3. 行列式与迹 (trace)

$A=(a_{ij})\in M_n(\mathbb{C})$ $\quad f_A(\lambda)=|\lambda I-A|=\overset{(*1)}{\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}}\overset{(*2)}{=}(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$ 代数学基本定理

(可能有重根)

$f_A(\lambda)=0\Rightarrow$ 特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n\in\mathbb{C}\leftarrow$ 未必是实数.

定理 $A\in M_n(\mathbb{C})$. 则 $trA\ (:=\sum_{i=1}^n a_{ii})=\sum_{i=1}^n\lambda_i$. $\quad |A|=\prod_{i=1}^n\lambda_i$.

pf. (1) 考虑(*1)及(*2)中 λ^{n-1} 的系数.

(*1): $-\sum_{i=1}^n a_{ii}$. $\quad$ (*2): $-\sum_{i=1}^n\lambda_i\Rightarrow\sum_{i=1}^n\lambda_i=trA$.

(2) 考虑(*1)及(*2)中的常数项(令 $\lambda=0$) $\Rightarrow|-A|=(-1)^n\prod_{i=1}^n\lambda_i\Rightarrow|A|=\prod_{i=1}^n\lambda_i$. #

注: ① 可用来验证特征值是否算错.

② $A\in M_n(\mathbb{C})$ 可逆 $\Leftrightarrow$ A的n个特征值全不为0.

若A可逆且 $A\vec{x}=\lambda\vec{x}$, $\vec{x}\neq\vec{0}$, 则 $\quad A^{-1}\vec{x}=\frac{1}{\lambda}\vec{x}$. $\quad\frac{1}{\lambda}$ 为 A^{-1} 特征值. same eigenvectors.

例 已知 $A\in M_n(\mathbb{C})$ 的n个特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 求 $2I-A$ 的特征值及 $|2I-A|$.

解: $f_A(\lambda)=|\lambda I-A|=\prod_{i=1}^n(\lambda-\lambda_i)$.

$f_{2I-A}(\lambda)=|\lambda I-(2I-A)|=|(\lambda-2)I+A|=(-1)^n|(2-\lambda)I-A|=(-1)^n f_A(2-\lambda)=(-1)^n\prod_{i=1}^n(2-\lambda-\lambda_i)$

$=\prod_{i=1}^n(\lambda-(2-\lambda_i))$

令 $f_{2I-A}(\lambda)=0\Rightarrow 2I-A$ 的特征值为 $2-\lambda_1, 2-\lambda_2,\cdots,2-\lambda_n$.

$\therefore\det(2I-A)=\prod_{i=1}^n(2-\lambda_i)$.

4. AB及A+B的特征值

设λ是A的特征值, β是B的特征值 $\not\Rightarrow$ λ+β是A+B的特征值

$\not\Rightarrow$ λβ是AB ····

$AB\vec{x}=\beta A\vec{x}$ (B的特征向量; 未必是A的特征向量)

定理 $A,B\in M_n$. A and B share the same n independent eigenvectors $\Leftrightarrow AB=BA$.

注: 设 $A\vec{x}=\lambda\vec{x}$, $B\vec{x}=\beta\vec{x}$, $\vec{x}\neq\vec{0}$, 则

$AB\vec{x}=A(B\vec{x})=A\beta\vec{x}=\beta(A\vec{x})=\lambda\beta\vec{x}$

$BA\vec{x}=B(\lambda\vec{x})=\lambda(B\vec{x})=\lambda\beta\vec{x}$

λβ是AB的特征值.

§6.2 相似对角化

<u>定义</u> $A, B \in M_n$. 若∃可逆矩阵 $X \in M_n$. s.t. $X^{-1}AX = B$. 则称 B 相似于 A, 记作 $B \sim A$.

注: (1) 自反性. $A \sim A$. $\forall A \in M_n$. $I^{-1}AI = A$

(2) 对称性. 若 $B \sim A$, 则 $A \sim B$. $X^{-1}AX = B \Rightarrow XBX^{-1} = A$ 称A与B相似

(3) 传递性. 若 $A \sim B$, $B \sim C$. 则 $A \sim C$. $A = X^{-1}BX$, $B = Q^{-1}CQ \Rightarrow A = (QX)^{-1}C(QX)$

<u>性质</u> $A, B \in M_n$.

(1) 若 $A \sim B$, 则 $A^m \sim B^m$, 其中m是正整数.

Pf. 若 $A \sim B$, 则∃可逆矩阵 $X \in M_n$. s.t. $A = X^{-1}BX$.

则 $A^m = (X^{-1}BX)^m = (X^{-1}BX)(X^{-1}BX)\cdots(X^{-1}BX) = X^{-1}B^mX$

$\therefore A^m \sim B^m$.

(2) 若 $A \sim B$. 且A可逆, 则B也可逆, 且 $A^{-1} \sim B^{-1}$.

Pf. $B = XAX^{-1}$, 可逆矩阵的乘积仍可逆 $\Rightarrow$ B可逆.

$B^{-1} = (XAX^{-1})^{-1} = XA^{-1}X^{-1}$, $\therefore B^{-1} \sim A^{-1}$.

(3) 相似矩阵有相同的特征值和相同的特征多项式.

Pf. 设 $A \sim B$. 则有 $A = X^{-1}BX$. 其中X可逆.

$f_A(\lambda) = |\lambda I - A| = |\lambda I - X^{-1}BX| = |X^{-1}(\lambda I - B)X| = \underbrace{|X^{-1}|\,|\lambda I - B|\,|X|}_{\text{相乘为}1} = |\lambda I - B| = f_B(\lambda)$.

$\therefore$ A与B有相同的特征多项式, 从而有相同的特征值.

(4) 相似矩阵有相同的迹和相同的行列式.

Pf. $\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i$. $|A| = \prod_{i=1}^{n} \lambda_i$.

<u>Q</u>: 什么样的矩阵与对角阵相似? 即可相似对角化?

好处: 若 $A = X^{-1}\Lambda X$, 其中 $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.

则 $A^m = X^{-1}\Lambda^m X = X^{-1}\operatorname{diag}(\lambda_1^m, \lambda_2^m, \cdots, \lambda_n^m)X$. 易求解.

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

Pf. "$\Rightarrow$". 设$A \in M_n$与$\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$相似

即$\exists$可逆矩阵$X \in M_n$. s.t. $X^{-1}AX = \Lambda$

则$AX = X\Lambda$. 令$X = [\vec{x_1}\ \vec{x_2} \cdots \vec{x_n}]$, 有

$$A[\vec{x_1}\ \vec{x_2} \cdots \vec{x_n}] = [\vec{x_1}\ \vec{x_2} \cdots \vec{x_n}]\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

$\therefore A\vec{x_i} = \lambda_i \vec{x_i}$. $i = 1, 2, \cdots, n$.

又$\because X$可逆. $\therefore \vec{x_i} \neq \vec{0}$, $i = 1, 2, \cdots, n$ 且$\vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}$线性无关.

$\therefore \vec{x_1}, \vec{x_2}, \cdots, \vec{x_n}$线性无关且分别是$A$的属于特征值$\lambda_1, \lambda_2, \cdots, \lambda_n$的特征向量.

"$\Leftarrow$". 若$A\vec{x_i} = \lambda_i \vec{x_i}$, $i = 1, 2, \cdots, n$ 且$\{\vec{x_i}, i = 1, \cdots, n\}$线性无关.

取$X = [\vec{x_1}\ \vec{x_2} \cdots \vec{x_n}]$, 则有$X$可逆, 且满足

$AX = X\, diag(\lambda_1, \cdots, \lambda_n) \Rightarrow X^{-1}AX = \Lambda$. #.

定理 设$\lambda_1, \lambda_2, \cdots, \lambda_s$是$A$的$s$个互异的特征值, $\vec{x}_{i1}, \vec{x}_{i2}, \cdots, \vec{x}_{i,m_i}$是$A$的属于$\lambda_i$的$m_i$个线性无关的特征向量 $i = 1, 2, \cdots, s$, 则$\vec{x}_{11}, \vec{x}_{12}, \cdots, \vec{x}_{1,m_1}, \vec{x}_{21}, \vec{x}_{22}, \cdots, \vec{x}_{2m_2}, \cdots, \vec{x}_{s1}, \vec{x}_{s2}, \cdots, \vec{x}_{s m_s}$也线性无关.

Pf. 对s作数学归纳法.

① $s = 1$时, $\vec{x}_{1}, \cdots, \vec{x}_{1,m_1}$线性无关 结论成立.

② 设结论对$s-1$的情形成立.

③ s的情形.

设$\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\vec{x}_{ij} = \vec{0}$. (*)

用A左乘(*), 又有$A\vec{x}_{ij} = \lambda_i \vec{x}_{ij}$. 得$\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\lambda_i \vec{x}_{ij} = \vec{0}$

用λ_s乘(*), 得$\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\lambda_s \vec{x}_{ij} = \vec{0}$

相减, 得. $\sum_{i=1}^{s-1}\sum_{j=1}^{m_i} k_{ij}(\lambda_i - \lambda_s)\vec{x}_{ij} = \vec{0}$.

由②, 有$\vec{x}_{ij}$, $i = 1, \cdots, s-1$, $j = 1, \cdots, m_i$ 线性无关.

$\therefore k_{ij}(\lambda_i - \lambda_s) = 0$. $i = 1, \cdots, s-1$, $j = 1, \cdots, m_i$.

而$\lambda_i \neq \lambda_s$ $(i = 1, \cdots, s-1)$. $\therefore k_{ij} = 0$, $i = 1, \cdots, s-1$, $j = 1, \cdots, m_i$

代入(*), 有$\sum_{j=1}^{m_s} k_{sj}\vec{x}_{sj} = \vec{0}$, 而$\vec{x}_{sj}$, $j = 1, \cdots, m_s$ 线性无关, $\therefore k_{sj} = 0$, $j = 1, \cdots, m_s$.

$\therefore \vec{x}_{ij}$, $i = 1, \cdots, s$, $j = 1, \cdots, m_i$ 线性无关.

由归纳假设法, 命题成立. #.

定理. 若$A\in M_n$有n个互异的特征值$\lambda_1, \lambda_2, \cdots, \lambda_n$, 则$A$可对角化, 且$A\sim \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

$A\in M_n(\mathbb{C})$, 一般情形: $f_A(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, $\lambda_i\neq\lambda_j$ for $i\neq j$.

代数重数: $n_1+n_2+\cdots+n_s = n$.

几何重数: $m_i = \dim V_{\lambda_i}$ 且 $m_i\leq n_i$, $i=1,2,\cdots,s$.

$\Big\} \Rightarrow \sum_{i=1}^{s} m_i \leq n$.

定理. $A\in M_n(\mathbb{C})$. $f_A(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, $\lambda_i\neq\lambda_j$ if $i\neq j$.

A可对角化 $\Leftrightarrow \sum_{i=1}^{s} m_i = n \Leftrightarrow m_i = n_i,\ i=1,2,\cdots,s \Leftrightarrow \mathrm{rank}(\lambda_i I - A) = n - n_i,\ i=1,2,\cdots,s$

$V_{\lambda_i} = N(\lambda_i I - A)$

$m_i = \dim V_{\lambda_i} = \dim N(\lambda_i I - A) = n - \mathrm{rank}(\lambda_i I - A)$

例. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ (§6.1 中例题)

解. Step 1. characteristic polynomial $f_A(\lambda) = |\lambda I - A| = (\lambda-4)(\lambda-1)^2$

令$f_A(\lambda)=0$, 得 eigenvalues: $\lambda_1 = 4\ (n_1=1)$; $\lambda_2 = 1\ (n_2=2)$.

Step 2. ① $\because n_1 = 1$. $\therefore$ 必有 $m_1 = n_1 = 1$. $(m_1\geq 1,\ m_1\leq n_1)$

② $\lambda_2 I - A = \begin{bmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{bmatrix}$. $\therefore \mathrm{rank}(\lambda_2 I - A) = 1$.

$\therefore m_2 = \dim N(\lambda_2 I - A) = 3 - \mathrm{rank}(\lambda_2 I - A) = 2 = n_2$ (或直接写 $\mathrm{rank}(\lambda_2 I - A) = 3 - n_2$)

综合①②, 知A可对角化.

Step 3. §6.1中已求得过: λ_1的特征向量 $\vec{x}_{11}$ (构成V_{λ_1}一组基)

λ_2 - - - - - - $\vec{x}_{21}, \vec{x}_{22}$ (构成V_{λ_2}一组基).

Step 4. 令 $X = [\vec{x}_{11}\ \ \vec{x}_{21}\ \ \vec{x}_{22}] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, 则

X可逆, 且 $AX = X\Lambda$, 其中 $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_2 \end{bmatrix} = \begin{bmatrix} 4 & & \\ & 1 & \\ & & 1 \end{bmatrix}$

$\Rightarrow X^{-1}AX = \Lambda$, 即 $A\sim\Lambda$.

Fibonacci Numbers

$0, 1, 1, 2, 3, 5, 8, 13, \cdots$, $F_{k+2} = F_{k+1} + F_k$, $F_0 = 0$.

What's F_{100}?

$$\begin{cases} F_{k+2} = F_{k+1} + F_k \\ F_{k+1} = F_{k+1} \end{cases} \Rightarrow \begin{bmatrix} F_{k+2} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix},$$ 记为 $\vec{u}_{k+1} = A\vec{u}_k$

$$\begin{bmatrix} F_{101} \\ F_{100} \end{bmatrix} = \vec{u}_{100} = A\vec{u}_{99} = A^2\vec{u}_{98} = \cdots = A^{100}\vec{u}_0, \quad \vec{u}_0 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

A可对角化 $\begin{cases} \text{特征值 } \lambda_1 = \frac{1+\sqrt{5}}{2}, & \text{对应特征向量 } \vec{x}_1 = [\lambda_1, 1]^T. \\ \text{- - - } \lambda_2 = \frac{1-\sqrt{5}}{2}, & \text{- - - - - - - } \vec{x}_2 = [\lambda_2, 1]^T \end{cases}$ $\vec{x}_1$ 与 $\vec{x}_2$ 线性无关

How to compute $A^k \vec{u}_0$ if A is diagonalizable? (general case) $A \in M_n$

设A有n个线性无关的特征向量 $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$，分别对应特征值 $\lambda_1, \cdots, \lambda_n$.

令 $X = [\vec{x}_1\ \vec{x}_2 \cdots \vec{x}_n]$, $\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$，则 $AX = X\Lambda$, $\Rightarrow A = X\Lambda X^{-1}$.

设 $\vec{u}_0 = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$，则

$$A^k\vec{u}_0 = c_1A^k\vec{x}_1 + c_2A^k\vec{x}_2 + \cdots + c_nA^k\vec{x}_n = c_1(\lambda_1)^k\vec{x}_1 + c_2(\lambda_2)^k\vec{x}_2 + \cdots + c_n(\lambda_n^k)\vec{x}_n$$

Step 1 Write $\vec{u}_0$ as a combination of $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n$.

$$\vec{u}_0 = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n = [\vec{x}_1 \cdots \vec{x}_n]\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = X\vec{c} \quad \Rightarrow \vec{c} = X^{-1}\vec{u}_0$$

Step 2 Multiply each c_i by $(\lambda_i)^k$.

$$\begin{bmatrix} c_1(\lambda_1)^k \\ c_2(\lambda_2)^k \\ \vdots \\ c_n(\lambda_n)^k \end{bmatrix} = \begin{bmatrix} (\lambda_1)^k & & & \\ & (\lambda_2)^k & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \Lambda^k\vec{c} = \Lambda^k X^{-1}\vec{u}_0$$

Step 3 Add up the pieces $c_i(\lambda_i)^k\vec{x}_i$ to get $A^k\vec{u}_0 = \sum_{i=1}^{n} c_i(\lambda_i)^k\vec{x}_i$

$$A^k\vec{u}_0 = [\vec{x}_1\ \vec{x}_2 \cdots \vec{x}_n]\begin{bmatrix} c_1(\lambda_1)^k \\ c_2(\lambda_2)^k \\ \vdots \\ c_n(\lambda_n)^k \end{bmatrix} = X\Lambda^k\vec{c} = X\Lambda^k X^{-1}\vec{u}_0$$

注：$A^k\vec{u}_0 = (X\Lambda X^{-1})^k\vec{u}_0 = X\Lambda^k X^{-1}\vec{u}_0$

P309 Example 3.