线性代数 第20讲

11月17日

第五章第2讲 对角化和谱分解

上一讲内容回顾

对角化与谱分解的概念

应用实例

矩阵可对角化的条件

矩阵的特征值和特征向量

定义5.2.1 (特征值) 给定 n 阶方阵 A, 如果对 $\lambda \in \mathbb{G}$, 存在非零向量 $x \in \mathbb{G}^n$, 使得 $Ax = \lambda x$, 则称数 λ 为方阵 A (在 \mathbb{G} 上)的一个特征值,而称非零向量 x 为方阵 A (在 \mathbb{G} 上)的一个属于特征值 λ 的特征向量.

- 1. 只有方阵才有特征值和特征向量;
- 2. 零向量不是特征向量;
- 3. 如果 x 是特征向量,则对任意 $k \neq 0$, kx 都是特征向量.

二元组 (λ, x) 常称为方阵 A 的一个特征对.

特别地,对实方阵 A,如果特征对 (λ, x) 满足 $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$,则分别称二者为 A 在 \mathbb{R} 上的特征值和特征向量,称该二元组为 A 在 \mathbb{R} 上的特征对.

1

矩阵的特征值和特征向量的计算

 (λ_0, x_0) 是一个特征对,当且仅当 $Ax_0 = \lambda_0 x_0$,即 $(\lambda_0 I - A)x_0 = 0$,注意到,特征向量 $x_0 \neq 0$,齐次线性方程组有非零解,因此 λ_0 是A的特征值当且仅当 $(\lambda_0 I - A)$ 不可逆,也等价于其行列式 $\det(\lambda_0 I - A) = 0$.

由此我们首先得到如下结论.

命题5.2.3 数 λ_0 是 A 的特征值, 当且仅当 $\det(\lambda_0 I_n - A) = 0$. 特别地, 0 是 A 的特征值当且仅当 A 不可逆.



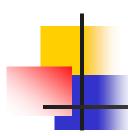
矩阵A的特征多项式

$$p_A(\lambda) := \det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \quad \text{称为矩阵 A 的特征多项式}.$$

$$A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}, \quad |\lambda I - A| = \begin{vmatrix} \lambda - 3 & 2 \\ -1 & \lambda \end{vmatrix} = \lambda^2 - 3\lambda + 2 \Rightarrow \lambda_1 = 1, \lambda_2 = 2$$

$$(\lambda_1 I - A) x = 0 \Rightarrow \begin{bmatrix} -2 & 2 \\ -1 & 1 \end{bmatrix} x = 0 \Rightarrow x_1 = k_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\left(\lambda_2 I - A\right) x = 0 \Rightarrow \begin{bmatrix} -1 & 2 \\ -1 & 2 \end{bmatrix} x = 0 \Rightarrow x_2 = k_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$



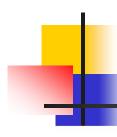
特征多项式和特征子空间

$$p_A(\lambda) := \det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \quad \text{称为矩阵 A 的特征多项式}.$$

定理 5.2.4 设 n 阶方阵 A 的特征多项式为 $p_A(\lambda)$, 那么

- 1. 数 λ_0 是 A 的特征值, 当且仅当 $p_A(\lambda_0) = 0$, 即 λ_0 是特征多项式 $p_A(\lambda)$ 的根.
- 2. 向量 $x_0 \in \mathbb{C}^n$ 是 A 的属于 λ_0 的特征向量, 当且仅当 $x_0 \in \mathcal{N}(\lambda_0 I_n A)$ 且 $x_0 \neq 0$, 即 x_0 是 $\lambda_0 I_n - A$ 的零空间中的非零向量.

解空间 $\mathcal{N}(\lambda_0 I_n - A)$ 称为 A 的属于 λ_0 的**特征子空间**. 注意到特征向量不为零, $\{A \text{ 的属于 } \lambda_0 \text{ 的特征向量}\} = \mathcal{N}(\lambda_0 I_n - A) \setminus \{\mathbf{0}\}.$



特征值的代数重数与几何重数

定义 5.2.10 (代数重数) 给定 n 阶方阵 A 及 A 的一个特征值 $\lambda_0 \in \mathbb{C}$,如果 λ_0 是特征多项式 $p_A(\lambda)$ 的 n_0 重根,则称 n_0 为 λ_0 作为 A 的特征值的代数重数(简称重数),称 λ_0 是 A 的一个 n_0 重特征值.

一个1重特征值,又称为单特征值.

命题 5.2.11 给定 n 阶<mark>实方阵</mark> A, 如果 λ_0 是它的一个非实数特征值,则 $\overline{\lambda_0}$ 也是它的特征值,且其代数重数和 λ_0 的代数重数相等. 进一步地,如果复向量 \boldsymbol{x}_0 是属于 λ_0 的特征向量,则 $\overline{\boldsymbol{x}}_0$ 是属于 $\overline{\lambda_0}$ 的特征向量.

定义 5.3.6 (几何重数) 给定 n 阶方阵 A 及其特征值 λ_0 ,称特征子空间 $\mathcal{N}(\lambda_0 I_n - A)$ 的维数为 λ_0 作为 A 的特征值的**几何重数**.

任意特征值的几何重数都不小于 1, 因此特征值至少对应一个特征向量.

定理 5.1.5 (Vieta 定理) 复系数一元 n 次多项式 $p(x) = a_0 x^n + a_1 x^{n-1} + \dots + a_n$ 的 n 个根(计重数) x_1, \dots, x_n 满足:

$$\begin{split} -\frac{a_1}{a_0} &= x_1 + x_2 + \dots + x_n, \\ & \vdots \\ (-1)^k \frac{a_k}{a_0} &= \sum_{1 \leqslant i_1 < \dots < i_k \leqslant n} x_{i_1} x_{i_2} \cdots x_{i_k}, \\ & \vdots \\ (-1)^n \frac{a_n}{a_0} &= x_1 x_2 \cdots x_n. \end{split}$$

命题 5.2.12 给定 n 阶方阵 A, 其特征多项式具有如下形式:

$$p_A(\lambda) = \lambda^n - \operatorname{trace}(A)\lambda^{n-1} + \dots + (-1)^n \det(A),$$

其中 $trace(A) = a_{11} + a_{22} + \dots + a_{nn}$ 是方阵 A 的对角元素的和, 称为方阵 A 的**迹**.

$$p_A(\lambda) := \det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{m1} & -a_{m2} & \cdots & \lambda - a_{mn} \end{vmatrix} \cdot \begin{vmatrix} \lambda = \mathbf{0} \text{ B} \\ \det(\lambda I_n - A) = (-\mathbf{1})^n |A| \end{vmatrix}$$

n 阶方阵有且只有 n 个特征值 (计重数). $\lambda_1 + \dots + \lambda_n = \operatorname{trace}(A), \quad \lambda_1 \dots \lambda_n = \det(A).$

多项式的零点与矩阵特征值

对任意首项系数为1的多项式

$$p(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1} + x^n,$$

$$C_p = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & 0 & 1 \\ -c_0 & \cdots & -c_{n-2} & -c_{n-1} \end{bmatrix}, \quad C_p \text{ 的特征多项式是 } p(x)$$

矩阵 C_p 称为多项式 p(x) 的友矩阵. 因此在实践中, 任意多项式的根的近似值, 就可以通过求解其友矩阵的特征值来得到.

练习 5.2.2 构造符合要求的矩阵 A.

1. A 的特征多项式为 $\lambda^2 - 9\lambda + 20$,构造三个不同的 A.

2.
$$A = \begin{bmatrix} 0 & 1 \\ * & * \end{bmatrix}$$
, 且 A 的特征值为 4,7.

3.
$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ * & * & * \end{bmatrix}$$
, 且 A 的特征值为 $1, 2, 3$.

- 练习 5.2.11 设 n 阶实方阵 A 满足 $A^{T}v = \lambda v$, 其中 $v \in \mathbb{R}^{n}, v \neq 0$.
 - 1. 设 $Aw = \mu w$, 且 $w \in \mathbb{R}^n$, $\lambda \neq \mu$, 证明 v, w 正交.
 - 2. 证明, 实对称矩阵的属于不同特征值的实特征向量正交.

命题5.3.3 方阵的属于不同特征值的特征向量线性无关.

证. 设方阵 A 有特征向量 x_1 , … , x_r , 分别属于特征值 λ_1 , … , λ_r , 且 λ_1 , … , λ_r 两两不同.

采用数学归纳法. 当 r=1 时,因为特征向量不为零,因此线性无关.

现假设任意 r-1 个不同特征值的特征向量都线性无关,

观察方程 $k_1x_1 + k_2x_2 + \cdots + k_rx_r = 0$. 等式两边左乘 A, 则有

 $\mathbf{0} = A(k_1x_1 + k_2x_2 + \cdots + k_rx_r) = k_1\lambda_1x_1 + k_2\lambda_2x_2 + \cdots + k_r\lambda_rx_r.$

再减去原方程的 λ_1 倍, $k_1\lambda_1x_1 + k_2\lambda_2x_2 + \cdots + k_r\lambda_rx_r - \lambda_1(k_1x_1 + k_2x_2 + \cdots + k_rx_r)$

就有 $k_2(\lambda_2 - \lambda_1) x_2 + \cdots + k_r(\lambda_r - \lambda_1) x_r = \mathbf{0}$.

根据归纳假设, x_2 , … , x_r 线性无关,于是 $k_i(\lambda_i - \lambda_1) = 0$, i = 2, … , r.

由于 λ_1 和 λ_2 , …, λ_r 不同, 因此 $k_i = 0$, i = 2, …, r.

又得 $k_1x_1 = 0$, 由特征向量不为零得 $k_1 = 0$. 故 x_1 , …, x_r 线性无关.



可对角化与谱分解

定义5.3.1 (谱分解) 对方阵 A, 如果存在可逆矩阵 X 使得 $X^{-1}AX = \Lambda$ 是对角矩阵,则称 A 是(在G 上)可对角化的, X 把 A 对角化,或 X 对角化 A.

如果方阵 A, X, Λ 都是实矩阵,则称 Λ 在 $\mathbb R$ 上可对角化. 当 Λ 可对角化时,分解 $\Lambda = X\Lambda X^{-1}$ 称为 Λ 的谱分解。 之所以称为谱分解,是因为特征值也称为谱.

如果矩阵 A 是可对角化的, 即 $A = X\Lambda X^{-1}$,则 $A'' = X\Lambda'' X^{-1}$



PageRank: Ranking Web Pages



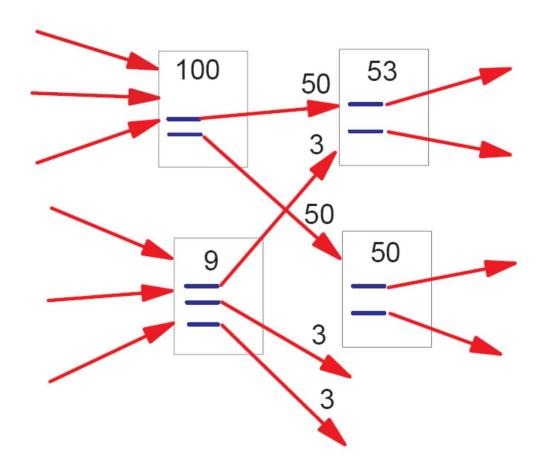
PageRank模型的原始思想:

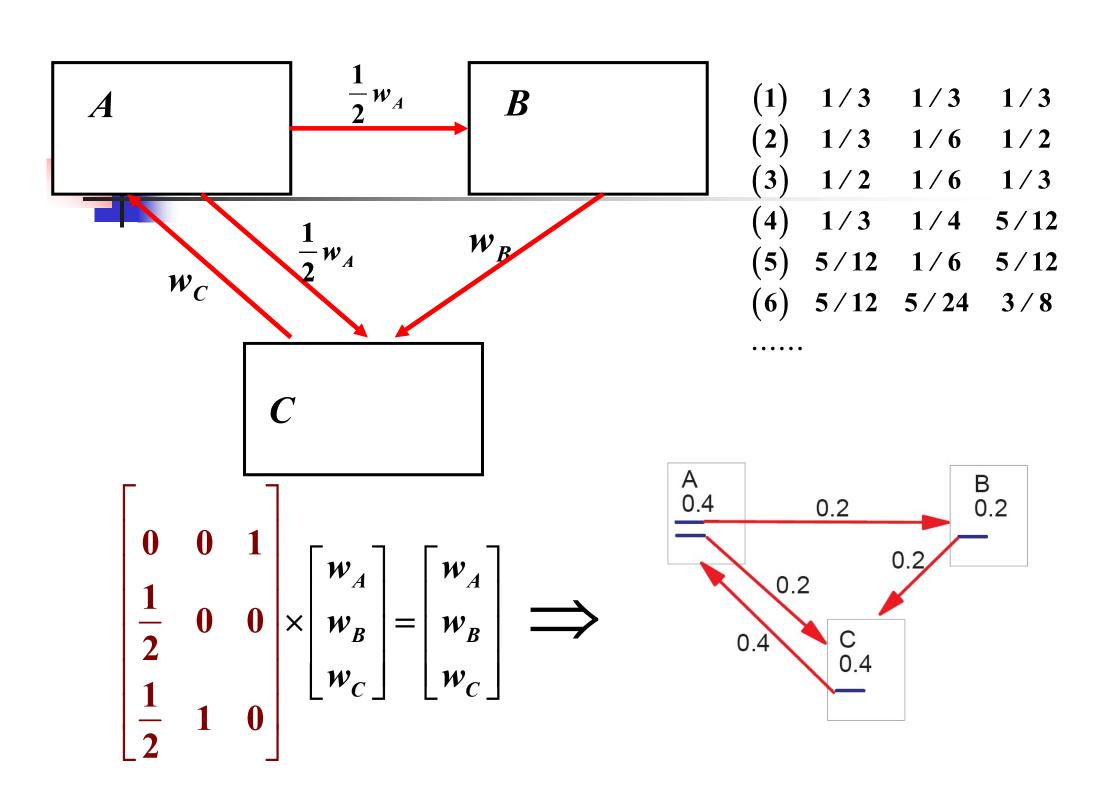
依赖于网页的拓扑结构 —网页间的相互链接关系

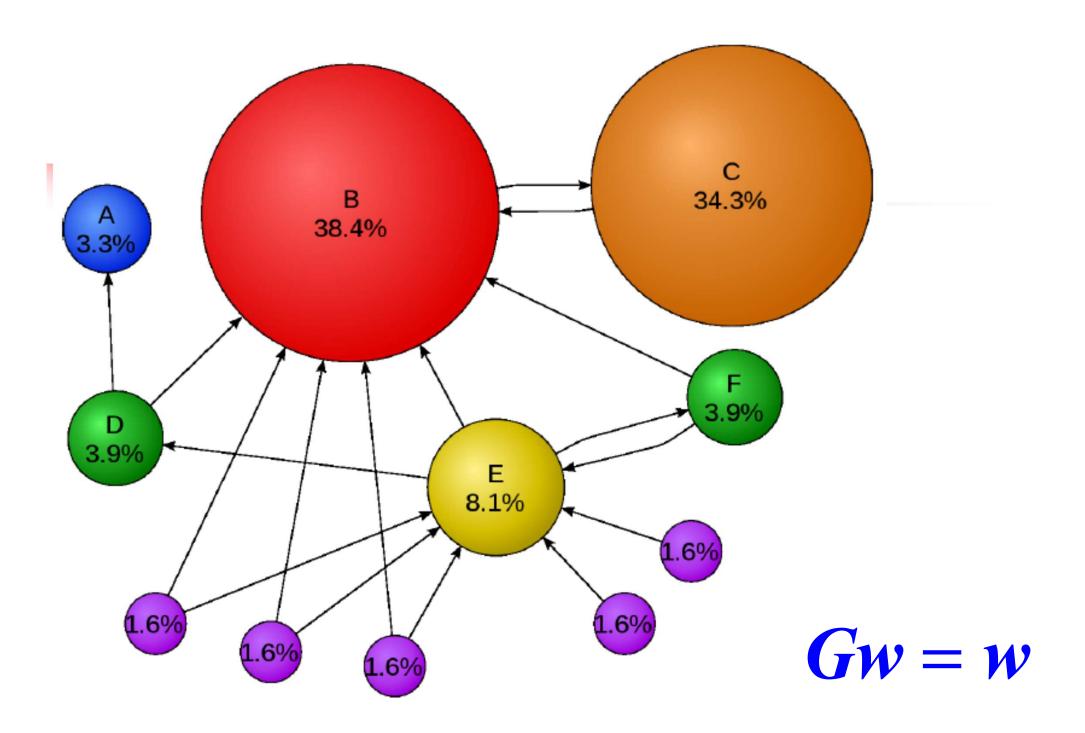


Cartoon illustrating basic principle of PageRank.

The size of each face is proportional to the total size of the other faces which are pointing to it.







 $A \in \mathbb{R}^{n \times n}$,假设其特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$,满足 $\left| \lambda_1 \right| > \left| \lambda_2 \right| \ge \cdots \ge \left| \lambda_n \right|$,且对应的特征向量为 x_1, x_2, \cdots, x_n , $Ax_k = \lambda x_k \left(k = 1, 2, \cdots, n \right)$

任取非零向量 $u \in \mathbb{R}^n$, $u = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$

$$Au = A(c_1x_1 + c_2x_2 + \dots + c_nx_n) = c_1\lambda_1x_1 + c_2\lambda_2x_2 + \dots + c_n\lambda_nx_n$$

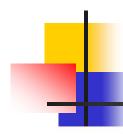
$$A^2u = A \cdot Au = A(c_1\lambda_1x_1 + c_2\lambda_2x_2 + \dots + c_n\lambda_nx_n) = c_1\lambda_1^2x_1 + c_2\lambda_2^2x_2 + \dots + c_n\lambda_n^2x_n$$
...

$$A^{n}u = c_{1}\lambda_{1}^{n}u_{1} + c_{2}\lambda_{2}^{n}u_{2} + \dots + c_{n}\lambda_{n}^{n}u_{n}$$

读
$$c_1 \neq 0$$
,则 $A^n u = \lambda_1^n \left(c_1 u_1 + c_2 \frac{\lambda_2^n}{\lambda_1^n} u_2 + \dots + c_n \frac{\lambda_n^n}{\lambda_1^n} u_n \right) \rightarrow \lambda_1^n c_1 u_1$

矩阵可对角化的条件

命题5.3.2 对 n 阶方阵 A, A 可对角化, 当且仅当 A 有 n 个线性无关的特征向量.



矩阵可对角化的一个充分条件

命题 5.3.3 方阵的属于不同特征值的特征向量线性无关.

矩阵可对角化的一个充分条件

推论5.3.4 有 n 个不同特征值的 n 阶方阵,即特征值都是单特征值的方阵,可对角化.

定义 5.3.6 (几何重数) 给定 n 阶方阵 A 及其特征值 λ_0 ,称特征子空间 $\mathcal{N}(\lambda_0 I_n - A)$ 的维数为 λ_0 作为 A 的特征值的**几何重数**.

任意特征值的几何重数都不小于 1, 因此特征值至少对应一个特征向量.

命题 5.3.7 方阵的特征值的几何重数不大于其代数重数.

命题 5.3.7 方阵的特征值的几何重数不大于其代数重数.

证. 设 λ_0 是 n 阶方阵 A 的几何重数是 r 的一个特征值,那么存在 r 个线性无关的特征向量 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r$,使得 $A\boldsymbol{x}_i=\lambda_0\boldsymbol{x}_i, i=1,\cdots,r$. 则 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r$ 可以扩充成 \mathbb{C}^n 的一组基 $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_r,\boldsymbol{y}_{r+1},\cdots,\boldsymbol{y}_n$.

记 $X_{\lambda_0} \coloneqq \begin{bmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_r \end{bmatrix}$, $Y \coloneqq \begin{bmatrix} \boldsymbol{y}_{r+1} & \cdots & \boldsymbol{y}_n \end{bmatrix}$, $X \coloneqq \begin{bmatrix} X_{\lambda_0} & Y \end{bmatrix}$, 而 X 可逆. 由于 X 的列是一组基,AX 的列能被其线性表示,于是存在方阵 M ,使得 AX = XM . 由于 X_{λ_0} 的列是 A 的属于 λ_0 的特征向量,所以 M 具有如下分块形式:

$$AX = A \begin{bmatrix} X_{\lambda_0} & Y \end{bmatrix} = \begin{bmatrix} X_{\lambda_0} & Y \end{bmatrix} \begin{bmatrix} \lambda_0 I_r & M_{12} \\ O & M_{22} \end{bmatrix} = XM.$$

而 A 和 $M = X^{-1}AX$ 具有相同的特征多项式,因为

$$\det(\lambda I_n - A) = \det(X^{-1}(\lambda I_n - A)X) = \det(\lambda I_n - M).$$

另一方面,

$$\det(\lambda I_n - M) = \det\left(\begin{bmatrix}\lambda I_r - \lambda_0 I_r & -M_{12} \\ \mathbf{0} & \lambda I_{n-r} - M_{22}\end{bmatrix}\right) = (\lambda - \lambda_0)^r \det(\lambda I_{n-r} - M_{22}).$$

这说明 λ_0 是 A 的一个代数重数至少 r 的特征值.

定义5.3.8 几何重数和代数重数相等的特征值, 称为半单特征值. 几何重数小于代数重数的特征值, 称为亏损特征值.

如果一个特征值的代数重数是1,那么由几何重数不小于1。可知,它是半单特征值,即单特征值是半单特征值.

定理5.3.9

- 1. n 阶方阵 A 可对角化, 当且仅当其特征值都半单(几何重数和代数重数相等).
- 2. n 阶实方阵 A 在 R 上可对角化, 当且仅当其特征多项式的根都是实根, 且其特征值都半单.

证. 第 1 条: "⇒": 如果 A 可对角化,设 $x_{1,1}, \cdots x_{1,k_1}, \cdots, x_{s,1}, \cdots x_{s,k_s}$ 是 n 个线性无关的特征向量,满足 $Ax_{i,j} = \lambda_i x_{i,j}$. 因此, $x_{i,1}, \cdots x_{i,k_i}$ 是特征子空间 $\mathcal{N}(\lambda_i I_n - A)$ 中一个线性无关的向量组,于是 k_i 小于等于 λ_i 的几何重数 m_i ,且 $\sum_{1 \leqslant i \leqslant s} k_i = n$. 设 n_i 是 λ_i 的代数重数,则 $\sum_{1 \leqslant i \leqslant s} n_i = n$. 由命题 5.3.7 可知, $k_i \leqslant m_i \leqslant n_i$. 比较两个和式可知 $k_i = m_i = n_i$,即任意特征值半单.

" \leftarrow ":如果 A 的特征值都半单,设 $\lambda_1, \cdots, \lambda_s$ 是所有特征值,则几何重数 m_i 等于代数重数 n_i ,因此 $\sum_{1\leqslant i\leqslant s}m_i=n$. 取每个特征子空间 $\mathcal{N}(\lambda_iI_n-A)$ 的一组基 $\boldsymbol{x}_{i,1},\cdots \boldsymbol{x}_{i,m_i}$,合并起来的向量组 $\boldsymbol{x}_{1,1},\cdots \boldsymbol{x}_{1,m_1},\cdots, \boldsymbol{x}_{s,1},\cdots \boldsymbol{x}_{s,m_s}$ 仍旧线性无关,证明与命题 5.3.3 的证明类似. 而合并得到的向量组中有 $\sum_{1\leqslant i\leqslant s}m_i=n$ 个向量,即 A 有 n 个线性无关的特征向量,因此可对角化.



不可对角化矩阵的例子

矩阵
$$S_k = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$
 $(k \neq 0)$,只有一个特征值 1,其代数重数是 2,几何重数是 1,

该矩阵不可对角化.

矩阵
$$J_n(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}_{n \times n}$$
 只有一个特征值 λ ,其代数重数是 n ,几何重数

是 1, 当 n > 1 时 λ 是亏损特征值. 形如 $J_n(\lambda)$ 的矩阵称为关于 λ 的 n 阶 **Jordan 块**. 当 n > 1 时,它不可对角化.

设
$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$$
,求 A^n

解:

(1) 求矩阵
$$A$$
 的特征值 $|\lambda I - A| = (\lambda - 1 - \sqrt{-2})(\lambda - 1 + \sqrt{-2})$, $\lambda_1 = 1 + \sqrt{-2}$, $\lambda_2 = 1 - \sqrt{-2}$,

(2) 求属于不同特征值的特征向量.

由
$$(\lambda_1 I - A)X = 0$$
 解得 $X_1 = (1, -\sqrt{-2})^T$,故 $AX_1 = (1 + \sqrt{-2})X_1$,由 $(\lambda_2 I - A)X = 0$ 解得 $X_2 = (1, \sqrt{-2})^T$,故 $AX_2 = (1 - \sqrt{-2})X_2$, $\therefore A(X_1, X_2) = (X_1, X_2) \begin{bmatrix} 1 + \sqrt{-2} \\ 1 - \sqrt{-2} \end{bmatrix}$,

$$\therefore A^{n} = P \begin{bmatrix} 1 + \sqrt{-2} & & \\ & 1 - \sqrt{-2} \end{bmatrix}^{n} P^{-1} = \begin{bmatrix} 1 & 1 \\ -\sqrt{-2} & \sqrt{-2} \end{bmatrix} \begin{bmatrix} (1 + \sqrt{-2})^{n} & & \\ & (1 - \sqrt{-2})^{n} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{-2}}{4} \\ \frac{1}{2} & -\frac{\sqrt{-2}}{4} \end{bmatrix}$$

$$= \begin{bmatrix} \left(1 + \sqrt{-2}\right)^n & \left(1 - \sqrt{-2}\right)^n \\ -\sqrt{-2}\left(1 + \sqrt{-2}\right)^n & \sqrt{-2}\left(1 - \sqrt{-2}\right)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{-2}}{4} \\ \frac{1}{2} & -\frac{\sqrt{-2}}{4} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2} \left[\left(1 + \sqrt{-2} \right)^n + \left(1 - \sqrt{-2} \right)^n \right] & \frac{\sqrt{-2}}{4} \left[\left(1 + \sqrt{-2} \right)^n - \left(1 - \sqrt{-2} \right)^n \right] \\ -\frac{\sqrt{-2}}{2} \left[\left(1 + \sqrt{-2} \right)^n + \left(1 - \sqrt{-2} \right)^n \right] & \frac{1}{2} \left[\left(1 + \sqrt{-2} \right)^n + \left(1 - \sqrt{-2} \right)^n \right] \end{bmatrix}$$

设矩阵
$$A = \begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}$$
, $\mathbb{I} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ 是 A 的特征向量,求 a,b,c,d,e,f .

设特征向量
$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
, $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, 对应的特征值分别为 λ , λ , λ ,

则由
$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}$$
 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ = λ_1 $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, 解得 λ_1 = 3, 同理解得 λ_2 = 0, λ_3 = 0

所以
$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

作业(11月17日)

练习5.3.

1, 2, 3, 4, 7(1, 4), 8, 10, 12, 13

11月22日提交