π

回顾上节课内容：

1. 给定$A \in M_n(\mathbb{C})$，如果对$\lambda \in \mathbb{C}$，存在非零向量$\boldsymbol{x} \in \mathbb{C}^n$，使得$A\boldsymbol{x} = \lambda\boldsymbol{x}$,则称$\lambda$为$A$（在$\mathbb{C}$上）的一个**特征值**, 称非零向量$\boldsymbol{x}$为$A$的一个属于特征值为$\lambda$的**特征向量**。$(\lambda, \boldsymbol{x})$称为**特征对**。

2. $p_A(\lambda) = |\lambda I - A|$称为$A$的特征多项式。$\lambda_0$是$A$的特征值当且仅当$p_A(\lambda_0) = 0$

3. 向量$\boldsymbol{x}_0 \in \mathbb{C}^n$是$A$的属于$\lambda_0$的特征向量当且仅当$\boldsymbol{x}_0 \neq \boldsymbol{0}$且$\boldsymbol{x}_0 \in \mathcal{N}(\lambda_0 I_n - A)$.

4.线性映射视角看待矩阵可对角化：$A = X\Lambda X^{-1}$ 得到$AX = X\Lambda$

令$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$, $A\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i$。因此，$\mathbb{R}^n$的一组基$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$满足$A\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i$

(d) 矩阵的特征值、特征向量与特征多项式的一些性质:

命题：设A为n阶矩阵，特征值为$\lambda_1, \dots, \lambda_n \in \mathbb{C}$。则

(1) 行列式=特征值之积，即$|A| = \lambda_1 \lambda_2 \dots \lambda_n$.

(2) 迹=特征值之和，即$\mathrm{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

(3) $p_A(\lambda) = \lambda^n - \mathrm{trace}(A)\lambda^{n-1} + \cdots + (-1)^n \det A$.

证明：设$\lambda_1, \dots, \lambda_n \in \mathbb{C}$为$A$的特征值。

$p_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$.

$p_A(0) = (-1)^n \lambda_1 \lambda_2 \dots \lambda_n = (-1)^n |A|$.

因此，$|A| = \lambda_1 \lambda_2 \dots \lambda_n$.

将$p_A(\lambda) = \begin{vmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{vmatrix}$沿第一行展开，得到

$$p_A(\lambda) = (\lambda - a_{11})(\lambda - a_{22}) \dots (\lambda - a_{nn}) + (\text{次数} \le n - 2\text{的项})$$

$$= \lambda^n - (a_{11} + \cdots + a_{nn})\lambda^{n-1} + (\text{次数} \le n - 2\text{的项})$$

另一方面，

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$$

$$= \lambda^n - (\lambda_1 + \cdots + \lambda_n)\lambda^{n-1} + (\text{次数} \le n - 2\text{的项})$$

对比得知，$a_{11} + \cdots + a_{nn} = \lambda_1 + \cdots + \lambda_n$.

代数重数与几何重数:

$A \in M_n(\mathbb{C})$, 由代数学基本定理,

$p_A(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \ldots (\lambda - \lambda_k)^{n_k}$, 其中$\lambda_i \in \mathbb{C}$互不相同,

$n_1, \ldots, n_k \geq 1, n_1 + \cdots + n_k = n.$

$AM(\lambda_i) = n_i$ 称为特征值λ_i的**代数重数** (<u>A</u>lgebraic <u>M</u>ultiplicity)

$GM(\lambda_i) = \dim \mathcal{N}(\lambda_i I_n - A)$称为$\lambda_i$的**几何重数**(<u>G</u>eometric <u>M</u>ultiplicity)

我们下小节会看到，对任意特征值λ，总有 $GM(\lambda) \leq AM(\lambda)$

例:

$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. p_Q(\lambda) = (\lambda - i)(\lambda + i).$

$AM(i) = GM(i) = 1, AM(-i) = GM(-i) = 1$

π

实矩阵的复特征值与复特征向量:

命题：$A \in M_n(\mathbb{R})$. $\lambda_0 \in \mathbb{C}\backslash\mathbb{R}$. $(\lambda_0, \boldsymbol{x}_0)$为$A$的特征对。那么$(\overline{\lambda_0}, \overline{\boldsymbol{x}_0})$也为$A$的特征对且$AM(\lambda_0) = AM\left(\overline{\lambda_0}\right), GM(\lambda_0) = GM\left(\overline{\lambda_0}\right)$.

证明：由于$A\boldsymbol{x}_0 = \lambda_0\boldsymbol{x}_0$，取共轭得到$\bar{A}\overline{\boldsymbol{x}_0} = \overline{\lambda_0}\overline{\boldsymbol{x}_0}$.

由于A为实矩阵，$\bar{A} = A$, 于是$A\overline{\boldsymbol{x}_0} = \overline{\lambda_0}\overline{\boldsymbol{x}_0}$. 因此$\left(\overline{\lambda_0}, \overline{\boldsymbol{x}_0}\right)$也为$A$的特征对。

设$n_0 = AM(\lambda_0)$，$n_0' = AM\left(\overline{\lambda_0}\right)$。由于$A$为实矩阵，$p_A(\lambda) \in \mathbb{R}[\lambda]$为实多项式，

由$p_A(\lambda) = (\lambda - \lambda_0)^{n_0}\left(\lambda - \overline{\lambda_0}\right)^{n_0'}\cdots$为实多项式，得到$n_0 = n_0'$

由于$\mathcal{N}(\lambda_0 I - A)$的线性无关的向量$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d$取共轭后得到线性无关的向量$\overline{\boldsymbol{x}_1}, \ldots, \overline{\boldsymbol{x}_d}$。因此 $GM(\lambda_0) = GM\left(\overline{\lambda_0}\right)$.

小结: 对于实矩阵，复特征值成对出现。“成对”意思是:

(1) λ是特征值则 $\bar{\lambda}$也是特征值

(2) λ与$\bar{\lambda}$的重数（代数重数、几何重数）相等

本节小结：

(1) A可对角化：有可逆矩阵X使得$A = X\Lambda X^{-1}$, Λ为对角阵

$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$, $A\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$

(2) 特征值和特征向量：

特征多项式$\det(\lambda I_n - A) = 0$求特征值；

解方程$(\lambda I_n - A)\boldsymbol{x} = \boldsymbol{0}$求特征向量

(3) $|A| = \lambda_1\lambda_2 \cdots \lambda_n$； $\mathrm{trace}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

(4) 对于实矩阵，复特征值成对出现

5.2 对角化和谱分解

主要内容：

(a) 对角化和谱分解的定义

(b) 可对角化的等价条件

(a) 对角化和谱分解的定义

对方阵A，如果存在可逆矩阵$X \in M_n(\mathbb{C})$使得$X^{-1}AX = \Lambda$为对角矩阵，则称A是（在$\mathbb{C}$上）**可对角化**的。

如果方阵A, X, Λ都是实矩阵，则称A在$\mathbb{R}$上可对角化。

当A可对角化时，分解$A = X\Lambda X^{-1}$称为A的**谱分解**。

可对角化矩阵的方幂:

可对角化的矩阵容易计算矩阵的幂：

如果A可对角化，即有可逆矩阵X使得$A = X\Lambda X^{-1}$.

那么，$A^2 = X\Lambda X^{-1}X\Lambda X^{-1} = X\Lambda^2 X^{-1}$　　　$A^k = X\Lambda^k X^{-1}$.

命题:

如果$X^{-1}AX = \Lambda$，Λ为对角矩阵，那么Λ对角线元素为A的特征值。

证明：设 $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

$$p_A(\lambda) = |\lambda I - A| = \left|\lambda I - X\Lambda X^{-1}\right| = |X||\lambda I - \Lambda||X^{-1}| = |\lambda I - \Lambda|$$

$$= (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

从证明中看到，$X^{-1}AX$与A具有相同的特征多项式、特征值

例:

$0 \neq \boldsymbol{u} \in \mathbb{R}^2, \|\boldsymbol{u}\| = 1, P = \boldsymbol{u}\boldsymbol{u}^T$. P为在$\mathbb{R}\boldsymbol{u}$上的正交投影矩阵。

$P(\boldsymbol{u}) = \boldsymbol{u}$. 因此，$\boldsymbol{u}$为特征值为1的特征向量。

取非零向量$\boldsymbol{u}^\perp \perp \boldsymbol{u}$，则$\boldsymbol{u}^\perp$为特征值为0的特征向量。

特征向量$\boldsymbol{u}$，$\boldsymbol{u}^\perp$构成$\mathbb{R}^2$的一组标准正交基

$$P[\boldsymbol{u}, \boldsymbol{u}^\perp] = [\boldsymbol{u}, \boldsymbol{u}^\perp]\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

令$X = [\boldsymbol{u}, \boldsymbol{u}^\perp] \in M_2(\mathbb{R})$，得到$PX = X\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$，即$X^{-1}PX = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

例:

设$\mathcal{M}$为$\mathbb{R}^m$的子空间，$P_{\mathcal{M}}$为在子空间$\mathcal{M}$上的正交投影矩阵

取$\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$为$\mathcal{M}$的一组标准正交基，

取$\boldsymbol{x}_{r+1}, \dots, \boldsymbol{x}_m$为$\mathcal{M}^{\perp}$的一组标准正交基。

则$P_{\mathcal{M}}\boldsymbol{x}_i = \boldsymbol{x}_i, 1 \le i \le r$; $P_{\mathcal{M}}\boldsymbol{x}_i = \boldsymbol{0}, r+1 \le i \le m$.

令$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_m], P_{\mathcal{M}}X = Xdiag(1, \dots, 1, 0 \dots, 0)$,

即，$X^{-1}P_{\mathcal{M}}X = diag(1, \dots, 1, 0 \dots, 0)$.

回顾:

$Q = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$在$\mathbb{R}$上不可对角化，在$\mathbb{C}$上可对角化

(b) 矩阵可对角化的等价条件

问题: 什么样的矩阵可以对角化?

矩阵可对角化的条件:

“有足够多”线性无关的特征向量

两个版本: 粗略的版本和精细的版本

代数重数与几何重数

可对角化的等价条件—粗略版本

定理：

n阶方阵A可对角化当且仅当A有n个线性无关特征向量。

证明：

如果A可对角化，则有可逆矩阵X使得$X^{-1}AX = \Lambda, \Lambda = diag(\lambda_1, \dots, \lambda_n), \lambda_i$为$A$的特征值

$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$可逆，$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$线性无关。

由$X^{-1}AX = \Lambda$得到$AX = X\Lambda$,

于是$A\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i, 1 \le i \le n$. 即$X$的列向量为$A$的特征向量。

反之，设$\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$为$A$的线性无关的特征向量，特征值分别为$\lambda_1, \dots, \lambda_n$

令$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]$. X为可逆方阵。

于是，$A[\boldsymbol{x}_1, \dots, \boldsymbol{x}_n] = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_n]\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$. 因此，$X^{-1}AX = \Lambda$.

命题: (从属于不同特征值的特征向量线性无关)

设$\lambda_1, \ldots, \lambda_k$为矩阵$A$互不相同的特征值，$\boldsymbol{x}_i$分别为属于$\lambda_i$的特征向量，那么$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$线性无关。

证明:

以$k=3$为例。设有c_1, c_2, c_3使得

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + c_3\boldsymbol{x}_3 = \mathbf{0}$$

分别作用A, A^2，得到

$$c_1\lambda_1\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2 + c_3\lambda_3\boldsymbol{x}_3 = \mathbf{0}$$

$$c_1\lambda_1^2\boldsymbol{x}_1 + c_2\lambda_2^2\boldsymbol{x}_2 + c_3\lambda_2^2\boldsymbol{x}_3 = \mathbf{0}$$

我们得到$[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]\begin{bmatrix} c_1 & c_1\lambda_1 & c_1\lambda_1^2 \\ c_2 & c_2\lambda_2 & c_2\lambda_2^2 \\ c_3 & c_3\lambda_3 & c_3\lambda_3^2 \end{bmatrix} = [\mathbf{0}, \mathbf{0}, \mathbf{0}]$

$$[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]\begin{bmatrix} c_1 & & \\ & c_2 & \\ & & c_3 \end{bmatrix}\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix} = [\mathbf{0}, \mathbf{0}, \mathbf{0}]$$

由于λ_i互异，$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 \\ 1 & \lambda_2 & \lambda_2^2 \\ 1 & \lambda_3 & \lambda_3^2 \end{bmatrix}$可逆。于是，

$$[c_1\boldsymbol{x}_1, c_2\boldsymbol{x}_2, c_3\boldsymbol{x}_3] = [\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]\begin{bmatrix} c_1 & & \\ & c_2 & \\ & & c_3 \end{bmatrix} = [\mathbf{0}, \mathbf{0}, \mathbf{0}].$$

由于$\boldsymbol{x}_i$非零，$c_i = 0$.

应用:

如果n 阶方阵A有n个互不相同特征值，那么A可对角化。

证明:

每个特征值一定有一个特征向量。而分属于这n个互不相同特征值的特征向量线性无关，因此，A有n个线性无关特征向量。

问题：如何判断n 阶方阵A有n个互不相同特征值呢?

A有n个互不相同特征值 $\Leftrightarrow p_A(\lambda)$无重根

$\Leftrightarrow p_A(\lambda)$与$p'_A(\lambda)$没有公因子

例:

判断$\begin{bmatrix} 5 & 1 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 9 \end{bmatrix}$是否可对角化?

可对角化，因为它有3个互不相同的特征值5，0，9.

例:

$A=\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$是否可对角化?

不可对角化。因为A的特征值为2，代数重数为2，

如果A可对角化，那么$A=X2I_2X^{-1}=2I_2$. 矛盾。

不可对角化的原因是：几何重数< 代数重数。

回顾：代数重数与几何重数:

设$\lambda_1, \dots, \lambda_k$分别为$A$互不相同的特征值，

$$p_A(\lambda) = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_k)^{n_k}, \qquad n_1 + \cdots + n_k = n.$$

$AM(\lambda_i) = n_i$为特征值λ_i代数重数

$GM(\lambda_i) = \dim \mathcal{N}(\lambda_i I_n - A)$为$\lambda_i$的几何重数

$GM(\lambda_i) =$ 属于特征值λ_i的线性无关的特征向量的个数。

代数重数v.s.几何重数:

命题:

对于特征值λ_i， $GM(\lambda_i) \leq AM(\lambda_i)$, 即几何重数≤ 代数重数。

证明:

设$\boldsymbol{x}_1, \dots \boldsymbol{x}_{d_i}$为特征子空间$\mathcal{N}(\lambda_i I_n - A)$的一组基, $d_i = GM(\lambda_i)$,

将它们扩充为$\mathbb{C}^n$的一组基$\boldsymbol{x}_1, \dots, \boldsymbol{x}_{d_i}, \boldsymbol{u}, \dots, \boldsymbol{v}$.

$$A[\boldsymbol{x}_1, \dots, \boldsymbol{x}_{d_i}, \boldsymbol{u}, \dots, \boldsymbol{v}] = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_{d_i}, \boldsymbol{u}, \dots, \boldsymbol{v}] \begin{bmatrix} \lambda_i I_{d_i} & B \\ O & A_1 \end{bmatrix}$$

令$X = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_{d_i}, \boldsymbol{u}, \dots, \boldsymbol{v}]$. $X^{-1}AX$与A有相同的特征多项式。因此,

$$p_A(\lambda) = \det\left(\lambda I_n - \begin{bmatrix} \lambda_i I_{d_i} & B \\ 0 & A_1 \end{bmatrix}\right) = \begin{vmatrix} (\lambda - \lambda_i) I_{d_i} & -B \\ 0 & \lambda I_{n-d_i} - A_1 \end{vmatrix}$$

$$= (\lambda - \lambda_i)^{d_i} \left|\lambda I_{n-d_i} - A_1\right|.$$

因此，$AM(\lambda_i) \geq d_i = GM(\lambda_i)$.

n阶方阵A, $\lambda_1, \dots, \lambda_k$为$A$的特征值，互不相同

特征值: λ_1 λ_2 … λ_k

特征多项式 = $(\lambda-\lambda_1)^{n_1}$ $(\lambda-\lambda_2)^{n_2}$ … $(\lambda-\lambda_k)^{n_k}$

$n = n_1 + n_2 + \cdots + n_k$

代数重数: n_1 n_2 … n_k

几何重数: d_1 d_2 … d_k

线性无关的特征向量总数 $d = d_1 + d_2 + \cdots + d_k$

回顾：方阵A可对角化 $\Leftrightarrow$ $d = n$

(由于$d_i \leq n_i$)

$\Leftrightarrow$ $d_i = n_i, \forall i$

可对角化的等价条件—代数重数v.s.几何重数版本:

定理:

A可对角化当且仅当对所有特征值$\lambda_i, GM(\lambda_i) = AM(\lambda_i)$.

即每个特征值都是有足够多的线性无关的特征向量。

例:

$\boldsymbol{u}, \boldsymbol{v} \neq \boldsymbol{0}$, 讨论秩1矩阵$A = \boldsymbol{u}\boldsymbol{v}^T = \left(u_i v_j\right)_{1\leq i,j\leq n}$是否可对角化。

由于A的秩为1, $\dim\mathcal{N}(A) = n - rank(A) = n - 1.$

因此,0是A的特征值, $GM(0) = n - 1.$

$trace(A) = u_1 v_1 + \cdots + u_n v_n = \boldsymbol{u}^T\boldsymbol{v}$。

因为$trace(A)$是A的特征值之和，且A已有至少$n-1$特征值为0，

$\boldsymbol{u}^T\boldsymbol{v}$也是$A$的特征值.

于是A的特征多项式$p_A(\lambda) = \lambda^{n-1}\left(\lambda - \boldsymbol{u}^T\boldsymbol{v}\right)$.

如果$\boldsymbol{u}^T\boldsymbol{v} = 0$， $AM(0) = n$， A不可对角化。

如果$\boldsymbol{u}^T\boldsymbol{v} \neq 0$， $GM(0) = n - 1 = AM(0)$， $GM(\boldsymbol{u}^T\boldsymbol{v}) = AM(\boldsymbol{u}^T\boldsymbol{v}) = 1$,$A$可对角化。

例:

形如$J_n(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & & \\ & \lambda_0 & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_0 \end{bmatrix}$的$n$阶方阵称为若尔当块（Jordan block).

讨论$J_n(\lambda_0)$是否可对角化。

$p_{J_n(\lambda_0)}(\lambda) = (\lambda - \lambda_0)^n$因此，$J_n(\lambda_0)$的特征值为$\lambda_0, AM(\lambda_0) = n$.

$$\lambda_0 I_n - J_n(\lambda_0) = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

$$\dim\mathcal{N}\big(\lambda_0 I_n - J_n(\lambda_0)\big) = n - rank\left(\begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}\right) = 1.$$

因此，如果$n \geq 2, J_n(\lambda_0)$不可对角化。

本节小结:

1. n阶方阵可对角化 $\Leftrightarrow$有n个线性无关的特征向量

2. 如果方阵有n个互异的特征值, 则可对角化

   原因: 属于不同特征值的特征向量线性无关

3. A可对角化当且仅当对所有特征值$\lambda_i, GM(\lambda_i) = AM(\lambda_i)$.