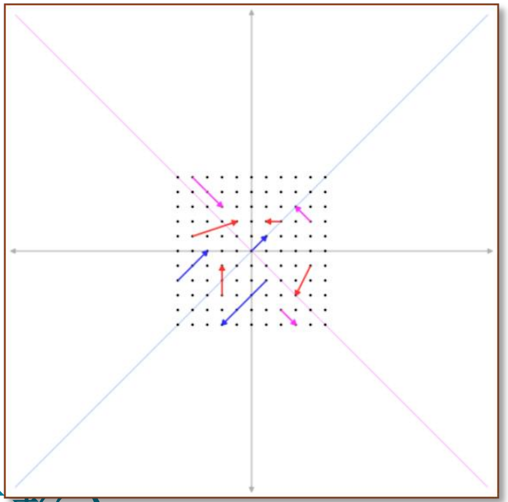
5 Eigenvalues and Eigenvectors（特征值与特征向量）

5.6



SIMILARITY TRANSFORMATIONS

Similar Matrices (相似矩阵)

Similarity Transformations (相似变换)

Triangularization and Diagonalization (三角化与对角化)

The Jordan Form (若当型)

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

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

independent

S: invertible matrix

$x_1, \ldots, x_n$: eigenvectors

$\lambda_1, \ldots, \lambda_n$: eigenvalues

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

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

orthonormal

Q: orthogonal matrix

$\lambda_1, \ldots, \lambda_n$: real

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

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

orthonormal

U: unitary matrix

$\lambda_1, \ldots, \lambda_n$: real

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

I. Similar Matrices (相似矩阵)

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

Remark 1 (1) $\boldsymbol{A}$ is similar to itself. (自反性)

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

(3) If $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are similar, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ are similar, then we can also conclude that $\boldsymbol{A}_1$ and $\boldsymbol{A}_3$ are similar. (传递性)

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

Moreover, k can be -1 if $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible.

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

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

$$|\boldsymbol{B} - \lambda\boldsymbol{I}| = |\boldsymbol{M}^{-1}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{M}|$$

$$= |\boldsymbol{M}^{-1}| \cdot |\boldsymbol{A} - \lambda\boldsymbol{I}| \cdot |\boldsymbol{M}| = |\boldsymbol{A} - \lambda\boldsymbol{I}|.$$

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

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

i.e., λ is an eigenvalue of $\boldsymbol{B}$, and an corresponding eigenvector is $\boldsymbol{M}^{-1}\boldsymbol{v}$.

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

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

Remark 3 If $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$, then

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

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

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

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

$|\boldsymbol{A} - \lambda\boldsymbol{I}| = |\boldsymbol{B} - \lambda\boldsymbol{I}| = (\lambda - 2)^2$, but $\boldsymbol{A}$ and $\boldsymbol{B}$ are not similar, since

for any invertible matrix $\boldsymbol{M}$, $\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \boldsymbol{M}^{-1}(2\boldsymbol{I})\boldsymbol{M} = 2\boldsymbol{I} = \boldsymbol{A} \neq \boldsymbol{B}$.

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 $\boldsymbol{A}$: $\boldsymbol{x} \longmapsto \boldsymbol{A}\boldsymbol{x}$.

The matrix $\boldsymbol{A}$ depends on the choice of **basis**.

We will see next:

Similarity Transformation ⇔ Change of Basis

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

(同一个线性变换在两组基下的表示矩阵$\boldsymbol{A}$和$\boldsymbol{B}$是相似的.)

We explain this for 2×2 matrices.

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

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

$$T(\boldsymbol{v}_1) = a_{11}\boldsymbol{v}_1 + a_{12}\boldsymbol{v}_2,$$
$$T(\boldsymbol{v}_2) = a_{21}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2.$$

(基向量的像可以被基向量线性表出)

Let

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

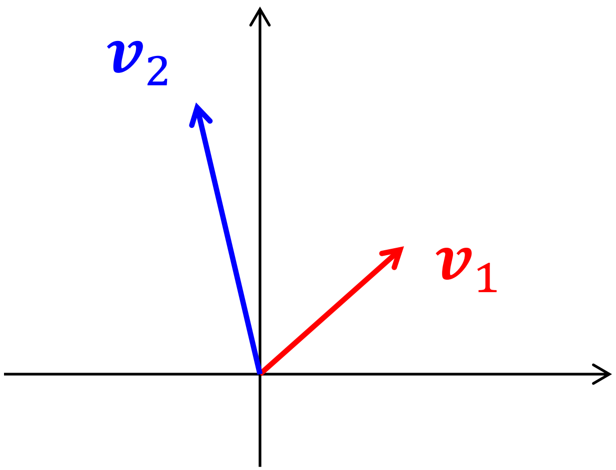
Then

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

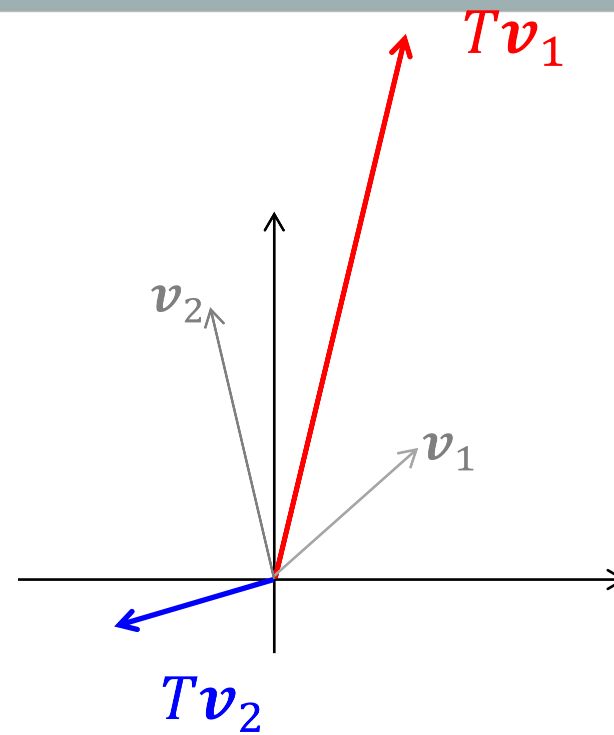
(The linear transformation T is represented by the matrix $\boldsymbol{A}$ with respect to the basis $\boldsymbol{v}_1, \boldsymbol{v}_2$: $\boldsymbol{A}$是线性变换T在一组基$\boldsymbol{v}_1, \boldsymbol{v}_2$下的矩阵)

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

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



$T\colon\ \mathbf{R}^2 \to \mathbf{R}^2$



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

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

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

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

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

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

Then

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

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

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

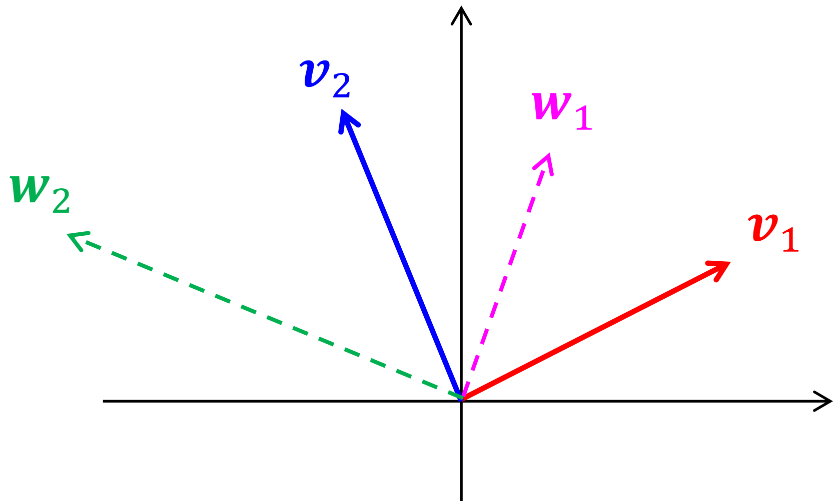
then $[T\boldsymbol{w}_1 \quad T\boldsymbol{w}_2] = [T\boldsymbol{v}_1 \quad T\boldsymbol{v}_2]\boldsymbol{M}$.

Proof. This is due to

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

Change of Basis

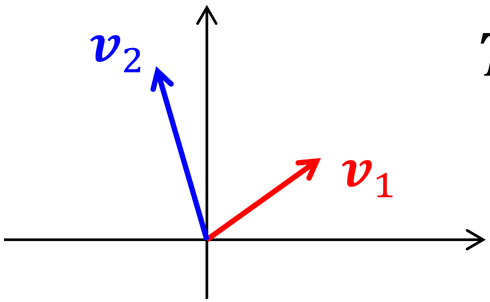
Basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ Basis $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$



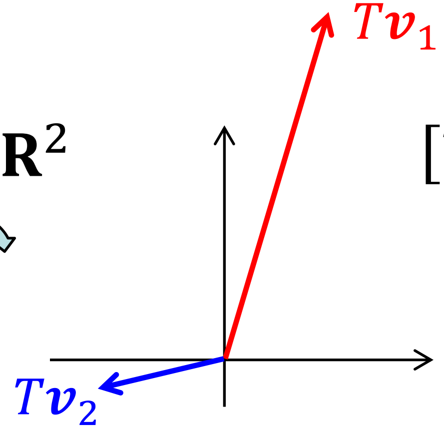
$$\left.\begin{aligned} \boldsymbol{w}_1 &= m_{11}\boldsymbol{v}_1 + m_{12}\boldsymbol{v}_2 \\ \boldsymbol{w}_2 &= m_{21}\boldsymbol{v}_1 + m_{22}\boldsymbol{v}_2 \end{aligned}\right\} \quad [\boldsymbol{w}_1 \quad \boldsymbol{w}_2] = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2]\boldsymbol{M}$$

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

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

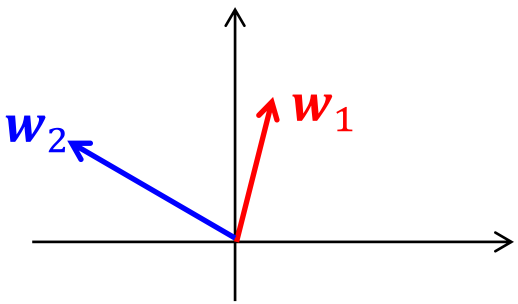


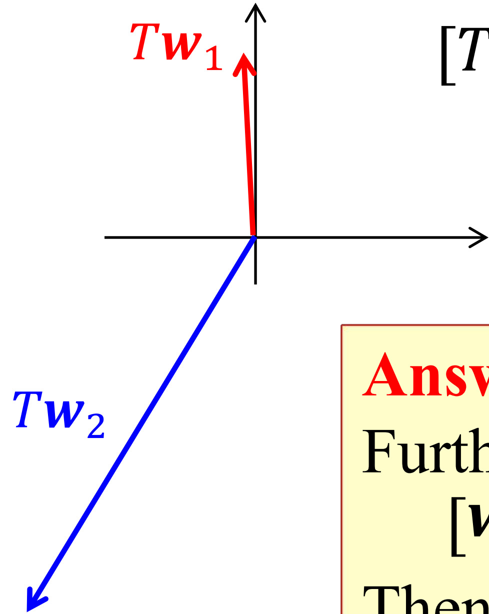
$T\colon\ \mathbf{R}^2 \to \mathbf{R}^2$



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

Basis $\{\boldsymbol{w}_1,\ \boldsymbol{w}_2\}$





$$[T\boldsymbol{w}_1 \quad T\boldsymbol{w}_2] = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2]\boldsymbol{B}$$

Question:

How is $\boldsymbol{A}$ related to $\boldsymbol{B}$?

Answer: $\boldsymbol{A}$ *is similar to* $\boldsymbol{B}$.

Furthermore, if

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

Then $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$.

T: 问题→答案

$$\text{English} \underset{M^{-1}}{\overset{M}{\rightleftarrows}} \text{汉语}$$

$$\boxed{B = M^{-1} A M}$$

M

M^{-1}

图片来自网络

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

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

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

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

$$[T\boldsymbol{w}_1\ \ T\boldsymbol{w}_2] = [T\boldsymbol{v}_1\ \ T\boldsymbol{v}_2]\boldsymbol{M} = [\boldsymbol{v}_1\ \ \boldsymbol{v}_2]\boldsymbol{A}\boldsymbol{M}$$

$$= [\boldsymbol{v}_1\ \ \boldsymbol{v}_2]\boldsymbol{M}\boldsymbol{B} = [\boldsymbol{w}_1\ \ \boldsymbol{w}_2]\boldsymbol{B}.$$

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

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

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

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

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

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

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

$$[\boldsymbol{w}_1 \;\; \boldsymbol{w}_2]\boldsymbol{B} = [T\boldsymbol{w}_1 \;\; T\boldsymbol{w}_2] = [T\boldsymbol{v}_1 \;\; T\boldsymbol{v}_2]\boldsymbol{M}$$
$$= [\boldsymbol{v}_1 \;\; \boldsymbol{v}_2]\boldsymbol{A}\boldsymbol{M} = [\boldsymbol{w}_1 \;\; \boldsymbol{w}_2]\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}.$$

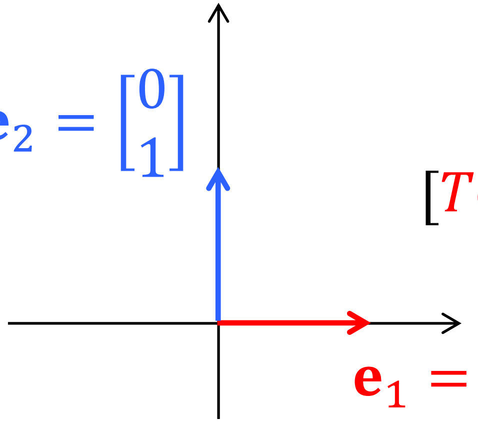
Therefore, $\boldsymbol{B} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$.

https://www.geogebra.org/m/uvyqfefh

Example 1

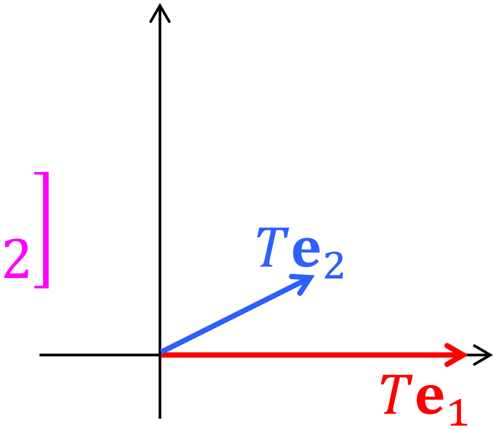
$T: \mathbb{R}^2 \to \mathbb{R}^2$

$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$



$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$$[T\mathbf{e}_1 \quad T\mathbf{e}_2] = [\mathbf{e}_1 \quad \mathbf{e}_2] \begin{bmatrix} 2 & 1 \\ 0 & 1/2 \end{bmatrix}$$

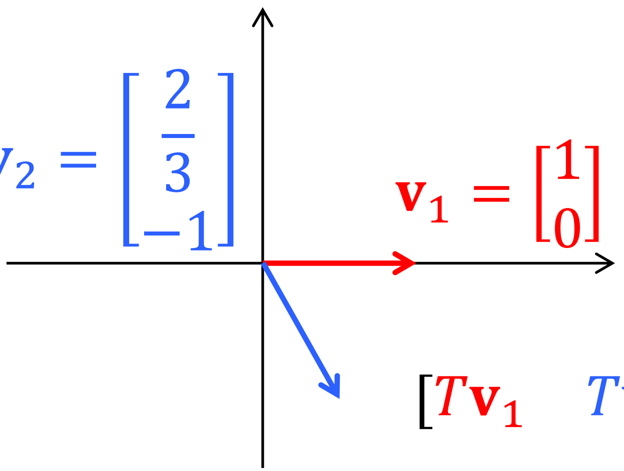


$T\mathbf{e}_2$

$T\mathbf{e}_1$

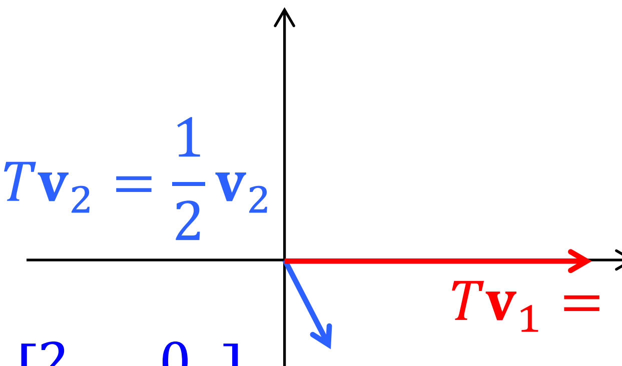
$\mathbf{v}_1, \mathbf{v}_2$ are two linearly independent eigenvectors of $\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 0 & 1/2 \end{bmatrix}$.

$\mathbf{v}_2 = \begin{bmatrix} \frac{2}{3} \\ -1 \end{bmatrix}$



$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$T\mathbf{v}_2 = \frac{1}{2}\mathbf{v}_2$



$T\mathbf{v}_1 = 2\mathbf{v}_1$

$$[T\mathbf{v}_1 \quad T\mathbf{v}_2] = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} 2 & 0 \\ 0 & 1/2 \end{bmatrix}$$

$\boldsymbol{M} = \begin{bmatrix} 1 & 2/3 \\ 0 & -1 \end{bmatrix}$, 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 2. Define T: $\mathbf{R}^2 \to \mathbf{R}^2$ by $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$.

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

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

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

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

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

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

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

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

Choose a basis consisting of eigenvectors.

Change of Basis ⇔ Similarity transformations

$$[\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \cdots \quad \boldsymbol{w}_n] = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n]\boldsymbol{M}$$

新基 旧基 过渡矩阵

- Any vector $\boldsymbol{v}$ in V can be expressed as a linear combination

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n = y_1\boldsymbol{w}_1 + y_2\boldsymbol{w}_2 + \cdots + y_n\boldsymbol{w}_n$$

i.e.

$$\boldsymbol{v} = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [\boldsymbol{w}_1 \quad \boldsymbol{w}_2 \quad \cdots \quad \boldsymbol{w}_n]\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \cdots \quad \boldsymbol{v}_n]\left(\boldsymbol{M}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}\right) \Rightarrow \boxed{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{M}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}$$

旧坐标 新坐标

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

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

Proof. Let $\boldsymbol{A}$ be a matrix of degree n, and assume that $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$, namely, λ_1 is an eigenvalue and $\boldsymbol{x}_1$ is a unit eigenvector.

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

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

$$\boldsymbol{A}\boldsymbol{U}_1 = \boldsymbol{A}[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_n] = [\boldsymbol{A}\boldsymbol{x}_1 \quad \boldsymbol{A}\boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{A}\boldsymbol{x}_n]$$

$$= \underbrace{[\boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \cdots \quad \boldsymbol{x}_n]}_{\boldsymbol{U}_1} \left[\begin{array}{c|ccc} \lambda_1 & b_{12} & \cdots & b_{1n} \\ \hline 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{array}\right] \leftarrow \boldsymbol{B}$$

This leads to $\boldsymbol{U}_1^{-1}\boldsymbol{A}\boldsymbol{U}_1 = \begin{bmatrix} \lambda_1 & * \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix}$, and $\boldsymbol{B}$ is of order $(n-1)$.

Let λ_2 be an eigenvalue of $\boldsymbol{B}$ and $\boldsymbol{y}_2$ a unit eigenvector.

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

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

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

$$= \begin{bmatrix} \lambda_1 & * \\ \mathbf{0} & \boldsymbol{M}_2^{-1}\boldsymbol{B}\boldsymbol{M}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ \mathbf{0} & \mathbf{0} & \boldsymbol{C} \end{bmatrix}.$$

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

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

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

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

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

But $\boldsymbol{A}$ is triangularizable ($\boldsymbol{A}$ can be triangularized by a unitary matrix).

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

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

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

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 $\boldsymbol{A}$ is Hermitian, i.e., $\boldsymbol{A} = \boldsymbol{A}^H$ (When $\boldsymbol{A}$ is real, it means $\boldsymbol{A} = \boldsymbol{A}^T$),

this triangular $\boldsymbol{T} = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ is also Hermitian:

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

Therefore, $\boldsymbol{T}$ must be diagonal.

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

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

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

The columns of $\boldsymbol{Q}$ *(or* $\boldsymbol{U}$*) consist of orthonormal eigenvectors of* $\boldsymbol{A}$.

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

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

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

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

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

Remark Split $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}}$ into:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \lambda_1 \boldsymbol{x}_1 \boldsymbol{x}_1^T + \lambda_2 \boldsymbol{x}_2 \boldsymbol{x}_2^T + \lambda_3 \boldsymbol{x}_3 \boldsymbol{x}_3^T$$

$$= (+1)\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} + (+1)\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1)\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= (+1)\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} + (-1)\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{bmatrix} = \lambda_1 \boldsymbol{P}_1 + \lambda_3 \boldsymbol{P}_3,$$

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

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

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

Some special matrices

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

These matrices are all diagonalizable.

Now we want the whole class -- called "normal".

NORMAL MATRICES

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

$$\boldsymbol{N}\boldsymbol{N}^H = \boldsymbol{N}^H\boldsymbol{N}.$$

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

(For example, if $\boldsymbol{A} = \boldsymbol{A}^H$, then $\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{A} = \boldsymbol{A}^2$;

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

We will show that, normal matrices are *exactly* the matrices which are diagonalizable by unitary matrices. (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 $\boldsymbol{A}$ is a normal matrix if and only if there exists a unitary matrix $\boldsymbol{U}$ such that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}$ is diagonal.*)

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

Proof. "➔" Let $\boldsymbol{A}$ be a matrix, and $\boldsymbol{U}$ a unitary matrix such that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{D}$ is diagonal.

Then $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1}$, and $\boldsymbol{A}^H = \boldsymbol{U}\boldsymbol{D}^H\boldsymbol{U}^{-1}$. Thus

$$\boldsymbol{A}\boldsymbol{A}^H = (\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1})(\boldsymbol{U}\boldsymbol{D}^H\boldsymbol{U}^{-1}) = \boldsymbol{U}\boldsymbol{D}\boldsymbol{D}^H\boldsymbol{U}^{-1}$$
$$= \boldsymbol{U}\boldsymbol{D}^H\boldsymbol{D}\boldsymbol{U}^{-1} = \boldsymbol{U}\boldsymbol{D}^H\boldsymbol{U}^{-1}\boldsymbol{U}\boldsymbol{D}\boldsymbol{U}^{-1} = \boldsymbol{A}^H\boldsymbol{A},$$

i.e., $\boldsymbol{A}$ is a normal matrix.

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

$$\boldsymbol{T}\boldsymbol{T}^H = (\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U})(\boldsymbol{U}^H\boldsymbol{A}^H\boldsymbol{U}) = \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{U} = \boldsymbol{U}^{-1}\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{U}$$
$$= (\boldsymbol{U}^{-1}\boldsymbol{A}^H\boldsymbol{U})(\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U}) = \boldsymbol{T}^H\boldsymbol{T},$$

and $\boldsymbol{T}\boldsymbol{T}^H = \boldsymbol{T}^H\boldsymbol{T}$, i.e., $\boldsymbol{T}$ is a normal matrix.

It follows that since $\boldsymbol{T}$ is triangular, $\boldsymbol{T}$ 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 $\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M}$ as *nearly diagonal as possible.*

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

$$\boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

$$\boldsymbol{J}_i - \lambda_i \boldsymbol{I} = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

The diagonal value λ_i is an eigenvalue of $\boldsymbol{J}_i$.

Since $\boldsymbol{J}_i - \lambda_i \boldsymbol{I}$ is of rank $k-1$, the nullspace of $\boldsymbol{J}_i - \lambda_i \boldsymbol{I}$ has dimension 1. In other words, the eigenspace of λ_i is of dimension 1.

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

$$\boldsymbol{J} = \boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \begin{bmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s \end{bmatrix},$$

upper triangular matrix

where each $\boldsymbol{J}_i$ *is a Jordan block, corresponding to an eigenvalue* λ_i *and only one independent eigenvector.*

The same λ_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 $\boldsymbol{J}$.

$$\boldsymbol{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

特征值对应若当块个数由几何重数决定.

Example 5 Find the Jordan form of $\boldsymbol{A}$, where $\boldsymbol{A}=\begin{bmatrix}0&1&2\\0&0&1\\0&0&0\end{bmatrix}$.

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

algebraic multiplicity : 3
geometric multiplicity : 1 $\longrightarrow J = J_{1\,(3\times 3)}$

$$\boldsymbol{B}=\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix},\quad \boldsymbol{C}=\left[\begin{array}{cc:c}0&1&0\\0&0&0\\ \hdashline 0&0&0\end{array}\right],\quad \boldsymbol{D}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}.$$

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

its Jordan form has only one block, and so the Jordan form is $\boldsymbol{B}$.

(Remark: As for $\boldsymbol{D}$= zero matrix, *it is in a family by itself*; the only matrix similar to $\boldsymbol{D}$ is $\boldsymbol{M}^{-1}\boldsymbol{0}\boldsymbol{M}=\boldsymbol{0}$.)

先确定J，后确定M

For $\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $\boldsymbol{J} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, how to find the matrix $\boldsymbol{M}$?

<u>Idea</u>: Since $\boldsymbol{AM} = \boldsymbol{MJ}$, therefore,

$$\boldsymbol{A}[\boldsymbol{x}_1 \ \ \boldsymbol{x}_2 \ \ \boldsymbol{x}_3] = [\boldsymbol{x}_1 \ \ \boldsymbol{x}_2 \ \ \boldsymbol{x}_3] \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

i.e., $\boldsymbol{A}\boldsymbol{x}_1 = \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{x}_2 = \boldsymbol{x}_1$, $\boldsymbol{A}\boldsymbol{x}_3 = \boldsymbol{x}_2$.

$\boldsymbol{A}$ has only one independent eigenvector $\boldsymbol{x}_1 = (1,0,0)^{\mathrm{T}}$, and

$$\boldsymbol{A}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \boldsymbol{0}, \qquad \boldsymbol{A}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad \boldsymbol{A}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}.$$

Finally, $\boldsymbol{x}_1 \ \ \boldsymbol{x}_2 \ \ \boldsymbol{x}_3$ go into $\boldsymbol{M} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and $\boldsymbol{M}^{-1}\boldsymbol{AM} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

Example 6 Find the Jordan form of $\boldsymbol{A}$, where $\boldsymbol{A} = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

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

$$\boldsymbol{B} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad \boldsymbol{C} = \left[\begin{array}{cc:c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hdashline 0 & 0 & 2 \end{array}\right], \quad \boldsymbol{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since $\boldsymbol{A}$ has <u>only one</u> independent eigenvector $(1,0,0)^{\mathrm{T}}$,

its Jordan form has <u>only one</u> block, and so the Jordan form is $\boldsymbol{B}$.

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

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

$$\boldsymbol{B} = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad \boldsymbol{C} = \left[\begin{array}{cc:c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hdashline 0 & 0 & 2 \end{array}\right], \quad \boldsymbol{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Since $\boldsymbol{A}$ has <u>two</u> independent eigenvectors $(1,0,0)^{\mathrm{T}}$ and $(0,2,-1)^{\mathrm{T}}$, thus $\boldsymbol{A}$ has <u>exactly two</u> Jordan blocks, and so the Jordan form of $\boldsymbol{A}$ is $\boldsymbol{C}$.

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

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

$\boldsymbol{A}^k = \boldsymbol{S\Lambda}^k\boldsymbol{S}^{-1}$.

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

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

Since $\boldsymbol{J} = \begin{bmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s \end{bmatrix}$, so $\boldsymbol{J}^k = \begin{bmatrix} \boldsymbol{J}_1^k & & & \\ & \boldsymbol{J}_2^k & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s^k \end{bmatrix}$.

For instance, if λ is a triple eigenvalue with a single eigenvector, then the 3×3 block $\boldsymbol{J}_i$ will enter, and

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

For instance, if λ is a triple eigenvalue with a single eigenvector, then the 3×3 block $\boldsymbol{J}_i$ will enter, and

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

These powers of $\boldsymbol{J}$ are a part of the solutions $\boldsymbol{u}_k$. The other part is the $\boldsymbol{M}$ that connects the original $\boldsymbol{A}$ to the more convenient matrix $\boldsymbol{J}$:

if $\boldsymbol{u}_{k+1} = \boldsymbol{A}\boldsymbol{u}_k$ then $\boldsymbol{u}_k = \boldsymbol{A}^k\boldsymbol{u}_0 = \boldsymbol{M}\boldsymbol{J}^k\boldsymbol{M}^{-1}\boldsymbol{u}_0$.

The Spectral Theorem for Real Symmetric Matrices

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

(一个对称的 $n \times n$ 实矩阵具有下面的特性)

a. $\boldsymbol{A}$ has n real eigenvalues, counting multiplicities. ($\boldsymbol{A}$有n个实特征值, 包含重复的特征值)

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation. (对于每一个特征值λ, 对应特征子空间的维数等于λ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. $\boldsymbol{A}$ is orthogonally diagonalizable. ($\boldsymbol{A}$可以正交对角化)

The Spectral Theorem for Hermitian Matrices

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

(一个 $n \times n$ 厄米特矩阵具有下面的特性)

a. $\boldsymbol{A}$ has n real eigenvalues, counting multiplicities. ($\boldsymbol{A}$有n个实特征值, 包含重复的特征值)

b. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation. (对于每一个特征值λ, 对应特征子空间的维数等于λ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. $\boldsymbol{A}$ can be diagonalized by a unitary matrix. ($\boldsymbol{A}$可以用酉矩阵对角化)

Similarity Transformations

1. $\boldsymbol{A}$ is ***diagonalizable***: The columns of $\boldsymbol{S}$ are eigenvectors and $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda}$.

2. $\boldsymbol{A}$ is ***arbitrary***: The columns of $\boldsymbol{M}$ include "generalized eigenvectors" of $\boldsymbol{A}$, and the Jordan form $\boldsymbol{M}^{-1}\boldsymbol{A}\boldsymbol{M} = \boldsymbol{J}$ is *block diagonal*.

3. $\boldsymbol{A}$ is ***arbitrary***: The unitary $\boldsymbol{U}$ can be chosen so that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{T}$ is *triangular*.

4. $\boldsymbol{A}$ is ***normal***, $\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{A}^H\boldsymbol{A}$: then $\boldsymbol{U}$ can be chosen so that $\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{\Lambda}$.

Special cases of normal matrices, all with orthonormal eigenvectors:

(a) If $\boldsymbol{A} = \boldsymbol{A}^H$ is Hermitian, then all λ_i are real.

(b) If $\boldsymbol{A} = \boldsymbol{A}^T$ is real symmetric, then $\boldsymbol{\Lambda}$ is real and $\boldsymbol{U} = \boldsymbol{Q}$ is orthogonal.

(c) If $\boldsymbol{A} = -\boldsymbol{A}^H$ is skew-Hermitian, then all λ_i are purely imaginary.

(d) If $\boldsymbol{A}$ 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