§ 矩阵的相似对角化

问题：如何判断一个方阵是否相似于对角矩阵.

相似于对角矩阵的方阵称为可对角化的. 也称相应的线性变换为可对角化的.

注：$A=\begin{pmatrix}2&1\\0&2\end{pmatrix}$ 不可对角化.

否则 $A=T\begin{pmatrix}2&\\&2\end{pmatrix}T^{-1}=2TT^{-1}=2I_2$ 矛.

§ 矩阵相似于对角矩阵的充要条件.

引理：$A\in F^{n\times n}$，属于 A 的不同特征值的特征向量线性无关.

即. $A\vec{x}_i=\lambda_i\vec{x}_i$ $(\vec{x}_i\neq 0,\ 1\leq i\leq k)$，则 $\vec{x}_1,\cdots\vec{x}_k$ 线性无关.

证明：设 $\mu_1\vec{x}_1+\mu_2\vec{x}_2+\cdots+\mu_k\vec{x}_k=0$. 则

$$\begin{aligned}0&=A^i(\mu_1\vec{x}_1+\mu_2\vec{x}_2+\cdots+\mu_k\vec{x}_k)\\&=\mu_1A^i\vec{x}_1+\mu_2A^i\vec{x}_2+\cdots+\mu_kA^i\vec{x}_k\\&=\lambda_1^i(\mu_1\vec{x}_1)+\lambda_2^i(\mu_2\vec{x}_2)+\cdots+\lambda_k^i(\mu_k\vec{x}_k)\\&=(\mu_1\vec{x}_1,\cdots,\mu_k\vec{x}_k)\begin{pmatrix}\lambda_1^i\\\vdots\\\lambda_k^i\end{pmatrix}\qquad\forall\ i\geq 0.\end{aligned}$$

$$\Rightarrow (\mu_1\vec{X_1},\cdots,\mu_k\vec{X_k})\underbrace{\begin{pmatrix}1&\lambda_1&\cdots&\lambda_1^{k-1}\\1&\lambda_2&\cdots&\lambda_2^{k-1}\\\cdots&\cdots&\cdots&\cdots\\1&\lambda_k&\cdots&\lambda_k^{k-1}\end{pmatrix}}_{\text{可逆！}}=0$$

$$\Rightarrow (\mu_1\vec{X_1},\cdots,\mu_k\vec{X_k})=0$$

$$\Rightarrow \quad \mu_i\vec{x_i}=0 \quad \forall i=1,\cdots,k$$

$$\Rightarrow \mu_1=\cdots=\mu_k=0$$

□

定理：$A\in F^{n\times n}$ 可对角化 $\Leftrightarrow$ A 有 n 个线性无关的特征向量.

证：$\Rightarrow$) $\quad A=T\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}T^{-1}$. 记 $T=(\eta_1,\cdots,\eta_n)$. 则

$\eta_1,\cdots,\eta_n$ 线性无关，且 $A\eta_i=\lambda_i\eta_i$

(因为 $A(\eta_1\cdots\eta_n)=AT=T\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}=(\eta_1,\cdots,\eta_n)\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}=(\lambda_1\eta_1,\cdots,\lambda_n\eta_n)$)

$\Leftarrow$) 若 $\eta_1\cdots\eta_n$ 线性无关，且 $A\eta_i=\lambda_i\eta_i$. 记 $T=(\eta_1\cdots\eta_n)$.

$$\left.\begin{array}{l}\eta_1\cdots\eta_n \text{ 线性无关} \Rightarrow T \text{ 可逆}\\ A\eta_i=\lambda_i\eta_i \Rightarrow A(\eta_1\cdots\eta_n)=(\eta_1\cdots\eta_n)\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\end{array}\right\}$$

$$\Rightarrow AT=T\,\mathrm{diag}(\lambda_1\cdots\lambda_n) \Rightarrow A\sim \mathrm{diag}(\lambda_1\cdots\lambda_n).$$

注：该对角矩阵在不记对角元的位置意义下唯一.

推论：若 $A\in F^{n\times n}$ 有 n 个两两不同的特征值，则 A 可对角化.

证：引理 + 定理. □

*§ 特征值的代数重数与几何重数

$A \in \mathbb{C}^{n\times n}$

- $P_A(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$

  $n_i = \lambda_i$ 的代数重数

- $V_A(\lambda_i) := \{\vec{x}\in\mathbb{C}^n \mid A\vec{x} = \lambda_i\vec{x}\}$

  $m_i := \dim_{\mathbb{C}} V_A(\lambda_i)$ 称为 λ_i 的几何重数.

定理：1) $1\leq$ 几何重数 $\leq$ 代数重数 (即 $m_i \leq n_i \ \forall i=1,\cdots,s$)

2) A 可对角化 $\Leftrightarrow$ 每个特征值的几何重数与代数重数相等

例： $A=\begin{pmatrix} 0 & 0 & x \\ 1 & 1 & y \\ 1 & 0 & 0 \end{pmatrix}$ 可对角化 $\Leftrightarrow$ x, y ??

解： $P_A(\lambda)=\begin{vmatrix} \lambda & 0 & -x \\ -1 & \lambda-1 & -y \\ -1 & 0 & \lambda \end{vmatrix} = \lambda^2(\lambda-1)-(\lambda-1)x = (\lambda-1)(\lambda^2-x)$.

1° $x\neq 0,1 \Rightarrow A$ 有3个不同特征值 $1, +\sqrt{x}, -\sqrt{x}$. $\Rightarrow$ 可对角化

2° $x=0 \Rightarrow A$ 有2个不同特征值 $0, 1$ 重数为 $2, 1$.

0的几何重数 $= 3-\mathrm{rank}(0\cdot I-A) = 1$

$\Rightarrow$ 不可对角化

3° $x=1$ $\Rightarrow$ A有2个不同特征值 $1,-1$ 重数为 $2,1$.

1的几何重数 $=3-\text{rank}(I-A)=\begin{cases}2 & y=1\\ 1 & y\neq 1\end{cases}$

$\Rightarrow A\begin{cases}\text{可对角化，若 } y=1\\ \text{不可对角化，若 } y\neq 1.\end{cases}$

综上，A可对角化 $\Leftrightarrow$ $x\neq 0,1$ 或 $\begin{cases}x=1\\ y=1\end{cases}$.

§ 相似于上三角形矩阵

不是每个矩阵都可以对角化（例 $\begin{pmatrix}2&1\\0&2\end{pmatrix}$），但可上三角化！

定理：1) 任何一个n阶复方阵A都相似于一个上三角形矩阵J，且

2) J的对角线上元素为A的全体特征值.

证：1) 对n归纳. $n=1$ ✓ 设结论对 $n-1$ 阶方阵成立.

· 任取A的特征值 λ_1 及特征向量 $\vec{x}_1$,

· 将 $\vec{x}_1$ 扩充为 $\mathbb{C}^n$ 的一组基 $\vec{x}_1,\cdots,\vec{x}_n$. 则

$$A(\vec{x}_1,\cdots,\vec{x}_n) = (\vec{x}_1,\cdots,\vec{x}_n)\begin{pmatrix}\lambda_1 & * \\ 0 & A_1\end{pmatrix}$$

记 $T=(\vec{x}_1,\cdots,\vec{x}_n)\in\mathbb{C}^{n\times n}$ 则 T 可逆 且

$$A = T\begin{pmatrix}\lambda_1 & * \\ & A_1\end{pmatrix}T^{-1} \overset{归纳}{=\!=} T\begin{pmatrix}\lambda_1 & * \\ & SJ_1S^{-1}\end{pmatrix}T^{-1}$$

$$= \underbrace{T\begin{pmatrix}1 & \\ & S_1\end{pmatrix}}_{=S}\underbrace{\begin{pmatrix}1 & \\ & S_1^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & * \\ & S_1J_1S_1^{-1}\end{pmatrix}\begin{pmatrix}1 & \\ & S_1\end{pmatrix}}\underbrace{\begin{pmatrix}1 & \\ & S_1^{-1}\end{pmatrix}T^{-1}}_{S^{-1}}$$

$$= S\begin{pmatrix}\lambda_1 & ** \\ 0 & J_1\end{pmatrix}^{=J} S^{-1} = SJS^{-1}$$

2) A与J有相同的特征值，且J的特征值为对角线上的元素。□

注：J的取法不唯一！例如 $\begin{pmatrix}1 & 1 \\ 0 & 1\end{pmatrix}$ 与 $\begin{pmatrix}1 & 2 \\ 0 & 1\end{pmatrix}$ 相似

$$\begin{pmatrix}2 & \\ & 1\end{pmatrix}\begin{pmatrix}1 & 1 \\ & 1\end{pmatrix}\begin{pmatrix}\frac{1}{2} & \\ & 1\end{pmatrix} = \begin{pmatrix}1 & 2 \\ 0 & 1\end{pmatrix}$$

例：$A\in\mathbb{C}^{2\times 2}$. $A^2=0 \Rightarrow A=0$ 或 A相似于 $\begin{pmatrix}0 & 1 \\ 0 & 0\end{pmatrix}$.

证：A的特征值均为0. $(Ax=\lambda x\ (x\neq 0) \Rightarrow 0=A^2x=\lambda^2 x \Rightarrow \lambda^2=0)$

定理 $\Rightarrow A\sim\begin{pmatrix}0 & a \\ 0 & 0\end{pmatrix}$

1° $a=0 \Rightarrow A\sim 0 \Rightarrow A=0$

2° $a\neq 0 \Rightarrow A\sim\begin{pmatrix}0 & a \\ 0 & 0\end{pmatrix} = \begin{pmatrix}a & \\ & 1\end{pmatrix}\begin{pmatrix}0 & 1 \\ 0 & 0\end{pmatrix}\begin{pmatrix}a^{-1} & \\ & 1\end{pmatrix} \sim \begin{pmatrix}0 & 1 \\ 0 & 0\end{pmatrix}$