In Sec 5.12, we found conditions that guaranteed the diagonalizability of an $n \times n$ matrix but we did not consider what class or classes of matrices might actually satisfy those conditions. In this chapter, we will show that every Symmetric matrix is Diagonalizable. This is an extremely important result because many applications utilize it in some essential way.

SEC 7.1 ORTHOGONAL MATRICES.

In this Section, we will discuss the class of matrices whose inverses can be obtained by transposition. Such matrices occur in a variety of applications and arise as well as transition matrices when one orthonormal basis is changed to another.

Definition: A square matrix 'A' is said to be Orthogonal if its transpose is the same as its inverse, that is, if $A^{-1} = A^T$

or, equivalently, if $AA^T = A^TA = I$ —— (1)

Example (1) A 3x3 Orthogonal Matrix

The matrix
$$A = \begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix}$$ is Orthogonal, since

$$A^TA = \begin{bmatrix} \frac{3}{7} & -\frac{6}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{3}{7} & \frac{6}{7} \\ \frac{6}{7} & \frac{2}{7} & -\frac{3}{7} \end{bmatrix}\begin{bmatrix} \frac{3}{7} & \frac{2}{7} & \frac{6}{7} \\ -\frac{6}{7} & \frac{3}{7} & \frac{2}{7} \\ \frac{2}{7} & \frac{6}{7} & -\frac{3}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

Example (2) Rotation and Reflection Matrices are Orthogonal

Recall from Sec 4.9 that the standard matrix for the counterclockwise rotation of R^2 through an angle θ is
$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

This matrix is Orthogonal for all choices of θ since

$$A^TA = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \cos\theta\sin\theta \\ -\sin\theta\cos\theta + \sin\theta\cos\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We can verify that Reflection matrices and Rotation matrices given in Tables of Sec 4.9 are all Orthogonal.

NOTE: Observe that for Orthogonal matrices in Example (1) & (2), both the row vectors and column vectors form orthonormal sets with respect to Euclidean inner product. This is a consequence of the following theorem —

THEOREM ① The following statements are equivalent for an $n \times n$ matrix 'A' —

(i) 'A' is Orthogonal.

(ii) The row vectors of 'A' form an orthonormal set in R^n with respect to Euclidean inner product.

(iii) The column vectors of 'A' form an orthonormal set in R^n with respect to Euclidean inner product.

PROPERTIES OF ORTHOGONAL MATRICES :- The following theorem lists three more fundamental properties of orthogonal matrices —

THEOREM ② (i) The Inverse of an orthogonal matrix is orthogonal.

(ii) The product of orthogonal matrices is orthogonal.

(iii) If matrix A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.

Example ③ $\det(A) = \pm 1$ FOR AN ORTHOGONAL MATRIX 'A'

The matrix $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ is orthogonal since its row (and column) vectors form orthonormal sets in R^2 with the Euclidean inner product.

$$\det(A) = \left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right) - \left(-\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\right)$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

Inchanging the rows of A produces an orthogonal matrix whose determinant is -1.

ORTHOGONAL MATRICES AS LINEAR OPERATORS - We observed in Example ② that standard matrices for the basic reflection and rotation operators on R^2 and R^3 are Orthogonal. The next theorem will explain why this is so.

THEOREM ③ If 'A' is an $n \times n$ matrix, then the following statements are equivalent —

(i) 'A' is Orthogonal.

(ii) $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ in R^n.

(iii) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in R^n.

Theorem ③ has a useful geometric interpretation when considered from the viewpoint of matrix transformations. If 'A' is an orthogonal matrix and $T_A : R^n \to R^n$ is multiplication by A, then we will call T_A an Orthogonal operator on R^n. It follows from parts (i) & (ii) of Theorem ③ that the orthogonal operators on R^n are precisely those operators that leave the lengths of all vectors unchanged. This explains why, in Example ②, we found the standard matrices for the basic reflections and rotations of R^2 and R^3 to be orthogonal.

Parts (i) & (iii) of Theorem ③ imply that orthogonal operators leave the angle between two vectors unchanged.

SEC (7.2) ORTHOGONAL DIAGONALIZATION

In this Section, we will be concerned with the problem of diagonalizing a symmetric matrix A. As we will see, this problem is closely related to that of finding an orthonormal basis for R^n that consists of eigenvectors of A. Problems of this type are important because many of the matrices that arise in applications are symmetric.

THE ORTHOGONAL DIAGONALIZATION PROBLEM

In a Definition of Sec (5.2) we defined two square matrices, A and B to be Similar if there is an invertible matrix P such that $P^{-1}AP = B$. In this Section, we will be concerned with the special case in which it is possible to find an orthogonal matrix P for which this relationship holds.

We begin with the following definition —

Definition: If A & B are square matrices, then we say that A & B are Orthogonally Similar if there is an orthogonal matrix P such that $P^TAP = B$.

If the matrix A is orthogonally similar to some diagonal matrix, say $P^TAP = D$ then we say that 'A' is orthogonally diagonalizable and that 'P' orthogonally diagonalizes 'A'.

Our first goal in this Section is to determine what conditions a matrix must satisfy to be orthogonally diagonalizable. As a first step, observe that there is no hope of orthogonally diagonalizing a matrix that is not symmetric.

CONDITIONS FOR ORTHOGONAL DIAGONALIZABILITY

The following theorem shows that every symmetric matrix is, in fact, orthogonally diagonalizable. In this theorem and for the remainder of this Section, orthogonal will mean orthogonal with respect to the Euclidean inner product on R^n.

THEOREM (1) If A is an $n \times n$ matrix, then the following statements are equivalent —

(i) 'A' is orthogonally diagonalizable.

(ii) 'A' has an orthonormal set of n eigenvectors.

(iii) 'A' is symmetric.

PROPERTIES OF SYMMETRIC MATRICES

Our next goal is to devise a procedure for orthogonally diagonalizing a symmetric matrix, but before we can do so, we need the following critical theorem about eigenvalues and eigenvectors of the symmetric matrices.

THEOREM (2) If A is a symmetric matrix, then

(i) The eigenvalues of A are all real numbers.

(ii) Eigenvectors from different eigenspaces are Orthogonal.

ORTHOGONALLY DIAGONALIZING AN $n \times n$ SYMMETRIC MATRICES.

<u>Step①</u> Find a basis for each eigenspace of matrix A.

<u>Step②</u> Apply Gram-Schmidt process to each of these bases to obtain an orthonormal basis for each eigenspace.

<u>Step③</u> Form the matrix P whose columns are the vectors constructed in Step②. This matrix will orthogonally diagonalize matrix A and the eigenvalues on the diagonal of $D = P^T A P$ will be in the same order as their corresponding eigenvectors in P.

<u>REMARK</u>: The justification of this procedure should be clear :- Theorem ② ensures that eigenvectors from different eigenspaces are orthogonal and applying the Gram-Schmidt process ensures that the eigenvectors within the same eigenspace are Orthonormal. It follows that the entire set of eigenvectors obtained by this procedure will be Orthonormal.

Example ① Orthogonally Diagonalizing a Symmetric Matrix

Find an Orthogonal matrix P that diagonalizes the matrix, $A = \begin{bmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{bmatrix}$.

Soln. The characteristic Equ. of A is

$$\det(\lambda I - A) = 0$$

$$\begin{vmatrix} \lambda-4 & -2 & -2 \\ -2 & \lambda-4 & -2 \\ -2 & -2 & \lambda-4 \end{vmatrix} = 0$$

$$\Rightarrow \quad (\lambda-4)\begin{vmatrix} \lambda-4 & -2 \\ -2 & \lambda-4 \end{vmatrix} - (-2)\begin{vmatrix} -2 & -2 \\ -2 & \lambda-4 \end{vmatrix} + (-2)\begin{vmatrix} -2 & \lambda-4 \\ -2 & -2 \end{vmatrix} = 0$$

$$\Rightarrow (\lambda-4)\left[(\lambda-4)^2 - (-2)(-2)\right] + 2\left[-2(\lambda-4) - (-2)(-2)\right] - 2\left[(-2)(-2) - (-2)(\lambda-4)\right] = 0$$

$$\Rightarrow \quad \lambda^3 - 12\lambda^2 + 36\lambda - 32 = 0$$

$$\Rightarrow \quad (\lambda-2)(\lambda^2 - 10\lambda + 16) = 0$$

$$\Rightarrow \quad (\lambda-2)(\lambda-2)(\lambda-8) = 0$$

$$\Rightarrow \quad \lambda = 2, 2, 8$$

Thus, the distinct eigenvalues of A are $\lambda = 2$ and $\lambda = 8$.

By the method used in Example ⑦ of Sec (5.1), it can be shown that

$$u_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \qquad \text{——— ①}$$

form a basis for the eigenspace corresponding to $\lambda = 2$.

Applying Gram-Schmidt process to $\{u_1, u_2\}$ yields the following orthonormal eigenvectors-

$$v_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \qquad \text{——— ②}$$

The eigenspace corresponding $\lambda = 8$ has $u_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as a basis.

Applying Gram-Schmidt process to $\{u_3\}$ (i.e., normalizing u_3) yields $V_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$.

Finally, using v_1, v_2 and v_3 as column vectors, we obtain

$$P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$ which orthogonally diagonalizes A.

As a check, we can confirm that $\quad P^T A P = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$

SEC (7.3) QUADRATIC FORMS

In this Section, we will use matrix methods to study real-valued functions of several variables in which each term is either the square of a variable or the product of two variables. Such functions arise in a variety of applications including geometry, vibrations of mechanical systems, statistics and electrical engineering.

DEFINITION OF A QUADRATIC FORM

Expressions of the form $\quad a_1x_1 + a_2x_2 + \cdots + a_nx_n$

occurred in our study of linear equations and linear systems. If $a_1, a_2, \cdots, a_n$ are treated as fixed constants, then this expression is a real-valued function of n-variables $x_1, x_2, \cdots, x_n$ and is called a Linear form on R^n. All variables in a linear form occur to the first power and there are no products of variables.

Here we will be concerned with Quadratic forms on R^n, which are functions of the form

$$a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 + \text{all possible terms } a_kx_ix_j \text{ in which } x_i \neq x_j$$

The terms of the form $a_kx_ix_j$ are called cross product terms. It is common to combine the cross product terms involving x_ix_j with those involving x_jx_i to avoid duplication.

Thus, a general quadratic form on R^2 would typically be expressed as —

$$a_1x_1^2 + a_2x_2^2 + 2a_3x_1x_2 \quad \text{—— ①}$$

and a general quadratic form on R^3 as —

$$a_1x_1^2 + a_2x_2^2 + a_3x_3^2 + 2a_4x_1x_2 + 2a_5x_1x_3 + 2a_6x_2x_3 \quad \text{—— ②}$$

If as usual, we do not distinguish between the no. 'a' and the 1×1 matrix $[a]$ and if we let '$\mathbf{x}$' be the column vector of the variables, then ① & ② can be expressed in matrix form as —

$$[x_1 \;\; x_2]\begin{bmatrix} a_1 & a_3 \\ a_3 & a_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}^TA\mathbf{x}$$

and

$$[x_1 \;\; x_2 \;\; x_3]\begin{bmatrix} a_1 & a_4 & a_5 \\ a_4 & a_2 & a_6 \\ a_5 & a_6 & a_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{x}^TA\mathbf{x}$$

Note that the matrix A in these formulas is symmetric, that it diagonal entries are the coefficients of the squared terms, and its off-diagonal entries are half the coefficients of cross product terms.

In general, if A is symmetric $n\times n$ matrix and '$\mathbf{x}$' is an $n\times1$ column vector of variables then we call the function $\quad Q_A(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, the Quadratic form associated with A.

In the case where A is a diagonal matrix, the quadratic form $\mathbf{x}^TA\mathbf{x}$ has no cross product terms; for example, if A has diagonal entries $\lambda_1, \lambda_2, \cdots, \lambda_n$, then

$$\mathbf{x}^TA\mathbf{x} = [x_1 \;\; x_2 \;\; \cdots \;\; x_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda_1x_1^2 + \lambda_2x_2^2 + \cdots + \lambda_nx_n^2$$

Example ① Expressing Quadratic Forms in Matrix Notation

In each part, express the quadratic form in the matrix notation $x^T A x$, where A is symmetric –

(i) $2x^2 + 6xy - 5y^2$

(ii) $x_1^2 + 7x_2^2 + 3x_3^2 + 4x_1x_2 - 2x_1x_3 + 8x_2x_3$

Soln – The diagonal entries of A are the coefficients of the squared terms and the off-diagonal entries are half the coefficients of the cross product terms, so

$$2x^2+6xy-5y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = x^T A x, \text{ where } A = \begin{bmatrix} 2 & 3 \\ 3 & -5 \end{bmatrix}$$

$$x_1^2+7x_2^2+3x_3^2+4x_1x_2-2x_1x_3+8x_2x_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x^T A x, \text{ where } A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 7 & 4 \\ -1 & 4 & 3 \end{bmatrix}$$

POSITIVE DEFINITE QUADRATIC FORMS

Definition ① A quadratic form $\mathbf{x}^T A \mathbf{x}$ is said to be —

(i) Positive Definite if $\mathbf{x}^T A \mathbf{x} > 0$ for $\mathbf{x} \neq 0$.

(ii) Negative Definite if $\mathbf{x}^T A \mathbf{x} < 0$ for $\mathbf{x} \neq 0$.

(iii) Indefinite if $\mathbf{x}^T A \mathbf{x}$ has both positive and negative values.

NOTE: The terminology in Definition ① also applies to the matrix A; that is, matrix A is positive definite, negative definite, or indefinite in accordance with whether the associated quadratic form has that property.

The following theorem provides a way of using eigenvalues to determine whether a matrix A and its associated quadratic form $x^T A x$ are positive definite, negative definite or indefinite —

THEOREM ① If A is a symmetric matrix, then —

(i) $\mathbf{x}^T A \mathbf{x}$ is positive definite iff all eigenvalues of A are positive.

(ii) $\mathbf{x}^T A \mathbf{x}$ is negative definite iff all eigenvalues of A are negative.

(iii) $\mathbf{x}^T A \mathbf{x}$ is Indefinite iff A has at least one positive and at least one negative eigenvalue.

NOTE:- The three classifications in Definition ① do not exhaust all of the possibilities. For example, a quadratic form for which $\mathbf{x}^T A \mathbf{x} \geq 0$ if $\mathbf{x} \neq 0$ is called positive semi-definite and one for which $\mathbf{x}^T A \mathbf{x} \leq 0$ if $\mathbf{x} \neq 0$ is called negative semi-definite.

Example ② Positive Definite Quadratic forms

It is not usually possible to tell from the signs of the entries in a symmetric matrix A whether that matrix is positive definite, negative definite or indefinite.

For example, the entries of the matrix $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$

are non-negative, but the matrix is indefinite since its eigenvalues are $\lambda = 1, 4, -2$.

To see this another way, let us write out the quadratic form as —

$$\mathbf{x}^T A \mathbf{x} = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 3x_1 + x_2 + x_3 \\ x_1 + \quad 2x_3 \\ x_1 + 2x_2 \end{bmatrix}$$

$$= x_1(3x_1 + x_2 + x_3) + x_2(x_1 + 2x_3) + x_3(x_1 + 2x_2)$$

$$\mathbf{x}^T A \mathbf{x} = 3x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$$

We can now see, for example that

for $x_1 = 0,\ x_2 = 1,\ x_3 = 1$, we have $\mathbf{x}^T A \mathbf{x} = 3(0) + 0 + 0 + 4(1)(1) = 4$

and for $x_1 = 0,\ x_2 = 1,\ x_3 = -1$, we have $\mathbf{x}^T A \mathbf{x} = 0 + 0 + 4(1)(-1) = -4$

CLASSIFYING CONIC SECTIONS USING EIGENVALUES.

If λ_1, λ_2 are the eigenvalues of a matrix A of size 2×2, then

(i) $x^T A x = 1$ represents an Ellipse if $\lambda_1 > 0$ & $\lambda_2 > 0$.

(ii) $x^T A x = 1$ represents no graph if $\lambda_1 < 0$ & $\lambda_2 < 0$.

(iii) $x^T A x = 1$ represents a Hyperbola if λ_1 & λ_2 have opposite signs.

THEOREM ② If A is a symmetric 2×2 matrix, then

(i) $x^T A x = 1$ represents an Ellipse if A is positive definite.

(ii) $x^T A x = 1$ has no graph if A is negative definite.

(iii) $x^T A x = 1$ represents a Hyperbola if A is Indefinite.

THEOREM ③ A symmetric matrix 'A' is Positive Definite iff the determinant of every principal submatrix is positive.

SEC-(7.5) HERMITIAN, UNITARY AND NORMAL MATRICES.

We know that every real symmetric matrix is orthogonally diagonalizable and that the real symmetric matrices are the only orthogonally diagonalizable matrices.

In this Section, we will consider the diagonalization problem for complex matrices.

HERMITION AND UNITARY MATRICES. The transpose operation is less important for complex matrices than for real matrices. A more useful operation for complex matrices is given in the following definition —

Conjugate Transpose of a Matrix: If 'A' is a complex matrix, then the Conjugate Transpose of 'A', denoted by A^*, is defined as —

$$A^* = (\bar{A})^T = (\overline{A^T}) \qquad \text{——— (1)}$$

NOTE: A^* is same as A^T for real matrices.

Example (1) **Conjugate Transpose**

Find the conjugate transpose of the matrix, $A = \begin{bmatrix} 1+i & -i & 0 \\ 2 & 3-2i & i \end{bmatrix}$.

Soln. We have $\quad \bar{A} = \begin{bmatrix} 1-i & i & 0 \\ 2 & 3+2i & -i \end{bmatrix}$

Hence $\quad A^* = (\bar{A})^T = \begin{bmatrix} 1-i & 2 \\ i & 3+2i \\ 0 & -i \end{bmatrix}$

The following theorem shows that the basic algebraic properties of the conjugate transpose operation are similar to those of the transpose —

THEOREM (1) If k is a complex scalar and if A, B & C are complex matrices whose sizes are such that the stated operations can be performed, then

(i) $(A^*)^* = A$

(ii) $(A+B)^* = A^* + B^*$

(iii) $(A-B)^* = A^* - B^*$

(iv) $(kA)^* = \bar{k}A^*$

(v) $(AB)^* = B^*A^*$

Hermitian and Unitary Matrices —

A square complex matrix 'A' is said to be Unitary if $A^{-1} = A^*$ —— ①

and is said to be Hermitian if $A^* = A$ —— ②

NOTE ①. A Unitary matrix can also be defined as a square matrix A for which

$$AA^* = A^*A = I.$$

NOTE ②. If A is a real matrix, then $A^* = A^T$,

in which case ① becomes $A^{-1} = A^T$ and ② becomes $A^T = A$.

Thus, the Unitary matrices are complex generalizations of real Orthogonal matrices and Hermitian matrices are complex generalizations of real Symmetric matrices.

Example ② Recognizing Hermitian Matrices

Hermitian matrices are easy to recognize because their diagonal entries are real, and the entries that are symmetrically positioned across main diagonal are complex conjugates.

Thus, for example, we can tell by inspection that the matrix

$$A = \begin{bmatrix} 1 & i & 1+i \\ -i & -5 & 2-i \\ 1-i & 2+i & 3 \end{bmatrix}$$ is Hermitian.

The fact that real symmetric matrices have real eigenvalues is a special case of the following general result about Hermitian matrices —

THEOREM ② The Eigenvalues of a Hermitian matrix are real numbers.

The fact that eigenvectors from different eigenspaces of a real symmetric matrix are orthogonal is a special case of the following more general result about Hermitian matrices —

THEOREM ③ If A is a Hermitian matrix, then eigenvectors from different eigenspaces are Orthogonal.

Example ③ Eigenvalues and Eigenvectors of a Hermitian Matrix

Confirm that the Hermitian matrix, $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$

has real eigenvalues and the eigenvectors from different eigenspaces are Orthogonal.

Soln. The characteristic Eqn. of matrix A is

$$\det(\lambda I - A) = 0$$

$$\det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}\right) = 0$$

$$\begin{vmatrix} \lambda-2 & -1-i \\ -1+i & \lambda-3 \end{vmatrix} = 0$$

$$\Rightarrow \quad (\lambda-2)(\lambda-3) - (-1+i)(-1-i) = 0$$

$$\Rightarrow \quad \lambda^2 - 3\lambda - 2\lambda + 6 - [(-1)^2 - (i)^2] = 0$$

$$\Rightarrow \quad \lambda^2 - 5\lambda + 6 - [1+1] = 0$$

$$\Rightarrow \quad \lambda^2 - 5\lambda + 4 = 0$$

$$\Rightarrow \quad (\lambda-1)(\lambda-4) = 0$$

$$\Rightarrow \quad \underline{\lambda = 1 \ \& \ \lambda = 4} \quad \text{which are Real.}$$

To find the Eigenvectors, we must solve the system

$$(\lambda I - A)\mathbf{x} = \mathbf{0}$$

$$\begin{bmatrix} \lambda-2 & -1-i \\ -1+i & \lambda-3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{———} \quad ①$$

With $\lambda = 1$, the system ① becomes

$$\begin{bmatrix} -1 & -1-i \\ -1+i & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1+i \\ -1+i & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad R_1 \to (-1)R_1$$

$$\begin{bmatrix} 1 & 1+i \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad R_2 \to R_2 - (-1+i)R_1$$

$$\Rightarrow \quad x_1 + (1+i)x_2 = 0 \quad \text{———} \quad ②$$

Let $x_2 = t$, then from ② $\quad x_1 = (-1-i)t$

$\therefore$ Eigenvector corresponding $\lambda = 1$ is $\quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$= \begin{bmatrix} (-1-i)t \\ t \end{bmatrix}$$

$$\mathbf{X} = t \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$$

Thus, basis for eigenspace corresponding to $\lambda = 1$ is $\quad V_1 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$

Now with $\lambda = 4$, the system ① becomes

$$\begin{bmatrix} 2 & -1-i \\ -1+i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \frac{-1-i}{2} \\ -1+i & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad , R_1 \to \tfrac{1}{2} R_1$$

$$\begin{bmatrix} 1 & \frac{-1-i}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad , R_2 \to R_2 - (-1+i) R_1$$

$$\Rightarrow \quad x_1 + \tfrac{1}{2}(-1-i)\, x_2 = 0 \quad \text{———} \quad ③$$

Let $x_2 = t$, then from ③ $\quad x_1 = \frac{1}{2}(1+i)\, t$

$\therefore$ Eigenvector corresponding to $\lambda = 4$ is $\quad \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$= \begin{bmatrix} \left(\frac{1+i}{2}\right) t \\ t \end{bmatrix}$$

$$\mathbf{X} = t \begin{bmatrix} \frac{1}{2}(1+i) \\ 1 \end{bmatrix}$$

Thus, basis for eigenspace corresponding to $\lambda = 4$ is $\quad V_2 = \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix}$

Now $V_1 . V_2 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix} . \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix}$

$$= (-1-i)\left(\overline{\frac{1+i}{2}}\right) + (1)(\overline{1})$$

$$= \tfrac{1}{2}(-1-i)(1-i) + 1$$

$$= \tfrac{1}{2}(-1 + i - i + i^2) + 1$$

$$= \tfrac{1}{2}(-1-1) + 1 = 0$$

$\therefore$ The vectors $\mathbf{V}_1$ and $\mathbf{V}_2$ are Orthogonal.

Unitary matrices are not usually easy to recognize by inspection. However the following analog of Theorems ① & ③ provides a way of ascertaining whether a matrix is Unitary without computing its inverse.

THEOREM ④ If 'A' is an $n \times n$ matrix with complex entries, then following are Equivalent –

(i) Matrix A is Unitary.

(ii) $||A\mathbf{x}|| = ||\mathbf{x}||$, for all $\mathbf{x}$ in C^n.

(iii) $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in C^n.

(iv) The column vectors of 'A' form Orthonormal set in C^n with respect to complex Euclidean inner product.

(v) The Row vectors of 'A' form Orthonormal set in C^n with respect to complex Euclidean inner product.

Example ④ A Unitary Matrix

Use Theorem ④ to show that $A = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix}$ is Unitary and then find A^{-1}.

Soln. We will show that the vectors $\mathbf{r}_1 = \begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \end{bmatrix}$ and $\mathbf{r}_2 = \begin{bmatrix} \frac{1-i}{2} & \frac{-1+i}{2} \end{bmatrix}$ are Orthonormal.

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = \left(\frac{1+i}{2}\right)\overline{\left(\frac{1-i}{2}\right)} + \left(\frac{1+i}{2}\right)\overline{\left(\frac{-1+i}{2}\right)}$$

$$= \left(\frac{1+i}{2}\right)\left(\frac{1+i}{2}\right) + \left(\frac{1+i}{2}\right)\left(\frac{-1-i}{2}\right)$$

$$= \frac{1}{4}(1+i+i+i^2) + \frac{1}{4}(-1-i-i-i^2)$$

$$= \frac{1}{4}(1+2i-1) + \frac{1}{4}(-1-2i+1) = \frac{i}{2} - \frac{i}{2} = 0$$

$$||\mathbf{r}_1|| = \sqrt{\left|\frac{1+i}{2}\right|^2 + \left|\frac{1+i}{2}\right|^2} = \sqrt{\frac{1}{4}|1+i|^2 + \frac{1}{4}|1+i|^2} = \sqrt{\frac{1}{4}(1^2+1^2) + \frac{1}{4}(1^2+1^2)} = 1$$

$$||\mathbf{r}_2|| = \sqrt{\left|\frac{1-i}{2}\right|^2 + \left|\frac{-1+i}{2}\right|^2} = \sqrt{\frac{1}{4}|1-i|^2 + \frac{1}{4}|-1+i|^2} = \sqrt{\frac{1}{4}[(1)^2+(-1)^2] + \frac{1}{4}[(-1)^2+(1)^2]} = 1$$

Since the row vectors of matrix A form Orthonormal set, so matrix A is Unitary.

It follows that

$$A^{-1} = A^*$$

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

$$= \begin{bmatrix} \frac{1-i}{2} & \frac{1-i}{2} \\ \frac{1+i}{2} & \frac{-1-i}{2} \end{bmatrix}^T = \begin{bmatrix} \frac{1-i}{2} & \frac{1+i}{2} \\ \frac{1-i}{2} & \frac{-1-i}{2} \end{bmatrix}$$

We can verify the validity of this result by showing that $AA^* = A^*A = I$.

UNITARY DIAGONALIZABILITY

: Since Unitary matrices are the complex analogs of the real Orthogonal matrices, the following defⁿ. is a natural generalization of orthogonal diagonalizability for real matrices.

Definition: A square complex matrix is said to be Unitarily Diagonalizable if there is a unitary matrix P such that $P^* A P = D$ is a complex diagonal matrix.

Any such matrix P is said to Unitarily diagonalize 'A'.

Recall that a real symmetric $n \times n$ matrix 'A' has an orthonormal set of n eigenvectors and is orthogonally diagonalized by any $n \times n$ matrix whose column vectors are an orthonormal set of eigenvectors of 'A'. Here is the complex analog of that result —

THEOREM (5) Every $n \times n$ Hermitian matrix 'A' has an orthonormal set of n eigenvectors and is unitarily diagonalized by any $n \times n$ matrix 'P' whose column vectors form an orthonormal set of eigenvectors of 'A'.

PROCEDURE FOR UNITARILY DIAGONALIZING A HERMITIAN MATRIX

Step (1) Find a basis for each eigenspace of matrix 'A'.

Step (2) Apply Gram-Schmidt process to each of these bases to obtain orthonormal bases for the eigenspaces.

Step (3) Form the matrix P whose column vectors are the basis vectors obtained in step (2). This will be a unitary matrix (Theorem (4)) and will unitarily diagonalize 'A'.

Example (5) Unitarily Diagonalization of a Hermitian matrix

Find a matrix P that unitarily diagonalizes the Hermitian matrix, $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$.

Solⁿ. We showed in Example (3) that the Eigenvalues of A are $\lambda = 1$ & $\lambda = 4$, and that bases for the corresponding eigenspaces are $v_1 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix}$

Since each eigenspace has only one basis vector, the Gram-Schmidt process is simply a matter of normalizing these basis vectors.

$$p_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{|-1-i|^2 + |1|^2}} \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$p_2 = \frac{v_2}{\|v_2\|} = \frac{1}{\sqrt{\left|\frac{1+i}{2}\right|^2 + |1|^2}} \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{\frac{1}{4}|1+i|^2 + 1}} \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{\frac{1}{2}+1}} \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix} = \frac{2}{\sqrt{6}} \begin{bmatrix} \frac{1+i}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1+i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

Thus, 'A' is unitarily diagonalized by matrix

$$P = [p_1 \;\; p_2] = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix}.$$

SKEW-SYMMETRIC AND SKEW-HERMITIAN MATRIX

A square matrix 'A' with real entries is defined to be Skew-symmetric if $A^T = -A$.

A Skew-symmetric matrix must have zeros on the main diagonal and each entry off the main diagonal must be the negative of its mirror image about main diagonal.

for example,

$$A = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 4 \\ 2 & -4 & 0 \end{bmatrix}$$

We can confirm that $A^T = -A$.

The complex analogs of Skew-symmetric matrices are the matrices for which $A^* = -A$. Such matrices are said to be Skew-Hermitian.

A Skew-Hermitian matrix must have zeros or pure imaginary numbers on the main diagonal and each entry off the main diagonal must be the negative of complex conjugate of its mirror image about the main diagonal.

for example,

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

NORMAL MATRICES:

Hermitian matrices enjoy many but not all of the properties of real symmetric matrices. For example, we know that real symmetric matrices are orthogonally diagonalizable and Hermitian matrices are unitarily diagonalizable. However, whereas the real symmetric matrices are the only orthogonally diagonalizable matrices, the Hermitian matrices do not constitute the entire class of unitarily diagonalizable complex matrices; that is, there exist unitarily diagonalizable matrices that are not Hermitian. Specifically, it can be proved that a square complex matrix 'A' is unitarily diagonalizable iff $AA^* = A^*A$.

Matrices with this property are said to be Normal. Normal matrices include Hermitian, Skew-Hermitian and Unitary matrices in the complex case and Symmetric, skew-symmetric, and Orthogonal matrices in the real case. The non-zero skew-symmetric matrices are interesting because they are examples of real matrices that are not orthogonally diagonalizable but are unitarily diagonalizable.