only if $\bar{A} = A^T$.

b. Show that the diagonal entries of any Hermitian matrix are real.

Problem 12 1. Show that every complex matrix Z can be written uniquely in the form $Z = A + iB$, where A and B are real matrices.

2. If $Z = A + iB$ as in (a), show that Z is Hermitian if and only if A is symmetric, and B is skew-symmetric (that is, $B^T = -B$).

Problem 13 If Z is any complex $n \times n$ matrix, show that ZZ^H and $Z + Z^H$ are Hermitian.

Problem 14 A complex matrix B is called skew-Hermitian if $B^H = -B$.

-
- Show that $Z - Z^H$ is skew-Hermitian for any square complex matrix Z .
 - If B is skew-Hermitian, show that B^2 and iB are Hermitian.
 - If B is skew-Hermitian, show that the eigenvalues of B are pure imaginary ($i\lambda$ for real λ).
 - Show that every $n \times n$ complex matrix Z can be written uniquely as $Z = A + B$, where A is Hermitian and B is skew-Hermitian.

Problem 15 Let U be a unitary matrix. Show that:

- $\|U\mathbf{x}\| = \|\mathbf{x}\|$ for all columns $\mathbf{x}$ in $\mathbb{C}^n$.
- $|\lambda| = 1$ for every eigenvalue λ of U .

Problem 16 1. If Z is an invertible complex matrix, show that Z^H is invertible and that $(Z^H)^{-1} = (Z^{-1})^H$.

- Show that the inverse of a unitary matrix is again unitary.
- If U is unitary, show that U^H is unitary.

Problem 17 Let Z be an $m \times n$ matrix such that $Z^H Z = I_n$ (for example, Z is a unit column in $\mathbb{C}^n$).

- Show that $V = ZZ^H$ is Hermitian and satisfies $V^2 = V$.
- Show that $U = I - 2ZZ^H$ is both unitary and Hermitian (so $U^{-1} = U^H = U$).

Problem 18 1. If N is normal, show that zN is also normal for all complex numbers z .

- Show that (a) fails if normal is replaced by Hermitian.

Problem 19 Show that a real 2×2 normal matrix is either symmetric or has the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

Problem 20 If A is Hermitian, show that all the coefficients of $c_A(z)$ are real numbers.

Problem 21 1. If $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, show that $U^{-1}AU$ is not diagonal for any invertible complex matrix U .

- If $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, show that $U^{-1}AU$ is not upper triangular for any real invertible matrix U .

Problem 22 If A is any $n \times n$ matrix, show that $U^H AU$ is lower triangular for some unitary matrix U .

Problem 23 If A is a 3×3 matrix, show that $A^2 = 0$ if and only if there exists a unitary matrix U such that $U^H AU$

has the form $\begin{bmatrix} 0 & 0 & u \\ 0 & 0 & v \\ 0 & 0 & 0 \end{bmatrix}$ or the form $\begin{bmatrix} 0 & u & v \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Problem 24 If $A^2 = A$, show that $\text{rank } A = \text{tr } A$. [Hint: Use Schur's theorem.]

Text Source

This section was adapted from Section 8.6 of Keith Nicholson's *Linear Algebra with Applications*. (CC-BY-NC-SA)
W. Keith Nicholson, *Linear Algebra with Applications*, Lyryx 2018, Open Edition, pp. 445–456.