

0.1 矩阵相似和可对角化的计算

0.1.1 相似初等变换及其应用

命题 0.1 (相似初等变换)

设 A 为 n 阶矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似初等变换**:

1. 对换 A 的第 i 行与第 j 行, 再对换第 i 列与第 j 列;
2. 将 A 的第 i 行乘以非零常数 c , 再将第 i 列乘以 c^{-1} ;
3. 将 A 的第 i 行乘以常数 c 加到第 j 行上, 再将第 j 列乘以 $-c$ 加到第 i 列上.

设 A 是具有相同行列分块方式的分块矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似分块初等变换**:

1. 对换 A 的第 i 分块行与第 j 分块行, 再对换第 i 分块列与第 j 分块列;
2. 将 A 的第 i 分块行左乘非异阵 M , 再将第 i 分块列右乘 M^{-1} ;
3. 将 A 的第 i 分块行左乘矩阵 M 加到第 j 分块行上, 再将第 j 分块列右乘 $-M$ 加到第 i 分块列上.

容易验证: 任一相似变换都是若干次相似初等变换的复合.

推论 0.1

设 $A = \text{diag}\{A_1, A_2, \dots, A_m\}$ 是分块对角矩阵, 其中 A_i 都是方阵, 求证: $\text{diag}\{A_1, A_2, \dots, A_m\}$ 相似于 $\text{diag}\{A_{i_1}, A_{i_2}, \dots, A_{i_m}\}$, 其中 $A_{i_1}, A_{i_2}, \dots, A_{i_m}$ 是 A_1, A_2, \dots, A_m 的一个排列.

证明 对换 A 的第 i 分块行与第 j 分块行, 再对换第 i 分块列与第 j 分块列. 这是一个相似变换, 变换的结果是将 A 的 (i, i) 分块和第 (j, j) 分块对换了位置. 又任一排列都可以通过若干次对换来实现, 因此 $\text{diag}\{A_1, A_2, \dots, A_m\}$ 和 $\text{diag}\{A_{i_1}, A_{i_2}, \dots, A_{i_m}\}$ 相似. \square

命题 0.2

若分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 满足 $r\begin{pmatrix} A & B \\ C & D \end{pmatrix} = r(A)$, 则 $r\begin{pmatrix} A & B \\ C & D \end{pmatrix} = r\begin{pmatrix} A \\ C \end{pmatrix} = r(A)$.

证明 由条件可得

$$r(A) \leq r\begin{pmatrix} A & B \\ C & D \end{pmatrix} = r\begin{pmatrix} A \\ C \end{pmatrix} \leq r\begin{pmatrix} A & B \\ C & D \end{pmatrix} = r(A).$$

故 $r\begin{pmatrix} A & B \\ C & D \end{pmatrix} = r\begin{pmatrix} A \\ C \end{pmatrix} = r(A)$. \square

例题 0.1 设 n 阶方阵 A, B 满足 $r(ABA) = r(B)$, 求证: AB 与 BA 相似.

注 不妨设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立. 则对于一般的满足条件的矩阵 A , 设 $r(A) = r$, 则存在可逆矩阵 P, Q , 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 再根据条件可知

$$r((PAQ)(Q^{-1}BP^{-1})(PAQ)) = r(PABAQ) = r(ABA) = r(B) = r(Q^{-1}BP^{-1}).$$

从而由假设可知, $(PAQ)(Q^{-1}BP^{-1})$ 与 $(Q^{-1}BP^{-1})(PAQ)$ 相似, 即 $PABP^{-1}$ 与 $Q^{-1}BAQ$ 相似. 又注意到 $PABP^{-1}$ 与 AB 相似, $Q^{-1}BAQ$ 与 BA 相似, 因此 AB 与 BA 也相似. 故不妨设成立.

证明 设 P, Q 为 n 阶非异阵, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r = r(A)$. 注意到问题的条件和结论在相抵变换: $A \mapsto$

$PAQ, B \mapsto Q^{-1}BP^{-1}$ 下保持不变, 故不妨从一开始就假设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 为

对应的分块, 则由 $r(ABA) = r(B)$ 可得 $r\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = r(B_{11})$. 从而由命题 0.2 可得 $r\begin{pmatrix} B_{11} & B_{12} \end{pmatrix} = r(B_{11})$ 以及 $r\begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} = r(B_{11})$, 再由定理 ?? 可知存在矩阵 M, N , 使得 $B_{11}N = B_{12}, MB_{11} = B_{21}$.

将 $AB = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}$ 的第二分块行左乘 N 加到第一分块行, 再将第一分块列右乘 $-N$ 加到第二分块列, 于是 AB 相似于 $\begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix}$. 将 $BA = \begin{pmatrix} B_{11} & O \\ B_{21} & O \end{pmatrix}$ 的第一分块行左乘 $-M$ 加到第二分块行, 再将第二分块列右乘 M 加到第一分块列, 于是 BA 相似于 $\begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix}$. 因此, AB 与 BA 相似. \square

0.1.2 利用相似不变量来判定矩阵不相似

定义 0.1 (相似不变量)

相似的矩阵具有相同的迹、行列式、特征多项式和极小多项式等, 故它们被称为矩阵相似关系下的不变量, 也称为相似不变量.

注 若两个矩阵的相似不变量不相同, 则它们必不相似. 利用这种方法来判断两个矩阵不相似是很简便的.

笔记 由矩阵迹、行列式、特征多项式的性质易知, 相似的矩阵具有相同的迹、行列式、特征多项式. 由命题 ?? 可知, 相似的矩阵具有相同的极小多项式. 因此上述定义是良定义的.

例题 0.2 设 A, B 为 n 阶方阵, 求证: $AB - BA$ 必不相似于 kI_n , 其中 k 是非零常数.

证明 注意到 $\text{tr}(AB - BA) = 0$, $\text{tr}(kI_n) = nk \neq 0$, 又矩阵的迹是相似不变量, 因此 $AB - BA$ 和 kI_n 必不相似. \square

例题 0.3 设 A, B 为 n 阶正交矩阵, 且线性方程组 $(A + B)x = 0$ 的解空间维数是奇数, 求证: A 和 B 必不相似.

证明 由假设可知 $n - r(A + B)$ 为奇数, 再由命题 ?? 可知 $|A| = -|B| \neq 0$, 又矩阵的行列式是相似不变量, 因此 A 和 B 必不相似. \square

0.1.3 过渡矩阵 P 的计算

首先, 我们介绍一下当矩阵相似于对角矩阵时求过渡矩阵的方法. 设 n 阶矩阵 A 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 可逆矩阵 $P = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 为其列分块, 且

$$P^{-1}AP = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

则

$$AP = P \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

即

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\lambda_1\alpha_1, \lambda_2\alpha_2, \dots, \lambda_n\alpha_n),$$

于是 $A\alpha_i = \lambda_i\alpha_i$, 这表明 α_i 就是属于特征值 λ_i 的特征向量. 因此 P 的 n 个列向量就是 A 的 n 个线性无关的特征向量. 注意: 因为特征向量不唯一, 所以过渡矩阵 P 也不唯一. 另外, P 的第 i 个列向量对应于 A 的第 i 个特征值.

例题 0.4 设三阶矩阵 A 的特征值为 1, 1, 4, 对应的特征向量依次为 $(2, 1, 0)'$, $(-1, 0, 1)'$, $(0, 1, 1)'$, 试求矩阵 A .

解 容易验证 A 的这 3 个特征向量线性无关, 故 A 必相似于对角矩阵, 即有

$$P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

根据上面的分析, 有

$$P = \begin{pmatrix} 2 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

于是

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 7 & -3 \\ -3 & 6 & -2 \end{pmatrix}.$$

□

0.1.4 可对角化判定的计算

例题 0.5 已知矩阵 $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & x & -2 \\ -3 & -3 & y \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & z \end{pmatrix}$ 相似.

(1) 求 x, y, z 的值;

(2) 求一个满足 $P^{-1}AP = B$ 的可逆矩阵 P .

解 (1) 显然 $z \neq 2$, 否则由 A 相似于 $2I_3$ 可知 $A = 2I_3$, 矛盾. 于是 A 的特征值为 2 (2 重), z (1 重). 因为 A 可对角化, 所以特征值 2 的几何重数也等于 2, 即线性方程组 $(A - 2I_3)x = O$ 的解空间维数也是 2, 也即 $3 - r(A - 2I_3) = 2$. 故有

$$r(A - 2I_3) = r \begin{pmatrix} -1 & -1 & 1 \\ 2 & x-2 & -2 \\ -3 & -3 & y-2 \end{pmatrix} = 1,$$

由此可得 $x = 4, y = 5$. 再由矩阵的迹等于特征值之和可得 $10 = \text{tr}(A) = 4 + z$, 故 $z = 6$.

(2) 通过求解线性方程组 $(\lambda I_3 - A)x = O (\lambda = 2, 6)$ 计算可得: 特征值 2 的两个线性无关的特征向量为 $\alpha_1 = (-1, 1, 0)'$, $\alpha_2 = (1, 0, 1)'$; 特征值 6 的特征向量为 $\alpha_3 = (1, -2, 3)'$. 因此

$$P = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

□

例题 0.6 设 $A = \begin{pmatrix} 3 & 2 & -2 \\ -k & -1 & k \\ 4 & 2 & -3 \end{pmatrix}$, 当 k 为何值时, 存在可逆矩阵 P , 使得 $P^{-1}AP$ 是对角矩阵? 求出 P 和对角矩阵.

解 经计算可得 $|\lambda I_3 - A| = (\lambda - 1)(\lambda + 1)^2$, 因此 A 的特征值为 1 (1 重), -1 (2 重). 对单特征值 1, 其几何重数与代数重数必相等 (因为几何重数总小于代数重数且一个特征值必存在一个对应的特征向量); 因此要使 A 可对角化, 特征值 -1 的几何重数必须等于 2 才行, 故有

$$r(A + I_3) = r \begin{pmatrix} 4 & 2 & -2 \\ -k & 0 & k \\ 4 & 2 & -2 \end{pmatrix} = 1,$$

于是 $k = 0$. 通过计算可得: 特征值 1 的特征向量为 $(1, 0, 1)'$; 特征值 -1 的两个线性无关的特征向量为 $(-1, 2, 0)'$, $(1, 0, 2)'$.


因此

$$P = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

□

0.1.5 可对角化矩阵的应用

例题 0.7 设矩阵 $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & -2 \\ -3 & -3 & 5 \end{pmatrix}$, 求 A^n .

 **笔记** 本题中的矩阵是一个可对角化矩阵, 因此可以使用下列方法: 先求出可逆矩阵 P , 使得 $P^{-1}AP = B$ 是对角矩阵. 因为对角矩阵的幂很容易求出, 故由 $A^n = PB^nP^{-1}$ 即可得到结果.

解 经计算可得 $|\lambda I_3 - A| = (\lambda - 2)^2(\lambda - 6)$, 因此 A 的特征值为 2 (2 重), 6 (1 重). 通过计算可得: 特征值 2 有两个线性无关的特征向量 $(-1, 1, 0)'$, $(1, 0, 1)'$; 特征值 6 有特征向量 $(1, -2, 3)'$, 于是 A 有完全的特征向量系, 从而可对角化. 注意到

$$P = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}, B = P^{-1}AP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

故

$$A^n = PB^nP^{-1} = \frac{1}{4} \begin{pmatrix} 5 \cdot 2^n - 6^n & 2^n - 6^n & 6^n - 2^n \\ 2 \cdot 6^n - 2^{n+1} & 2 \cdot 6^n + 2^{n+1} & 2^{n+1} - 2 \cdot 6^n \\ 3 \cdot 2^n - 3 \cdot 6^n & 3 \cdot 2^n - 3 \cdot 6^n & 2^n + 3 \cdot 6^n \end{pmatrix}.$$

□

例题 0.8 下列数列称为 Fibonacci 数列: $a_0 = 0, a_1 = 1, a_2 = 1, a_3 = 2, a_4 = 3, a_5 = 5, a_6 = 8, \dots$, 通用递推式来表示为 $a_{n+2} = a_{n+1} + a_n$, 试求 Fibonacci 数列通项的显式表达式.

注 本例为用递推式定义的数列的通项提供了一个一般性的方法. 在三对角行列式中, 我们用技巧性相当高的方法求出了用下列递推式定义的数列的通项:

$$D_n = aD_{n-1} - bcD_{n-2}, D_1 = a, D_2 = a^2 - bc.$$

现在请读者自己用上面的方法来求出通项表达式.

证明 首先用矩阵来表示递推式:

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}.$$

令 $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, 则

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = A \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = A^2 \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = \dots = A^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

只要求出 A^n 就可以算出 a_n 来. A 的特征多项式为 $|\lambda I_2 - A| = \lambda^2 - \lambda - 1$, 解得

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

对应于特征值 λ_1, λ_2 的特征向量分别是:

$$\alpha_1 = \begin{pmatrix} \frac{1 + \sqrt{5}}{2} \\ 1 \end{pmatrix}, \alpha_2 = \begin{pmatrix} \frac{1 - \sqrt{5}}{2} \\ 1 \end{pmatrix}.$$

若记

$$\mathbf{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix},$$

则

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

因此

$$\mathbf{A} = \mathbf{P} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \mathbf{P}^{-1}, \mathbf{A}^n = \mathbf{P} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}^n \mathbf{P}^{-1}.$$

经计算可得

$$\mathbf{A}^n = \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \\ \left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n & \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{pmatrix},$$

由此可得

$$a_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right).$$

□