

0.1 矩阵分解

定理 0.1 (Jordan 分解)

设 \mathbb{F} 是一代数闭域, $n \in \mathbb{N}$.

1. 对 $A \in \mathbb{F}^{n \times n}$, 存在 $C, B \in \mathbb{F}^{n \times n}$ 使得:

(1): $A = C + B$;

(2): C 是幂零矩阵, B 是可对角化矩阵;

(3): $BC = CB$;

(4): 存在没有常数项的多项式 $p(x), q(x) \in \mathbb{F}[x]$, 使得 $C = q(A), B = p(A)$.

且满足 (1)–(3) 的分解 B, C 是唯一的.

2. 对可逆矩阵 $A \in \mathbb{F}^{n \times n}$, 存在 $C, B \in \mathbb{F}^{n \times n}$ 使得:

(1): $A = CB$;

(2): C 是特征值全为单位元 1 的矩阵, B 是可对角化的可逆矩阵;

(3): $BC = CB$;

(4): 存在没有常数项的多项式 $p(x), q(x) \in \mathbb{F}[x]$, 使得 $C = q(A), B = p(A)$.

且满足 (1)–(3) 的分解 B, C 是唯一的.



注 要证满足条件 (1)(2)(3) 的分解是唯一的等价于: 设 B, C 满足 (1)(2)(3)(4), 再设 B', C' 满足 (1)(2)(3), 只要证明 $b = B', C = C'$ 即可.

证明

1. **存在性:** 注意到相似变换是同一线性变换在不同基下的表示罢了, 因此不妨设

$$A = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix},$$

其中 $J_i, i = 1, 2, \dots, s$ 是同一个特征值对应所有 Jordan 块排在一起的分块对角矩阵, 设其对应特征值为 $\lambda_i, i = 1, 2, \dots, s$. 我们有

$$A = \begin{pmatrix} J_1 - \lambda_1 I & 0 & \cdots & 0 \\ 0 & J_2 - \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s - \lambda_s I \end{pmatrix} + \begin{pmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_s I \end{pmatrix} = C + B.$$

由中国剩余定理, 设 $s_i \in \mathbb{N}$ 是 J_i 的阶数, 则存在 $p \in \mathbb{F}[\lambda]$, 使得

$$p(\lambda) \equiv \lambda_i \pmod{(\lambda - \lambda_i)^{s_i}}, \quad i = 1, 2, \dots, s. \quad (1)$$

如果式 (1) 中某个 $\lambda_i = 0$, 此时显然有解 p , 从而 p 没有常数项. 如果 $\lambda_i \neq 0, i = 1, 2, \dots, s$, 我们可以在 (1) 中补一个方程 $p(\lambda) = 0 \pmod{\lambda}$ 使得 $p(0) = 0$, 从而 p 没有常数项. 注意到

$$(J_i - \lambda_i I)^{s_i} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{s_i \times s_i} = O.$$

于是 $p(J_i) = k_i (J_i) (J_i - \lambda_i I)^{s_i} + \lambda_i I = \lambda_i I, i = 1, 2, \dots, s$. 因此 $p(A) = B$. 令 $q = x - p$, 则

$$q(A) = A - p(A) = A - B = C.$$

这就证明了存在性.

唯一性: 若还有 $A = C' + B'$ 且满足条件 (1)–(3), 则 $C' - C = B - B'$. 又 C' 和 C', B' 可交换, 所以 C' 和 A 可交换, 又 $C = q(A)$, 所以 C' 和 C 可交换. 类似的 B 和 B' 可交换. 因此对充分大的 n , 利用 C, C' 为幂零矩阵和可交换的矩阵同时对角化, 我们有

$$0 = \sum_{k=0}^n C_n^k (C')^k C^{n-k} = (C' - C)^n = (B - B')^n \implies B = B', C = C',$$

这就证明了唯一性.

2. **存在性:** 当 A 可逆时, 注意到相似变换只是同一线性变换在不同基下的表示, 因此不妨设

$$A = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_s \end{pmatrix},$$

其中 $J_i, i = 1, 2, \dots, s$ 是同一个特征值对应所有 Jordan 块排在一起的分块对角矩阵, 设其对应特征值为 λ_i , 且 $\lambda_i \neq 0, i = 1, 2, \dots, s$. 于是我们有

$$A = \begin{pmatrix} \frac{J_1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{J_2}{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{J_s}{\lambda_s} \end{pmatrix} \begin{pmatrix} \lambda_1 I & 0 & \cdots & 0 \\ 0 & \lambda_2 I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_s I \end{pmatrix} = CB.$$

显然上述 C, B 满足 (1)(2)(3), 下证 (4). 由中国剩余定理, 设 $s_i \in \mathbb{N}$ 是 J_i 的阶数, 则存在 $p \in \mathbb{F}[\lambda]$, 使得

$$p(\lambda) \equiv \lambda_i \pmod{(\lambda - \lambda_i)^{s_i}}, i = 1, 2, \dots, s,$$

$$p(\lambda) \equiv 0 \pmod{\lambda}.$$

从而 $p(0) = 0$, 即 p 没有常数项. 注意到

$$(J_i - \lambda_i I)^{s_i} = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{s_i \times s_i}^{s_i} = O.$$

于是 $p(J_i) = k_i (J_i) (J_i - \lambda_i I)^{s_i} + \lambda_i I = \lambda_i I, i = 1, 2, \dots, s$. 因此 $p(A) = B$. 因为逆矩阵是原矩阵的多项式, 所以有 $(p(A))^{-1} = h(A), h \in \mathbb{F}[\lambda]$. 令 $q = x \cdot h \in \mathbb{F}[\lambda]$, 则

$$q(a) = Ah(A) = A(p(A))^{-1} = AB^{-1} = C.$$

这就证明了存在性.

唯一性: 对 $A = C'B'$ 且满足条件 (1)–(3), 我们有 $C^{-1}C' = B(B')^{-1}$. 又 C' 和 C', B' 可交换, 所以 C' 和 A 可交换, 又 $C = q(A)$, 所以 C' 和 C 可交换. 类似的 B 和 B' 可交换. 利用 C, C' 可同时上三角化, 因此 $C^{-1}C'$ 特征值全部是 1, 故 $B(B')^{-1}$ 特征值全为 1. 利用 B, B' 可同时对角化, 因此 $B(B')^{-1} = E$, 于是 $B = B', C = C'$, 这就证明了唯一性.

□

定理 0.2 (LU 分解)

设 \mathbb{F} 是一域, 则对 $A \in \mathbb{F}^{n \times n}$, 存在 $C, B \in \mathbb{F}^{n \times n}$ 使得 $A = BC$ 且 B 为主对角元都为 1 的下三角矩阵, C 为可逆上三角矩阵的充要条件是 A 的顺序主子式全不为 0. 并且这种分解是唯一的.



证明 充分性: 当 $n = 1$ 问题是显然的. 设命题对 $n - 1$ 时成立, 当 n 时, 设

$$A = \begin{pmatrix} A_{n-1} & \alpha \\ \beta^T & a_{nn} \end{pmatrix}, A_{n-1} \in \mathbb{F}^{(n-1) \times (n-1)}, \alpha, \beta \in \mathbb{F}^{n-1}, a_{nn} \in \mathbb{F}.$$

利用归纳假设, 存在 $C_{n-1}, B_{n-1} \in \mathbb{F}^{(n-1) \times (n-1)}$ 使得 $A_{n-1} = B_{n-1}C_{n-1}$ 且 B_{n-1} 为主对角元都为 1 的下三角矩阵, C_{n-1} 为可逆上三角矩阵. 注意到

$$A = \begin{pmatrix} A_{n-1} & \alpha \\ \beta^T & a_{nn} \end{pmatrix} = \begin{pmatrix} B_{n-1}C_{n-1} & \alpha \\ \beta^T & a_{nn} \end{pmatrix} \rightarrow \begin{pmatrix} B_{n-1}C_{n-1} & \alpha \\ 0 & a_{nn} - \beta^T C_{n-1}^{-1} B_{n-1}^{-1} \alpha \end{pmatrix},$$

于是

$$\begin{pmatrix} E_{n-1} & 0 \\ -\beta^T C_{n-1}^{-1} B_{n-1}^{-1} & 1 \end{pmatrix} A = \begin{pmatrix} B_{n-1}C_{n-1} & \alpha \\ 0 & a_{nn} - \beta^T C_{n-1}^{-1} B_{n-1}^{-1} \alpha \end{pmatrix}.$$

于是我们有

$$A = \begin{pmatrix} E_{n-1} & 0 \\ -\beta^T C_{n-1}^{-1} B_{n-1}^{-1} & 1 \end{pmatrix}^{-1} \begin{pmatrix} B_{n-1} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} C_{n-1} & B_{n-1}^{-1} \alpha \\ 0 & a_{nn} - \beta^T C_{n-1}^{-1} B_{n-1}^{-1} \alpha \end{pmatrix} = BC,$$

因为 $\det A \neq 0$, 所以 C 是可逆矩阵. 这样我们就完成了 n 时的证明, 从而由数学归纳法即证.

必要性: 当 $n = 1$ 命题显然成立, 假设命题对 $n - 1$ 时成立, 则当 n 时, 设 $A = BC$ 且 B 为主对角元都为 1 的下三角矩阵, C 为可逆上三角矩阵.

假设

$$B = \begin{pmatrix} B'_{n-1} & 0 \\ \gamma^T & 1 \end{pmatrix}, C = \begin{pmatrix} C'_{n-1} & \eta \\ 0 & c_{nn} \end{pmatrix}, B'_{n-1}, C'_{n-1} \in \mathbb{F}^{(n-1) \times (n-1)}, c_{nn} \in \mathbb{F} \setminus \{0\}.$$

现在就有

$$A = BC = \begin{pmatrix} B'_{n-1}C'_{n-1} & B'_{n-1}\eta + c_{nn} \\ \gamma^T C'_{n-1} & \gamma^T \eta + c_{nn} \end{pmatrix}.$$

由归纳假设知 $B'_{n-1}C'_{n-1}$ 顺序主子式不为 0. 显然 $\det A \neq 0$, 这就证明了 n 的情况, 从而由数学归纳法即证.

唯一性: 若还有 $A = BC = B'C'$, 这里 B' 为主对角元都为 1 的下三角矩阵, C' 为可逆上三角矩阵. 于是

$$C(C')^{-1} = B^{-1}B'.$$

注意到上式左边是上三角矩阵, 右边是对角线全为 1 的下三角矩阵, 故

$$C(C')^{-1} = B^{-1}B' = E_n,$$


从而 $C = C', B = B'$, 这就完成了证明. □

定理 0.3 (奇异值分解)

设 $A \in \mathbb{R}^{m \times n}(\mathbb{C}^{m \times n})$, 则存在实正交矩阵 (酉矩阵) $U \in \mathbb{R}^{m \times m}(\mathbb{C}^{m \times m})$ 和 $V \in \mathbb{R}^{n \times n}(\mathbb{C}^{n \times n})$, 使得

$$U^T A V = \begin{pmatrix} \Lambda & 0 \\ 0 & 0 \end{pmatrix},$$

这里 Λ 是 $r(A)$ 阶对角矩阵且对角元都为正数.

 **笔记** 如果这个定理已经被证明, 那么

$$(U^T A V)^T U^T A V = V^T A^T A V = \begin{pmatrix} \Lambda^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

由于 $A^T A$ 半正定, 所以 Λ 的对角线元素是 $A^T A$ 的非 0 特征值的算术平方根. 由 CD 和 DC 有完全一样 (包括重数) 的非 0 特征值, 我们知道 Λ 的对角线元素也是 AA^T 的非 0 特征值的算术平方根.

注 奇异值分解证明过程也是其计算过程.

证明 仅证明实矩阵的情形, 因为 $A^T A$ 半正定, 所以存在正交矩阵 $V \in \mathbb{R}^{n \times n}$ 使得

$$V^T A^T A V = \begin{pmatrix} \Lambda^2 & 0 \\ 0 & 0 \end{pmatrix}. \quad (2)$$

我们设

$$V = (V_1 \ V_2), V_1 \in \mathbb{R}^{n \times r}, V_2 \in \mathbb{R}^{n \times (n-r)},$$

则有

$$V^T A^T A V = \begin{pmatrix} V_1^T A^T A V_1 & V_1^T A^T A V_2 \\ V_2^T A^T A V_1 & V_2^T A^T A V_2 \end{pmatrix}. \quad (3)$$

对比 (2) 和 (3) 得

$$0 = V_2^T A^T A V_2 = (A V_2)^T (A V_2) \xrightarrow{\text{命题??}} A V_2 = 0, V_1^T A^T A V_1 = \Lambda^2.$$

待定

$$U = (U_1 \ U_2), U_1 \in \mathbb{R}^{m \times r}, U_2 \in \mathbb{R}^{m \times (m-r)},$$

则

$$U^T A V = \begin{pmatrix} U_1^T A V_1 & U_1^T A V_2 \\ U_2^T A V_1 & U_2^T A V_2 \end{pmatrix} = \begin{pmatrix} U_1^T A V_1 & 0 \\ U_2^T A V_1 & 0 \end{pmatrix}.$$

目标是

$$U_1^T A V_1 = \Lambda \iff U_1 = A V_1 \Lambda^{-1}.$$

现在取定满足上述条件的 U_1 , 则

$$U_1^T U_1 = \Lambda^{-1} V_1^T A^T A V_1 \Lambda^{-1} = E.$$

现在 U_1 列向量组是标准正交向量组, 我们可以把 U_1 扩充为正交矩阵, 此时

$$U_2^T A V_1 = U_2^T U_1 \Lambda = 0,$$

这就完成了证明. □

定理 0.4 (半正定矩阵开方唯一)

设 A 是实 n 阶半正定矩阵, 则存在唯一的实 n 阶半正定矩阵 B , 使得 $A = B^2$.



笔记 当然, 对于复版本的半正定 Hermite 矩阵, 这个结果也是对的. 显然, 证明的方法就是不妨设标准技巧, 即实对称矩阵的正交相似对角化.

注 对于这个定理中的 B , 存在至多 $n-1$ 次多项式 $p \in \mathbb{R}[x]$ 使得 $B = p(A)$. 这是因为由 Lagrange 插值定理, 存在至多 $n-1$ 次多项式使得

$$p(\lambda_i) = \sqrt{\lambda_i}, i = 1, 2, \dots, n.$$

证明 取正交矩阵 T , 使得

$$T^T A T = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}, \lambda_i \geq 0, i = 1, 2, \dots, n.$$

直接取半正定矩阵 $B = T \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}\} T^T$, 直接验证就有 $A = B^2$.

下面证明最为困难的唯一性. 事实上, 若还有半正定矩阵 C , 有 $A = C^2 = B^2$. 再取正交矩阵 S, Q , 使得

$$K \triangleq S^T B S = \text{diag}\{\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_n}\} = Q^T C Q,$$

于是

$$B = S K S^T, C = Q K Q^T \implies S K^2 S^T = B^2 = C^2 = Q K^2 Q^T \implies Q^T S K^2 = K^2 Q^T S.$$

现在假设

$$K = \begin{pmatrix} \mu_1 E & & \\ & \mu_2 E & \\ & & \ddots \\ & & & \mu_s E \end{pmatrix}, \text{ 这里 } \mu_i \text{ 互不相同 } (\mu_i \in \{\lambda_1, \dots, \lambda_n\}).$$

于是把 $Q^T S$ 对应分块为 (X_{ij}) , 那么就有

$$\mu_j^2 X_{ij} = X_{ij} \mu_i^2.$$

于是对任何 $i \neq j$ 都有 $X_{ij} = 0$. 现在

$$Q^T S = \text{diag}\{X_{11}, X_{22}, \dots, X_{ss}\} \implies Q^T S K = K Q^T S \implies B = S K S^T = Q K Q^T = C,$$

这就完成了证明. □

命题 0.1


设 $A \in \mathbb{R}^{n \times n}$ 是半正定的. 若 $\alpha \in \mathbb{R}^n$ 使得 $\alpha^T A \alpha = 0$, 则 $A \alpha = 0$. ♣

证明 由半正定矩阵开方唯一得半正定矩阵 $C \in \mathbb{R}^{n \times n}$ 使得 $A = C^2$. 于是

$$0 = \alpha^T A \alpha = (C \alpha)^T (C \alpha) = 0 \implies C \alpha = 0 \implies A \alpha = C^2 \alpha = 0.$$
□

定理 0.5 (极分解)

1. 设 $A \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$, 则存在正交矩阵 (酉矩阵) $Q \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$ 和半正定矩阵 (Hermite 半正定矩阵) $T \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$ 使得 $A = QT$ 且 T 是唯一的. 特别的当 A 可逆, 则 Q 也是唯一的.
2. 设 $A \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$, 则存在正交矩阵 (酉矩阵) $Q \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$ 和半正定矩阵 (Hermite 半正定矩阵) $T \in \mathbb{R}^{n \times n} (\mathbb{C}^{n \times n})$ 使得 $A = TQ$ 且 T 是唯一的. 特别的当 A 可逆, 则 Q 也是唯一的. ♥

 **笔记** 极分解对非方阵也有结果, 此时要用不完全的正交矩阵来代替, 因为不常用, 我们略去.

证明 只证明实的情况, 复的情况类似.

1. 由奇异值分解, 我们知道存在正交矩阵 $U, V \in \mathbb{R}^{n \times n}$ 和奇异标准型 Λ , 使得 $A = U \Lambda V$. 现在

$$A = UV \cdot V^T \Lambda V = Q \cdot T.$$

若还有 $A = Q' T'$, 这里 $Q', T' \in \mathbb{R}^{n \times n}$ 分别为正交矩阵和半正定矩阵. 则

$$A^T A = (T')^2 = T^2.$$

由命题??(2) 可知 $A^T A$ 是半正定阵. 又由半正定矩阵开方唯一知 $T = T'$.

当 A 可逆, 由于 T 正定且唯一, 故

$$A = Q' T' = Q' T = Q T \implies Q' = Q.$$


2. 和第一问类似. □

定理 0.6 (满秩分解)

设 \mathbb{F} 是域, 则

1. 设 $A \in \mathbb{F}^{m \times n}$ 且秩为 $r \in \mathbb{N}_0$, 证明存在列满秩矩阵 P 和行满秩矩阵 Q , 使得 $A = PQ$.
2. 设 $A \in \mathbb{F}^{m \times n}$ 且秩为 n , 证明存在可逆矩阵 $P \in \mathbb{F}^{m \times m}$, 使得 $PA = \begin{pmatrix} E_n \\ 0 \end{pmatrix}$.
3. 设 $A \in \mathbb{F}^{m \times n}$ 且秩为 m , 证明存在可逆矩阵 $Q \in \mathbb{F}^{n \times n}$, 使得 $AQ = (E_m \ 0)$.
4. 设 $A, B \in \mathbb{F}^{m \times n}$ 是列满秩矩阵, 则存在可逆矩阵 $C \in \mathbb{F}^{m \times m}$ 使得 $A = CB$.

5. 设 $A, B \in \mathbb{F}^{m \times n}$ 是行满秩矩阵, 则存在可逆矩阵 $C \in \mathbb{F}^{n \times n}$ 使得 $A = BC$.

 **笔记** 本定理的二三条可以直接使用.

证明

1. 设可逆矩阵 $P_1 \in \mathbb{F}^{m \times m}, Q_1 \in \mathbb{F}^{n \times n}$ 使得

$$P_1 A Q_1 = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

注意到

$$A = P_1^{-1} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_1^{-1} = P_1^{-1} \begin{pmatrix} E_r \\ 0 \end{pmatrix} \cdot (E_r \quad 0) Q_1^{-1} = P \cdot Q,$$

这就证明了第一问.

2. 这是课本定理, 通过行变换化阶梯得到的.

3. 对 A^T 用第二问的结论即可得到行的结论.

4. 由第二问结论, 设可逆矩阵 $P_A, P_B \in \mathbb{F}^{m \times m}$ 使得 $P_B B = P_A A = \begin{pmatrix} E_n \\ 0 \end{pmatrix}$, 则

$$A = P_A^{-1} P_B B = C B.$$

这就证明了第四问.

5. 和第四问同理.

□

定理 0.7 (幂等分解)

设 \mathbb{F} 是域, $A \in \mathbb{F}^{n \times n}$, 证明存在幂等矩阵 $T_1, T_2 \in \mathbb{F}^{n \times n}$, 可逆矩阵 $S_1, S_2 \in \mathbb{F}^{n \times n}$, 使得 $A = T_1 S_1 = S_2 T_2$.

□

证明 注意到

$$A = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q = P \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} P^{-1} P Q = T_1 S_1,$$

这里 $P, Q \in \mathbb{F}^{n \times n}$ 是可逆矩阵. 类似的可得另一分解.

□

定理 0.8 (秩一分解)

设 \mathbb{F} 是域, $r \in \mathbb{N}$, 则秩 r 矩阵 $A \in \mathbb{F}^{m \times n}$ 可以分解为 r 个 $\mathbb{F}^{m \times n}$ 中秩 1 矩阵之和.

□

证明 设可逆矩阵 $P_1 \in \mathbb{F}^{m \times m}, Q_1 \in \mathbb{F}^{n \times n}$ 使得

$$P_1 A Q_1 = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}.$$

注意到

$$A = P_1^{-1} \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} Q_1^{-1} = P_1^{-1} \underbrace{\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} Q_1^{-1} + P_1^{-1} \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 0 & \\ & & & \ddots \end{pmatrix} Q_1^{-1} + \cdots}_{\text{合计 } r \text{ 个}},$$

这就完成了证明.

□

定理 0.9 (施密特正交化对应的矩阵分解, QR 分解)

1. 设列满秩矩阵 $A \in \mathbb{R}^{m \times n}$, $m \geq n$, 则存在唯一的列向量为正交单位向量组的 $Q \in \mathbb{R}^{m \times n}$, 主对角元都为正数的上三角矩阵 $R \in \mathbb{R}^{n \times n}$, 使得 $A = QR$.
2. 此外, 在标准内积下, 对任何 $\beta \in \mathbb{R}^n$, $R^{-1}Q^T\beta$ 是矩阵方程 $A^T Ax = A^T \beta$ 的唯一解.
3. 设矩阵 $A \in \mathbb{R}^{m \times n}$, $m \geq n$, 则存在列向量为正交单位向量组的 $Q \in \mathbb{R}^{m \times n}$, 主对角元都为非负数的上三角矩阵 $R \in \mathbb{R}^{n \times n}$, 使得 $A = QR$.



笔记 本结果对复列满秩矩阵, 则正交矩阵改为酉矩阵, 实上三角矩阵改为复上三角矩阵之后命题也对. 此外注意这里是任一个内积而不只是标准内积.

证明 不妨设就是标准内积, 一般情况可类似讨论.

1. 先证明存在性: 设 $A = (\alpha_1 \alpha_2 \cdots \alpha_n)$. 因为 A 列满秩, 所以对 A 列向量组做施密特正交化得

$$\begin{cases} \beta_1 = \alpha_1 \\ \beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} \beta_1 \\ \beta_3 = \alpha_3 - \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)} \beta_1 - \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)} \beta_2 \\ \vdots \end{cases}$$

于是有

$$\begin{aligned} (\alpha_1 \alpha_2 \cdots \alpha_n) &= \left(\frac{\beta_1}{\|\beta_1\|} \quad \frac{\beta_2}{\|\beta_2\|} \quad \cdots \quad \frac{\beta_n}{\|\beta_n\|} \right) \begin{pmatrix} \|\beta_1\| & \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} & \cdots & \frac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)} \\ & \|\beta_2\| & \cdots & \frac{(\alpha_n, \beta_2)}{(\beta_2, \beta_2)} \\ & & \ddots & \vdots \\ & & & \|\beta_n\| \end{pmatrix} \\ &= \left(\frac{\beta_1}{\|\beta_1\|} \quad \frac{\beta_2}{\|\beta_2\|} \quad \cdots \quad \frac{\beta_n}{\|\beta_n\|} \right) R = QR. \end{aligned}$$

再证明唯一性: 若还有另外一种分解 $A = QR = Q'R'$, 则

$$Q = Q' \cdot R'R^{-1} = Q' \cdot T,$$

这里 T 是主对角元都为正数的上三角矩阵. 从而 Q, Q' 的列向量组等价, 于是 Q, Q' 的列向量张成空间的是一致的, 并考虑这个空间的正交补空间, 即得

$$(Q \tilde{Q}) = (Q' \tilde{Q}) \begin{pmatrix} T & 0 \\ 0 & E_{m-n} \end{pmatrix}$$

且 $(Q \tilde{Q}), (Q' \tilde{Q})$ 都是 m 阶正交矩阵. 现在

$$(Q' \tilde{Q})^{-1}(Q \tilde{Q}) = \begin{pmatrix} T & 0 \\ 0 & E_{m-n} \end{pmatrix}$$

且左边是正交矩阵. 直接计算矩阵列向量的内积我们知道上三角的正交矩阵一定是单位矩阵. 因此

$$(Q' \tilde{Q})^{-1}(Q \tilde{Q}) = E_m \implies \begin{cases} Q' = Q \\ T = E_n \end{cases} \implies \begin{cases} Q' = Q \\ R' = R \end{cases},$$

这就证明了第一问.

2. 注意到

$$r(A) = r(A^T) \stackrel{\text{命题??(2)}}{\geq} r\left(A^T \begin{pmatrix} A & \beta \end{pmatrix}\right) = r\left(\begin{pmatrix} A^T A & A^T \beta \end{pmatrix}\right) \geq r(A^T A) \stackrel{\text{命题??}}{=} r(A),$$

我们知道 $r\left(\begin{pmatrix} A^T A & A^T \beta \end{pmatrix}\right) = r(A^T A)$, 从而线性方程 $A^T Ax = A^T \beta$ 有唯一解. 显然

$$A^T AR^{-1}Q^T\beta = A^T QRR^{-1}Q^T\beta = A^T \beta,$$

这就完成了第二问的证明.

3. 第三问和第一问整体是类似的, 但需要细微的修改施密特正交化过程. 事实上, 设 $A = (\alpha_1 \alpha_2 \cdots \alpha_n)$, 定义

$$\begin{cases} \beta_1 = \frac{\alpha_1}{\|\alpha_1\|}, & \alpha_1 \neq 0 \\ \beta_1 = 0, & \alpha_1 = 0 \end{cases},$$

然后假设 $\beta_1, \beta_2, \dots, \beta_{k-1}$ 已定义好. 我们定义

$$\beta_k = \begin{cases} \frac{\alpha_k - \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j}{\left\| \alpha_k - \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j \right\|}, & \alpha_k - \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j \neq 0 \\ 0, & \alpha_k - \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j = 0 \end{cases}.$$

于是有

$$\alpha_k = \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j + \left\| \alpha_k - \sum_{j=1}^{k-1} (\alpha_k, \beta_j) \beta_j \right\| \beta_k, k = 1, 2, \dots, n,$$

这给出了一个矩阵表示

$$A = (\beta_1, \beta_2, \dots, \beta_n) R. \quad (4)$$

将 $\beta_1, \beta_2, \dots, \beta_n$ 中不为 0 的全部拿出来扩充为 n 个标准正交单位向量, 并把为 0 的 β_k 替换成扩充时加入的新向量. 此时得标准正交向量组 $\widetilde{\beta}_1, \widetilde{\beta}_2, \dots, \widetilde{\beta}_n$ 且等式 (4) 变成

$$A = (\widetilde{\beta}_1, \widetilde{\beta}_2, \dots, \widetilde{\beta}_n) R = QR,$$

这里 R 没有变换. 这就完成了证明.

我们知道正定矩阵 A 总能做 $A = C^T C$. 我们现在可以给 C 一个更好的刻画. □

定理 0.10 (正定矩阵分解为上三角)

1. 设 $A \in \mathbb{R}^{n \times n}$ 是正定矩阵, 则存在对角元都为正数的可逆上三角矩阵 $C \in \mathbb{R}^{n \times n}$ 使得 $A = C^T C$.
2. 设 $A \in \mathbb{R}^{n \times n}$ 是半正定矩阵, 则存在对角元都为非负数的上三角矩阵 $C \in \mathbb{R}^{n \times n}$ 使得 $A = C^T C$.



笔记 本结果对 Hermite 正定矩阵也对, 此时 $C \in \mathbb{C}^{n \times n}$.

证明

1. 由 A 正定, 存在可逆 $D \in \mathbb{R}^{n \times n}$ 使得 $A = D^T D$. 由定理 23.10, 我们知道存在实 n 阶正交矩阵 T 和主对角元都为正数的实 n 阶上三角矩阵 C 使得 $D = TC$, 于是

$$A = C^T T^T T C = C^T C,$$

这就完成了证明.

2. 由 A 半正定, 存在 $D \in \mathbb{R}^{n \times n}$ 使得 $A = D^T D$. 由定理 0.9, 我们知道存在实 n 阶正交矩阵 T 和主对角元都为非负数的实 n 阶上三角矩阵 C 使得 $D = TC$, 于是

$$A = C^T T^T T C = C^T C,$$

这就完成了证明. □

定理 0.11

设 $A \in \mathbb{C}^{n \times n}$ 满足 $A\bar{A} = I_n$, 则存在可逆复矩阵 $B \in \mathbb{C}^{n \times n}$, 使得 $A = B(\bar{B})^{-1}$. ♥

证明 首先设 A 的所有单位圆周上互不相同的特征值为

$$-e^{i\theta_j}, \theta_j \in [0, 2\pi), j = 1, 2, \dots, s.$$

取 $\theta \in [0, 2\pi)$ 使得 $\theta \neq \frac{\theta_j}{2} + k\pi, j = 0, 1, 2, \dots, s, k \in \mathbb{Z}$.

现在考虑 $B = e^{i\theta}I + e^{-i\theta}A$, 则对任何 A 的特征值 λ , 若 $e^{i\theta} + e^{-i\theta}\lambda = 0$, 则 $e^{i2\theta} = -\lambda$, 两边取模知 $\lambda = -e^{i2\theta}, j \in \{1, 2, \dots, s\}$, 所以存在 $k \in \mathbb{Z}$, 使得 $\theta = \frac{\theta_j}{2} + k\pi$, 这和 θ 定义矛盾! 因此 B 可逆, 此时

$$A\bar{B} = A(e^{-i\theta}I + e^{i\theta}\bar{A}) = e^{i\theta}I + e^{-i\theta}A = B,$$

因此我们证明了存在可逆复矩阵 B , 使得 $A = B(\bar{B})^{-1}$. □

命题 0.2

设 $A \in \mathbb{C}^{n \times n}$ 是对称酉矩阵, 证明存在可逆矩阵 $Q \in \mathbb{C}^{n \times n}$ 使得 $A = \bar{Q}Q^{-1}$. ♠

证明 注意到 \bar{A} 满足定理 0.11, 所以存在可逆复矩阵 $Q \in \mathbb{C}^{n \times n}$, 使得 $A = Q(\bar{Q})^{-1}$, 从而 $A = \bar{Q}Q^{-1}$. □

定理 0.12

设 A 是 n 阶可逆复矩阵, 则 $A^{-1} \sim A^*$ 的充要条件是存在可逆 n 阶复矩阵 B 使得 $A = B^{-1}B^*$. ♡

证明 充分性: 若存在可逆 n 阶复矩阵 B 使得 $A = B^{-1}B^*$, 则

$$A^{-1} = (B^{-1})^*B \sim B^*(B^{-1})^*B(B^{-1})^* = B(B^{-1})^* = A^*.$$

必要性: 设可逆 n 阶复矩阵 P 使得 $A^{-1} = P^{-1}A^*P$. 因为 $|A^*P + \lambda^2P^*|$ 在 $\lambda = 0$ 不为 0, 所以不是 0 多项式, 所以 $|A^*P + \lambda^2P^*|$ 在单位圆周上不能处处为 0, 因此设

$$|A^*P + \lambda_0^2P^*| \neq 0, |\lambda_0| = 1, \lambda_0 \in \mathbb{C},$$


于是利用 $\lambda_0^{-1} = \bar{\lambda}_0$, 我们有

$$\begin{aligned} & (\lambda_0^{-1}A^*P + \lambda_0P^*)^{-1}(\lambda_0^{-1}A^*P + \lambda_0P^*)^* = A \\ & \Leftrightarrow \lambda_0P^*A + \bar{\lambda}_0P = \lambda_0^{-1}A^*PA + \lambda_0P^*A \\ & \Leftrightarrow \lambda_0P^*A + \bar{\lambda}_0P = \lambda_0^{-1}PA^{-1}A + \lambda_0P^*A \\ & \Leftrightarrow \lambda_0P^*A + \bar{\lambda}_0P = \lambda_0^{-1}P + \lambda_0P^*A, \end{aligned}$$

故可取 $B \triangleq \lambda_0^{-1}A^*P + \lambda_0P^*$. □

命题 0.3

任何复矩阵都能分解为两个对称矩阵之积且其中一个可逆. ♠

 **笔记** 证明的想法即利用使得 Jordan 转置的矩阵.

证明 对复矩阵 $A \in \mathbb{C}^{n \times n}$, 取可逆复矩阵 $P \in \mathbb{C}^{n \times n}$ 使得 $A = PJP^{-1}$, J 是 A 的 Jordan 标准型. 注意到若 $J = H_1H_2$ 使得 $H_1, H_2 \in \mathbb{C}^{n \times n}$ 且对称, 并且其中一个可逆, 则

$$A = PH_1P^T \cdot (P^T)^{-1}H_2P^{-1}$$

为所求分解.

我们设

$$J = \text{diag}\{J_1, J_2, \dots, J_s\}, J_i \text{ 是一个 Jordan 块}, i = 1, 2, \dots, s,$$

则回忆矩阵 H (引理??), 我们有

$$\begin{aligned} J &= \text{diag}\{J_1, J_2, \dots, J_s\} = \text{diag}\{HHJ_1, HHJ_2, \dots, HHJ_s\} \\ &= \text{diag}\{H, H, \dots, H\} \cdot \text{diag}\{HJ_1, HJ_2, \dots, HJ_s\}. \end{aligned}$$

注意到 $\text{diag}\{H, H, \dots, H\}$ 可逆以及

$$(\text{diag}\{HJ_1, HJ_2, \dots, HJ_s\})^T = \text{diag}\{J_1^T H, J_2^T H, \dots, J_s^T H\} = \text{diag}\{HJ_1, HJ_2, \dots, HJ_s\},$$

我们就完成了证明. \square

定理 0.13

设域 \mathbb{F} 上对称 n 阶矩阵 A 顺序主子式全不为 0, 那么存在 \mathbb{F} 上主对角元全为 1 的 n 阶上三角矩阵 B , 主对角元全不为 0 的对角矩阵 D , 使得 $A = B^T D B$, 并且这种分解是唯一的. \heartsuit

证明 存在性: 当 $n = 1$ 命题显然成立, 假设命题对 $n - 1$ 成立, 当 n 时, 设

$$A = \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & c \end{pmatrix}, A_{n-1} \in \mathbb{F}^{(n-1) \times (n-1)}, c \in \mathbb{F}, \alpha \in \mathbb{F}^{n-1},$$

且 A_{n-1} 是对称矩阵. 注意到

$$\begin{pmatrix} E_{n-1} & 0 \\ -\alpha^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \alpha \\ \alpha^T & c \end{pmatrix} \begin{pmatrix} E_{n-1} & -A_{n-1}^{-1} \alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} A_{n-1} & 0 \\ 0 & c - \alpha^T A_{n-1}^{-1} \alpha \end{pmatrix}.$$

由 $\det A \neq 0$ 知 $c - \alpha^T A_{n-1}^{-1} \alpha \neq 0$. 利用归纳假设, 我们有

$$A_{n-1} = K^T D_{n-1} K, K, D_{n-1} \in \mathbb{F}^{(n-1) \times (n-1)},$$

且 K 是主对角元全为 1 的上三角矩阵, D_{n-1} 是主对角元不为 0 的对角矩阵. 于是

$$\begin{aligned} \begin{pmatrix} A_{n-1} & \alpha \\ \beta^T & c \end{pmatrix} &= \begin{pmatrix} E_{n-1} & 0 \\ -\alpha^T A_{n-1}^{-1} & 1 \end{pmatrix}^{-1} \begin{pmatrix} K^T D_{n-1} K & 0 \\ 0 & c - \alpha^T A_{n-1}^{-1} \alpha \end{pmatrix} \begin{pmatrix} E_{n-1} & -A_{n-1}^{-1} \alpha \\ 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} E_{n-1} & 0 \\ \alpha^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} K^T D_{n-1} K & 0 \\ 0 & c - \alpha^T A_{n-1}^{-1} \alpha \end{pmatrix} \begin{pmatrix} E_{n-1} & A_{n-1}^{-1} \alpha \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} E_{n-1} & 0 \\ \alpha^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} K^T & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D_{n-1} & 0 \\ 0 & c - \alpha^T A_{n-1}^{-1} \alpha \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E_{n-1} & A_{n-1}^{-1} \alpha \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

取

$$B = \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} E_{n-1} & A_{n-1}^{-1} \alpha \\ 0 & 1 \end{pmatrix}, D = \begin{pmatrix} D_{n-1} & 0 \\ 0 & c - \alpha^T A_{n-1}^{-1} \alpha \end{pmatrix},$$

则有 $A = B^T D B$, 这就证明了存在性.

唯一性: 若还有满足条件的分解 $A = B_2^T D_2 B_2$, 则

$$B^T D B = B_2^T D_2 B_2 \implies (B_2^T)^{-1} B^T D = D_2 B_2 B^{-1}.$$

又 $(B_2^T)^{-1} B^T$ 主对角元全为 1 的下三角矩阵, $B_2 B^{-1}$ 主对角元全为 1 的上三角矩阵. 设

$$(B_2^T)^{-1} B^T = (b_{ij}), B_2 B^{-1} = (b'_{ij}), D = \text{diag}\{a_1, a_2, \dots, a_n\}, D_2 = \text{diag}\{a'_1, a'_2, \dots, a'_n\},$$

则

$$b_{ij} a_j = a_i b'_{ij}, \forall 1 \leq i, j \leq n.$$

当 $i < j$, 我们有 $0 = a_i b'_{ij} \implies b'_{ij} = 0$, 故 $B_2 B^{-1}$ 只能为对角矩阵, 因此只能有

$$B_2 B^{-1} = E \implies B_2 = B \implies D = D_2,$$

这就证明了唯一性. \square

定理 0.14 (CS 分解)

设 $Q_1 \in \mathbb{R}^{m_1 \times n_1}, Q_2 \in \mathbb{R}^{m_2 \times n_1}, Q_1^T Q_1 + Q_2^T Q_2 = I_{n_1}$, 这里 $\min\{m_1, m_2\} \geq n_1$. 则存在正交矩阵 $V_1 \in \mathbb{R}^{n_1 \times n_1}, U_1 \in$

$\mathbb{R}^{m_1 \times m_1}, U_2 \in \mathbb{R}^{m_2 \times m_2}$ 和 $0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \theta_{n_1} \leq \frac{\pi}{2}$ 使得

$$\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}^T \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} V_1 = \begin{pmatrix} C \\ S \end{pmatrix}, \quad (5)$$

这里 $C \in \mathbb{R}^{m_1 \times n_1}, S \in \mathbb{R}^{m_2 \times n_1}$. 且

$$C = \begin{pmatrix} \cos \theta_1 & & & \\ & \cos \theta_2 & & \\ & & \ddots & \\ & & & \cos \theta_{n_1} \end{pmatrix}, S = \begin{pmatrix} \sin \theta_1 & & & \\ & \sin \theta_2 & & \\ & & \ddots & \\ & & & \sin \theta_{n_1} \end{pmatrix}.$$

♡

注 我们把上述 C, S 的 \cos, \sin 分别称为 C, S 的对角线, 并记作 (即使不是方阵)

$$C = \text{diag}\{\cos \theta_1, \dots, \cos \theta_{n_1}\}, S = \text{diag}\{\sin \theta_1, \dots, \sin \theta_{n_1}\}.$$

证明 由 $Q_1^T Q_1 + Q_2^T Q_2 = I_{n_1}$ 知 Q_1, Q_2 的奇异值分解的非 0 元都在 $[0, 1]$ 之间. 假设

$$C \triangleq U_1^T Q_1 V_1 = \begin{pmatrix} I_t & 0 \\ 0 & \Sigma \end{pmatrix}.$$

再设 $U_1^T Q_1 V_1$ 的对角线为

$$1 = c_1 = c_2 = \dots = c_t > c_{t+1} \geq \dots \geq c_{n_1} \geq 0.$$

再设 $Q_2 V_1 = (W_1 \ W_2), W_1 \in \mathbb{R}^{m_2 \times t}$, 我们有

$$\begin{pmatrix} U_1 & 0 \\ 0 & I_{m_2} \end{pmatrix}^T \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} V_1 = \begin{pmatrix} U_1^T Q_1 V_1 \\ Q_2 V_1 \end{pmatrix} = \begin{pmatrix} I_t & 0 \\ 0 & \Sigma \\ W_1 & W_2 \end{pmatrix}.$$

注意到

$$\begin{pmatrix} U_1^T Q_1 V_1 \\ Q_2 V_1 \end{pmatrix}^T \begin{pmatrix} U_1^T Q_1 V_1 \\ Q_2 V_1 \end{pmatrix} = V_1^T Q_1^T Q_1 V_1 + V_1^T Q_2^T Q_2 V_1 = I_{n_1},$$

即 $\begin{pmatrix} I_t & 0 \\ 0 & \Sigma \\ W_1 & W_2 \end{pmatrix}$ 的列向量应该是单位向量, 从而 $W_1 = 0$ 且 W_2 的列非 0 并满足

$$W_2^T W_2 = I_{n_1-t} - \Sigma^T \Sigma = \text{diag}\{1 - c_{t+1}^2, \dots, 1 - c_{n_1}^2\}$$

是可逆矩阵.

设 $s_k = \sqrt{1 - c_k^2}, k = 1, 2, \dots, n_1$. 然后矩阵 $Z = W_2 \text{diag}\left\{\frac{1}{s_{t+1}}, \dots, \frac{1}{s_{n_1}}\right\} = W_2 D$ 的列向量是正交规范向量组, 因此我们可以把 Z 扩充为正交矩阵 $U_2 = (Z \ Z') \in \mathbb{R}^{m_2 \times m_2}$. 现在注意到

$$(Z')^T Z = 0 \implies (Z')^T W_2 D = 0 \implies (Z')^T W_2 = 0,$$

于是我们有

$$\begin{aligned} S &\triangleq U_2^T Q_2 V_1 = (0, U_2^T W_2) = \begin{pmatrix} 0 & D^T W_2^T W_2 \\ 0 & (Z')^T W_2 \end{pmatrix} \\ &= \begin{pmatrix} 0 & D^T W_2^T W_2 \\ 0 & 0 \end{pmatrix} = \text{diag}\{s_1, \dots, s_{n_1}\} \end{aligned}$$

满足要求, 于是我们完成了证明. □