Week 11 讲稿

[Def] $\mathcal{A}: V\to V$ 是一个线性变换．若 $\exists\ 0\neq v\in V$, $\lambda\in\mathbb{R}$ s.t. $\mathcal{A}v=\lambda v$

则称 v 是关于特征值 λ 的特征向量．

Rmk: ①注意，特征值和特征向量是对线性空间上的线性变换定义的．

②选定基后，$\mathcal{A}$ 可用矩阵表示．注意，方阵才有特征值（$\mathcal{A}$ 是 $V\to V$ 两个同维空间之间的 map）

③ 特征向量非零

④ 特征值可以是 0

⑤ 矩阵 A 的特征值 & 特征向量是 $Av=\lambda v$.

• 特征值 & 向量：$Av=\lambda v \Rightarrow (A-\lambda I)\underset{}{\overset{\text{选基后的坐标}}{v}}=0$． $A-\lambda I$ 存在非零解 v，故 $A-\lambda I$ 不可逆， v是特征向量，故非零

即 $\det(A-\lambda I)=0$. $\det(A-\lambda I)$ 称为特征多项式．$(A-\lambda I)v=0$ 的解 v 是特征向量．

若 $A\in M_{n\times n}(\mathbb{R})$ 则 $|A-\lambda I|$ 是关于 λ 的 n 次实系数方程．

若 $A\in M_{n\times n}(\mathbb{C})$ 则 $|A-\lambda I|$ 是关于 λ 的 n 次复系数方程．

△Question：A 是 n×n 矩阵，则 A 是否有 n 个特征值？

[Exp] $f(x)=(x-1)^2$ 则此时 $x=1$ 是 $f(x)=0$ 的二重根，根 1 被算入两次．

同理，若 $g(x)=(x-2)^2(x-3)^3$ 则 $x=2$ 是 $g(x)=0$ 的二重根，$x=3$ 是 $g(x)=0$ 的三重根．
我们称 $g(x)=0$ 在计入重数意义下有 5 个根 $(2+3=5)$.

[Thm]（代数学基本定理）设 $f(x)$ 是关于 x 的 n 次复系数多项式，i.e.,
$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $a_i\in\mathbb{C}$．则 $f(x)=0$ 必有 n 个复根（在计算重根意义下）

* 不建议大家看此 thm 证明，用到复变函数的内容远超此课程要求．

[Application] $A\in M_{n\times n}(\mathbb{C})$ 则 $|A-\lambda I|$ 是关于 λ 的 n 次复系数多项式．
故 $|A-\lambda I|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$ 必有 n 个复根．（其中 $\lambda_1,\cdots,\lambda_n$ 可以有相同值，例如 λ_1 可以等于 λ_2，这表示 $\lambda_1=\lambda_2$ 是 $|A-\lambda I|=0$ 的二重根）

[Exp] 对于 $A\in M_{n\times n}(\mathbb{R})$，A 必有 n 个复特征值，但可以有少于 n 个实特征值．

考虑 $A=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$ 是旋转 90° 矩阵．

$|A-\lambda I|=\lambda^2+1=0$ A 有 0 个实特征值但 A 有 2 个复特征值 $\pm i$

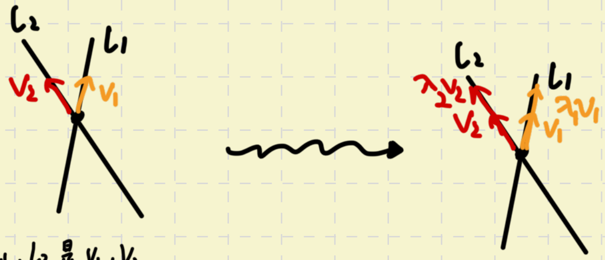
△ 特征值与特征向量的几何解释

$\mathcal{A}: V\to V$ 在选定基下用矩阵 A 表示．为了能画图，设 $\dim V=2$ 即 A 是 2×2 矩阵．

设 A 有两个特征向量 v_1,v_2，分别对应特征值 λ_1,λ_2，即 $Av_1=\lambda_1v_1$, $Av_2=\lambda_2v_2$.

设 $\lambda_1,\lambda_2\in\mathbb{R}$.

则A所表示的线性变换是：



l_1, l_2是v_1, v_2所在直线

l_1, l_2线上的所有向量分别拉伸λ_1与λ_2倍

[Exp] 1) $A=\begin{bmatrix}1 & \\ & 2\end{bmatrix}$ $\quad \lambda_1=1 \quad v_1=\begin{bmatrix}1\\0\end{bmatrix}$, $\lambda_2=2 \quad v_2=\begin{bmatrix}0\\1\end{bmatrix}$

(2) $A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$是旋转矩阵，所有向量都被转$90^\circ$而没有拉伸，因此$A$没有实特征值.

• [Thm] 若$n\times n$矩阵A有n个特征值，则 $tr(A)=\sum_{i=1}^{n}\lambda_i$ ， $|A|=\prod_{i=1}^{n}\lambda_i$

＊注意此公式计入重根. 例如若$\lambda_1=\lambda_2=2$，$\lambda_3=-2$，则 $tr(A)=2+2+(-2)=2$
$|A|=2\times 2\times(-2)=-8$.

pf: A有n个特征值，即$|A-\lambda I|=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$.

取$\lambda=0$得 $|A|=\lambda_1\lambda_2\cdots\lambda_n$

$$\det(A-\lambda I)=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda\end{vmatrix}$$ 分别计算 L.H.S. 与 R.H.S.

中$(-\lambda)^{n-1}$的系数.

R.H.S. $\det(A-\lambda I)=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda)$

$$=\sum_{i=1}^{n}\lambda_i(-\lambda)^{n-1}+\cdots$$

L.H.S.

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda\end{vmatrix}$$

$$\overset{\text{按第1行}}{\underset{\text{展开}}{=\!=}}(a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & \cdots & a_{2n}\\ \vdots & & \\ a_{n2} & \cdots & a_{nn}-\lambda\end{vmatrix}-a_{12}M_{1,2}+a_{13}M_{1,3}-\cdots+(-1)^{1+n}a_{1n}M_{1,n}$$

$M_{1,2}, M_{1,3}, \cdots, M_{1,n}$中只含有$n-2$个$\lambda$，因此$M_{1,2}, \cdots, M_{1,n}$是关于$\lambda$的$n-2$次多项式，$(-\lambda)^{n-1}$只存在于第一项中.

$$(a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = (a_{11}-\lambda)\left[(a_{22}-\lambda)B_{1,1} - a_{23}B_{1,2} + \cdots + (-1)^{n+1}a_{2n}B_{1,n}\right]$$

其中$B_{1,i}$是$\begin{vmatrix} a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$的余子式.

$B_{1,2}, B_{1,3} \cdots B_{1,n}$中只含有$n-3$个$\lambda$，因此$B_{12} \cdots B_{1n}$是关于$\lambda$的$n-3$次多项式，可以舍去. 故$(-\lambda)^{n-1}$只在$(a_{11}-\lambda)(a_{22}-\lambda)B_{11}$中. 同理可得

$(-\lambda)^{n-1}$只在$(a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$中. 故 $(-\lambda)^{n-1}$系数为$\sum_{i=1}^{n} a_{ii} = tr(A)$.

故 $\sum_{i=1}^{n} \lambda_i = tr(A)$.

[Application] A有0特征值 $\Leftrightarrow$ A不可逆.

$|A| = \prod_i \lambda_i = 0$

- 一些特殊矩阵的特征值

△ $Av = \lambda v \Rightarrow A^n v = A^{n-1}Av = A^{n-1}\lambda v = \lambda A^{n-1} v = \lambda A^{n-2} Av = \lambda^2 A^{n-2} v$
$= \cdots = \lambda^n v$.

△ $Av = \lambda v \Leftrightarrow A^{-1}v = \lambda^{-1}v$

△ A与A^T有相同特征值. $|A-\lambda I| = 0 \Leftrightarrow |(A-\lambda I)^T| = 0 \Leftrightarrow |A^T - \lambda I| = 0$

同一个λ满足$|A-\lambda I|=0$与$|A^T-\lambda I|=0$ 故λ既是A的eigval 又是A^T的eigval.

* $(A+B)^T = A^T + B^T$, $|A+B| \neq |A| + |B|$

△ 设P是投影矩阵，即P是满足$P^2=P$的对称矩阵，则P的每个特征值是0或1.

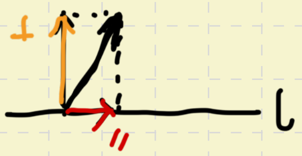
$$\begin{cases} P^2 v = \lambda^2 v \\ Pv = \lambda v \\ P^2 = P \end{cases} \Rightarrow \lambda^2 v = \lambda v \Leftrightarrow (\lambda^2 - \lambda)v = 0$$ 由于特征向量$v \neq 0$，有$\lambda^2 - \lambda = 0$

故 $\lambda = 0$或1.

* Maybe这里的"或"字很令人疑惑，它的含义如下：设P有n个特征向量$\lambda_1, \lambda_2, \cdots, \lambda_n$. 则$\lambda_i = 0$或1. 可以全是1，例如单位阵I，也可全是0，例如零矩阵

* How can we understand? 把 $\mathbb{R}^n$ 往子空间A上做投影，含义

恰为：所有A上向量不动(乘1)，所有与A垂直向量投到0.



⊥分量乘上0

//分量乘上1

• 对角化 ：设 A 是 $n\times n$ 矩阵.

△ 设A有n个线性无关特征向量 $v_1,\dots,v_n$ 对应特征值 $\lambda_1,\dots,\lambda_n$，即

$Av_i = \lambda_i v_i$，$i\in[\![1,n]\!]$. 表示 $i=1,2,\dots,n$

$$A[v_1\cdots v_n] = [Av_1\ \ Av_2\cdots Av_n]$$

$$= [\lambda v_1\ \ \lambda v_2\ \cdots \lambda v_n]$$

$$= [v_1\ \ v_2\ \cdots v_n]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$$

记 $S := [v_1\cdots v_n]$. 由于 $\{v_1,\dots,v_n\}$ L. ind., S可逆.

故 $AS = S\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}$ 亦写作 $A = S\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}S^{-1}$ 称为A的对角化.

△ 从对A作对角化的过程可以见出 A可作对角化 $\Leftrightarrow$ A有n个L. ind. 特征向量.

* Rmk：注意A可作对角化与特征值是否含0无关！

[Application] 矩阵对角化可以用来计算乘积

$$A^n = (S\Lambda S^{-1})^n = S\Lambda S^{-1}S\Lambda S^{-1}\cdots S\Lambda S^{-1} = S\Lambda^n S^{-1}.$$

• 判断A是否可对角化方法：

[prop] 若 $x_1, x_2,\dots,x_k$ 是矩阵A 特征值 $\lambda_1,\lambda_2,\dots,\lambda_k$ 的特征向量，则这些特征向量 L. ind.

Pf： We prove by induction (数学归纳法). 当 $k=2$. 考虑 $c_1x_1+c_2x_2=0$. (*)

$$\left.\begin{array}{ll} A(*) \Rightarrow c_1\lambda_1x_1+c_2\lambda_2x_2=0 & (1) \\ \lambda_1(*) \Rightarrow c_1\lambda_1x_1+c_2\lambda_1x_2=0 & (2)\end{array}\right\}\ (1)-(2)\text{ 得 } c_2(\lambda_2-\lambda_1)x_2=0$$

由 $\lambda_2-\lambda_1\neq 0$, $x_2\neq 0$ 得 $C_2=0$. 将 $C_2=0$ 代入(*) 得 $C_1x_1=0$. $x_1\neq 0$ 故 $C_1=0$.

假设对 $k-1$ 个向量成立，即任意 $k-1$ 不同特征值对应的特征向量 L.ind.

设 $C_1x_1+C_2x_2+\cdots+C_kx_k=0$ (*)，两边同时左乘 A，有

$C_1Ax_1+C_2Ax_2+\cdots+C_kAx_k=0$ 即 $C_1\lambda_1x_1+C_2\lambda_2x_2+\cdots+C_k\lambda_kx_k=0$ (1)

(*) 两边同乘 λ_1 得 $C_1\lambda_1x_1+C_2\lambda_1x_2+\cdots+C_k\lambda_1x_k=0$ (2)

(1)-(2) 得 $C_2(\lambda_2-\lambda_1)x_2+\cdots+C_k(\lambda_k-\lambda_1)x_k=0$ 由 $x_2,\cdots,x_k$ L.ind.,

得 $C_2(\lambda_2-\lambda_1)=\cdots=C_k(\lambda_k-\lambda_1)=0$. 又由 $\lambda_i-\lambda_1\neq 0$ 得 $C_2=\cdots=C_k=0$.

将 $C_2=\cdots=C_k=0$ 代回(*)，得 $C_1=0$. 故 $x_1,\cdots,x_k$ L.ind..

[Application] 若 A 有 n 个互不相同特征值，则 A 有 n 个 L.ind. 特征向量，则 A 可对角化.

- 特征子空间:

回顾求特征向量的过程 $(A-\lambda I)v=0$ 的解是特征向量，因此所有关于 λ 的特征向量构成一个线性空间，即 $N(A-\lambda I)$，称为特征子空间，记作 $E(A,\lambda)$

* $\dim E(A,\lambda)$ 是对应 λ 特征值的极大 L.ind. 特征向量组的个数.

[Fact] 设 A 可对角化，则 $\mathbb{R}^n=\bigoplus_{i=1}^{k}E(A,\lambda_i)$ for some k.

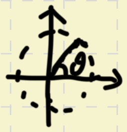
[Corollary] $n=\dim\mathbb{R}^n=\sum_{i=1}^{k}\dim E(A,\lambda_i)$

复数补充

△ $\mathbb{C}=\{a+ib\,|\,a,b\in\mathbb{R}\}$.

$\mathbb{C}$作为$\mathbb{C}$-vector space是1维的，基是$\{1\}$；$\mathbb{C}$作为$\mathbb{R}$-vector space是2维的，基是$\{1,i\}$.

△欧拉公式 $e^{i\theta}=\cos\theta+i\sin\theta$



$$e^{i(\theta+2\pi n)}=\cos(\theta+2n\pi)+i\sin(\theta+2n\pi)=\cos\theta+i\sin\theta.$$

△$\forall z\in\mathbb{C}$，z形如 $z=re^{i\theta}$. 其中r称为复数的模长，θ称为复数的幅角.

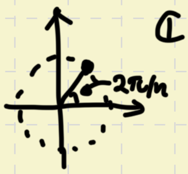
△若$z=a+ib$，则称$\bar{z}:=a-ib$是z的共轭.（若$z=re^{i\theta}$，$\bar{z}=re^{-i\theta}$）

△常用公式：$z\bar{z}=|z|^2$

• 复数开根

△Question: $x^n=1$ 在$\mathbb{C}$中的解？

Answer: $e^{i\frac{2\pi}{n}k}$，$k=0,1,\cdots,n-1$ 共n个根，称为本原n次单位根.



$\mathbb{C}$ 复平面.

* Rmk: 本原n次单位根的解构成循环群$\mathbb{Z}_n$.

△Question: 给定$z\in\mathbb{C}$，求$x^n=z$的解？

设$z=re^{i\theta}$. 则$x^n=re^{i\theta}$. 它的解为 $r^{\frac{1}{n}}e^{i\frac{2\pi}{n}k}$，$k=0,1,2,\cdots,n-1$.

Naively，解是 $r^{\frac{1}{n}}e^{\frac{i\theta}{n}}$ —解不够 但 $x^n=re^{i\theta}\cdot 1$ 其中1开n次根有解$e^{i\frac{2\pi}{n}k}$

故 $x^n=re^{i\theta}$的解是 $x=r^{\frac{1}{n}}e^{i\frac{\theta}{n}}e^{i\frac{2\pi}{n}k}$，$k=1,2,\cdots,n-1$

Check: $\left(r^{\frac{1}{n}}e^{i\frac{\theta}{n}}e^{i\frac{2\pi}{n}k}\right)^n=re^{i\theta}e^{i2\pi k}=re^{i\theta}$.

• Big picture

$$\underset{\mathbb{R}^2\text{上的积分}}{\iint f(x)\,dx\,dy}\xrightarrow{\text{推广}}\underset{\mathbb{C}\text{上的积分}}{\int f(z)\,dz}\xrightarrow[\text{流形}]{\mathbb{C}\text{是特殊的}}\underset{\text{流形上的积分}}{\int\ldots d\omega}\sim n\text{-form}$$

*复值函数有比实值函数更好的性质，因此往往让问题变简单.

40. There are six 3 by 3 permutation matrices P. What numbers can be the *determinants* of P? What numbers can be *pivots*? What numbers can be the *trace* of P? What *four numbers* can be eigenvalues of P?

(i) 设 P 只在 $a_{1i_1}, a_{2i_2}, a_{3i_3}$ 处为 1，其中 i_1, i_2, i_3 是 $\{1,2,3\}$ 的一个排列

$$|P| = \sum_{j_1 j_2 j_3} a_{1j_1} a_{2j_2} a_{3j_3} (-1)^{\tau(j_1 j_2 j_3)} = a_{1i_1} a_{2i_2} a_{3i_3} (-1)^{\tau(i_1 i_2 i_3)} = (-1)^{\tau(i_1 i_2 i_3)}$$

(ii) 非零主元是 1.

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}$$

(iii) $tr(P) =$ 0 或 1 或 3

(iv) 6 个置换矩阵除了 I 之外都是二阶元或三阶元.

对于二阶元 即 $P^2 = I$ 有 $P^2 v = \lambda^2 v = v$ 即 $\lambda^2 = 1$ 故 $\lambda = \pm 1$

对于三阶元 即 $P^3 = I$ 有 $P^3 v = \lambda^3 v = v$ 即 $\lambda^3 = 1$ 故 $\lambda = 1, e^{i\frac{2\pi}{3}}, e^{i\frac{4\pi}{3}}$

因此 4 种可能特征值是 $\pm 1, e^{i\frac{2\pi}{3}}, e^{i\frac{4\pi}{3}}$

4. 设 A 为 $m \times n$ 矩阵, B 为 $n \times m$ 矩阵, λ 为一个未定元. 证明:

$$\lambda^n |\lambda I_m - AB| = \lambda^m |\lambda I_n - BA|.$$

pf: Claim: AB 与 BA 有相同的非零特征值.

pf for the claim: 设 $ABx = \lambda x$, $\lambda \neq 0$. 左乘 B 得 $BA(Bx) = \lambda Bx$

$\lambda \neq 0$，故 $Bx \neq 0$.

故 λ 是 BA 的特征值. 同理任意 BA 非零特征值是 AB 非零特征值.

设 AB 特征值为 $\underbrace{\lambda_1, \lambda_2, \cdots, \lambda_k, 0, 0, \cdots, 0}_{m}$

则 BA 特征值为 $\underbrace{\lambda_1, \cdots, \lambda_k, 0, \cdots, 0}_{n}$

$$\begin{aligned} \lambda^n |\lambda I_m - AB| &= \lambda^n (\lambda - \lambda_1) \cdots (\lambda - \lambda_k) \lambda^{m-k} \\ &= \lambda^m (\lambda - \lambda_1) \cdots (\lambda - \lambda_k) \lambda^{n-k} \\ &= \lambda^m |\lambda I_n - BA| \end{aligned}$$

5. 设 A, B, C, D 都为 n 阶矩阵, 且 $AC = CA$. 令

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

证明: $|M| = |AD - CB|$. 注意: 这里矩阵 A 不一定是可逆的.

pf:
(i) Special case: A 可逆.

用分块矩阵版本的行变换 $\begin{bmatrix} I & \\ -C & A \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A & B \\ AC-CA & AD-CB \end{bmatrix}$

$= \begin{bmatrix} A & B \\ 0 & AD-CB \end{bmatrix}$

两边取行列式 $|A||M| = |A||AD-CB|$ 即

$|M| = |AD-CB|$

(ii) General case: A 不可逆.

A 不可逆 $\Rightarrow$ A 有 0 特征值 $\Rightarrow$ $|A-\lambda I| = 0$ 有零解

Claim: 任意足够小 $\varepsilon > 0$, $A + \varepsilon I$ 可逆.

pf for the claim: $|A+\varepsilon I - \lambda' I| = |A - (\lambda' - \varepsilon) I| = 0$.

比较 $|A-\lambda I|$ 与 $|A-(\lambda'-\varepsilon)I|$, 得 $\lambda' - \varepsilon = \lambda$, 即任何 $|A+\varepsilon I - \lambda' I| = 0$ 的解 λ', 有 $\lambda' - \varepsilon$ 是 $|A-\lambda I| = 0$ 的解. 因此 $A+\varepsilon I$ 的所有特征值都是 A 特征平移 ε, 因此不存在零解, 即 $A+\varepsilon I$ 可逆.

$A_\varepsilon C = (A+\varepsilon I)C = AC + \varepsilon C = CA + \varepsilon C = C A_\varepsilon$

考虑 $M_\varepsilon = \begin{bmatrix} A_\varepsilon & B \\ C & D \end{bmatrix}$. 由 case (i), 有 $|M_\varepsilon| = |A_\varepsilon D - CB|$

$$\lim_{\varepsilon \to 0} |M_\varepsilon| = |M|$$

$$\parallel$$

$$\lim_{\varepsilon \to 0} |A_\varepsilon D - CB| = |AD - CB|.$$