Lecture 14. Orthogonal Diagonalization

December 2022

Hermitian Matrices

DEFINITION 14.1 The **conjugate transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^H obtained by interchanging the rows and columns and then taking the complex conjugate of each entry:

$$A^H = \bar{A}^T = \overline{A^T}.$$

Observe that if A is real, then $A^H = A^T$.

EXAMPLE 14.2

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

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 (3).

THEOREM 14.3 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) $(AB)^H = B^H A^H$.

The most natural generalization of the real symmetric matrices is the following.

DEFINITION 14.4 A square complex matrix A is called **Hermitian** if $A^H = A$, equivalently $\bar{A} = A^T$.

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

EXAMPLE 14.5 $A = \begin{pmatrix} 1 & -2+i \\ -2-i & 3 \end{pmatrix}$ is Hermitian, whereas $B = \begin{pmatrix} 1 & i \\ i & 3 \end{pmatrix}$ and $C = \begin{pmatrix} -2 & -i \\ i & i \end{pmatrix}$ are not.

From now on, let $f : V \to V$ be a linear operator on an inner product complex vector space V.

THEOREM 14.6 An operator f is self-adjoint if and only if the matrix A of f with respect to an orthonormal basis is Hermitian.

Proof. If A is Hermitian, we have $A^T = \bar{A}$. If $\mathbf{z}, \mathbf{w} \in V$, then $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{z}^T \bar{\mathbf{w}}$, so

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

Therefore, the operator f is self-adjoint.

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. □

THEOREM 14.7 Let $f : V \to V$ be a self-adjoint operator.

1. The eigenvalues of f are real.
2. Eigenvectors of f corresponding to distinct eigenvalues are orthogonal.

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

$$\lambda\langle \mathbf{z}, \mathbf{w} \rangle = \langle \lambda\mathbf{z}, \mathbf{w} \rangle = \langle f\mathbf{z}, \mathbf{w} \rangle = \langle \mathbf{z}, f\mathbf{w} \rangle = \langle \mathbf{z}, \mu\mathbf{w} \rangle = \bar{\mu}\langle \mathbf{z}, \mathbf{w} \rangle \tag{1}$$

If we put $\mu = \lambda$ 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 Part 1.

Similarly, μ is real, so equation (1) 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 Part 2. □

COROLLARY 14.8 The eigenvalues of Hermitian matrices are always real.

Orthogonal and Unitary Matrices

THEOREM 14.9 The following conditions are equivalent for an $n \times n$ matrix P.

(1) P is invertible and $P^{-1} = P^T$.
(2) The rows of P are orthonormal with respect to standard inner product.
(3) The columns of P are orthonormal with respect to standard inner product.

Proof. First recall that condition (1) is equivalent to $PP^T = I$. Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ denote the rows of P. Then $\mathbf{x}_j^T$ is the jth column of P^T, so the (i, j)-entry of PP^T is $\mathbf{x}_i \cdot \mathbf{x}_j$. Thus $PP^T = I$ means that $\mathbf{x}_i \cdot \mathbf{x}_j = 0$ if $i \neq j$ and $\mathbf{x}_i \cdot \mathbf{x}_j = 1$ if $i = j$. Hence condition (1) is equivalent to (2). The proof of the equivalence of (1) and (3) is similar. □

DEFINITION 14.10 An $n \times n$ matrix P is called an **orthogonal matrix** if it satisfies one (and hence all) of the conditions in Theorem 14.9.

EXAMPLE 14.11 The rotation matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is orthogonal for any θ.

These orthogonal matrices have the virtue that they are easy to invert — simply take the transpose. But they have many other important properties as well. If $f : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator, it can be proved that f is distance preserving if and only if its matrix is orthogonal. In particular, the matrices of rotations and reflections about the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$ are all orthogonal.

THEOREM 14.12 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 $\mathbb{C}^n$ with respect to standard Hermitian inner product.
(3) The columns of A are an orthonormal set in $\mathbb{C}^n$ with respect to standard Hermitian inner product.

Proof is similar to that of Theorem 14.9. □

DEFINITION 14.13 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 14.14 The matrix $A = \begin{pmatrix} 1+i & 1 \\ 1-i & i \end{pmatrix}$ has orthogonal columns, but the rows are not orthogonal. Normalizing the columns gives the unitary matrix $\frac{1}{2}\begin{pmatrix} 1+i & \sqrt{2} \\ 1-i & \sqrt{2}i \end{pmatrix}$.

Spectral Theorem

Recall that an $n\times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors. Moreover, the matrix P with these eigenvectors as columns is a diagonalizing matrix for A, that is $P^{-1}AP$ is diagonal. As we have seen, the really nice bases of $\mathbb{R}^n$ are the orthogonal ones, so a natural question is: which $n \times n$ matrices have an orthogonal basis of eigenvectors? These turn out to be precisely the symmetric matrices, and this is the main result of this lecture.

DEFINITION 14.15 A linear operator $f : V \to V$ of a finite-dimensional inner product space V is called **orthogonally diagonalizable** if there exists an orthonormal basis of V in which the matrix of f is diagonal.

An $n \times n$ matrix A is said to be **orthogonally diagonalizable** when an orthogonal matrix P can be found such that $P^{-1}AP = P^TAP$ is diagonal.

This condition turns out to characterize the symmetric matrices.

THEOREM 14.16 (Spectral Theorem) Suppose a linear operator $f : V \to V$ has a matrix A with respect to an orthonormal basis. Then the following conditions are equivalent.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

REMARK 14.17 Because the eigenvalues of a (real) symmetric matrix are real, Theorem 14.16 is also called the **real spectral theorem**, and the set of distinct eigenvalues is called the **spectrum** of the matrix.

Now the procedure for diagonalizing a symmetric $n \times n$ matrix is clear.

1. Find the distinct eigenvalues (all real).
2. Find orthonormal bases for each eigenspace (the Gram–Schmidt algorithm may be needed).

3. Then the set of all these basis vectors is orthonormal and contains n vectors.

Here is an example.

EXAMPLE 14.18 Orthogonally diagonalize the symmetric matrix $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

Solution. The characteristic polynomial is

$$p_A(\lambda) = \begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = \lambda(\lambda - 2)$$

Hence the distinct eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = 2$. By Gaussian elimination we obtain eigenvectors:

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

These vectors are orthogonal to each other. Normalizing gives orthonormal vectors

$$\frac{1}{\sqrt{2}}\mathbf{x}_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \qquad \frac{1}{\sqrt{2}}\mathbf{x}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

so

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}}\mathbf{x}_1 & \frac{1}{\sqrt{2}}\mathbf{x}_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

is an orthogonal matrix such that $P^{-1}AP$ is diagonal.

THEOREM 14.19 (Spectral Theorem in Complex Vector Spaces) If A is Hermitian, there is a unitary matrix U such that $U^H AU$ is diagonal.