

线性代数习题课讲义

陈文轩

更新: August 1, 2026

目录

1 矩阵的相似	1
2 一些其他结论 *	4
2.1 Caylay-Hamilton 定理	4
2.2 Gershgorin 圆盘定理	4
2.3 同时对角化问题	5

1 矩阵的相似

与相抵类似, $\forall A \in \mathbb{C}^{n \times n}$, A 的相抵标准形是 $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, 为了给出相似的标准型和最简单的代表元, 需要使用到以下特征值的知识。

定义 1.1 (矩阵的特征值和特征向量) 对 $A \in \mathbb{F}^{n \times n}$, 若 $\exists 0 \neq x \in \mathbb{F}^n, \lambda \in \mathbb{F}, s.t. Ax = \lambda x$, 则称 λ 是 A 的特征值, x 是属于特征值 λ 的对应特征向量。

注 由于矩阵可以看作在特定基下的线性变换, 此时 Ax 是向量 x 在线性变换 $\mathcal{A}$ 下的像, 此时 $Ax = \lambda x$ 是指线性变换 $\mathcal{A}$ 对 x 的作用是将 x 拉伸 λ 倍。

定义 1.2 (线性变换的特征值和特征向量) 对 $\mathcal{A} \in \mathcal{L}(V, V)$, 若 $\exists 0 \neq \alpha \in V, \lambda \in \mathbb{F}, s.t. \mathcal{A}\alpha = \lambda\alpha$, 则称 λ 是 $\mathcal{A}$ 的特征值, α 是属于特征值 λ 的对应特征向量。

定理 1.1 (线性变换的特征值与其在一组基下特征值的关系) 设 $\mathcal{A} \in \mathcal{L}(V, V), \dim V = n, \mathcal{A}(\alpha_1, \dots, \alpha_n) = (\alpha_1, \dots, \alpha_n)A, A \in \mathbb{F}^{n \times n}$, 则 $\mathcal{A}\alpha = \lambda\alpha \Leftrightarrow Ax = \lambda x$, 其中 x 为 α 在基 $\{\alpha_1, \dots, \alpha_n\}$ 下的坐标。

注 该定理给出了一般线性变换与矩阵的特征值与特征向量的关系, 从而以下只讨论矩阵的特征值。

算法 1.2 (特征值与特征向量的计算) 对 $A \in \mathbb{C}^{n \times n}$, 可以通过以下步骤计算其特征值与特征向量:

1. 计算特征多项式 $p_A(\lambda) = \det(\lambda I - A) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}$, $\sum_{i=1}^s \lambda_i = n$, 这是关于 λ 的一个 n 次多项式, 从而有 n 个根;
2. 令 $p_A(\lambda) = 0$, 得到特征值 $\lambda_1(n_1 \text{ 重}), \lambda_2(n_2 \text{ 重}) \cdots \lambda_s(n_s \text{ 重})$, 这些是 A 的特征值;
3. 对每一个 $\lambda_i (1 \leq i \leq s)$, 求出线性方程组 $(\lambda_i I - A)x = 0$ 的非零解, 即为 λ_i 的特征向量;

注 1. $(\lambda_i I - A)x = 0$ 的解空间 V_i 称为特征值 λ_i 对应的特征子空间。

2. λ_i 作为 $p_A(\lambda) = 0$ 的根的重数称为 λ_i 的代数重数; $\dim V_i := m_i$ 称为 λ_i 的几何重数
3. 可以证明 $m_i \leq n_i$, 即几何重数小于等于代数重数。

命题 1.3 (特征值的性质)

- 若 λ 是 A 的特征值, x 是对应特征向量, $f(\lambda)$ 是多项式, 则 $f(\lambda)$ 是 $f(A)$ 的特征值, x 是 $f(A)$ 的对应特征向量;
- 若 λ 是 A 的特征值, x 是对应特征向量, 则 λ 是 A^T 的特征值, 但 x 不一定是对应特征向量;
- 若 λ 是 A 的特征值, 且 $A^T A = I$, 则 $|\lambda| = 1$;
- 若 λ 是实矩阵 A 的特征值, 且 $A = A^T$, 则 λ 是实数, 且存在 P 满足 $P^T P = I$, 使得 $P^{-1} A P$ 是对角矩阵;
- 若 λ 是实矩阵 A 的特征值, 且 $A = -A^T$, 则 λ 是纯虚数或零。

定理 1.4 特征值和特征向量是相似不变量, 即若 A, B 相似, 则 A, B 有相同的特征值。

定理 1.5 若 $\lambda_1, \cdots, \lambda_n$ 是 A 的全部特征值, 则 $\operatorname{tr} A = \sum_{i=1}^n \lambda_i, \det A = \prod_{i=1}^n \lambda_i$ 。

现在考虑相似标准形的一种简单情形, 即可相似于对角矩阵。一般情形, 即相似于 Jordan 标准形会在后面提及。

从定义出发, A 可对角化 $\Leftrightarrow \exists T \in GL_n(\mathbb{C}), \Lambda$ 是对角矩阵, 使得 $A = T \Lambda T^{-1}$ 。设 $\Lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_n), T = (x_1, \cdots, x_n)$, 则相似等价于 $AT = \Lambda T = (\lambda_1 x_1, \cdots, \lambda_n x_n)$ 。故只需要找到一组线性无关的特征向量排成 T 即可。如此我们也找到了求出相似对角化的过渡矩阵的方法。

引理 1.6 对 $A \in \mathbb{F}^{n \times n}$, A 属于不同特征值的特征向量线性无关。

定理 1.7 $A \in \mathbb{F}^{n \times n}$ 可相似对角化当且仅当 A 有 n 个线性无关的特征向量。

因此可以得到一些可对角化的条件:

命题 1.8 • 若 n 阶方阵 A 有 n 个不同的特征值, 则 A 可对角化;

- 若方阵 A 的所有特征值的代数重数等于几何重数, 则 A 可对角化;

例 1.1 记 $\mathbb{R}_n[x] = \left\{ f(x) \mid f(x) = \sum_{i=0}^n a_i x^i, a_i \in \mathbb{R} \right\}$, 定义线性变换 $\mathcal{D}: \mathbb{R}_n[x] \rightarrow \mathbb{R}_n[x], \mathcal{D}f(x) = f'(x)$, 求一组

基使得 $\mathcal{D}$ 在该基下的矩阵为 $J_n(0) = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$ 。

解答 设 $\{f_0, f_1, \dots, f_{n-1}\}$ 是满足题意的基, 则 $\mathcal{D}f_0 = 0, \mathcal{D}f_1 = f_0, \mathcal{D}f_2 = f_1, \dots, \mathcal{D}f_{n-1} = f_{n-2}$, 从而其关系是从前往后的逐项积分关系, 且 f_0 是常数项非零的常数多项式, 取 $f_0 = c_0$, 则 $f_1 = c_0 x + c_1, f_2 =$

$\frac{c_0}{2}x^2 + c_1x + c_2, \dots, f_{n-1} = \frac{c_0}{(n-1)!}x^{n-1} + \frac{c_1}{(n-2)!}x^{n-2} + \dots + c_{n-2}x + c_{n-1}$, 其中 $c_0, c_1, \dots, c_{n-1}$ 为任意常数, 故 $\{f_0, f_1, \dots, f_{n-1}\}$ 是满足题意的基。

例 1.2 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times m}$, 证明: AB 与 BA 的非零特征值相同。

证明.

$$\begin{aligned} \lambda^n \det(\lambda I_m - AB) &= \det \begin{bmatrix} \lambda I_n & O \\ A & \lambda I_m - AB \end{bmatrix} = \det \left(\begin{bmatrix} \lambda I_n & B \\ A & \lambda I_m \end{bmatrix} \begin{bmatrix} \lambda I_n & -B \\ O & \lambda I_m \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} \lambda I_n & -B \\ O & \lambda I_m \end{bmatrix} \begin{bmatrix} \lambda I_n & B \\ A & \lambda I_m \end{bmatrix} \right) = \det \begin{bmatrix} \lambda I_n - BA & O \\ A & \lambda I_m \end{bmatrix} = \lambda^m \det(\lambda I_n - BA) \end{aligned}$$

故 $\lambda^n p_{AB}(\lambda) = \lambda^m p_{BA}(\lambda)$, 即 AB 与 BA 的非零特征值相同

例 1.3 $A = \begin{bmatrix} 4 & 2 & 2 \\ 0 & 4 & 0 \\ 0 & -2 & 2 \end{bmatrix}$, 求 A^n 。

解答 $p_A(\lambda) = \det(\lambda I_n - A) = (\lambda - 2)(\lambda - 4)^2$, 故 A 有单重特征值 2 和二重特征值 4。

$(2I - A)x = 0$ 有非零解 $(1, 0, -1)^\top$, $(4I - A)x = 0$ 有线性无关非零解 $(1, 0, 0)^\top$ 和 $(0, 1, -1)^\top$,

$$\text{记 } T = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix}, \Lambda = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \text{ 有 } A = T\Lambda T^{-1},$$

$$\text{故 } A^n = T\Lambda^n T^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2^n & 0 & 0 \\ 0 & 4^n & 0 \\ 0 & 0 & 4^n \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 4^n & 4^n - 2^n & 4^n - 2^n \\ 0 & 4^n & 0 \\ 0 & 2^n - 4^n & 2^n \end{bmatrix}.$$

例 1.4 设 f 是 $\mathbb{R}^{n \times n}$ 上的线性函数, 且 $\forall A, B \in \mathbb{R}^{n \times n}, f(AB) = f(BA)$, 证明: $\exists \lambda \in \mathbb{R}, s.t. f(A) = \lambda \operatorname{tr} A, \forall A \in \mathbb{R}^{n \times n}$ 。

证明. 不妨设 $n \geq 2$, 则 $\forall i \neq j, f(E_{ij}) = f(E_{ii}E_{ij}) = f(O) = 0, f(E_{ii}) = f(E_{i1}E_{1i}) = f(E_{1i}E_{i1}) = f(E_{11})$, 故 $f(A) = f(E_{11}) \operatorname{tr} A$ 。

例 1.5 判断以下命题是否正确。

1. 若 $AB = O$, 则 $\det(xI - A) \det(xI - B) = x^n \det(xI - A - B)$;
2. AB 的特征值总能写成 A 的特征值和 B 的特征值的乘积;
3. 若 $\operatorname{rank} A = 1$, 则 $\det(xI - A) = x^n - \operatorname{tr} A \cdot x^{n-1}$ 。

解答

1. 正确。 $\det(xI - A) \det(xI - B) = \det(x^2I - x(A + B) + AB) = \det(x^2I - x(A + B)) = x^n \det(xI - A - B)$;
2. 错误。反例: $A = E_{12}, B = E_{21}$, 则 $AB = E_{11}, BA = E_{22}$, AB 的特征值为 1, 0, 而 A, B 的特征值均为 0;
3. 正确。设 $A = uv^\top, u, v \in \mathbb{R}^n$, 则 $\det(xI - A) = \det(xI - uv^\top) = x^n \det(I - \frac{1}{x}uv^\top) = x^n (1 - \frac{1}{x}v^\top u) = x^n - (v^\top u)x^{n-1} = x^n - \operatorname{tr} A \cdot x^{n-1}$ 。

2 一些其他结论 *

2.1 Caylay-Hamilton 定理

定理 2.1 (Caylay-Hamilton 定理) 对 $A \in \mathbb{F}^{n \times n}$, $p_A(A) = 0$, 其中 $p_A(\lambda) = \det(\lambda I - A)$ 。

证明. 设 $\lambda I - A$ 的伴随方阵是 B , 则 B 中元素都是次数至多为 $n-1$ 的多项式, 记 $b_{ij} = \sum_{k=0}^{n-1} b_{ij}^{(k)} \lambda^k$, 记 $B_k = (b_{ij}^{(k)})_{n \times n}$,

则 $B = \sum_{k=0}^{n-1} B_k \lambda^k$, 故 $p_A(\lambda)I_n = (\lambda I - A)B = \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - AB_{n-1}) + \cdots + \lambda(B_0 - AB_1) - AB_0$,

设 $p_A(\lambda) = \lambda^n + \sum_{k=0}^{n-1} a_k \lambda^k$, 比较系数有 $B_{n-1} = I_n, B_{n-2} - AB_{n-1} = a_{n-1}I_n, \cdots, B_0 - AB_1 = a_1I_n, -AB_0 = a_0I_n$,

以上等式乘 $A^n, A^{n-1}, \cdots, A, I_n$ 后相加, 得到 $p_A(A) = A^n + \sum_{k=0}^{n-1} a_k A^k = 0$ 。

推论 对 $A \in \mathbb{F}^{n \times n}$, A^n 可以用 $A^{n-1}, \cdots, A^2, A, I$ 线性表示, 即 $I, A, A^2, \cdots$ 张成空间的维数至多为 n 。

推论 Caylay-Hamilton 定理也给出了计算逆矩阵的一种方法: 若 A 可逆, 则由 $p_A(A) = 0$ 可得 $A^{-1} = -\frac{1}{a_0}(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_2A + a_1I)$ 。

2.2 Gershgorin 圆盘定理

矩阵的特征值的估计是一个重要的问题, Gershgorin 定理给出了特征值的一个上界。为了方便叙述, 我们先给出以下定义:

定义 2.1 对方阵 $A = (a_{ij})_{n \times n} \in \mathbb{F}^{n \times n}$,

- 若 $\forall 1 \leq i \leq n, |a_{ii}| > \sum_{j \neq i} |a_{ij}|$, 则称 A 是严格行对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$, 则称 A 是行对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| > \sum_{j \neq i} |a_{ji}|$, 则称 A 是严格列对角占优方阵;
- 若 $\forall 1 \leq i \leq n, |a_{ii}| \geq \sum_{j \neq i} |a_{ji}|$, 则称 A 是列对角占优方阵。

定理 2.2 (Levy-Desplanques) 若 A 是行或列严格对角占优矩阵, 则 $\det A \neq 0$ 。

证明. 只证明行严格对角占优矩阵的情形, 对列严格对角占优矩阵, 考虑其转置。

对 $A = (a_{ij})_{n \times n}$, 若 $\det A = 0$, 则 $Ax = 0$ 有非零解 $x = (x_1, \cdots, x_n)^T$, 设 $\max\{|x_1|, \cdots, |x_n|\} = |x_k| > 0$,

则 $\sum_{j=1}^n a_{kj}x_j = 0$, 即 $a_{kk}x_k = -\sum_{j \neq k} a_{kj}x_j$, 两边取模, $|a_{kk}||x_k| = |\sum_{j \neq k} a_{kj}x_j| \leq \sum_{j \neq k} |a_{kj}||x_j| \leq |x_k| \sum_{j \neq k} |a_{kj}|$,

$\Rightarrow |a_{kk}| \leq \sum_{j \neq k} |a_{kj}|$, 与 A 行严格对角占优矛盾, 故 $\det A \neq 0$ 。

定理 2.3 Gershgorin 圆盘定理 n 阶复方阵 $A = (a_{ij})_{n \times n}$ 的特征值落在复平面上 n 个圆盘 $|z - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$ 的并集内, 其中 $1 \leq i \leq n$ 。

证明. 设 λ 是 A 的特征值, 则 $\det(\lambda I - A) = 0$, 故由 Levy-Desplanques 定理有 $\lambda I - A$ 不是行严格对角占优矩阵。

故 $\exists 1 \leq i \leq n$, 使得 $|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$, 即特征值落在圆盘内。

2.3 同时对角化问题

定理 2.4 A, B 为 n 阶方阵, 且 A, B 均可对角化, 则 $\exists P \in GL_n(\mathbb{C})$ s.t. $P^{-1}AP, P^{-1}BP$ 均为对角矩阵当且仅当 $AB = BA$ 。

证明. $(\Rightarrow)$

由于对角矩阵乘法可交换, 故 $P^{-1}APP^{-1}BP = P^{-1}BPP^{-1}AP$, 即 $P^{-1}ABP = P^{-1}BAP$ 。由 P 可逆有 $AB = BA$ 。

$(\Leftarrow)$

由 A 可对角化有 $\exists Q \in GL_n(\mathbb{C})$ s.t. $Q^{-1}AQ = \begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix}$, 其中 λ_i 互不相同, $\sum_{i=1}^s n_i = n$ 。

由 $AB = BA$ 有 $Q^{-1}AQQ^{-1}BQ = Q^{-1}BQQ^{-1}AQ$ 。对 $Q^{-1}BQ$ 分块为 $\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix}$

$$\begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss} \end{bmatrix} \begin{bmatrix} \lambda_1 I_{n_1} & & \\ & \lambda_2 I_{n_2} & \\ & & \ddots \\ & & & \lambda_s I_{n_s} \end{bmatrix}$$

计算得到对 $i \neq j$ 均有 $B_{ij} = 0$, 即 $Q^{-1}BQ = \begin{bmatrix} B_{11} & & \\ & B_{22} & \\ & & \ddots \\ & & & B_{ss} \end{bmatrix}$ 为准对角矩阵。

由 B 可对角化有 $Q^{-1}BQ$ 可对角化, 故 B_{ii} 可对角化, 即 $\exists N_i$ 可逆 s.t. $N_i^{-1}B_{ii}N_i = \Sigma_i$ 为对角矩阵。

阵, $\begin{bmatrix} N_1^{-1} & & \\ & N_2^{-1} & \\ & & \ddots \\ & & & N_s^{-1} \end{bmatrix} Q^{-1}AQ \begin{bmatrix} N_1 & & \\ & N_2 & \\ & & \ddots \\ & & & N_s \end{bmatrix} = \begin{bmatrix} \lambda_1 N_1^{-1} N_1 & & \\ & \lambda_2 N_2^{-1} N_2 & \\ & & \ddots \\ & & & \lambda_s N_s^{-1} N_s \end{bmatrix}$ 是对角矩阵,

$$\text{而 } \begin{bmatrix} N_1^{-1} & & & \\ & N_2^{-1} & & \\ & & \ddots & \\ & & & N_s^{-1} \end{bmatrix} Q^{-1} B Q \begin{bmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_s \end{bmatrix} = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & \ddots & \\ & & & \Sigma_s \end{bmatrix} \text{ 也是对角矩阵。}$$

$$\text{记 } N = \begin{bmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_s \end{bmatrix}, \text{ 取 } P = QN \text{ 即为所求。}$$

例 2.1 ((2020~2021 期末)) 对 A, B 为 n 阶实对称方阵, 且 $AB = BA$, 证明: $\exists P \in GL_n(\mathbb{C})$ s.t. $P^{-1}AP, P^{-1}BP$ 均为对角矩阵。

解答 这是定理的直接推论。注意到实对称矩阵均可对角化即可。