

# *Linear Algebra*



Instructor: Jing YAO

## 5

# Eigenvalues and Eigenvectors (特征值与特征向量)

## 5.6

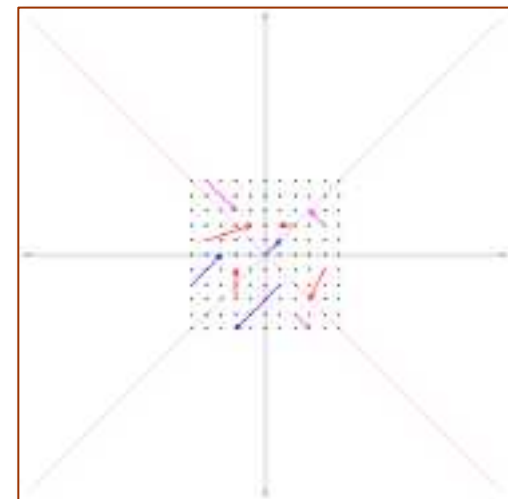
## SIMILARITY TRANSFORMATIONS

Similar Matrices(相似矩阵)

Similarity Transformations(相似变换)

Triangularization and Diagonalization

The Jordan Form



When  $A$  is **diagonalizable**:  $S^{-1}AS = \Lambda$

$$S = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$S$ : invertible matrix

$x_1, \dots, x_n$ : eigenvectors

*independent*

$\lambda_1, \dots, \lambda_n$ : eigenvalues

When  $A$  is **real symmetric**:  $Q^{-1}AQ = \Lambda$

$$Q = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$Q$ : orthogonal matrix

*orthonormal*

$\lambda_1, \dots, \lambda_n$ : real

When  $A$  is **Hermitian**:  $U^{-1}AU = \Lambda$

$$U = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \\ | & | & \cdots & | \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

$U$ : unitary matrix

The family of  
 $M^{-1}AM$ ?

*orthonormal*

$\lambda_1, \dots, \lambda_n$ : real

# I. Similar Matrices (相似矩阵)

**Definition 1** Two matrices  $A$  and  $B$  are said to be **similar (相似)** if there is an invertible matrix  $M$  such that  $B = M^{-1}AM$  (also denoted by  $A \sim B$ ).

**Remark 1** (1)  $A$  is similar to itself. (自反性)

(2) If  $A$  is similar to  $B$ , then  $B$  must be similar to  $A$ . (对称性)

(3) If  $A_1$  and  $A_2$  are similar,  $A_2$  and  $A_3$  are similar, then we can also conclude that  $A_1$  and  $A_3$  are similar. (传递性)

**Remark 2** If  $A$  and  $B$  are similar, then  $A^k$  and  $B^k$  ( $k$  is a positive integer) are also similar.

Moreover,  $k$  can be  $-1$  if  $A$  and  $B$  are invertible.

**Theorem 1** Assume that  $\mathbf{B} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$ . Then  $\mathbf{A}$  and  $\mathbf{B}$  have **the same eigenvalues**. A vector  $\mathbf{v}$  is an eigenvector of  $\mathbf{A}$  if and only if  $\mathbf{M}^{-1}\mathbf{v}$  is an eigenvector of  $\mathbf{B}$ .

**Proof.**  $\mathbf{B} - \lambda\mathbf{I} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M} - \lambda\mathbf{I} = \mathbf{M}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{M}$ , and so

$$\begin{aligned} |\mathbf{B} - \lambda\mathbf{I}| &= |\mathbf{M}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{M}| \\ &= |\mathbf{M}^{-1}| \cdot |\mathbf{A} - \lambda\mathbf{I}| \cdot |\mathbf{M}| = |\mathbf{A} - \lambda\mathbf{I}|. \end{aligned}$$

Thus the characteristic polynomials  $|\mathbf{A} - \lambda\mathbf{I}|$  and  $|\mathbf{B} - \lambda\mathbf{I}|$  are equal and have the same roots. So the eigenvalues of  $\mathbf{A}$  and  $\mathbf{B}$  are the same.

Suppose that  $\mathbf{v}$  is an eigenvector, i.e.,  $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$  for an eigenvalue  $\lambda$ . Then  $\mathbf{M}\mathbf{B}\mathbf{M}^{-1}\mathbf{v} = \lambda\mathbf{v}$ , and  $\mathbf{B}\mathbf{M}^{-1}\mathbf{v} = \lambda(\mathbf{M}^{-1}\mathbf{v})$ , i.e.,  $\lambda$  is an eigenvalue of  $\mathbf{B}$ , and an corresponding eigenvector is  $\mathbf{M}^{-1}\mathbf{v}$ .

**Remark 1** Every  $M^{-1}AM$  has the same eigenvalues as  $A$ .

**Remark 2** Every  $M^{-1}AM$  has the same number of independent eigenvectors as  $A$ . (Each eigenvector is multiplied by  $M^{-1}$ ).

**Remark 3** If  $B = M^{-1}AM$ , then

$$|A| = |B|, \text{ and } \text{trace}(A) = \text{trace}(B).$$

**Remark 4** If  $B = M^{-1}AM$ , then  $\text{rank}(A) = \text{rank}(B)$ .

**Remark 5** If  $B = M^{-1}AM$ , then  $A$  and  $B$  have the same characteristic polynomial. However, if  $A$  and  $B$  have the same characteristic polynomial, they are *not necessarily* similar. For example,

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}.$$

$|A - \lambda I| = |B - \lambda I| = (\lambda - 2)^2$ , but  $A$  and  $B$  are not similar, since for any invertible matrix  $M$ ,  $M^{-1}AM = M^{-1}(2I)M = 2I = A \neq B$ .

**Example 1.**  $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$  has eigenvalues 1 and 0.

If  $M = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$ , then

$B = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}$ : triangular with  $\lambda = 1$  and 0.

If  $M = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ , then

$B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$ : projection with  $\lambda = 1$  and 0.

If  $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ , then

$B$  = an arbitrary matrix with  $\lambda = 1$  and 0.

## II. Similarity Transformation (相似变换)

*Recall that:*

Every linear transformation is represented by a matrix: any linear transformation  $T$  from  $\mathbf{R}^n$  to  $\mathbf{R}^m$  can be implemented via left-multiplication by a matrix  $A$ :  $\mathbf{x} \mapsto A\mathbf{x}$ .

The matrix  $A$  depends on the choice of **basis**.

*We will see next:*

### Similarity Transformation $\Leftrightarrow$ Change of Basis

*If we change the basis by  $M$ , we change the matrix  $A$  to a **similar** matrix  $B$ , and  $B = M^{-1}AM$ .*

(同一个线性变换在两组基下的表示矩阵 $A$ 和 $B$ 是相似的.)

We explain this for  $2 \times 2$  matrices.



Let  $V = \mathbf{R}^2$ , and let  $T$  be a transformation of  $V$ .

Given a basis  $\mathbf{v}_1, \mathbf{v}_2$ , there exist scalars  $a_{ij}$  such that

$$T(\mathbf{v}_1) = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2,$$

$$T(\mathbf{v}_2) = a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2.$$

(基向量的像可以被基向量线性表出)

Let

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}.$$

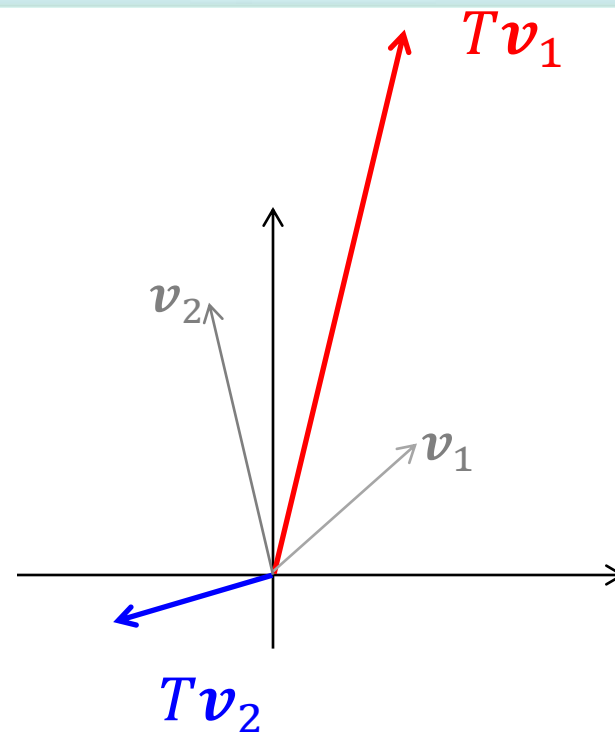
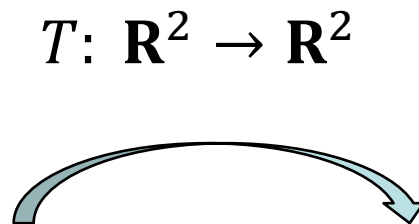
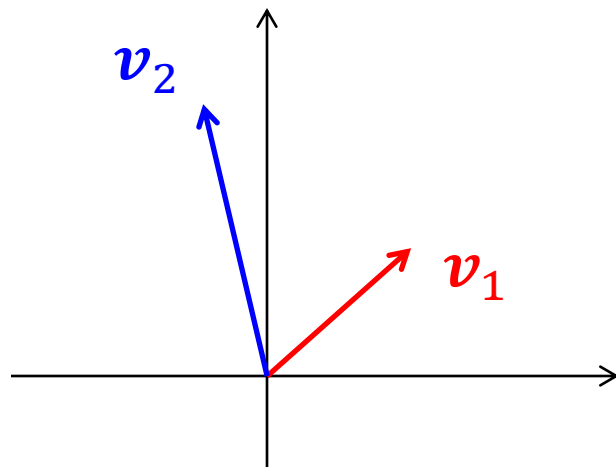
Then

$$[T(\mathbf{v}_1) \quad T(\mathbf{v}_2)] = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = [\mathbf{v}_1 \quad \mathbf{v}_2] \mathbf{A}.$$

(The linear transformation  $T$  is represented by the matrix  $\mathbf{A}$  with respect to the basis  $\mathbf{v}_1, \mathbf{v}_2$ :  $\mathbf{A}$ 是线性变换 $T$ 在一组基 $\mathbf{v}_1, \mathbf{v}_2$ 下的矩阵)

(For simplicity, we will write  $T(\mathbf{x})$  as  $T\mathbf{x}$ .)

Basis  $\{\mathbf{v}_1, \mathbf{v}_2\}$



$$T\mathbf{v}_1 = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2$$

$$T\mathbf{v}_2 = a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2$$

$$[T\mathbf{v}_1 \quad T\mathbf{v}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} = [\mathbf{v}_1 \quad \mathbf{v}_2] \mathbf{A}$$

(The linear transformation  $T$  is represented by the matrix  $\mathbf{A}$  with respect to the basis  $\{\mathbf{v}_1, \mathbf{v}_2\}$ :  $\mathbf{A}$  是线性变换  $T$  在一组基  $\{\mathbf{v}_1, \mathbf{v}_2\}$  下的矩阵)

Let  $\mathbf{w}_1, \mathbf{w}_2$  be another basis. Then there exist scalars  $m_{ij}$  such that

$$\begin{cases} \mathbf{w}_1 = m_{11}\mathbf{v}_1 + m_{12}\mathbf{v}_2, \\ \mathbf{w}_2 = m_{21}\mathbf{v}_1 + m_{22}\mathbf{v}_2. \end{cases}$$

Then

$$[\mathbf{w}_1 \quad \mathbf{w}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \end{bmatrix} = [\mathbf{v}_1 \quad \mathbf{v}_2] \mathbf{M}.$$

( $\mathbf{M}$ 是从一组基 $\mathbf{v}_1, \mathbf{v}_2$ 到另一组基 $\mathbf{w}_1, \mathbf{w}_2$ 的过渡矩阵:  
transition matrix)

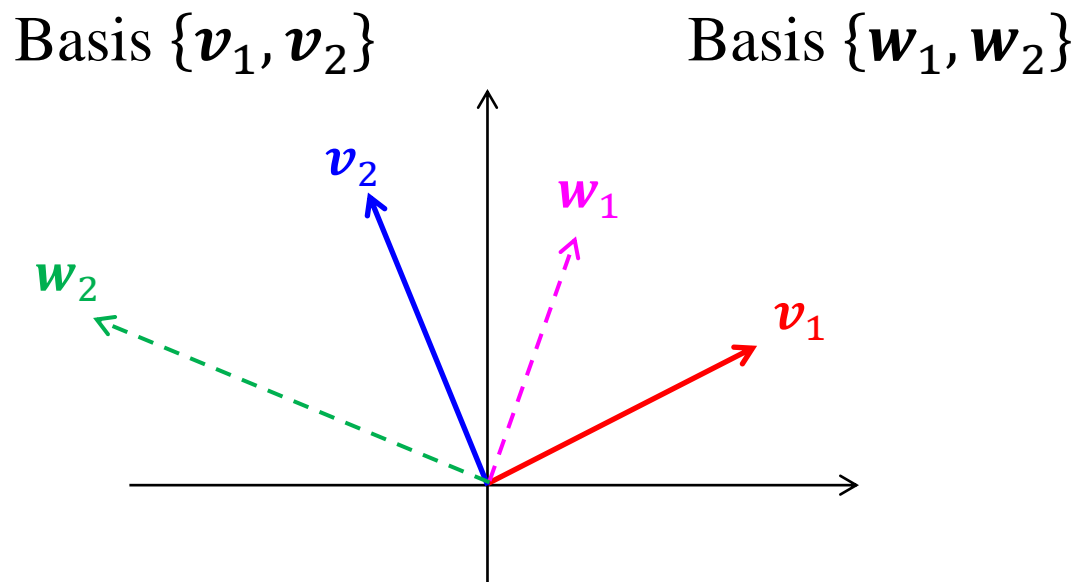
**Lemma 1** Let  $[\mathbf{w}_1 \quad \mathbf{w}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2] \mathbf{M}$ ,

then  $[T\mathbf{w}_1 \quad T\mathbf{w}_2] = [T\mathbf{v}_1 \quad T\mathbf{v}_2] \mathbf{M}$ .

**Proof.** This is due to

$$\begin{cases} T\mathbf{w}_1 = m_{11}T\mathbf{v}_1 + m_{12}T\mathbf{v}_2, \\ T\mathbf{w}_2 = m_{21}T\mathbf{v}_1 + m_{22}T\mathbf{v}_2. \end{cases}$$

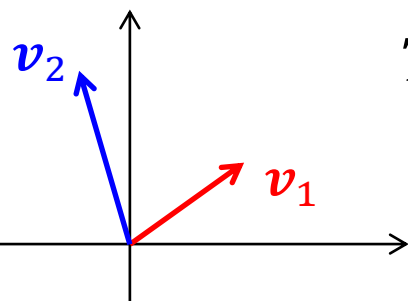
# Change of Basis



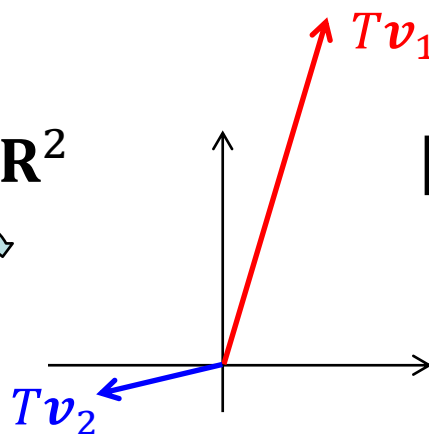
$$\left. \begin{aligned} w_1 &= m_{11}v_1 + m_{12}v_2 \\ w_2 &= m_{21}v_1 + m_{22}v_2 \end{aligned} \right\} [w_1 \quad w_2] = [v_1 \quad v_2]M$$

( $M$ 是从一组基 $\{v_1, v_2\}$ 到另一组基 $\{w_1, w_2\}$ 的过渡矩阵:  
transition matrix)

Basis  $\{v_1, v_2\}$

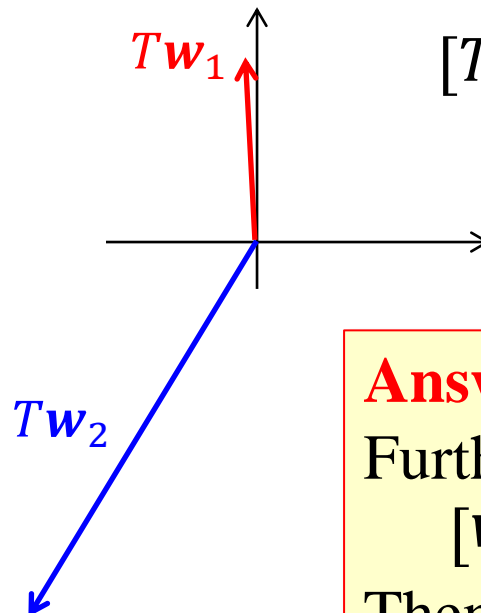
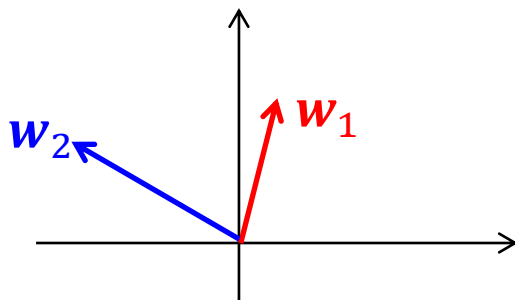


$$T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$$



$$[Tv_1 \quad Tv_2] = [v_1 \quad v_2]A$$

Basis  $\{w_1, w_2\}$



$$[Tw_1 \quad Tw_2] = [w_1 \quad w_2]B$$

**Question:**

How is  $A$  related to  $B$ ?

**Answer:**  $A$  is similar to  $B$ .

Furthermore, if

$$[w_1 \quad w_2] = [v_1 \quad v_2]M$$


Then  $B = M^{-1}AM$ .

**Theorem 2** *Two matrices represent the same linear transformation (with respect to different bases) if and only if they are similar.*

**Proof.** (We only state our proof for  $n = 2$ .)

Let  $\mathbf{A}$  be a matrix of degree 2, and let  $T$  be a linear transformation defined as below, where  $\mathbf{v}_1, \mathbf{v}_2$  is a basis,

$$[T\mathbf{v}_1 \quad T\mathbf{v}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{A}.$$

(1) “” Let  $\mathbf{M}$  be an invertible matrix, and let  $\mathbf{B} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$ . Let  $\mathbf{w}_1, \mathbf{w}_2$  be a basis defined by  $[\mathbf{w}_1 \quad \mathbf{w}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{M}$ .

Then, as  $\mathbf{A}\mathbf{M} = \mathbf{M}\mathbf{B}$ , we have

$$\begin{aligned} [T\mathbf{w}_1 \quad T\mathbf{w}_2] &= [T\mathbf{v}_1 \quad T\mathbf{v}_2]\mathbf{M} = [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{A}\mathbf{M} \\ &= [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{M}\mathbf{B} = [\mathbf{w}_1 \quad \mathbf{w}_2]\mathbf{B}. \end{aligned}$$

Thus the linear transformation  $T$  is represented by the matrix  $\mathbf{B}$  with respect to the basis  $\mathbf{w}_1, \mathbf{w}_2$ .

**Theorem 2** *Two matrices represent the same linear transformation (with respect to different bases) if and only if they are similar.*

**Proof.** (We only state our proof for  $n = 2$ .)

Let  $\mathbf{A}$  be a matrix of degree 2, and let  $T$  be a linear transformation defined as below, where  $\mathbf{v}_1, \mathbf{v}_2$  is a basis,

$$[T\mathbf{v}_1 \quad T\mathbf{v}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{A}.$$

(2) “” Assume that  $\mathbf{B}$  is matrix representing the linear transformation  $T$  relative to a basis  $\mathbf{w}_1, \mathbf{w}_2$ .

Let  $\mathbf{M}$  be the matrix such that

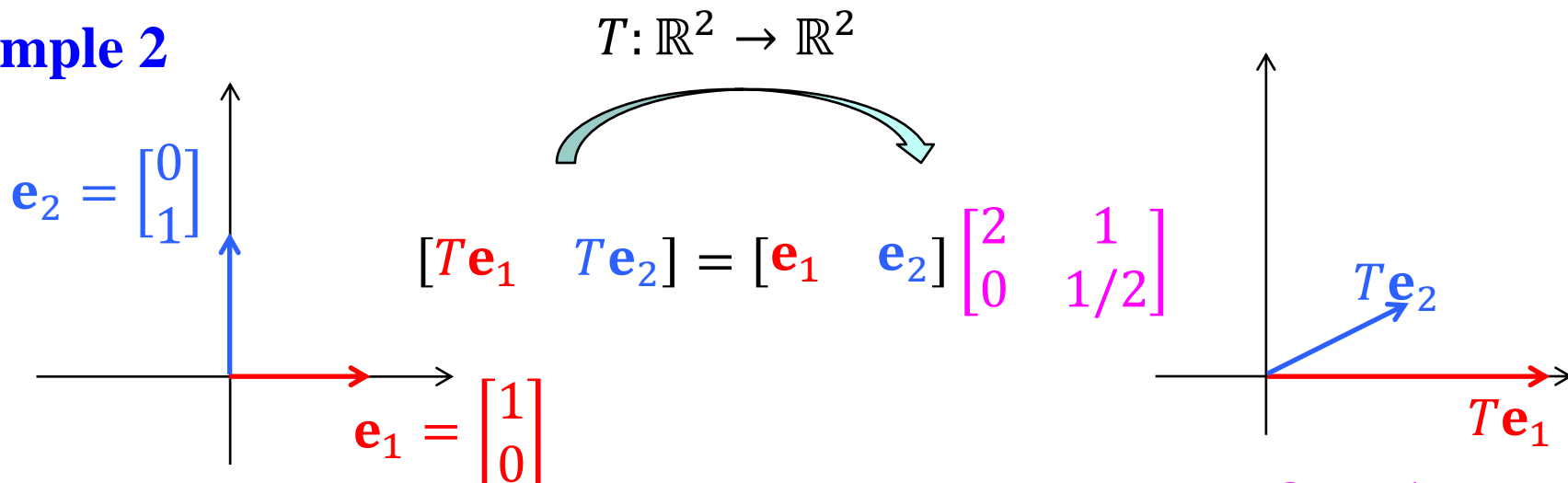
$$[\mathbf{w}_1 \quad \mathbf{w}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{M}.$$

Then,  $[\mathbf{v}_1 \quad \mathbf{v}_2] = [\mathbf{w}_1 \quad \mathbf{w}_2]\mathbf{M}^{-1}$ , and

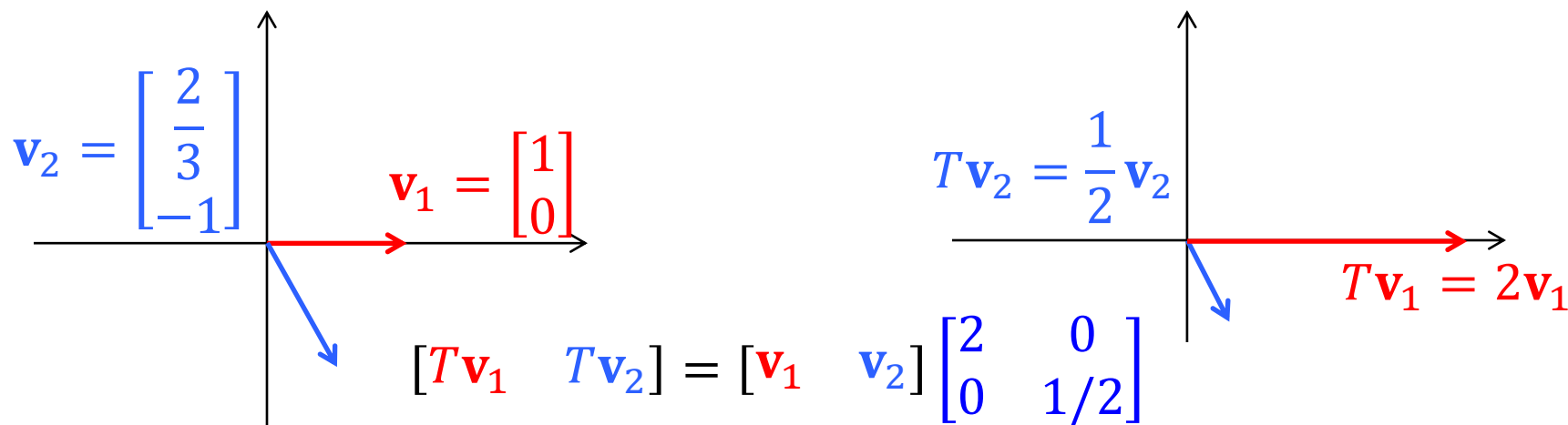
$$\begin{aligned} [\mathbf{w}_1 \quad \mathbf{w}_2]\mathbf{B} &= [T\mathbf{w}_1 \quad T\mathbf{w}_2] = [T\mathbf{v}_1 \quad T\mathbf{v}_2]\mathbf{M} \\ &= [\mathbf{v}_1 \quad \mathbf{v}_2]\mathbf{A}\mathbf{M} = [\mathbf{w}_1 \quad \mathbf{w}_2]\mathbf{M}^{-1}\mathbf{A}\mathbf{M}. \end{aligned}$$

Therefore,  $\mathbf{B} = \mathbf{M}^{-1}\mathbf{A}\mathbf{M}$ .

## Example 2



$\mathbf{v}_1, \mathbf{v}_2$  are two linearly independent eigenvectors of  $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 0 & 1/2 \end{bmatrix}$ .



$$\mathbf{M} = \begin{bmatrix} 1 & 2/3 \\ 0 & -1 \end{bmatrix}, \text{ and } \begin{bmatrix} 2 & 1 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 2/3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 2/3 \\ 0 & -1 \end{bmatrix}^{-1}$$



**Example 3.** Define  $T: \mathbf{R}^2 \rightarrow \mathbf{R}^2$  by  $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ , where  $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$ .

Find a basis (denoted by  $\mathbf{B}$ ) for  $\mathbf{R}^2$  with the property that representation matrix for  $T$  is a diagonal matrix.

**Solution** The eigenvalues of  $\mathbf{A}$  are distinct: 5 and 3, so  $\mathbf{A}$  is diagonalizable.

By diagonalizing  $\mathbf{A}$  into  $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$ , where

$$\mathbf{S} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \text{ and } \mathbf{\Lambda} = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}.$$

The columns of  $\mathbf{S}$ , call them  $\mathbf{b}_1$  and  $\mathbf{b}_2$ , are **eigenvectors** of  $\mathbf{A}$ .

By Theorem 2,  $\mathbf{\Lambda}$  is the representation matrix for  $T$  when  $\mathbf{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ .

The mappings  $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$  and  $\mathbf{u} \mapsto \mathbf{\Lambda}\mathbf{u}$  describe the same linear transformation, relative to different bases.

**Remark:** The way to simplify that matrix  $\mathbf{A}$ —in fact to diagonalize it—is to find its eigenvectors. In the language of linear transformations:

*Choose a basis consisting of eigenvectors.*

### III. Triangularization and Diagonalization (三角化与对角化)

Not every matrix can be diagonalized (对角化又称作相似对角化),

for instance,  $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ .

However, the following theorem tells us that each matrix can be triangularized by a unitary matrix.(并不是所有矩阵都可以对角化, 但每个矩阵都可以被酉矩阵三角化)

**Theorem 3 (Schur's lemma)** *For a matrix  $\mathbf{A}$  of degree  $n$ , there exists a **unitary** matrix  $\mathbf{U}$  of degree  $n$  such that  $\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{T}$  is **triangular**. The eigenvalues of  $\mathbf{A}$  appear along the diagonal of this similar matrix  $\mathbf{T}$ .*

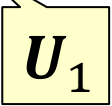
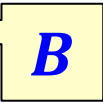
**Theorem 3** For a matrix  $A$  of degree  $n$ , there exists a *unitary* matrix  $U$  of degree  $n$  such that  $U^{-1}AU = T$  is *triangular*. The eigenvalues of  $A$  appear along the diagonal of this similar matrix  $T$ .

**Proof.** Let  $A$  be a matrix of degree  $n$ , and assume that  $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ , namely,  $\lambda_1$  is an eigenvalue and  $\mathbf{x}_1$  is a unit eigenvector.

( $A$  has at least one eigenvalue, in the worst case it could be repeated  $n$  times. And  $A$  has at least one unit eigenvector  $\mathbf{x}_1$ )

Then, using Gram-Schmidt process, there exists an orthonormal basis  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ , so  $U_1 = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$  is a *unitary matrix*.

$$\begin{aligned}
 AU_1 &= A[\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] = [A\mathbf{x}_1 \ A\mathbf{x}_2 \ \dots \ A\mathbf{x}_n] \\
 &= [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \begin{bmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \dots & b_{nn} \end{bmatrix}
 \end{aligned}$$

This leads to  $U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{bmatrix}$ , and  $B$  is of order  $(n - 1)$ .

Let  $\lambda_2$  be an eigenvalue of  $\mathbf{B}$  and  $\mathbf{y}_2$  a unit eigenvector.

Let  $\mathbf{M}_2$  be a unitary matrix with first column equal to  $\mathbf{y}_2$ . Then similarly we have  $\mathbf{M}_2^{-1}\mathbf{B}\mathbf{M}_2 = \begin{bmatrix} \lambda_2 & * \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$ .

Let  $\mathbf{U}_2 = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 \end{bmatrix}$ . Then  $\mathbf{U}_2$  is unitary, and

$$\begin{aligned} \mathbf{U}_2^{-1}(\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1)\mathbf{U}_2 &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 \end{bmatrix}^{-1} \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & \mathbf{M}_2^{-1}\mathbf{B}\mathbf{M}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ \mathbf{0} & \mathbf{0} & \mathbf{C} \end{bmatrix}. \end{aligned}$$

Notice that  $\mathbf{U}_1\mathbf{U}_2$  is still a unitary matrix.

Repeating this process produces a unitary matrix  $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2 \dots \mathbf{U}_{n-1}$ , such that  $\mathbf{U}^{-1}\mathbf{A}\mathbf{U}$  is a triangular matrix.

**Example 4** Let  $A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$ .

Then  $A$  has the eigenvalue  $\lambda = 1$  (algebraic multiplicity of  $\lambda$  is 2).

The only line of eigenvectors goes through  $[1, 1]^T$  (geometric multiplicity of  $\lambda$  is 1). So  $A$  is not diagonalizable.

But  $A$  is triangularizable ( $A$  can be triangularized by a unitary matrix).

After dividing by  $\sqrt{2}$ , this is the first column of  $U$ .

We choose  $U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$ , and the triangular

$$T = U^{-1}AU = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

has the eigenvalues on its diagonal.

This triangular form will show that any **Hermitian** matrix—whether its eigenvalues are *distinct or not* — has a **complete set** of orthonormal eigenvectors.

**When  $A$  is Hermitian**, i.e.,  $A = A^H$  (When  $A$  is real, it means  $A = A^T$ ), this triangular  $T = U^{-1}AU$  is also Hermitian:

$$T^H = (U^{-1}AU)^H = U^H A^H (U^{-1})^H = U^{-1}AU = T.$$

Therefore,  **$T$  must be diagonal**.

*This **finally** completes the proof of the **Spectral Theorem**.*

(1) *Every real symmetric matrix  $A$  can be diagonalized by an orthogonal matrix  $Q$ :  $Q^{-1}AQ = \Lambda$  ( $A = Q\Lambda Q^T$ ).*

(2) *Every Hermitian matrix  $A$  can be diagonalized by a unitary matrix  $U$ :  $U^{-1}AU = \Lambda$  ( $A = U\Lambda U^H$ ).*

*The columns of  $Q$  (or  $U$ ) consist of orthonormal eigenvectors of  $A$ .*

**Example 5** Let  $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ .

The spectral theorem says that this  $\mathbf{A} = \mathbf{A}^T$  can be diagonalized.

$\mathbf{A}$  has repeated eigenvalues  $\lambda_1 = \lambda_2 = 1$  and  $\lambda_3 = -1$ .

$\lambda_1 = \lambda_2 = 1$  has a plane of eigenvectors, and we pick an orthonormal

pair  $\mathbf{x}_1$  and  $\mathbf{x}_2$ :  $\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ ,  $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ .

and  $\mathbf{x}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$  for  $\lambda_3 = -1$ .

Therefore  $\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$  and  $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{\Lambda} = \text{diag}(1, 1, -1)$ .

**Remark** Split  $A = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$  into:

$$\begin{aligned}
 A &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 \mathbf{x}_1 \mathbf{x}_1^T + \lambda_2 \mathbf{x}_2 \mathbf{x}_2^T + \lambda_3 \mathbf{x}_3 \mathbf{x}_3^T \\
 &= (+1) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + (+1) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1) \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 &= (+1) \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1) \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \lambda_1 \mathbf{P}_1 + \lambda_3 \mathbf{P}_3,
 \end{aligned}$$

where  $\mathbf{P}_1$  is a projection of rank 2 (onto the plane of eigenvectors).

*Every Hermitian matrix with  $k$  different eigenvalues has a **spectral decomposition** into  $A = \lambda_1 \mathbf{P}_1 + \cdots + \lambda_k \mathbf{P}_k$ , where  $\mathbf{P}_i$  is the projection onto the eigenspace for  $\lambda_i$ . Since there is a full set of eigenvectors, the projections add up to the identity. And since the eigenspaces are orthogonal, two projections produce zero:  $\mathbf{P}_j \mathbf{P}_i = \mathbf{0}$ .*



**An important question:** For which matrices is  $T = \Lambda$  ?

### Some special matrices

Real matrices	Complex matrices	Eigenvalues
Symmetric $A^T = A$	Hermitian $A^H = A$	All $\lambda$ 's are real (on the real axis)
Skew-symmetric $A^T = -A$	Skew-Hermitian $A^H = -A$	All $\lambda$ 's are imaginary (including 0 sometimes) (on the imaginary axis)
Orthogonal $Q^T = Q^{-1}$	Unitary $U^H = U^{-1}$	all $ \lambda  = 1$ (on the unit circle)

*These matrices are all diagonalizable.*

Now we want the whole class -- called “normal”.

# NORMAL MATRICES

**Definition 2** A matrix  $N$  is called a **normal matrix** (正规矩阵) if

$$NN^H = N^H N.$$

Normal matrices include *symmetric, Hermitian, orthogonal, unitary, skew-symmetric, skew-Hermitian matrices*.

(For example, if  $A = A^H$ , then  $AA^H = A^H A = A^2$ ;

If  $U^H = U^{-1}$ , then  $UU^H = U^H U = I$ .)

We will show that, normal matrices are *exactly* the matrices which are diagonalizable. (Normal matrices are *exactly* the matrices that have a complete set of orthonormal eigenvectors.)

**Theorem 4** A matrix is *diagonalized* by a *unitary matrix* if and only if it is a *normal matrix*.

(In other words, A matrix  $A$  is a normal matrix if and only if there exists a unitary matrix  $U$  such that  $U^{-1}AU$  is diagonal.)

**Theorem 4** *A matrix is diagonalized by a unitary matrix if and only if it is a normal matrix.*

**Proof.** “ $\rightarrow$ ” Let  $A$  be a matrix, and  $U$  a unitary matrix such that  $U^{-1}AU = D$  is diagonal.

Then  $A = UDU^{-1}$ , and  $A^H = UD^H U^{-1}$ . Thus

$$\begin{aligned} AA^H &= (UDU^{-1})(UD^H U^{-1}) = UDD^H U^{-1} \\ &= UD^H DU^{-1} = UD^H U^{-1}UDU^{-1} = A^H A, \end{aligned}$$

i.e.,  $A$  is a normal matrix.

“ $\leftarrow$ ” Conversely, let  $A$  be a normal matrix. Let  $U$  be a unitary matrix such that  $U^{-1}AU = T$  is triangular. Then  $T^H = U^H A^H U$ , thus

$$\begin{aligned} TT^H &= (U^{-1}AU)(U^H A^H U) = U^{-1}AA^H U = U^{-1}A^H AU \\ &= (U^{-1}A^H U)(U^{-1}AU) = T^H T, \end{aligned}$$

and  $TT^H = T^H T$ , i.e.,  $T$  is a normal matrix.

It follows that since  $T$  is triangular,  $T$  is diagonal.

*(All normal triangular matrices are diagonal.— Exercise #19,20)*

## IV. The Jordan Form (若当形)

Although not every matrix is diagonalizable, every matrix can be converted into *Jordan form*.

The Jordan form of a matrix is important. However, we will not be able to study it in details, and instead we will only give a simple introduction.

We will systematically study it in *Advanced Linear Algebra (线性代数精讲)*.

**The goal:** to make  $M^{-1}AM$  as *nearly diagonal as possible*.

**Definition 3** A **Jordan block (若当块)** is a matrix of degree  $k$  with the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

The diagonal value  $\lambda_i$  is an eigenvalue of  $J_i$ .

Since  $J_i - \lambda_i I$  is of rank  $k - 1$ , the nullspace of  $J_i - \lambda_i I$  has dimension 1. In other words, the eigenspace of  $\lambda_i$  is of dimension 1.

**Theorem 5** If a matrix  $A$  has  $s$  independent eigenvectors, then it is similar to a matrix in the **Jordan form** (若当形) with  $s$  blocks:

$$J = M^{-1}AM = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix},$$

where each  $J_i$  is a Jordan block, corresponding to an eigenvalue  $\lambda_i$  and only one independent eigenvector.

The same  $\lambda_i$  will appear in several blocks, if it has several independent eigenvectors.

Moreover, two matrices are similar if and only if they share the same Jordan form  $J$ .

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

**Example 6** Check that  $\mathbf{T} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$  and  $\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$  and  $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$  all lead to Jordan form  $\mathbf{J} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ .

These four matrices have eigenvalues 1 and 1 with only *one* eigenvector—so  $\mathbf{J}$  consists of *one block*.

(T) From  $\mathbf{T}$  to  $\mathbf{J}$ , a diagonal  $\mathbf{M}$  will do it:

$$\mathbf{M}^{-1}\mathbf{T}\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \mathbf{J}.$$

(B) From  $\mathbf{B}$  to  $\mathbf{J}$ , a permutation  $\mathbf{P}$  does that:

$$\mathbf{P}^{-1}\mathbf{B}\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \mathbf{J}.$$

(A) From  $\mathbf{A}$  to  $\mathbf{J}$ , let  $\mathbf{U} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$ , then

$$\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \mathbf{T} \text{ and then } \mathbf{M}^{-1}\mathbf{T}\mathbf{M} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \mathbf{J}.$$

**$\mathbf{T}, \mathbf{A}, \mathbf{B}, \mathbf{J}$**   
are similar:  
they all  
belong to  
the same  
family.

**Example 7** Find the Jordan form of  $\mathbf{A}$ , where  $\mathbf{A} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ .

The eigenvalues of  $\mathbf{A}$  are all 0's (triple eigenvalue), so it will appear in all their Jordan blocks. Thus the Jordan form of  $\mathbf{A}$  is one of the following :

$$\mathbf{J}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{J}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{J}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since  $\mathbf{A}$  has only one independent eigenvector  $(1,0,0)^T$ ,

its Jordan form has only one block, and so the Jordan form is  $\mathbf{J}_1$ .

As for  $\mathbf{J}_3 =$  zero matrix, *it is in a family by itself*; the only matrix similar to  $\mathbf{J}_3$  is  $\mathbf{M}^{-1}\mathbf{0M} = \mathbf{0}$ .

**Example 8** Find the Jordan form of  $A$ , where  $A = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ .

The eigenvalues of  $A$  are equal to 2 (triple eigenvalue). Thus the Jordan form of  $A$  is one of the following:

$$J_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since  $A$  has only one independent eigenvector  $(1,0,0)^T$ ,

its Jordan form has only one block, and so the Jordan form is  $J_1$ .



**Example 9** Find the Jordan form of  $\mathbf{A}$ , where  $\mathbf{A} = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ .

The eigenvalues of  $\mathbf{A}$  are equal to 2 (triple eigenvalue). Thus the Jordan form of  $\mathbf{A}$  is one of the following:

$$J_1 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since  $\mathbf{A}$  has two independent eigenvectors  $(1,0,0)^T$  and  $(0,2,-1)^T$ , thus  $\mathbf{A}$  has exactly two Jordan blocks, and so the Jordan form of  $\mathbf{A}$  is  $J_2$ .

**Remark: Power of  $A$ .**

If  $A$  can be diagonalized, the powers of  $A = S\Lambda S^{-1}$  are easy:  
 $A^k = S\Lambda^k S^{-1}$ .

In general case, we have Jordan's similarity  $A = MJM^{-1}$ , so now we need the powers of  $J$ :

$$A^k = (MJM^{-1})(MJM^{-1}) \dots (MJM^{-1}) = MJ^k M^{-1}.$$

$$\text{Since } J = \begin{bmatrix} J_1 & & \\ & J_2 & \\ & & \ddots \\ & & & J_s \end{bmatrix}, \text{ so } J^k = \begin{bmatrix} J_1^k & & \\ & J_2^k & \\ & & \ddots \\ & & & J_s^k \end{bmatrix}.$$

For instance, if  $\lambda$  is a triple eigenvalue with a single eigenvector, then the  $3 \times 3$  block  $J_i$  will enter, and

$$(J_i)^k = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \frac{1}{2}k(k-1)\lambda^{k-2} \\ 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & \lambda^k \end{bmatrix}.$$

## *The Spectral Theorem for Real Symmetric Matrices*

An  $n \times n$  real symmetric matrix  $A$  ( $A \in \mathbf{R}^{n \times n}$  and  $A = A^T$ ) has the following properties:

(一个对称的  $n \times n$  实矩阵具有下面的特性)

- a.  $A$  has  $n$  real eigenvalues, counting multiplicities. ( $A$ 有 $n$ 个实特征值, 包含重复的特征值)
- b. The dimension of the eigenspace for each eigenvalue  $\lambda$  equals the multiplicity of  $\lambda$  as a root of the characteristic equation. (对于每一个特征值 $\lambda$ , 对应特征子空间的维数等于 $\lambda$ 作为特征方程的根的重数, 即: 几何重数=代数重数)
- c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal. (特征子空间相互正交, 这种正交性是在特征向量对应不同特征值的意义下成立的)
- d.  $A$  is orthogonally diagonalizable. ( $A$ 可以正交对角化)

## *The Spectral Theorem for Hermitian Matrices*

An  $n \times n$  Hermitian matrix  $A$  ( $A \in \mathbb{C}^{n \times n}$  and  $A = A^H$ ) has the following properties:

(一个  $n \times n$  厄米特矩阵具有下面的特性)

- a.  $A$  has  $n$  real eigenvalues, counting multiplicities. ( $A$ 有 $n$ 个实特征值, 包含重复的特征值)
- b. The dimension of the eigenspace for each eigenvalue  $\lambda$  equals the multiplicity of  $\lambda$  as a root of the characteristic equation. (对于每一个特征值 $\lambda$ , 对应特征子空间的维数等于 $\lambda$ 作为特征方程的根的重数, 即: 几何重数=代数重数)
- c. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal. (特征子空间相互正交, 这种正交性是在特征向量对应不同特征值的意义下成立的)
- d.  $A$  can be diagonalized by a unitary matrix. ( $A$ 可以用酉矩阵对角化)

## Similarity Transformations

1.  $A$  is *diagonalizable*: The columns of  $S$  are eigenvectors and  $S^{-1}AS = \Lambda$ .
2.  $A$  is *arbitrary*: The columns of  $M$  include “generalized eigenvectors” of  $A$ , and the Jordan form  $M^{-1}AM = J$  is *block diagonal*.
3.  $A$  is *arbitrary*: The unitary  $U$  can be chosen so that  $U^{-1}AU = T$  is *triangular*.
4.  $A$  is *normal*,  $AA^H = A^H A$ : then  $U$  can be chosen so that  $U^{-1}AU = \Lambda$ .

*Special cases of normal matrices, all with orthonormal eigenvectors:*

- (a) If  $A = A^H$  is Hermitian, then all  $\lambda_i$  are real.
- (b) If  $A = A^T$  is real symmetric, then  $\Lambda$  is real and  $U = Q$  is orthogonal.
- (c) If  $A = -A^H$  is skew-Hermitian, then all  $\lambda_i$  are purely imaginary.
- (d) If  $A$  is orthogonal or unitary, then all  $|\lambda_i| = 1$  are on the unit circle.

**Key words:**

*Similar Matrices*

*Similarity Transformations*

*Triangularization and Diagonalization; Normal matrices*

*The Jordan Form*

**Homework**

**See Blackboard**

