# Slide 24-Diagonalization and Spectral Theorem

## MAT2040 Linear Algebra

SSE, CUHK(SZ)

# Motivation

In some applications, given $A \in \mathbb{R}^{n \times n}$, it will be useful and important to compute $A^k$.

Let $A \in \mathbb{R}^{n \times n}$, if there exists another matrix $Q$ such that $Q^{-1}AQ = \Lambda$ ($\Lambda$ is a diagonal matrix), then $A = Q\Lambda Q^{-1}$ and $A^k = Q\Lambda^k Q^{-1}$. In this case, $A$ is called diagonalizable.

**Definition 24.1** (**Diagonalizable**) A $n \times n$ matrix $A$ is said to be diagonalizable if there exists a nonsingular matrix $X$ and a diagonal matrix $D$ such that

$$X^{-1}AX = D$$

Thus, $A = XDX^{-1}$, and $X$ diagonalizes $A$.

We will give a necessary and sufficient condition for the existence of such a decomposition.

First, we will show that eigenvectors belonging to distinct eigenvalues are linearly independent.

**Theorem 24.2** (**Eigenvectors belonging to distinct eigenvalues are linearly independent**)
If $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct eigenvalues of the $n \times n$ matrix $A$, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ are the eigenvectors corresponding to $\lambda_1, \lambda_2, \cdots, \lambda_k$, respectively, then $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k$ are linearly independent.

**Proof.** Suppose that there exists $c_1, \ldots, c_k \in \mathbb{R}$, such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \cdots + c_{k-1}\mathbf{x}_{k-1} + c_k\mathbf{x}_k = \mathbf{0} \tag{1}$$

Then

$$A(c_1\mathbf{x}_1 + \cdots + c_n\mathbf{x}_k) = c_1\lambda_1\mathbf{x}_1 + c_2\lambda_2\mathbf{x}_2 + \cdots + c_k\lambda_k\mathbf{x}_k = \mathbf{0}. \tag{2}$$

(2)-$\lambda_k\times$ (1) to cancel $\mathbf{x}_k$:

$$c_1(\lambda_1 - \lambda_k)\mathbf{x}_1 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\mathbf{x}_{k-1} = \mathbf{0}. \tag{3}$$

$$
\begin{aligned}
&A(c_1(\lambda_1-\lambda_k)\mathbf{x}_1+\cdots+c_{k-1}(\lambda_{k-1}-\lambda_k)\mathbf{x}_{n-1})\\
=&c_1(\lambda_1-\lambda_k)\lambda_1\mathbf{x}_1+\cdots+c_{k-1}(\lambda_{k-1}-\lambda_k)\lambda_{k-1}\mathbf{x}_{k-1}=\mathbf{0}.
\end{aligned}
\tag{4}
$$

(4)-$\lambda_{k-1}\times$ (3) to cancel $\mathbf{x}_{k-1}$:

$$
\begin{aligned}
&c_1(\lambda_1-\lambda_k)(\lambda_1-\lambda_{k-1})\mathbf{x}_1+\cdots+c_{k-2}(\lambda_{k-2}-\lambda_k)(\lambda_{k-2}-\lambda_{k-1})\mathbf{x}_{k-2}\\
=&\mathbf{0}.
\end{aligned}
\tag{5}
$$

We can continue this process to cancel $\mathbf{x}_{k-2},\ldots,\mathbf{x}_2$ to get:

$$
c_1(\lambda_1-\lambda_k)(\lambda_1-\lambda_{k-1})\ldots(\lambda_1-\lambda_2)\mathbf{x}_1=\mathbf{0} \qquad \text{which forces } c_1=0.
$$

Similarly, $c_i=0$ for $i=2,\ldots,k$. Thus, $c_1=c_2=\cdots=c_k=0$.

Thus, $\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_k$ are linearly independent.

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

The characteristic equation is

$$p_A(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -3 \\ 2 & -5-\lambda \end{vmatrix} = (2-\lambda)(-5-\lambda) + 6 = 0.$$

The eigenvalues are $\lambda_1 = 1, \lambda_2 = -4$.

Through some calculations and choose $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ as the eigenvector w.r.t $\lambda_1 = 1$. And choose $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as the eigenvector w.r.t $\lambda_2 = -4$.

Then $\mathbf{x}_1, \mathbf{x}_2$ are linearly independent.

**Theorem 24.3 (Sufficient and Necessary Condition for Diagonalization)** A $n \times n$ matrix $A$ is diagonalizable if and only if $A$ has $n$ linearly independent eigenvectors.

**Proof.** $\Leftarrow$ Suppose that $A$ has $n$ linearly independent eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$, then we have

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i$$

where the associated eigenvalue for the eigenvector $\mathbf{x}_i$ is $\lambda_i$ $(i = 1, \cdots, n)$. Thus

$$\begin{aligned}
&A[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] \\
=&[\lambda_1 \mathbf{x}_1, \lambda_2 \mathbf{x}_2, \cdots, \lambda_n \mathbf{x}_n] \\
=&[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)
\end{aligned}$$

Let $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, then $X$ is nonsingular since $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are linearly independent. Thus $X^{-1}AX = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$. Thus $A$ is diagonalizable.

$\Rightarrow$ Suppose that $A$ is diagonalizable, there exists a nonsingular matrix $X = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, s.t. $X^{-1}AX = D = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, then

$$A[\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n] = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$$

Thus

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad i = 1, \cdots, n.$$

Thus, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ are eigenvectors. Moreover, since $X$ is nonsingular, thus $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ is linearly independent.

**Remark**

1. If $A$ is diagonlizable, then the column vectors of the diagonalizing matrix $X$ are eigenvectors of $A$ and the diagonal elements of $D$ are the corresponding eigenvalues.

2. The diagonalizing matrix $X$ is not unique, reordering the columns of a given diagonalizing matrix or multiplying nonzero scalars for columns will product a new diagonalizing matrix.

3. $X^{-1}AX = D \Rightarrow A = XDX^{-1}$, $A^2 = XD^2X^{-1}$ and $A^k = XD^kX^{-1}$

**Example 24.4** Let $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$ The eigenvalues are $\lambda_1 = 1, \lambda_2 = -4$, the corresponding eigenvectors are chosen as

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

respectively.

Let $X = [\mathbf{x}_1, \mathbf{x}_2]$ and $D = \begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix}$, then

$$X^{-1}AX = D, \quad A = XDX^{-1}$$

Thus

$$A^{100} = XD^{100}X^{-1} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1^{100} & 0 \\ 0 & (-4)^{100} \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1}$$

The eigenvectors are not unique. Let

$$\mathbf{y}_1 = 2\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}6\\2\end{bmatrix}, \quad \mathbf{y}_2 = -1\begin{bmatrix}1\\2\end{bmatrix} = \begin{bmatrix}-1\\-2\end{bmatrix}$$

be the eigenvectors w.r.t. $\lambda = 1$ and $\lambda = -4$, respectively.
Let $Y_1 = [\mathbf{y}_1, \mathbf{y}_2]$ and $D = \begin{bmatrix}1 & 0\\0 & -4\end{bmatrix}$, then

$$Y_1^{-1} A Y_1 = D, \quad A = Y_1 D Y_1^{-1}$$

Let $Y_2 = [\mathbf{y}_2, \mathbf{y}_1]$ and $\Lambda = \begin{bmatrix}-4 & 0\\0 & 1\end{bmatrix}$, then

$$Y_2^{-1} A Y_2 = \Lambda, \quad A = Y_2 \Lambda Y_2^{-1}$$

**Corollary** Let $A$ be a matrix with size $n \times n$, if $A$ has $n$ distinct eigenvalues and the eigenvectors belonging to the $n$ distinct eigenvalues are $\mathbf{x}_1, \cdots, \mathbf{x}_n$, then $\mathbf{x}_1, \cdots, \mathbf{x}_n$ are linearly independent eigenvectors (by using the Theorem 24.2). Thus, $A$ is diagonalizable (by using the Theorem 24.3) and $P = [\mathbf{x}_1, \cdots, \mathbf{x}_n]$ diagonalizes $A$ ($P^{-1}AP = \Lambda$, where $\Lambda$ is a diagonal matrix).

**Example 24.4** For $A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}$. The eigenvalues are $\lambda_1 = 1, \lambda_2 = -4$, which are distinct. Thus, $A$ is diagonalizable.

However, if $A$ has eigenvalues with multiplicity $\geq 2$, then $A$ may or may not be diagonalizable, depending on whether $A$ has $n$ linearly independent eigenvectors or not.

**Example 24.5** Let

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

It has the eigenvalues $\lambda = 1$ with multiplicity 2. The eigenspace for $\lambda = 1$ is **Span** $\left\{ \left[ \begin{array}{c} 1 \\ 0 \end{array} \right] \right\}$, it does not has two linearly independent eigenvectors. Thus $A$ is not diagonalizable.

**Remark.** Square matrices that is not diagonalizable is called **defective** matrices.

**Example 24.6**

(1) $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{bmatrix}$, $p_A(\lambda) = (2-\lambda)^2(4-\lambda), \lambda_1 = 4, \lambda_2 = 2$.

The eigenspace for $\lambda_1 = 4$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$. $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ is the eigenvector w.r.t $\lambda_1 = 4$.

The eigenspace for $\lambda_2 = 2$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is the eigenvector w.r.t $\lambda_2 = 2$.

By Theorem 24.2, $\mathbf{x}_1, \mathbf{x}_2$ are linearly independent, but we cannot find three linearly independent eigenvectors. Thus, $A$ is not diagonalizable, it is defective.

(2) $A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & 6 & 2 \end{bmatrix}$, $\quad p_A(\lambda) = (2-\lambda)^2(4-\lambda), \lambda_1 = 4, \lambda_2 = 2.$

The eigenspace for $\lambda_1 = 4$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \right\}$, $\quad \mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ is the eigenvector w.r.t $\lambda_1 = 4$.

The eigenspace for $\lambda_2 = 2$ is **Span** $\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$, $\quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ and

$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent eigenvectors w.r.t $\lambda_2 = 2$.

It can be easily checked that $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ are linearly independent. There are three linearly independent eigenvectors. Thus, $A$ is diagonalizable.

**Observation:** The real square matrix may have complex eigenvalues, and the real square matrix may or may not be diagonalizable.

**However, the real symmetric matrix has very good properties. The real symmetric matrix must have real eigenvalues, and real symmetric matrix must be diagonalizable.**

Without given the detail proof, we will provide the following two important theorems for real symmetric matrices.

**Theorem 24.7** Let $A \in \mathbb{R}^{n\times n}$ be the real symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then
(1) $\lambda_i \in \mathbb{R}, \ \forall\ i = 1, \cdots, n$. (**The eigenvalues of real symmetric matrices are real numbers.**)
(2) If $\lambda_i \neq \lambda_j$, $\mathbf{x}_i$ is the eigenvectors w.r.t $\lambda_i$, $\mathbf{x}_j$ is the is the eigenvectors w.r.t $\lambda_j$, then $\mathbf{x}_i$, $\mathbf{x}_j$ are orthogonal.(**For real symmetric matrices, the eigenvectors belonging to different eigenvalues are orthogonal.**)

**Remark:** For any real square matrix, the eigenvalue may not be the real number. But if the matrix is real symmetric, then the eigenvalues of the matrix must be the real numbers.

Proof.

Only show the second part.

(2) $A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \ A\mathbf{x}_j = \lambda_j \mathbf{x}_j$

$\Rightarrow \mathbf{x}_j^T A\mathbf{x}_i = \lambda_i \mathbf{x}_j^T \mathbf{x}_i$, $\mathbf{x}_i^T A\mathbf{x}_j = \lambda_j \mathbf{x}_i^T \mathbf{x}_j$

$\Rightarrow (\mathbf{x}_j^T A\mathbf{x}_i)^T = \mathbf{x}_i^T A^T \mathbf{x}_j = \mathbf{x}_i^T A\mathbf{x}_j$

$\Rightarrow \lambda_i \mathbf{x}_i^T \mathbf{x}_j = (\lambda_i \mathbf{x}_j^T \mathbf{x}_i)^T = \lambda_j \mathbf{x}_i^T \mathbf{x}_j$ ($\lambda_i$ are real numbers)

$\Rightarrow (\lambda_i - \lambda_j)\mathbf{x}_i^T \mathbf{x}_j = 0$ ($\lambda_i \neq \lambda_j$)

$\Rightarrow \mathbf{x}_i^T \mathbf{x}_j = 0$ ($\mathbf{x}_i$ and $\mathbf{x}_j$ are orthogonal).

**Theorem 24.8** (**Spectral Theorem (eigen decomposition theorem) for Real Symmetric Matrix**) If $A$ is a real symmetric matrix, then there exists an orthogonal matrix $Q$ that diagonalizes $A$. i.e.,

$$Q^{-1}AQ = Q^{T}AQ = \Lambda$$

where $\Lambda$ is the diagonal matrix.

**Proof.** Skipped. This theorem is a special case of Hermitian matrix. See the theorem 24.24 for results of the Hermitian matrix.

Suppose the symmetric matric $A \in \mathbb{R}^{n \times n}$ and $Q^{-1}AQ = \Lambda$ ($\Lambda$ is a diagonal matrix), where $Q$ is an orthogonal matrix and $\Lambda$ is the diagonal matrix($\Lambda = diag(\lambda_1, \cdots, \lambda_n)$, $\lambda_1, \cdots, \lambda_n$ are real numbers), then

$$A = Q\Lambda Q^T = [\mathbf{q}_1, \cdots, \mathbf{q}_n] diag(\lambda_1, \cdots, \lambda_n) \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$$

Thus

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

This is called the **eigendecomposition** or **eigenvalue decomposition**.

**Example 24.9** Let

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

Find an orthogonal matrix $U$ that diagonalizes $A$.

**Solution**

$$p_A(\lambda) = \det(A - \lambda I_3) = \det\left(\begin{bmatrix} -\lambda & 2 & -1 \\ 2 & 3-\lambda & -2 \\ -1 & -2 & -\lambda \end{bmatrix}\right)$$

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

$$= -\lambda(-\lambda(3-\lambda) - 4) - 2(-2\lambda - 2) - (-4 + 3 - \lambda)$$

$$= -\lambda^3 + 3\lambda^2 + 9\lambda + 5 = -(\lambda^2 + 2\lambda + 1)\lambda + 5\lambda^2 + 10\lambda + 5$$

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

Eigenvalues are $\lambda_1 = \lambda_2 = -1, \lambda_3 = 5$.
The eigenspace for $\lambda = -1$ is

**Span** $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \right\}$. The eigenvectors $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ are not orthogonal.

Using Gram-Schmidt process, one has: $\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$\mathbf{u}_2' = \mathbf{u}_2 - < \mathbf{u}_2, \mathbf{v}_1 > \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{\mathbf{u}'_2}{\|\mathbf{u}'_2\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

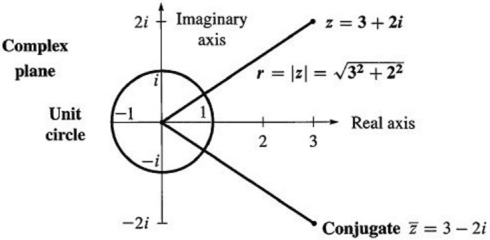
$\mathbf{v}_1, \mathbf{v}_2$ are orthonormal eigenvectors w.r.t. $\lambda = -1$.

The eigenspace for $\lambda = 5$ is **Span**$\left\{ \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \right\}$. $\mathbf{u}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ is the eigenvector w.r.t $\lambda = 5$.
The normalization of this vector is

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

$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is an orthonormal set in $\mathbb{R}^3$. Let
$U = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$ be the orthogonal matrix. Then
$U^{-1}AU = U^TAU = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

# Complex Number Field and Inner Product




$\mathbb{C}$ is the complex number field. And

$$\mathbb{C}^n = \left\{ \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \middle| u_i \in \mathbb{C} \right\}$$

# Complex Number Field and Inner Product

For $\alpha = a + bi \in \mathbb{C}$ ($i = \sqrt{-1}$ is the complex unit), the complex conjugate is given by $\bar{\alpha} = a - bi$, and the modulus of $\alpha$ is given by $|\alpha| = \sqrt{a^2 + b^2} = \sqrt{\alpha\bar{\alpha}}$.
For $\mathbf{z} = [z_1, z_2, \cdots, z_n]^T \in \mathbb{C}^n$, the length of $\mathbf{z}$ is given by

$$
\begin{aligned}
\| \mathbf{z} \| &= \sqrt{|z_1|^2 + \cdots + |z_n|^2} \\
&= \sqrt{\bar{z}_1 z_1 + \cdots + \bar{z}_n z_n} \\
&= \sqrt{\bar{\mathbf{z}}^T \mathbf{z}} \\
&= \sqrt{\mathbf{z}^H \mathbf{z}}
\end{aligned}
$$

where the notation $\mathbf{z}^H = \bar{\mathbf{z}}^T$ is the **conjugate transpose** of $\mathbf{z}$.

The scalar product in $\mathbb{C}^n$ is defined by $\mathbf{w}^H\mathbf{z}$ for $\mathbf{z}, \mathbf{w} \in \mathbb{C}^n$. It will be an excise to check that

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

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

| $\mathbb{R}^n$ | $\mathbb{C}^n$ |
|---|---|
| $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$ | $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^H\mathbf{x}$ |
| $\mathbf{y}^T\mathbf{x} = \mathbf{x}^T\mathbf{y}$ | $\mathbf{y}^H\mathbf{x} = \overline{\mathbf{x}^H\mathbf{y}}$ |
| $\Vert \mathbf{x} \Vert^2 = \mathbf{x}^T\mathbf{x}$ | $\Vert \mathbf{x} \Vert^2 = \mathbf{x}^H\mathbf{x}$ |
| $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \mathbf{x}^T\mathbf{y} = 0$ | $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \mathbf{y}^H\mathbf{x} = 0$ |

**Example 24.11** Let

$$\mathbf{u} = \begin{bmatrix} 2+i \\ -2+3i \end{bmatrix}, \quad \mathbf{z} = \begin{bmatrix} 5+i \\ 1-3i \end{bmatrix}$$

Then

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

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

$\mathbf{u}$ and $\mathbf{z}$ are orthogonal, i.e., $\mathbf{u} \perp \mathbf{z}$.

$$\mathbf{z}^H\mathbf{z} = |5+i|^2 + |1-3i|^2 = (25+1) + (1+9) = 36, \quad \| \mathbf{z} \| = 6$$

$$\mathbf{u}^H\mathbf{u} = |2+i|^2 + |-2+3i|^2 = (2^2+1) + ((-2)^2+3^2) = 18, \quad \| \mathbf{u} \| = 3\sqrt{2}$$

# Property 24.12

Let $A, B \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{n\times l}$, then

(1) $(A^H)^H = A$.

(2) $(\alpha A + \beta B)^H = \bar{\alpha} A^H + \bar{\beta} B^H$.

(3) $(AC)^H = C^H A^H$.

**Definition 24.13 (Hermitian Matrix)** Let $M = A + iB = (m_{ij})_{m\times n}$, where $A, B \in \mathbb{R}^{n\times n}$, then conjugate of $M$ is the matrix $\bar{M} = A - iB = (\bar{m}_{ij})_{m\times n}$. Define

$$M^H = (\bar{M})^T$$

If $M^H = M$, then we say that $M$ is a Hermitian matrix.
E.g.

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

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

# Unitary Matrix

**Definition 24.14** (**Unitary Matrix**) A matrix $U \in \mathbb{C}^{n\times n}$ is said to be **unitary matrix** if its column vectors form an orthonormal set in $\mathbb{C}^n$. In particular, if $U \in \mathbb{R}^{n\times n}$ and its column vectors form an orthonormal set in $\mathbb{R}^n$, then $U$ is called the **orthogonal matrix**.

**Remark.** $U$ is unitary$\Leftrightarrow$ $U^H U = I \Leftrightarrow U^{-1} = U^H$.
A real unitary matrix is the orthogonal matrix.

**Theorem 24.15** Let $M \in \mathbb{C}^{n\times n}$ be the Hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then
(1) $\lambda_i \in \mathbb{R},\ \forall\ i = 1, \cdots, n$.
(2) If $\lambda_i \neq \lambda_j$, then the eigenvectors $\mathbf{x}_i, \mathbf{x}_j$ w.r.t $\lambda_i, \lambda_j$ are orthogonal.

Proof is omitted.

**Remark:** Although the eigenvalues of the Hermitian matrix $A \in \mathbb{C}^{n\times n}$ are real numbers, the corresponding eigenvectors have complex entries and belong to $\mathbb{C}^n$.

**Proof.** (1) Let $\mathbf{x}$ be an eigenvector w.r.t. $\lambda$, then $M\mathbf{x} = \lambda\mathbf{x}$. Multiplying both sides by $\mathbf{x}^H$, one has

$$\mathbf{x}^H M\mathbf{x} = \lambda\mathbf{x}^H\mathbf{x} = \lambda \parallel \mathbf{x} \parallel^2$$

Thus

$$(\mathbf{x}^H M\mathbf{x})^H = (\lambda\mathbf{x}^H\mathbf{x})^H = \bar{\lambda}\mathbf{x}^H\mathbf{x} = \bar{\lambda} \parallel \mathbf{x} \parallel^2$$

On the other hand,

$$(\mathbf{x}^H M\mathbf{x})^H = \mathbf{x}^H M^H(\mathbf{x}^H)^H = \mathbf{x}^H M\mathbf{x}$$

Thus

$$\lambda \parallel \mathbf{x} \parallel^2 = \bar{\lambda} \parallel \mathbf{x} \parallel^2$$

In addition, $\mathbf{x}^H\mathbf{x} > 0$ since $\mathbf{x} \neq \mathbf{0}$. Thus $\lambda = \bar{\lambda}$, $\lambda$ must be a real number.

(2) $M\mathbf{x}_i = \lambda_i \mathbf{x}_i, \ M\mathbf{x}_j = \lambda_j \mathbf{x}_j$
$\Rightarrow \mathbf{x}_j^H M\mathbf{x}_i = \lambda_i \mathbf{x}_j^H \mathbf{x}_i$, $\mathbf{x}_i^H M\mathbf{x}_j = \lambda_j \mathbf{x}_i^H \mathbf{x}_j$
$\Rightarrow (\mathbf{x}_j^H M\mathbf{x}_i)^H = \mathbf{x}_i^H M^H \mathbf{x}_j = \mathbf{x}_i^H M\mathbf{x}_j$
$\Rightarrow \lambda_i \mathbf{x}_i^H \mathbf{x}_j = (\lambda_i \mathbf{x}_j^H \mathbf{x}_i)^H = \lambda_j \mathbf{x}_i^H \mathbf{x}_j$ ($\lambda_i$ are real numbers)
$\Rightarrow (\lambda_i - \lambda_j)\mathbf{x}_i^H \mathbf{x}_j = 0$ ($\lambda_i \neq \lambda_j$)
$\Rightarrow \mathbf{x}_i^H \mathbf{x}_j = 0$ ($\mathbf{x}_i$ and $\mathbf{x}_j$ are orthogonal).

**Corollary 24.16** If the eigenvalues of a Hermitian matrix $A \in \mathbb{C}^{n\times n}$ are distinct, then there exists a unitary matrix $U$ that diagonalizes $A$.

**Proof.** Let $\mathbf{x}_i$ be an eigenvector w.r.t. $\lambda_i$ $(i = 1, \cdots, n)$. Let $\mathbf{u}_i = \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}$, then $\mathbf{u}_i$ is the unit eigenvector w.r.t. eigenvalue $\lambda_i$. Since all $\lambda_i$ $(i = 1, \cdots, n)$ are distinct, then all $\mathbf{u}_i (i = 1, \cdots, n)$ are orthogonal by above theorem 24.20. Thus, $\mathbf{u}_i (i = 1, \cdots, n)$ is an orthonormal set in $\mathbb{C}^n$. Then the matrix whose column vectors are $\mathbf{u}_i (i = 1, \cdots, n)$ is the unitary matrix and diagonalize $A$.

**Remark:** The above theorem is still valid only under the condition the matrix is Hermitian. See the following theorem 24.24.

**Example 24.17** Let

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

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

thus $\lambda_1 = 3, \lambda_2 = 0$ are eigenvalues. The corresponding eigenvectors are $\mathbf{x}_1 = [1-i, 1]^T$ and $\mathbf{x}_2 = [-1, 1+i]^T$

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

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

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

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

**Theorem 24.18** (**Schur's decomposition** ) For each $n \times n$ matrix $A$, there exists a unitary matrix $U$ such that $U^H AU$ is upper triangular. i.e., $A = UTU^H$($T$ is an upper triangular matrix.)

**Proof.** We prove it by induction. $n = 1$ obviously valid.
Suppose that the conclusion is valid for $n = k$, then we look for $n = k + 1$. Let $A \in \mathbb{C}^{(k+1)\times(k+1)}$ and $\lambda_1$ be an eigenvalue of $A$ with the corresponding eigenvector $\mathbf{w}_1$ with unit length $\|\mathbf{w}_1\| = 1$. Start with $\mathbf{w}_1$, one can first find a set of $k + 1$ linearly independent vectors containing $\mathbf{w}_1$, and then use the Gram-Schmidt process, one can construct $\mathbf{w}_2, \cdots, \mathbf{w}_{k+1}$ such that $\{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{k+1}\}$ is an orthonormal set. Let

$$W = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{k+1}]$$

then $A\mathbf{w}_1 = \lambda_1 \mathbf{w}_1$. Thus
$AW = A[\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{k+1}] = [\lambda \mathbf{w}_1, A\mathbf{w}_2, \cdots, A\mathbf{w}_{k+1}]$

$$W^H A W = \begin{bmatrix} \mathbf{w}_1^T \\ \vdots \\ \mathbf{w}_n^T \end{bmatrix} [\lambda \mathbf{w}_1, A\mathbf{w}_2, \cdots, A\mathbf{w}_{k+1}] = \left[\begin{array}{c|ccc} \lambda_1 & \times & \cdots & \times \\ \hline 0 & & & \\ \vdots & & M & \\ 0 & & & \end{array}\right]$$

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

Now let

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

then

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

which is also the upper triangular matrix.
Let $U = WV$, then $U^H U = V^H W^H W V = I$. Thus $U$ is the unitary matrix.
$U^H A U = T$ is referred as the **Schur decomposition**.

**Theorem 24.19** (**Spectral Theorem for Complex Matrix**) If $A$ is Hermitian, then there exists a unitary matrix $U$ that diagonalizes $A$.

**Proof.** By above theorem, there is a unitary matrix $U$ such that $U^{-1}AU = U^H AU = T$, where $T$ is an upper triangular matrix. Furthermore,

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

Thus $T$ is Hermitian, $T$ must be a diagonal matrix. Moreover, the diagonal entries of $T$ are real numbers, which corresponding to the eigenvalues of $A$.

**Remark.** For the real symmetric matrix $A \in \mathbb{R}^{n\times n}$, its eigenvalues are real numbers, and the corresponding eigenvectors must have real entries and belong to $\mathbb{R}^n$, this is already given theorem 24.8.

**Theorem 24.20** (**Spectral Theorem (eigen decomposition theorem) for Real Symmetric Matrix**) If $A$ is a real symmetric matrix, then there exists an orthogonal matrix $U$ that diagonlizes $A$.

Suppose the symmetric matric $A \in \mathbb{R}^{n \times n}$ and $Q^{-1}AQ = \Lambda$, where $Q$ is an orthogonal matrix and $\Lambda$ is the diagonal matrix, then

$$A = Q\Lambda Q^T = [\mathbf{q}_1, \cdots, \mathbf{q}_n] diag(\lambda_1, \cdots, \lambda_n) \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix}$$

Thus

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

This is called the **eigendecomposition** or **eigenvalue decomposition**.

Recall: for two vectors $\mathbf{x}, \mathbf{q} \in \mathbb{R}^n$, where $\mathbf{q}$ is the unit vector, the projection of $\mathbf{x}$ onto $\mathbf{q}$ is $\mathbf{p} = \frac{<\mathbf{x},\mathbf{q}>}{<\mathbf{q},\mathbf{q}>}\mathbf{q} = \frac{\mathbf{q}^T\mathbf{x}}{\|\mathbf{q}\|^2}\mathbf{q} = \mathbf{q}\mathbf{q}^T\mathbf{x}$. Thus, the projection matrix onto $\mathbf{q}$ is $\mathbf{q}\mathbf{q}^T$. Therefore,

$$A = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^T + \cdots + \lambda_n \mathbf{q}_n \mathbf{q}_n^T$$

states that the real symmetric matrix $A$ is a linear combination of projection matrices $(\mathbf{q}_1\mathbf{q}_1^T, \cdots, \mathbf{q}_n\mathbf{q}_n^T)$.

If we write $\boldsymbol{A}$ as a combination of projection matrix, we can have a deep understanding for $\boldsymbol{Ax}$:

For any $\mathbf{x} \in \mathbb{R}^n$, let $\mathbf{x} = \sum_{i=1}^n c_i \mathbf{q}_i$ ($\{\mathbf{q}_1, \cdots, \mathbf{q}_n\}$ are orthonormal basis of $\mathbb{R}^n$.)

$$\boldsymbol{Ax} = \sum_{j=1}^n \lambda_j \mathbf{q}_j \mathbf{q}_j^T \sum_{i=1}^n c_i \mathbf{q}_i = \sum_{i=1}^n \sum_{j=1}^n \lambda_j c_i \mathbf{q}_j (\mathbf{q}_j^T \mathbf{q}_i) = \sum_{i=1}^n \lambda_i c_i \mathbf{q}_i.$$

since $\mathbf{q}_j^T \mathbf{q}_i = \delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$

If we set $n = 2$, it's clear to find that

$$\boldsymbol{x} = c_1\mathbf{q}_1 + c_2\mathbf{q}_2 \implies \boldsymbol{Ax} = \lambda_1 c_1 \boldsymbol{q}_1 + \lambda_2 c_2 \boldsymbol{q}_2$$

Showing in graph, we have

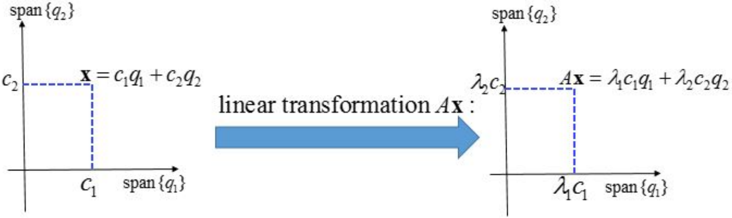



Figure: Linear transformation of $\boldsymbol{A}$.

**Theorem 24.21** Let $A \in \mathbb{R}^{n\times n}$ be the real symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then
(1) $\lambda_i \in \mathbb{R}, \ \forall\ i = 1, \cdots, n$.
(2) If $\lambda_i \neq \lambda_j$, then the eigenvectors $\mathbf{x}_i, \mathbf{x}_j$ w.r.t $\lambda_i, \lambda_j$ are orthogonal.

This theorem is a special case of Theorem 24.15.

**Remark** Theorem 24.20 and 24.21 are two important results for real symmetric matrices.

**Example 24.22** Let

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

Find an orthogonal matrix $U$ that diagonalize $A$.
First compute the eigenvalues, they are $\lambda_1 = \lambda_2 = -1, \lambda_3 = 5$. The eigenvector space for $\lambda = -1$ is

$$< [1, 0, 1]^T, [-2, 1, 0]^T >$$

$[1, 0, 1]^T, [-2, 1, 0]^T$ are not orthogonal, Using Gram-Schmidt process, one has:

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}}[1, 0, 1]^T$$

$$\mathbf{u}_2' = \mathbf{u}_2 - < \mathbf{u}_2, \mathbf{v}_1 > \mathbf{v}_1 = [-1, 1, 1]^T$$

$$\mathbf{v}_2 = \frac{1}{\sqrt{3}}[-1, 1, 1]^T$$

The eigenvector space for $\lambda = 5$ is

$$< [-1, -2, 1]^T >$$

The normalization of this vector is

$$\mathbf{v}_3 = \frac{1}{\sqrt{6}}[-1, -2, 1]^T$$

Let

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

Then $U^T A U = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

**Definition 24.23** (**Geometric multiplicity**)
Let $\lambda_0$ be an eigenvalue of the $n \times n$ square matrix $A$. The **geometric multiplicity** w.r.t. the eigenvalue $\lambda_0$ is the Nullity of $A - \lambda_0 I$, denoted by $\gamma_A(\lambda_0)$, i.e., $\gamma_A(\lambda_0) = \dim(Null(A - \lambda_0 I))$.

**Definition 24.24** (**Algebraic multiplicity**)
Let $\lambda_0$ be an eigenvalue of the $n \times n$ square matrix $A$ and the characteristic polynomial is $p_A(\lambda)$.

The **algebraic multiplicity** w.r.t. the eigenvalue $\lambda_0$ is the integer $\mu_A(\lambda_0)$ such that $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)}$ is a factor of $p_A(\lambda)$ while $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)+1}$ is not a factor of $p_A(\lambda)$. $\mu_A(\lambda_0)$ is the highest power such that $(\lambda_0 - \lambda)^{\mu_A(\lambda_0)}$ is a factor of $p_A(\lambda)$.

**Remark:** For the $n \times n$ square matrix $A$, suppose that $\lambda_1, \cdots, \lambda_s$ are all distinct eigenvalues of $A$ and
$p_A(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)^{m_1}(\lambda_2 - \lambda)^{m_2} \cdots (\lambda_s - \lambda)^{m_s}$, where $m_1, \cdots, m_s$ are positive integers and $m_1 + \cdots + m_s = n$. Then
$\mu_A(\lambda_1) = m_1, \cdots, \mu_A(\lambda_s) = m_s$.

**Example**

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

$p_A(\lambda) = (1 - \lambda)^2$, then $\mu_A(1) = 2$.

The eigenspace w.r.t. $\lambda = 2$ is **Span** $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$, $\gamma_A(1) = 1$.

One has $\mu_A(1) = 2$ but $\gamma_A(1) = 1$.

**Theorem 24.25** Let $\lambda$ be an eigenvalue of $n \times n$ matrix $A$, then

$$1 \leq \gamma_A(\lambda) \leq \mu_A(\lambda) \leq n$$

**Proof.** Skipped.

**Theorem 24.26** (**Sufficient and Necessary Condition for Diagonalization**) Let $A$ be a square matrix with size $n \times n$, then $A$ is diagonalizable $\Leftrightarrow \gamma_A(\lambda) = \mu_A(\lambda)$ for any eigenvalue $\lambda$.
**Proof.** Skipped.

**Example 24.27**

$$(1) A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{bmatrix}, \quad p_A(\lambda) = (2-\lambda)^2(4-\lambda), \text{ then}$$

$\mu_A(4) = 1, \mu_A(2) = 2.$

The eigenspace w.r.t. $\lambda_1 = 4$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}, \gamma_A(4) = 1 = \mu_A(4).$

The eigenspace w.r.t. $\lambda_2 = 2$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}, \gamma_A(2) = 1 < \mu_A(2) = 2.$

Thus, $A$ cannot be diagonalizable, it is defective.

$$(2) A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 4 & 0 \\ -3 & 6 & 2 \end{bmatrix}, \quad p_A(\lambda) = (2-\lambda)^2(4-\lambda), \text{ then}$$

$\mu_A(4) = 1, \mu_A(2) = 2.$

The eigenspace w.r.t. $\lambda_1 = 4$ is **Span** $\left\{ \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \right\}$, $\gamma_A(4) = 1 = \mu_A(4)$.

The eigenspace w.r.t. $\lambda_2 = 2$ is **Span** $\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$,

$\gamma_A(2) = 2 = \mu_A(2)$.

Thus, $A$ can be diagonalizable.

# Normal matrix

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

Examples:

1. Unitary matrices, Hermitian matrices.
2. Orthogonal symmetric and skew symmetric matrices.

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

is a normal matrix.

# Normal matrix

**Theorem: A matrix $A$ is diagonalizable by a unitary matrix if and only if $A$ is normal.**

**Proof.** We have shown that necessary condition and now prove sufficiency by Schur's theorem. There exists a unitary matrix $U$ such that $T = U^H AU$ is upper triangular.

$$T^H T = U^H A^H UU^H AU = U^H A^H AU$$
$$TT^H = U^H AUU^H A^H U = U^H AA^H U$$

Since $A^H A = AA^H$, we have $T^H T = TT^H$. The proof is completed by noting that a normal triangular matrix is diagonal.