Chapter 6

Eigenvalues

Eigenvalues

- If the equation $A\mathbf{x} = \lambda\mathbf{x}$ has a nonzero solution **x**, then λ is said to be an *eigenvalue* of A, and **x** is said to be an *eigenvector* belonging to λ.
- We can view eigenvalues (λ) as natural frequencies associated with linear transformations (A).
  - If A is an $n \times n$ matrix, we can think of A as representing a linear transformation from R^n to itself
  - If $\lambda > 0$, the effect of the operator (A) on any eigenvector belonging to λ is simply a stretching or shrinking by a constant factor.

6.1 Eigenvalues and Eigenvectors

- Many application problems involve applying a linear transformation repeatedly to a given vector
- The key to solving these problems is to choose a coordinate system or basis that is in some sense natural for the operator and simpler to do calculations involving the operator.
- These new **basis vectors** (***eigenvectors***) are associated with scaling factors (*eigenvalues*) that represent the natural frequencies of the operator.

Definition

Let A be an $n \times n$ matrix. A scalar λ is said to be an **eigenvalue** or a **characteristic value** of A if there exists a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$.

The vector $\mathbf{x}$ is said to be an **eigenvector** or a **characteristic vector** belonging to λ.

Example 2

If $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

then $A\mathbf{x} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 3\mathbf{x}$

$\Rightarrow$ λ = 3 is an eigenvalue of *A* and $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector belonging to λ = 3

- Actually, any nonzero multiple of **x** will be an eigenvector since

$$A(\underline{\alpha\mathbf{x}}) = \alpha(A\mathbf{x}) = \alpha(\lambda\mathbf{x}) = \lambda(\underline{\alpha\mathbf{x}})$$

$\Rightarrow$ $\alpha\mathbf{x}$ is also an eigenvector belonging to λ

Example 2-2

If $A = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$,

then $A\mathbf{x} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 12 \\ 6 \end{bmatrix} = 3 \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 3\mathbf{x}$

- If $A\mathbf{x} = \lambda\mathbf{x} \Rightarrow A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \Rightarrow (A - \lambda I)\mathbf{x} = \mathbf{0}$ **(1)**
- λ is an eigenvalue of A if and only if (1) has a nontrivial solution
- The set of solutions to (1) is $N(A - \lambda I)$, which is a subspace of R^n.

Example 2-2

- If λ is an eigenvalue of $\boldsymbol{A}$, then $N(\boldsymbol{A} - \lambda \boldsymbol{I}) \neq \{\mathbf{0}\}$ and any nonzero vector in $N(\boldsymbol{A} - \lambda \boldsymbol{I})$ is an eigenvector belonging to λ.
- The subspace $N(A - \lambda I)$ is called the **eigenspace** corresponding to the eigenvalue λ.
- Eq. (1) will have a **nontrivial solution** if and only if $(A - \lambda I)$ is singular, or

$$\det(A - \lambda I) = 0 \qquad (2)$$

which is called the **characteristic equation** for the matrix $\boldsymbol{A}$.

Example 2-2

- If Eq. (2) is expanded, we obtain an nth-degree polynomial in the variable λ,

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

  This polynomial is called the **characteristic polynomial**.
- The **roots** of the characteristic polynomial are the eigenvalues of A.
- There are n eigenvalues since there are n roots for an nth-degree polynomial

Summary

Let A be an $n \times n$ matrix and λ be a scalar. The following statements are equivalent.

(a) λ is an eigenvalue of A

(b) $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution

(c) $N(A - \lambda I) \neq \{\mathbf{0}\}$

(d) $(A - \lambda I)$ is singular (non-invertible)

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

Example 3

- Find the eigenvalues and the corresponding eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$$

Sol: The characteristic equation is: det($A - \lambda I$) = 0, then

$$\left| \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0 \Rightarrow \begin{vmatrix} 3-\lambda & 2 \\ 3 & -2-\lambda \end{vmatrix} = 0$$

Example 3

$\Rightarrow$ (3–λ) (-2–λ) – 6 = 0

$\Rightarrow$ $\lambda^2 - \lambda - 12 = 0$

$\Rightarrow$ (λ–4) (λ+3) = 0

$\Rightarrow$ $\lambda_1 = 4$ and $\lambda_2 = -3$

(1) for $\lambda_1 = 4$, we have to find $N(A - 4I)$:

$$\begin{bmatrix} 3-4 & 2 \\ 3 & -2-4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \Rightarrow \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

$\Rightarrow$ $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2\alpha \\ \alpha \end{bmatrix} = \alpha\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ is an eigenvector belonging to $\lambda_1 = 4$

Example 3

(2) for λ_2 = -3, we have to find $N(A + 3I)$:

$$\begin{bmatrix} 3+3 & 2 \\ 3 & -2+3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0} \Rightarrow \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

$\Rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -\alpha \\ 3\alpha \end{bmatrix} = \alpha \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ is an eigenvector belonging to λ_2 = -3

Example 3

(2) for $\lambda_2 = -3$, we have to find $N(A + 3I)$:

$\Rightarrow$ is an eigenvector belonging to $\lambda_2 = -3$

Example 4

- Let $A = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -2 & 1 \\ 1 & -3 & 2 \end{bmatrix}$, find the eigenvalues and the corresponding eigenspaces.

Sol: det(A – λI) = 0, then

$$\begin{vmatrix} 2-\lambda & -3 & 1 \\ 1 & -2-\lambda & 1 \\ 1 & -3 & 2-\lambda \end{vmatrix} = -\lambda(\lambda\text{-}1)^2 = 0$$

$\Rightarrow \lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 1$

Example 4

(1) for $\lambda_1 = 0$: $(A - \lambda_1 I)\mathbf{x} = \mathbf{0} \Rightarrow A\mathbf{x} = \mathbf{0} \Rightarrow$ the eigenspace is $N(A)$

$$\left[\begin{array}{ccc|c} 2 & -3 & 1 & 0 \\ 1 & -2 & 1 & 0 \\ 1 & -3 & 2 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha \\ \alpha \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$\Rightarrow$ the eigenspace corresponding to "$\lambda_1 = 0$" is $\alpha \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Example 4

(2) for $\lambda_2 = 1$: $(A - \lambda_1 I)\mathbf{x} = \mathbf{0} \Rightarrow (A - I)\mathbf{x} = \mathbf{0}$

$$A - I = \begin{bmatrix} 2-1 & -3 & 1 \\ 1 & -2-1 & 1 \\ 1 & -3 & 2-1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 1 \\ 1 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & -3 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Example 4

$$\Rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3\alpha - \beta \\ \alpha \\ \beta \end{bmatrix} = \alpha \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$\Rightarrow$ the eigenspace corresponding to "$\lambda = 1$" is

$$\alpha \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Example 5

- Find the eigenvalues and eigenspaces of the matrix

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

Sol:

$$\begin{vmatrix} 1-\lambda & 2 \\ -2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 4 = 0$$

$\Rightarrow \lambda_1 = 1+2i$ and $\lambda_2 = 1-2i$ Complex eigenvalues

Example 5

(1) for $\lambda_1 = 1+2i$ $$A-\lambda_1 I = \begin{bmatrix} -2i & 2 \\ -2 & -2i \end{bmatrix} \rightarrow \begin{bmatrix} -2i & 2 \\ 0 & 0 \end{bmatrix}$$

$\Rightarrow \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \alpha \\ \alpha i \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ i \end{bmatrix}$ is the eigenspace corresponding to $\lambda_1 = 1+2i$

$\Rightarrow \begin{bmatrix} 1 \\ i \end{bmatrix}$ is a basis for the eigenspace corresponding to $\lambda_1 = 1+2i$

Example 5

(2) for λ_2 = 1–2*i*

$$A-\lambda_2 I=\begin{bmatrix}2i & 2\\ -2 & 2i\end{bmatrix}\rightarrow\begin{bmatrix}2i & 2\\ 0 & 0\end{bmatrix}$$

$\Rightarrow \mathbf{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}\alpha\\ -\alpha i\end{bmatrix}=\alpha\begin{bmatrix}1\\ -i\end{bmatrix}$ is the eigenspace

corresponding to λ_2 = 1–2*i*

$\Rightarrow \begin{bmatrix}1\\ -i\end{bmatrix}$ is a basis for the eigenspace corresponding

to λ_2 = 1–2*i*

Complex Eigenvalues

- If A is an $n \times n$ matrix with real entries, then the characteristic polynomial of A will have **real** coefficients.
- All of its **complex roots** must occur in **conjugate pairs**. That is, if $\lambda = a + bi$ ($b \neq 0$) is an eigenvalue of A, then $\bar{\lambda} = a - bi$ must also be an eigenvalue of A.

Note

(1) If $A = (a_{ij})$ is a matrix with complex entries, then $\overline{A} = (\overline{a_{ij}})$ is the matrix formed from A by conjugating each of its entries $\overline{AB} = \overline{A}\ \overline{B}$

(2) If A and B are matrices with complex entries and the multiplication AB is possible, then

Complex Eigenvalues

- The eigenvectors also occur in conjugate pairs, if λ is a complex eigenvalue of a real $n \times n$ matrix A and $\mathbf{z}$ is an eigenvector belonging to λ, then

$$A\overline{\mathbf{z}} = \overline{A}\,\overline{\mathbf{z}} = \overline{A\mathbf{z}} = \overline{\lambda\mathbf{z}} = \overline{\lambda}\,\overline{\mathbf{z}}$$

Thus, $\overline{\mathbf{z}}$ is an eigenvector of A belonging to $\overline{\lambda}$

The Product and Sum of the Eigenvalues

- If $P(\lambda)$ is the characteristic polynomial of an $n\times n$ matrix A, then

$$P(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} \qquad (4)$$

- Expanding along the first column, we get

$$\det(A - \lambda I) = (a_{11} - \lambda)\det(M_{11}) + \sum_{i=2}^{n} a_{i1}(-1)^{i+1}\det(M_{i1})$$

- where the minor M_{i1} does not contain the two diagonal elements $(a_{11} - \lambda)$ and $(a_{ii} - \lambda)$

- Expanding det(M_{11}), we conclude that

  (5) $$(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$$

  is the only term in the expansion of det $(A - \lambda I)$ involving a product of more than n - 2 of the diagonal elements $\sum_{i=1}^{n} a_{ii}$
- the coefficient of λ^n is $(-1)^n$

  the coefficient of $(-\lambda)^{n-1}$ is $\boxed{\sum_{i=1}^{n}\lambda_i = \sum_{i=1}^{n} a_{ii}}$

Formulae

- If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A, then

$$P(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n)$$
$$= (\lambda_1 - \lambda)(\lambda_2 - \lambda) \ldots (\lambda_n - \lambda) \qquad (6)$$

$$\boxed{p(0) = \det(A) = \lambda_1 \lambda_2 \ldots \lambda_n}$$

- The coefficient of $(-\lambda)^{n-1}$ is $\boxed{\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}}$

- The sum of the diagonal elements of A is called the **trace** of A and is denoted by $\mathrm{tr}(A)$

Example 6

If $A = \begin{bmatrix} 5 & -18 \\ 1 & -1 \end{bmatrix}$

then det(A) = -5 + 18 = 13 and tr(A) = 5 – 1 = 4

The characteristic polynomial of A is given by

$$\begin{vmatrix} 5-\lambda & -18 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 13$$

and hence the eigenvalues of A are $\lambda_1 = 2 + 3i$ and $\lambda_2 = 2 - 3i$. Note that

$$\lambda_1 + \lambda_2 = 4 = \mathrm{tr}(A)$$
$$\lambda_1\lambda_2 = 13 = \det(A)$$

Theorem 6.1.1

Let A and B be $n \times n$ matrices. If B is similar to A ($B = S^{-1}AS$), then A and B have **the same characteristic polynomial** and consequently both have the same **same eigenvalues.**

- *Proof:*

$$
\begin{aligned}
p_B(\lambda) &= \det(B - \lambda I) \\
&= \det(S^{-1}AS - \lambda I) \\
&= \det(S^{-1}AS - S^{-1}\lambda I S) \\
&= \det(S^{-1}(A - \lambda I)\, S) \\
&= \det(S^{-1}) \det(A - \lambda I) \det(S) \\
&= \det(A - \lambda I) \\
&= p_A(\lambda)
\end{aligned}
$$

Example 7

Given $T = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ and $S = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$

$\Rightarrow$ $\det(T - \lambda I) = (2 - \lambda)\,(3 - \lambda) = 0$

$\Rightarrow$ The eigenvalues of T are $\lambda_1 = 2$, $\lambda_2 = 3$

Let $A = S^{-1}TS = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 6 & 6 \end{bmatrix}$

$\Rightarrow$ $\det(A - \lambda I) = (-1 - \lambda)\,(6 - \lambda) + 12$

$= \lambda^2 - 5\lambda + 6 = (\lambda - 2)\,(\lambda - 3) = 0$

$\Rightarrow$ The eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = 3$

6.3 Diagonalization

- Factoring an $n \times n$ matrix A into a product of the form XDX^{-1}, where D is diagonal

Theorem 6.3.1

If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of an $n \times n$ matrix A with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$., then $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are linearly independent.

▶ **Proof.** Let r be the dimension of the subspace of R^n spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_k$ and suppose that $r < k$. We may assume (reordering the $\mathbf{x}_i$'s and λ_i's if necessary) that $\mathbf{x}_1, \ldots, \mathbf{x}_r$ are linearly independent. Since $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{x}_{r+1}$ are linearly dependent, there exist scalars $c_1, \ldots, c_r, c_{r+1}$ not all zero such that

$$(1) \qquad c_1\mathbf{x}_1 + \cdots + c_r\mathbf{x}_r + c_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$$

Note that c_{r+1} must be nonzero; otherwise, $\mathbf{x}_1, \ldots, \mathbf{x}_r$ would be dependent. So $c_{r+1}\mathbf{x}_{r+1} \neq \mathbf{0}$ and hence $c_1, \ldots, c_r$ cannot all be zero. Multiplying (1) by A, we get

$$c_1A\mathbf{x}_1 + \cdots + c_rA\mathbf{x}_r + c_{r+1}A\mathbf{x}_{r+1} = \mathbf{0}$$

or

$$(2) \qquad c_1\lambda_1\mathbf{x}_1 + \cdots + c_r\lambda_r\mathbf{x}_r + c_{r+1}\lambda_{r+1}\mathbf{x}_{r+1} = \mathbf{0}$$

Subtracting λ_{r+1} times (1) from (2) gives

$$c_1(\lambda_1 - \lambda_{r+1})\mathbf{x}_1 + \cdots + c_r(\lambda_r - \lambda_{r+1})\mathbf{x}_r = \mathbf{0}$$

This contradicts the independence of $\mathbf{x}_1, \ldots, \mathbf{x}_r$. Therefore, r must equal k. ◀

Definition

An $n \times n$ matrix A is said to be **diagonalizable** if there exists a nonsingular matrix X and a diagonal matrix D such that

$$X^{-1}AX = D \quad (\text{i.e., } A = XDX^{-1})$$

We say that X **diagonalizes** A.

Theorem 6.3.2

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

Pf: (1) independent eigenvectors $\Rightarrow$ diagonalizable:

Suppose A has n linearly independent eigenvectors $\mathbf{x}_1$, $\mathbf{x}_2, \ldots, \mathbf{x}_n$

Let λ_i be the eigenvalue of A corresponding to $\mathbf{x}_i$ for each I

Let X be the matrix whose jth column vector is $\mathbf{x}_j$ for $j = 1, 2, \ldots, n$

Theorem 6.3.2 *proof*

$\Rightarrow A\mathbf{x}_j = \lambda_j\mathbf{x}_j$ is the *j*th column vector of *AX*

$$
\begin{aligned}
\Rightarrow AX &= (A\mathbf{x}_1, A\mathbf{x}_2, \cdots, A\mathbf{x}_n) \\
&= (\lambda_1\mathbf{x}_1, \lambda_2\mathbf{x}_2, \cdots, \lambda_n\mathbf{x}_n) \\
&= [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \\
&= XD
\end{aligned}
$$

Theorem 6.3.2 *proof*

Since X has n linearly independent column vectors

$\Rightarrow$ ***X*** is nonsingular

$\Rightarrow X^{-1}(AX) = X^{-1}(XD) = D$

$\Rightarrow A$ is diagonalizable

(2) diagonalizable $\Rightarrow$ independent eigenvectors:

Suppose **A** is diagonalizable

$\Rightarrow$ There exists a nonsingular matrix X such that

$$X^{-1}AX = D \text{ (i.e., } AX = XD)$$

If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the column vectors of X (i.e., $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$)

Theorem 6.3.2 *proof*

$$\Rightarrow AX = (A\mathbf{x}_1, A\mathbf{x}_2, \cdots, A\mathbf{x}_n) = XD$$

$$= (\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)\begin{bmatrix} d_{11} & & & \\ & d_{22} & & \\ & & \ddots & \\ & & & d_{nn} \end{bmatrix} = (d_{11}\mathbf{x}_1, d_{22}\mathbf{x}_2, \cdots, d_{nn}\mathbf{x}_n)$$

Theorem 6.3.2 *proof*

$\Rightarrow A\mathbf{x}_j = \lambda_j\mathbf{x}_j$ $(\lambda_j = d_{jj})$ for each j

$\Rightarrow \lambda_j$ is an eigenvalue of A and $\mathbf{x}_j$ is an eigenvector belonging to λ_j

Since the column vectors of X are linearly independent

$\Rightarrow A$ has n linearly independent eigenvectors

Remarks

1. If A is diagonalizable, then the column vectors of the diagonalizing matrix X are eigenvectors of A, and the diagonal elements of D are the corresponding eigenvalues of A.
2. The diagonalizing matrix X is not unique. *Reordering* the columns of X or *multiplying them by nonzero scalars* will produce a new diagonalizing matrix.

Remarks

3. If A is $n \times n$ and A has *n distinct eigenvalues*, then A is diagonalizable. If the eigenvalues are not distinct, then A may or may not be diagonalizable depending on whether A has n linearly independent eigenvectors.
4. If A is diagonalizable, then A can be factored into a product XDX^{-1}

- From remark 4 ($A = XDX^{-1}$)

  $A^2 = (XDX^{-1})\,(XDX^{-1}) = XD^2X^{-1}$

- In general,

$$A^k = XD^k X^{-1} = X \begin{bmatrix} (\lambda_1)^k & & & \\ & (\lambda_2)^k & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{bmatrix} X^{-1}$$

- Once we have a factorization $A = XDX^{-1}$, it is easy to compute powers of A

Example 1

- Let $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$

$\Rightarrow$ The eigenvalues of *A* are $\lambda_1 = 1$, $\lambda_2 = -4$
Corresponding to λ_1 and λ_2, the eigenvectors are

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \text{ and } \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$\Rightarrow$ Let $X = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \quad \Rightarrow \quad X^{-1} = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$

Example 1

$$\Rightarrow X^{-1}AX = \frac{1}{5}\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

$$= \frac{1}{5}\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -4 \\ 1 & -8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$$

and

$$\Rightarrow XDX^{-1} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}\begin{bmatrix} 2/5 & -1/5 \\ -1/5 & 3/5 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix} = A$$

Example 2

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

  $\Rightarrow$ The eigenvalues of *A* are $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 1$
- The eigenvector corresponding to $\lambda_1 = 0$ is ,
- $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$
- and the eigenvectors corresponding to $\lambda = 1$ are $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Example 2

$$\Rightarrow \text{Let } X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad \Rightarrow \quad X^{-1} = \begin{bmatrix} -2 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

$$\Rightarrow \quad XDX^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 & 2 \\ 1 & 0 & -1 \\ 2 & -1 & -1 \end{bmatrix}$$

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

Example 2

- Even though $\lambda = 1$ is a multiple eigenvalue, *A* is still diagonalizable since there are three linearly independent eigenvectors.

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

and

$$D^k = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = D$$

$A^k = XD^kX^{-1} = XDX^{-1} = A$ for any $k \geq 1$.

Definition

An $n \times n$ matrix A is said to be **defective** if A has fewer than n linearly independent eigenvectors.

- From Theorem 6.3.2, *a defective matrix is not diagonalizable*

Example 3

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

  $\Rightarrow$ The eigenvalues of *A* are $\lambda_1 = 1$, $\lambda_2 = 1$

- The eigenvector corresponding to $\lambda=1$ are

  $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

  $\Rightarrow$ *A* is defective and is not diagonalizable

Example 4

- Let $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & 6 & 2 \end{bmatrix}$

  $\Rightarrow$ The eigenvalues of *A* and B are of the same:

  $\lambda_1 = 4$, $\lambda_2 = \lambda_3 = 2$

  The eigenvector of *A* corresponding to λ_1=4 is $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2$

  and the eigenvectors corresponding to λ=2 is $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{e}_3$

  $\Rightarrow$ *A* is defective

Example 4

- The eigenvector of *B* corresponding to λ_1=4 is $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$

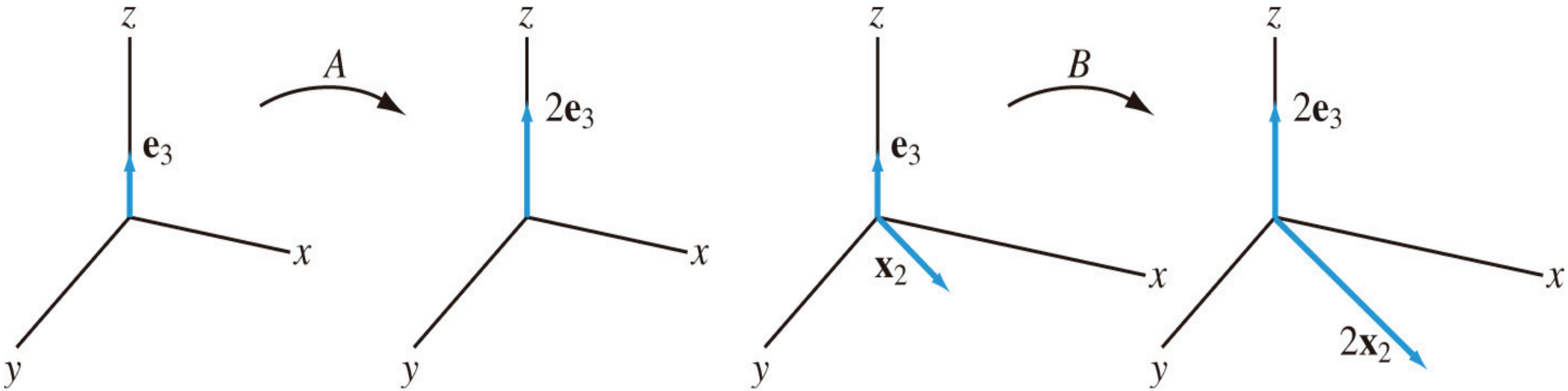
  and the eigenvectors corresponding to λ=2 is $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_3$

- $\Rightarrow$ **B** is diagonalizable

Geometrically interpretation

- The matrix B has the effect of stretching two linearly independent vectors by a factor of 2
- The eigenvalue λ=2 has *geometric multiplicity* 2 since the dimension of the eigenspace $N(B - 2I)$ is 2
- The matrix A only stretches the vectors along the z axis by a factor of 2
- The eigenvalue λ=2 has *algebraic multiplicity* 2, but dim $N(A - 2I) = 1$, *so its geometric multiplicity is only* 1

Figure 6.3.1



The Exponential of a Matrix

- Given a scalar a, the exponential e^a can be expressed in terms of a power series:

$$e^a = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} a^k$$

- Similarly, for any $n \times n$ matrix A, the *matrix exponential* e^A can be defined in terms of the convergent power series:

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots$$

- For a diagonal matrix $D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$, the matrix

  exponential is easy to compute:

$$e^D = \lim_{m \to \infty} \left(\mathrm{I} + \mathrm{D} + \frac{1}{2!} D^2 + \frac{1}{3!} D^3 + \cdots + \frac{1}{m!} D^m \right)$$

$$= \lim_{m \to \infty} \left(\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} \lambda_1^2 & & & \\ & \lambda_2^2 & & \\ & & \ddots & \\ & & & \lambda_n^2 \end{bmatrix} + \cdots + \frac{1}{m!} \begin{bmatrix} \lambda_1^m & & & \\ & \lambda_2^m & & \\ & & \ddots & \\ & & & \lambda_n^m \end{bmatrix} \right)$$

$$= \lim_{m \to \infty} \begin{bmatrix} \sum_{k=0}^{m} \frac{1}{k!} \lambda_1^k & & & \\ & \sum_{k=0}^{m} \frac{1}{k!} \lambda_2^k & & \\ & & \ddots & \\ & & & \sum_{k=0}^{m} \frac{1}{k!} \lambda_n^k \end{bmatrix}$$

$$= \begin{bmatrix} e^{\lambda_1} & & & \\ & e^{\lambda_2} & & \\ & & \ddots & \\ & & & e^{\lambda_n} \end{bmatrix}$$

- It is more difficult to compute the matrix exponential for a general $n \times n$ matrix A
- If A is diagonalizable ($A = XDX^{-1}$), then

$$A^k = XD^kX^{-1} \quad \text{for } k = 1, 2, \ldots$$

$$\begin{aligned}
e^A &= I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots \\
&= XIX^{-1} + XDX^{-1} + \frac{1}{2!}XD^2X^{-1} + \frac{1}{3!}XD^3X^{-1} + \cdots \\
&= X(I + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \cdots)X^{-1} \\
&= Xe^DX^{-1}
\end{aligned}$$

Example 6

- Compute e^A for $A = \begin{bmatrix} -2 & -6 \\ 1 & 3 \end{bmatrix}$

- *Sol:* The eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = 0$ with corresponding eigenvectors:

$$\mathbf{x}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$\Rightarrow \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $e^D = \begin{bmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{bmatrix} = \begin{bmatrix} e^1 & 0 \\ 0 & e^0 \end{bmatrix} = \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix}$

Example 6

$\Rightarrow$ *A* is diagonalizable,

$$A = XDX^{-1} = \begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix}$$

and

$$e^{A} = Xe^{D}X^{-1}$$

$$= \begin{bmatrix} -2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 3-2e & 6-6e \\ e-1 & 3e-2 \end{bmatrix}$$

6.4 Hermitian Matrix

- Let C^n denote the vector space of all n-tuples of complex numbers

Complex Inner Products

- If $\alpha = a + bi$ ($i = \sqrt{-1}$) is a complex number, the length of α is given by

$$|\alpha| = \sqrt{\bar{\alpha}\alpha} = \sqrt{(a - bi)(a + bi)} = \sqrt{a^2 + b^2}$$

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

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

- We write $\mathbf{z}^H$ for the transpose of for notational convenience, thus $\bar{\mathbf{z}}^T = \mathbf{z}^H$ and $\|\mathbf{z}\| = (\mathbf{z}^H \mathbf{z})^{1/2}$

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 **z** and **w** in *V* a complex number <**z**, **w**> satisfying the following conditions:

I. <**z**, **z**> ≥ 0 with equality if and only if **z** = **0**

II. $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ for all **z** and **w** in *V*

III. $\langle \alpha\mathbf{z} + \beta\mathbf{w}, \mathbf{u} \rangle = \alpha \langle \mathbf{z}, \mathbf{u} \rangle + \beta \langle \mathbf{w}, \mathbf{u} \rangle$

Recall: Theorem 5.5.2

If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an orthonormal basis for an 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 complex inner product space, if $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ is an orthonormal basis and

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

then

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

Definition

The inner product on C^n is defined by:

$$\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z} \qquad (1)$$

for all **z** and **w** in C^n.

- The complex inner product space C^n is similar to the real inner product space R^n
- The main difference is that in the complex case it is necessary to conjugate before transposing when taking an inner product:

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

Example 1

- If $\mathbf{z} = \begin{bmatrix} 5+i \\ 1-3i \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2+i \\ -2+3i \end{bmatrix}$

  then
$$\mathbf{w}^H \mathbf{z} = \begin{pmatrix} 2-i & -2-3i \end{pmatrix} \begin{bmatrix} 5+i \\ 1-3i \end{bmatrix}$$
$$= (2-i)(5+i) + (-2-3i)(1-3i)$$
$$= (11-3i) + (-11+3i)$$
$$= 0$$

$\Rightarrow$ **z** and **w** are <u>orthogonal</u>

Example 1

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

$$= (5-i)(5+i) + (1+3i)(1-3i)$$

$$= 26 + 10 = 36$$

$\Rightarrow$ **||z|| = 6**

$$\mathbf{w}^H\mathbf{w} = \begin{pmatrix} 2-i & -2-3i \end{pmatrix} \begin{bmatrix} 2+i \\ -2+3i \end{bmatrix}$$

$$= (2-i)(2+i) + (-2-3i)(-2+3i)$$

$$= 5 + 13 = 18$$

$$\Rightarrow \quad \|\mathbf{w}\| = \sqrt{18} = 3\sqrt{2}$$

Hermitian Matrices

- Let $M = (m_{ij})$ be an $m \times n$ matrix with $m_{ij} = a_{ij} + i\, b_{ij}$ 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.
- The conjugate of M is defined as follows:

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

- $\overline{M}$ 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 $m \times n$ matrix with complex entries is denoted by $C^{m \times n}$
- If A and B are elements of $C^{m \times n}$ and $C \in C^{m \times n}$, then the following rules are easily verified:

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

- Is the matrix $M = \begin{bmatrix} 3 & 2-i \\ 2+i & 4 \end{bmatrix}$ Hermitian?
- *Sol:*

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

$\Rightarrow$ *M* is Hermitian

Note

- If M is a matrix with real entries, then $M^H = M^T$.
- If M is a real symmetric matrix, then M is Hermitian
- We may view Hermitian matrix as the complex analogue of real symmetric matrix

Theorem 6.4.1

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

- *Pf:* (1) Let A be a Hermitian matrix
  Let λ be an eigenvalue of A and $\mathbf{x}$ be an eigenvector belonging to λ
  If $\alpha = \mathbf{x}^H A\mathbf{x}$, then

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

Theorem 6.4.1 *proof*

$\Rightarrow \alpha$ is real

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

$$\Rightarrow \quad \lambda = \frac{\alpha}{\|\mathbf{x}\|^2} \text{ is real}$$

(2) If $\mathbf{x}_1$ and $\mathbf{x}_2$ be eigenvectors belonging to distinct eigenvalue λ_1 and λ_2, then

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

and

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

Theorem 6.4.1 *proof*

$$\Rightarrow \lambda_{\mathbf{2}}\mathbf{x}_1^H\mathbf{x}_2 = \lambda_{\mathbf{1}}\mathbf{x}_1^H\mathbf{x}_2$$

Since $\lambda_{\mathbf{1}} \neq \lambda_{\mathbf{2}}$

$$\Rightarrow \mathbf{x}_1^H\mathbf{x}_2 = 0$$

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

Definition

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

- Thus, U is unitary if and only if $U^H U = 1$.
- If U is unitary, then, since the column vectors are orthonormal, U must have rank n
- $U^{-1} = IU^{-1} = U^H U U^{-1} = U^H$
- 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*.

- *Pf:* 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.
  From Theorem 6.4.1, $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is an orthonormal set in C^n
  Let $U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n)$, then U is unitary and U diagonalizes A

Note

$$U^{-1}AU = U^H AU = U^H A(\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n) = U^H (A\mathbf{u}_1, A\mathbf{u}_2, \cdots, A\mathbf{u}_n)$$

$$= (\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n)^H (\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \cdots, \lambda_n\mathbf{u}_n)$$

$$= \begin{bmatrix} \overline{\mathbf{u}}_1^T \\ \overline{\mathbf{u}}_2^T \\ \vdots \\ \overline{\mathbf{u}}_n^T \end{bmatrix} (\lambda_1\mathbf{u}_1, \lambda_2\mathbf{u}_2, \cdots, \lambda_n\mathbf{u}_n)$$

$$= \begin{bmatrix} \lambda_{\mathbf{1}}\overline{\mathbf{u}}_1^T\mathbf{u}_{\mathbf{1}} & \lambda_{\mathbf{2}}\overline{\mathbf{u}}_1^T\mathbf{u}_{\mathbf{2}} & \cdots & \lambda_n\overline{\mathbf{u}}_1^T\mathbf{u}_n \\ \lambda_{\mathbf{1}}\overline{\mathbf{u}}_2^T\mathbf{u}_{\mathbf{1}} & \lambda_{\mathbf{2}}\overline{\mathbf{u}}_2^T\mathbf{u}_{\mathbf{2}} & \cdots & \lambda_n\overline{\mathbf{u}}_2^T\mathbf{u}_n \\ \vdots & & & \\ \lambda_{\mathbf{1}}\overline{\mathbf{u}}_n^T\mathbf{u}_{\mathbf{1}} & \lambda_{\mathbf{2}}\overline{\mathbf{u}}_n^T\mathbf{u}_{\mathbf{2}} & & \lambda_n\overline{\mathbf{u}}_n^T\mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \lambda_{\mathbf{1}} & 0 & \cdots & 0 \\ 0 & \lambda_{\mathbf{2}} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \lambda_n \end{bmatrix}$$

Example 3

- Let $A = \begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix}$, find a unitary matrix U that diagonalizes A.
- *Sol:* (1) Find the eigenvalues and eigenvectors of A:

$$\det(A-\lambda I) = \begin{vmatrix} 2-\lambda & 1-i \\ 1+i & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)-(1-i)(1+i)$$

$$= (\lambda^2-3\lambda+2)-(2) = \lambda^2-3\lambda = 0$$

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

Example 3

(2) Find the unitary matrix *U*

Let $\mathbf{u}_1 = \frac{1}{\|\mathbf{x}_1\|}\mathbf{x}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1-i \\ 1 \end{bmatrix}$ and $\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2\|}\mathbf{x}_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1+i \end{bmatrix}$

Thus $U = (\mathbf{u}_1, \mathbf{u}_2) = \frac{1}{\sqrt{3}}\begin{bmatrix} 1-i & -1 \\ 1 & 1+i \end{bmatrix}$

and

$$U^H A U = \frac{1}{\sqrt{3}}\begin{bmatrix} 1+i & 1 \\ -1 & 1-i \end{bmatrix}\begin{bmatrix} 2 & 1-i \\ 1+i & 1 \end{bmatrix}\begin{bmatrix} 1-i & -1 \\ 1 & 1+i \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$$

- Actually, **Corollary 6.4.2** is valid even if the eigenvalues are not distinct

Theorem 6.4.3

(Schur's Theorem)

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

Pf: The proof is by induction on n

(1) The result is obvious if $n = 1$

(2) 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.

Theorem 6.4.3 *proof*

- Using the Gram-Schmidt process, construct $\mathbf{w}_2, \ldots, \mathbf{w}_{k+1}$ such that $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k+1}\}$ is an orthonormal basis for C^{k+1}. Let W be the matrix whose ith column vector is $\mathbf{w}_i$ for $i = 1, 2, \ldots, k+1$. Thus, by construction, W is unitary. Then

$$\begin{aligned} W^H A W &= W^H A(\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{k+1}) \\ &= \left(W^H A\mathbf{w}_1, W^H A\mathbf{w}_2, \cdots, W^H A\mathbf{w}_n\right) \end{aligned}$$

Theorem 6.4.3 *proof*

$\Rightarrow$ The first column of $W^H AW$ will be $W^H A\mathbf{w}_1$

$$W^H A\mathbf{w}_1 = W^H \lambda_1 \mathbf{w}_1 = \lambda_1 W^H \mathbf{w}_1 = \lambda_1 \begin{bmatrix} \overline{\mathbf{w}}_1^T \\ \overline{\mathbf{w}}_2^T \\ \vdots \\ \overline{\mathbf{w}}_{k+1}^T \end{bmatrix} \mathbf{w}_1$$

$$= \lambda_1 \begin{bmatrix} \overline{\mathbf{w}}_1^T \mathbf{w}_1 \\ \overline{\mathbf{w}}_2^T \mathbf{w}_1 \\ \vdots \\ \overline{\mathbf{w}}_{k+1}^T \mathbf{w}_1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 \mathbf{e}_1$$

Theorem 6.4.3 *proof*

Thus $W^H AW$ is the 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

$$V_1^H M V_1 = T_1, \text{ where } T_1 \text{ is triangular}$$

Theorem 6.4.3 *proof*

Let

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

Thus *V* is unitary and

$$V^H \underline{\underline{W^H A W}} V$$

$$= \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ \vdots & & V_1^H & & \\ 0 & & & & \end{array}\right] \left[\begin{array}{c|cccc} \lambda_1 & \times & \times & \cdots & \times \\ \hline 0 & & & & \\ \vdots & & M & & \\ 0 & & & & \end{array}\right] \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ \vdots & & V_1 & & \\ 0 & & & & \end{array}\right]$$

Theorem 6.4.3 *proof*

$$=\left[\begin{array}{c|cccc} \lambda_1 & \times & \times & \cdots & \times \\ \hline 0 & & & & \\ \vdots & & V_1^H M & & \\ 0 & & & & \end{array}\right]\left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ \vdots & & V_1 & & \\ 0 & & & & \end{array}\right]$$

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

- Let $U = WV$, the matrix U is unitary since

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

and $U^H AU = T$

Definition

The factorization $A = UTU^H$ is often referred to as the ***Schur decomposition*** of A.

- In the case that A is Hermitian ($A = A^H$), 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.

- *Pf:* By **Theorem 6.4.3**, there is a unitary matrix U such that $U^H AU = T$, where T is upper triangular. Then

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

Therefore, T is Hermitian and consequently must be diagonal

- In the case that A is real and symmetric, its eigenvalues and eigenvectors must be real. Thus, the diagonalizing matrix U must be orthogonal.

- If A is a real and symmetric matrix, then there is an orthogonal matrix U that diagonalizes A. That is, $U^H A U = D$, where D is diagonal.

Example 4

- Let $A=\begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$, find an orthogonal matrix *U* that diagonalizes *A*.

Sol: (1) Find the eigenvalues and eigenvectors of *A*:

$$\det(A-\lambda I)=\begin{vmatrix} 0-\lambda & 2 & -1 \\ 2 & 3-\lambda & -2 \\ -1 & -2 & 0-\lambda \end{vmatrix}=\begin{vmatrix} -\lambda & 2 & -1 \\ 2 & 3-\lambda & -2 \\ -1 & -2 & -\lambda \end{vmatrix}$$

$$=(-\lambda)(3-\lambda)(-\lambda)+(2)(-2)(-1)+(2)(-2)(-1)$$

$$-(-1)(3-\lambda)(-1)-(-2)(-2)(-\lambda)-(2)(2)(-\lambda)$$

$$=-\lambda^3+3\lambda^2+9\lambda+5=(1+\lambda)^2(5-\lambda)=0$$

Example 4

$\Rightarrow$ The eigenvalues of A are $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 5$

The eigenvectors corresponding to $\lambda = -1$ are $\mathbf{x}_1 = (1, 0, 1)^T$ and $\mathbf{x}_2 = (-2, 1, 0)^T$ will form a basis for the eigenspace $N(A+I)$

(2) Find the unitary matrix U:

Apply the Gram-Schmidt process, we can obtain an orthonormal basis for the eigenspace corresponding to $\lambda_1 = \lambda_2 = -1$:

Example 4

$$\mathbf{u_1} = \frac{1}{\|\mathbf{x_1}\|}\mathbf{x_1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}$$ and

$$\mathbf{p} = (\mathbf{x_2}^T\mathbf{u_1})\mathbf{u_1} = \left(\begin{bmatrix}-2 & 1 & 0\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix}\right)\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix} = \frac{-2}{\sqrt{2}}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}-1\\0\\-1\end{bmatrix}$$

$$\mathbf{x_2} - \mathbf{p} = (\mathbf{x_2}^T\mathbf{u_1})\mathbf{u_1} = \begin{bmatrix}-2\\1\\0\end{bmatrix} - \begin{bmatrix}-1\\0\\-1\end{bmatrix} = \begin{bmatrix}-1\\1\\1\end{bmatrix}$$

Example 4

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{x}_2 - \mathbf{p}\|}(\mathbf{x}_2 - \mathbf{p}) = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

- 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 $\mathbf{u}_2$, we need only normalize

$$\mathbf{u}_3 = \frac{1}{\|\mathbf{x}_3\|}\mathbf{x}_3 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

Example 4

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

$$U = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

diagonalizes A

- 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
  $\Rightarrow$ the diagonal elements of D are eigenvalues of A and the column vectors of U are eigenvectors of A.
- A cannot be defective, that is, A has a complete set of eigenvectors that form an orthonormal basis for C^n.

- If A has an orthonormal set of eigenvectors eigenvectors $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_n\mathbf{u}_n$, then

$$\begin{aligned} A\mathbf{x} &= A(c_1\mathbf{u_1} + c_2\mathbf{u_2} + \cdots + c_n\mathbf{u}_n) \\ &= c_1 A\mathbf{u_1} + c_2 A\mathbf{u_2} + \cdots + c_n A\mathbf{u}_n \\ &= c_1\lambda_1\mathbf{u_1} + c_2\lambda_2\mathbf{u_2} + \cdots + c_n\lambda_n\mathbf{u}_n \end{aligned}$$

  and

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

$$\Rightarrow \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^H \mathbf{x} \\ \mathbf{u}_2^H \mathbf{x} \\ \vdots \\ \mathbf{u}_n^H \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^H \\ \mathbf{u}_2^H \\ \vdots \\ \mathbf{u}_n^H \end{bmatrix} \mathbf{x} = U^H \mathbf{x}$$

$$\begin{aligned} \Rightarrow \quad A\mathbf{x} &= c_1\lambda_1\mathbf{u_1} + c_2\lambda_2\mathbf{u_2} + \cdots + c_n\lambda_n\mathbf{u}_n \\ &= \lambda_1 c_1\mathbf{u_1} + \lambda_2 c_2\mathbf{u_2} + \cdots + \lambda_n c_n\mathbf{u}_n \\ &= \lambda_1(\mathbf{u}_1^H\mathbf{x})\mathbf{u_1} + \lambda_2(\mathbf{u}_2^H\mathbf{x})\mathbf{u_2} + \cdots + \lambda_n(\mathbf{u}_n^H\mathbf{x})\mathbf{u}_n \end{aligned}$$

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

$$T = \begin{pmatrix} B_1 & \times & \cdots & \times \\ & B_2 & & \times \\ & O & \ddots & \\ & & & B_j \end{pmatrix} \qquad (2)$$

  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 R^n that is invariant under A.

Definition

A subspace S of 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 λ_1 = $a + b_i$ (where a and b are real and $b \neq 0$), and let $z_1 = x + i_y$ (where x and y are vectors in R^n) be an eigenvector belonging to λ_1. If S = Span($\mathbf{x}$, $\mathbf{y}$), then dim S = 2 and S is invariant under A.

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).

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, $Q^TAQ = D$, where D is diagonal.

Normal Matrices

- If A is any matrix with a complete 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, $D^H \neq D$.

$$\Rightarrow \quad A^H = (UDU^H)^H = UD^H U^H \neq A$$

- However, $AA^H = (UDU^H)(UD^H U) = UDD^H U^H$

and $A^H A = (UD^H U^H)(UDU^H) = UD^H DU^H$

- Since

$$D^H D = \begin{bmatrix} \overline{\lambda}_1 & 0 & \cdots & 0 \\ 0 & \overline{\lambda}_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \overline{\lambda}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \lambda_n \end{bmatrix} = \begin{bmatrix} \|\lambda_1\|^2 & 0 & \cdots & 0 \\ 0 & \|\lambda_2\|^2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \|\lambda_n\|^2 \end{bmatrix}$$

and

$$DD^H = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \overline{\lambda}_1 & 0 & \cdots & 0 \\ 0 & \overline{\lambda}_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \overline{\lambda}_n \end{bmatrix} = \begin{bmatrix} \|\lambda_1\|^2 & 0 & \cdots & 0 \\ 0 & \|\lambda_2\|^2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & & \|\lambda_n\|^2 \end{bmatrix}$$

$$\Rightarrow \quad AA^H = A^H A$$

Definition

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

Theorem 6.4.6

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

Pf: (1)We have shown that if a matrix has a complete set of orthonormal eigenvectors, then it is normal.

(2)From **Theorem 6.4.3**, for each $n \times n$ matrix A, there exists a unitary matrix U and a triangular matrix T such that $T = U^H A U$. Then,

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

And

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

Theorem 6.4.6 *proof*

- Since A is normal, thus $A^HA = AA^H$

  $\Rightarrow T^HT = TT^H$

  $\Rightarrow T$ is also normal (therefore, the diagonal elements of T^HT and TT^H are identical)
- Since

$$TT^H = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & & & \\ 0 & 0 & & t_{nn} \end{bmatrix} \begin{bmatrix} \bar{t}_{11} & 0 & \cdots & 0 \\ \bar{t}_{12} & \bar{t}_{22} & \cdots & 0 \\ \vdots & & & \\ \bar{t}_{1n} & \bar{t}_{2n} & & \bar{t}_{nn} \end{bmatrix}$$

Theorem 6.4.6 *proof*

and

$$T^H T = \begin{bmatrix} \bar{t}_{11} & 0 & \cdots & 0 \\ \bar{t}_{12} & \bar{t}_{22} & \cdots & 0 \\ \vdots & & & \\ \bar{t}_{1n} & \bar{t}_{2n} & & \bar{t}_{nn} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & & & \\ 0 & 0 & & t_{nn} \end{bmatrix}$$

Comparing the diagonal elements of $T^H T$ and TT^H, we see that

$$\|t_{11}\|^2 + \|t_{12}\|^2 + \|t_{13}\|^2 + \cdots + \|t_{1n}\|^2 = \|t_{11}\|^2$$

$$\|t_{22}\|^2 + \|t_{23}\|^2 + \cdots + \|t_{1n}\|^2 = \|t_{12}\|^2 + \|t_{22}\|^2$$

$$\vdots$$

$$\|t_{1n}\|^2 = \|t_{1n}\|^2 + \|t_{2n}\|^2 + \|t_{3n}\|^2 + \cdots + \|t_{nn}\|^2$$

Theorem 6.4.6 *proof*

$\Rightarrow t_{ij} = 0 \text{ whenever } i \neq j$

$\Rightarrow$ *U* Diagonalizes *A* and the column vectors of *U* are eigenvectors of *A*