Gram-Schmidt 标准正交化

定理. 设 $A=(\alpha_1,\cdots,\alpha_n)\in GL(n,\mathbb{R})$

则
$$\begin{cases} \beta_k=\alpha_k-\sum_{i=1}^{k-1}(\gamma_i^T\alpha_k)\gamma_i=\alpha_k-\sum_{i=1}^{k-1}\frac{\beta_i^T\alpha_k}{\beta_i^T\beta_i}\beta_i \\ \gamma_k=\frac{1}{\|\beta_k\|}\beta_k \end{cases}$$

$Q=(\gamma_1,\cdots,\gamma_n)$ 为正交方阵 ($O(n)$)

且有 A的 QR 分解

$$(\alpha_1\cdots\alpha_n)=(\gamma_1\cdots\gamma_n)\begin{pmatrix} \|\beta_1\| & \gamma_1^T\alpha_2 & \cdots & \gamma_1^T\alpha_n \\ & \|\beta_2\| & \cdots & \gamma_2^T\alpha_n \\ & & \ddots & \vdots \\ & & & \|\beta_n\| \end{pmatrix}$$

注1. 对 $A\in\mathbb{C}^{n\times n}$ 也有类似操作

注2. G-S 正交化 实际说明
在内积空间中, 基 $\xrightarrow{\text{G-S 正交化}}$ 标准正交基.

正交方阵. 酉方阵

$AA^T = I_n.$ $\quad$ $AA^H = I_n$

$(A \in \mathbb{R}^{n\times n})$ $\quad$ $A \in \mathbb{C}^{n\times n}$

· 正交相似.

(并未假定 A 对称)

定理. 设 $A \in \mathbb{R}^{n\times n}$. 其特征值为 $\lambda_1, \cdots, \lambda_n \in \mathbb{C}$.

且 $1 \le k \le s$ 时 $\lambda_{2k} = \overline{\lambda_{2k-1}} \notin \mathbb{R}$

$2s+1 \le k \le n$ 时 $\lambda_k \in \mathbb{R}$

则 A 可正交相似成如下方阵

$$\begin{pmatrix} A_1 & * & \cdots & * \\ & A_2 & & \\ & & \ddots & \\ & & A_s & * & \cdots \\ & & & \lambda_{2s+1} & \\ & & & & \ddots \\ & & & & & \lambda_n \end{pmatrix}$$

其中 A_k 为 2 阶方阵. $(1 \le k \le s)$

特征值为 $\lambda_{2k}, \lambda_{2k-1}$

A_k 形如 $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$

证明. $n=1$ 显然, . 下设 $n \ge 2$

Case 1. $s=0$ 即. 无 $\lambda \in \mathbb{C} \setminus \mathbb{R}$

设 λ_1. α_1 $\quad \|\alpha_1\| = 1$

将 α_1 扩充为正交方阵 P_1

则 $A P_1 = P_1 \begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix}$ $\quad$ B 的特征值为 $\lambda_2, \cdots, \lambda_n$

由归纳假设 $\exists P_2 \in O(n-1)$

$$P_2^{-1} B P_2 = \begin{pmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

故 $P = P_1 \begin{pmatrix} 1 & \\ & P_2 \end{pmatrix}$ 即为所求

Case 2 Sto. 设 $\lambda_1 = a_1 + b_1 i$ 对应特征向量 $u + iv$

则 $A(u \ v) = (u \ v)\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$

且 u, v 线性无关

将 u, v 标准正交化为 α_1, α_2

$(|\alpha_1| = |\alpha_2| = 1, \ \alpha_1^T\alpha_2 = 0)$

则 $A(\alpha_1 \ \alpha_2) = (\alpha_1 \ \alpha_2)\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}$

再扩充为正交阵 P_1

$\Rightarrow A P_1 = P_1 \begin{pmatrix} A_1 & * \\ 0 & B \end{pmatrix}$

由归纳假设 $P_2^{-1} B P_2 = \begin{pmatrix} A_2 & & \\ & \ddots & \\ & & A_{r} \end{pmatrix}$

$\Rightarrow P = P_1 \begin{pmatrix} I_2 & \\ & P_2 \end{pmatrix}$ 即为所求 $\square$

由上述定理可得：

定理 正交方阵可正交相似成

$$\mathrm{diag}\left(\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}, \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \cdots \begin{pmatrix} & \\ & \end{pmatrix}, I, -I \right)$$

证明. 这是因为 $\forall$ A 的特征值 $\lambda \in \mathbb{C}$, $|\lambda| = 1$ $\square$

☆ 定理. $\forall$ 对称实数方阵 可正交相似成 对角阵.

证明. 设 $P^TAP = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = P^TA^TP = (P^TAP)^T$
(P 正交阵)

$\Rightarrow * = 0 \Rightarrow$ 对角 □

定理 反对称方阵 可正交相似成

$$\mathrm{diag}\left(\begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & b_s \\ -b_s & 0 \end{pmatrix}, O \right)$$

证明 $P^TAP = \begin{pmatrix} A_1 & & & & * \\ & \ddots & & & \\ & & A_s & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$

$(P^TAP)^T = -P^TAP \Rightarrow \begin{cases} * = 0 \\ A_i \text{ 为反对称} \\ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \to a = 0 \end{cases}$ □

定义. $A \in \mathbb{R}^{n\times n}$ 若 $AA^T = A^TA$, 则称 A 规范方阵.

定理. $\forall$ 规范实方阵 可正交相似成 准对角

$$\mathrm{diag}\left(\begin{pmatrix} a_1 & b_1 \\ -b_1 & a_1 \end{pmatrix}, \cdots, \begin{pmatrix} a_s & b_s \\ -b_s & a_s \end{pmatrix}, \lambda_{2s+1}, \cdots, \lambda_n \right)$$

证明. 设 $P^TAP = \begin{pmatrix} A_1 & & & & & * \\ & A_2 & & & & \\ & & \ddots & & & \\ & & & A_s & & \\ & & & & \lambda_{2s+1} & \\ & & & & & \ddots & \\ & & & & & & \lambda_n \end{pmatrix} = B$. 由 A 规范 $\Rightarrow$ B 规范

设 $B=\begin{pmatrix} B_{11} & B_{12} \\ & B_{22} \end{pmatrix}$，$B^TB=BB^T$

$$\Rightarrow \begin{pmatrix} B_{11}^T B_{11} & B_{11}^T B_{12} \\ B_{12}^T B_{11} & B_{12}^T B_{12}+B_{22}^T B_{22} \end{pmatrix} = \begin{pmatrix} B_{11}B_{11}^T+B_{12}B_{12}^T & B_{12}B_{22}^T \\ B_{22}B_{12}^T & B_{22}B_{22}^T \end{pmatrix}$$

$$\Rightarrow \begin{cases} B_{11}^T B_{11} = B_{11}B_{11}^T + B_{12}B_{12}^T \\ B_{12}^T B_{12} + B_{22}^T B_{22} = B_{22}B_{22}^T \end{cases}$$

两边取 trace $\Rightarrow B_{12}=0 \Rightarrow \begin{cases} B_{11}^T B_{11} = B_{11}B_{11}^T \\ B_{22}^T B_{22} = B_{22}B_{22}^T \end{cases}$

$B_{11}=\begin{pmatrix} A_1 & & A_{ij} \\ & A_2 & \\ & & \ddots \\ & & & A_s \end{pmatrix}$. A_{ij} 2阶方块 $\quad A_i$ 记成 A_{ii}

则 $\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ & \ddots & & \\ & & & A_{ss} \end{pmatrix}\begin{pmatrix} A_{11}^T & & & \\ A_{12}^T & A_{22}^T & & \\ \vdots & & \ddots & \\ A_{1s}^T & & & A_{ss}^T \end{pmatrix} = \begin{pmatrix} A_{11}^T & & \\ & \ddots & \\ * & & A_{ss}^T \end{pmatrix}\begin{pmatrix} \ddots & * \\ & \ddots \end{pmatrix}$

逐次考虑两边 (i,i) 处的值即可.

($i=1$, $(左)_{11} = A_{11}A_{11}^T + A_{12}A_{12}^T + \cdots + A_{1s}A_{1s}^T$

$(右)_{11} = A_{11}^T A_{11}$ 两边 trace $\Rightarrow A_{1j}=0$ $\; j\geq 2$

同理考虑 $i=2$, 利用 $A_{1j}=0 \Rightarrow A_{2j}=0 \quad j\geq 3$

$\Rightarrow \cdots$)

则 B_{11} 为准对角阵.

同理, B_{22} 为对角阵.

$\square$

- 正交相抵.

- 定理. $\forall A \in \mathbb{R}^{m\times n}$, $\exists$ 正交方阵 P, Q. s.t.

$$A = P\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}Q. \qquad \Sigma = \mathrm{diag}(\sigma_1,\cdots,\sigma_r) \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

$$r = \mathrm{rank}\, A.$$

证明. 由 A^TA 对称. 故 $\exists$ 正交阵 Q

$A^TA = Q^T \mathrm{diag}(\lambda_1,\cdots,\lambda_n)\, Q$. 记 $Q^{-1} = (\alpha_1,\cdots,\alpha_n)$

由 A^TA 半正定. $\lambda_i \geq 0$ 不妨 $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$

设 $AQ^{-1} = (\beta_1 \cdots \beta_n)$ 则由 $(AQ^{-1})^T AQ^{-1} = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$

$\Rightarrow$ $\beta_1,\cdots,\beta_r$ 两两正交 且 $\|\beta_k\| = \sqrt{\lambda_k} =: \sigma_k \quad 1\leq k\leq r$

$\beta_k = 0 \quad k > r$

将 $\sigma_1^{-1}\beta_1,\cdots,\sigma_r^{-1}\beta_r$ 扩充为标准正交阵 P

则 $AQ^{-1} = (\beta_1,\cdots,\beta_r \ \ 0) = (\sigma_1^{-1}\beta_1,\cdots,\sigma_r^{-1}\beta_r, *)\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}$

$= P\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}$ $\square$

注意: $\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}$ 由 A 唯一决定, 但 P, Q 一般不唯一.

作业答案.(部分).

4. 证明：n 阶可逆复对称方阵在复数域上相合于

$$\begin{pmatrix} O & I_{(m)} \\ I_{(m)} & O \end{pmatrix}（当 n=2m） 或 \begin{pmatrix} O & I_{(m)} & \\ I_{(m)} & O & \\ & & 1 \end{pmatrix}（当 n=2m+1）.$$

证明. $\forall$ 复对称可逆矩阵 A

其相合标准形为 I_n.

故对

$\begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}$ 及 $\begin{pmatrix} 0 & I_m & \\ I_m & 0 & \\ & & 1 \end{pmatrix}$ 来说.

由于均可逆. $\Rightarrow$ 也相合于 I_n

$\Rightarrow$ $A \underset{相合}{\sim} I_n \overset{相合}{\sim} \begin{pmatrix} 0 & I_m \\ I_m & 0 \end{pmatrix}$ $n=2m$

或 $\begin{pmatrix} 0 & I_m & \\ I_m & 0 & \\ & & 1 \end{pmatrix}$ $n=2m+1$ □

6. 设 A 是 n 阶实对称矩阵，且 $\det A<0$，证明：必存在 n 维向量 $X\neq 0$，使 $X^TAX<0$.

证明. A 实对称.阵

相合于 $\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}$

$A = P^T\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix} P$ P 可逆

由 $\det A<0 \Rightarrow (\det P)^2 \det\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}<0$

$\Rightarrow$ q 为奇. 无 0. 特别地 $q\geq 1$

即 $A= P^T\begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} P$

取 $X = P^{-1} e_n$ $\left(e_n=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right)$

则 $X^T A X = e_n^T \begin{pmatrix} I_p & \\ & -I_q \end{pmatrix} e_n = -1$.即为所求

6. 求实函数 $f(x)=xSx^{T}+2\beta x^{T}+b$ 的最大值或最小值，其中 S 是 n 阶正定实对称方阵，$\beta=(b_1,b_2,\cdots,b_n)$ 与 $x=(x_1,x_2,\cdots,x_n)$ 是 n 维实行向量，b 是实数.

解. 写成矩阵形式

$$f(x)=\begin{pmatrix} x & 1\end{pmatrix}\begin{pmatrix} S & \beta^T \\ \beta & b\end{pmatrix}\begin{pmatrix} x^T \\ 1\end{pmatrix}$$

注意到 S 正定. 记 $S=P^TP$ （P 可逆）

则
$$\begin{pmatrix} I_n & 0 \\ -(P^TP)^{-1}\beta^T & 1\end{pmatrix}\begin{pmatrix} P^TP & \beta^T \\ \beta & b\end{pmatrix}\begin{pmatrix} I_n & -(P^TP)^{-1}\beta^T \\ 0 & 1\end{pmatrix}$$

$$=\begin{pmatrix} P^TP & \\ & b-\beta S^{-1}\beta^T\end{pmatrix}$$

则令
$$(y_1,\ y_2)=\begin{pmatrix} x & 1\end{pmatrix}\begin{pmatrix} I_n & \\ -(P^TP)^{-1}\beta^T & 1\end{pmatrix}^{-1}$$

$$=\begin{pmatrix} x & 1\end{pmatrix}\begin{pmatrix} I & \\ S^{-1}\beta^T & 1\end{pmatrix}$$

$\Rightarrow$ $y_2=1$. y_1 为 n 维行向量，与 x 一一对应.

$$\Rightarrow f(x)=\begin{pmatrix} y_1 & 1\end{pmatrix}\begin{pmatrix} S & \\ & b-\beta S^{-1}\beta^T\end{pmatrix}\begin{pmatrix} y_1^T \\ 1\end{pmatrix}$$

$$=y_1Sy_1^T+b-\beta S^{-1}\beta^T \geqslant b-\beta S^{-1}\beta^T$$

取 $y_1=0$ 可求出对应的 x_0.

此时 $f(x_0)=b-\beta S^{-1}\beta^T$ 为最小值.

无最大值.

□

10. 设 $\boldsymbol{A}$，$\boldsymbol{B}$ 均为 n 阶实对称正定矩阵，证明：如果 $\boldsymbol{A}-\boldsymbol{B}$ 正定，则 $\boldsymbol{B}^{-1}-\boldsymbol{A}^{-1}$ 亦正定.

证明. 由 A 正定. 设 $A=P^TP$

则 $A-B=P^T\left(I_n-(P^T)^{-1}BP^{-1}\right)P$.

记 $\hat{B}=(P^T)^{-1}BP^{-1}$

则 $A-B>0 \iff I_n-\hat{B}>0$　　(*)

而 $\hat{B}$ 与 B 相合 $\overset{B>0}{\Rightarrow}$ ∃ Q 正交阵

$\hat{B}=Q^T\operatorname{diag}(\lambda_1,\cdots,\lambda_n)Q$.　　$\lambda_i>0$

由 (*)　$I_n-\hat{B}=Q^T\left(I_n-\operatorname{diag}(\lambda_1,\cdots,\lambda_n)\right)Q$

$=Q^T\operatorname{diag}(1-\lambda_1,\cdots,1-\lambda_n)Q>0$

$\Rightarrow \forall i\ \ 0<\lambda_i<1$　　(**)

而 $B^{-1}-A^{-1}=B^{-1}-P^{-1}(P^T)^{-1}$

$=P^{-1}(\hat{B}^{-1}-I_n)(P^T)^{-1}$

$\overset{相合}{\sim}\ \hat{B}^{-1}-I_n$

$=Q^T\left(\operatorname{diag}(\lambda_1^{-1},\cdots,\lambda_n^{-1})-I_n\right)Q$

$\overset{相合}{\sim}\ \operatorname{diag}(\lambda_1^{-1}-1,\cdots,\lambda_n^{-1}-1)$

由于 $0<\lambda_i<1 \Rightarrow \lambda_i^{-1}-1>0 \Rightarrow$ 正定

□

注. 此题与 10.19. 课堂例题

若 $A\geqslant B>0$ 则 $\det A\geqslant\det B$ 证明思路类似.

3. 设 $S=(s_{ij})_{n\times n}$是正定实对称方阵. 求证:

$$Q(x_1,\cdots,x_n)=\begin{vmatrix} s_{11} & \cdots & s_{1n} & x_1 \\ \vdots & & \vdots & \vdots \\ s_{n1} & \cdots & s_{nn} & x_n \\ x_1 & \cdots & x_n & 0 \end{vmatrix}$$

是负定二次型.

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \det A \cdot \det(D - CA^{-1}D) \quad A \text{ 可逆.}$$

证明. $Q(x) = \begin{vmatrix} S & x \\ x^T & 0 \end{vmatrix} = \det S \cdot \underbrace{\det(0 - x^T S^{-1} x)}_{\in \mathbb{R}^{1\times 1}, \text{ 数}}$

$= -x^T S^{-1} x \cdot \det S$

故 $Q \le 0$. "=" 取到 $\Leftrightarrow$ $x=0$ □

5. $S_i\ (1\leqslant i\leqslant m)$是同阶实对称方阵. 求证: $S_1^2+\cdots+S_m^2=O \Leftrightarrow S_1=\cdots=S_m=O$.

证明. "$\Leftarrow$" 显然

"$\Rightarrow$" $S_1^2 = S_1^T S_1$

$\sum S_i^2 = 0 \Rightarrow \sum_{i=1}^{n} tr(S_i^T S_i) = 0 \quad (*)$

对 $\forall A = (a_{jk})_{n\times n}$ $tr(A^T A) = \sum_{j,k} a_{jk}^2 \geqslant 0$

设 $S_i = (a_{jk}^{(i)})_{j,k}$

则 $\sum_{i,j,k} (a_{jk}^{(i)})^2 = 0 \Rightarrow a_{jk}^{(i)} = 0 \ \forall i,j,k$

$\Rightarrow S_i = 0 \ \forall i$ □

6. 证明: 实方阵的每个奇异值都是特征值的充分必要条件是 $A\geqslant 0$.

证明 "$\Leftarrow$" 设 $A\geqslant 0$ 则 $\exists$ 正交方阵 Q

$A = Q^T \operatorname{diag}(\lambda_1,\cdots,\lambda_n) Q$

$\lambda_i \geqslant 0$ 为特征值.

$\Rightarrow A^T A = Q^T \operatorname{diag}(\lambda_1^2,\cdots,\lambda_n^2) Q$

奇异值为 $\sqrt{\lambda_i^2} = \lambda_i$ 为特征值.

"$\Rightarrow$" 设奇异值 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$

为特征值 ($\sigma_i = \lambda_i$).

则. 设 $A = P\ \mathrm{diag}(\sigma_1, \cdots, \sigma_n)\ Q$

P, Q 正交阵.

$$A^T A = Q^T \mathrm{diag}(\sigma_1^2, \cdots, \sigma_n^2)\ Q$$

再注意到 $\exists$ R 正交阵.

$$A = R^T \begin{pmatrix} \lambda_1 & \ddots & a_{ij} \\ 0 & \ddots & \lambda_n \end{pmatrix} R$$

则. $\mathrm{tr}(A^T A) = \mathrm{tr}\left(R^T (\quad)^2 R \right)$

$$= \sum_{i=1}^n \lambda_i^2 + \sum a_{ij}^2$$

又 $\mathrm{tr}(A^T A) = \mathrm{tr}\left(Q^T \mathrm{diag}(\lambda_1^2, \cdots, \lambda_n^2)\ Q \right)$

$$= \sum_{i=1}^n \lambda_i^2$$

$\Rightarrow$ $\sum a_{ij}^2 = 0$ $\Rightarrow$ $a_{ij} = 0$

$\Rightarrow$

$A = R^T\ \mathrm{diag}(\lambda_1, \cdots, \lambda_n)\ R$ 故 $A \geq 0$ $\square$

注. Schur 不等式: $A \in \mathbb{C}^{n\times n}$ 复方阵, λ_i 为特征值.

则 $\sum_{i=1}^n |\lambda_i|^2 \leq \mathrm{tr}(A^H A)$

"=" 取到 iff (当且仅当) A 为规范方阵.

本题可利用奇异值分解证明 Schur 不等式刚好取等. 再利用规范方阵标准型. 即可.

9.（M－P广义逆）设 A 是 $m\times n$ 复矩阵．如果 $n\times m$ 复矩阵 B 同时满足以下 4 个条件，就称 B 是 A 的一个 Morre－Penrose 广义逆，简称 M－P 广义逆：

（1）$ABA=A$；（2）$BAB=B$；（3）$(AB)^*=AB$；（4）$(BA)^*=BA$．

试通过以下步骤研究 Morre－Penrose 逆的存在性，唯一性及与解方程组的关系．

（1）选择 m 阶酉方阵 U_1 和 n 阶酉方阵 U_2 将 A 酉相抵到标准形

$$A_1=U_1AU_2=\begin{pmatrix}\mu_1 & & & \\ & \ddots & & \\ & & \mu_r & \\ & & & O\end{pmatrix}$$

其中 $\mu_1\geqslant\cdots\geqslant\mu_r>0$．令 $B_1=U_2^*BU_1^*$．求证：

B 是 A 的 M－P 广义逆 $\Leftrightarrow B_1$ 是 A_1 的 M－P 广义逆；

（2）证明：A_1 存在唯一的 M－P 广义逆，从而 A 存在唯一的 M－P 广义逆．我们将 A 的唯一的 M－P 广义逆记作 A^+；

（3）对任意 $\beta\in\mathbb{C}^{m\times1}$，求证：$X=A^+\beta$ 是线性方程组 $AX=\beta$ 的最小二乘解．如果最小二乘解不唯一，则 $X=A^+\beta$ 还是模 $|X|$ 最小的最小二乘解．

（4）设 $A\in F^{m\times n}$ 的秩为 r，则可写 $A=BC$ 使 $B\in F^{m\times r}$，$C\in F^{r\times n}$，$\operatorname{rank}B=\operatorname{rank}C=r$．已经知道 B^*B、CC^* 可逆．试验证

$$A^+=C^*(CC^*)^{-1}(B^*B)^{-1}B^*$$

满足 M－P 广义逆的 4 个条件．

（此处 A^* 指 $\bar{A}^T=A^H$）

解：(1) $A_1=U_1AU_2$，$B_2=U_2^HBU_1^H$

"$\Rightarrow$"，"$\Leftarrow$" 代入验证即可．

(2) 设 A 有两个广义逆 $B,\tilde{B}$

则 $B=BAB=(BA)^HB=A^HB^HB$

$=A^H\tilde{B}^H\underline{A^HB^H}B=A^H\tilde{B}^HBAB$

$=\tilde{B}AB=\tilde{B}B^HA^H=\tilde{B}B^HA^H\tilde{B}^HA^H$

$=\tilde{B}ABA\tilde{B}=\tilde{B}A\tilde{B}=\tilde{B}$．故唯一．

另证．$A_1=\begin{pmatrix}\mu_1 & & & \\ & \ddots & & \\ & & \mu_r & \\ & & & O\end{pmatrix}$ 设 $B_1=\begin{pmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{pmatrix}$

利用 $\begin{cases}A_1B_1A_1=A_1 & ①\\ B_1A_1B_1=B_1 & ②\\ (A_1B_1)^H=A_1B_1 & ③\\ (B_1A_1)^H=B_1A_1 & ④\end{cases}$

记 $\begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_r \end{pmatrix} = A_{11}$

由 ①、②、③、④，及 A_{11} 可逆，

可求出，$B_{11} = A_{11}^{-1}$，B_{12}，B_{21}，B_{22} 为 0

故 M-P 广义逆唯一

(3) 8.（复系数线性方程组的最小二乘解）设 $A \in \mathbf{C}^{m\times n}$，$\beta \in \mathbf{C}^{m\times 1}$，如果非齐次线性方程组 $AX=\beta$ 无解，我们求 $X_0 \in \mathbf{C}^{n\times 1}$ 使 $\det(AX_0-\beta)$ 最小，这样的 X_0 称为 $AX=\beta$ 的最小二乘解．求证：X_0 是 $AX=\beta$ 的最小二乘解 $\Leftrightarrow X_0$ 是方程组 $A^H AX = A^H\beta$ 的解．

证明：固定 X_0 为 $A^H A X = A^H \beta$ 的解.

$\forall\ X \in \mathbb{C}^{n\times 1}$

由于 $AX-\beta = A(X-X_0) + AX_0-\beta$

而 $(A(X-X_0))^H (AX_0-\beta)$

$= (X-X_0)^H A^H (AX_0-\beta)$

$= (X-X_0)^H (A^H A X_0 - A^H \beta) = 0$

故 $\|AX-\beta\|^2 = \|A(X-X_0)+AX_0-\beta\|^2$

$= [A(X-X_0)+AX_0-\beta]^H \cdot [A(X-X_0)+AX_0-\beta]$

$= \|A(X-X_0)\|^2 + \|AX_0-\beta\|^2$

$\geqslant \|AX_0-\beta\|^2$

"=" 取到当且仅当 $AX = AX_0$

故若 X 为最小二乘解 $\Leftrightarrow AX = AX_0$

下证 $AX = AX_0 \Leftrightarrow A^H A X = A^H \beta$

"$\Rightarrow$" $AX = AX_0 \Rightarrow A^H A X = A^H A X_0 = A^H \beta$

"$\Leftarrow$" 若 $A^H A X = A^H \beta$

而 $A^H A X_0 = A^H \beta$

$\Rightarrow A^H A(X-X_0) = 0$

$$\Rightarrow \quad (X-X_0)^H A^H A(X-X_0) = 0$$

即 $\|A(X-X_0)\|^2 = 0$

$$\Rightarrow \quad AX = AX_0 \qquad \square$$

回到 (3) 要证 $A^+\beta$ 为最小二乘.

即验证 $A^H A A^+ \beta = A^H \beta$

LHS (左式) $= A^H A A^+ \beta$

$$= A^H (AA^+)^H \beta$$

$$= A^H (A^+)^H A^H \beta = (AA^+A)^H \beta = A^H\beta$$

最小性, $\forall\ Y$, $A^H A Y = A^H \beta$

要证 $\|A^+\beta\| \le \|Y\|$

而 由 $A^H A A^+ \beta = A^H \beta$

$$\Rightarrow \quad A^H A (Y - A^+\beta) = 0$$

$$\Rightarrow \quad (Y - A^+\beta)^H A^H A (Y - A^+\beta) = 0$$

$$\Rightarrow \quad \|A(Y - A^+\beta)\|^2 = 0$$

$$\Rightarrow \quad AY = AA^+\beta$$

则 $\|Y\|^2 = \|Y - A^+\beta + A^+\beta\|^2$

$$= \|Y - A^+\beta\|^2 + (Y - A^+\beta)^H A^+\beta$$

$$+ (A^+\beta)^H (Y - A^+\beta) + \|A^+\beta\|^2$$

由于 $(Y - A^+\beta)^H (A^+\beta)$

$$= Y^H A^+\beta - \|A^+\beta\|^2$$

而 $Y^H A^+\beta = Y^H A^+ A A^+ \beta$

$$= Y^H (A^+A)^H A^+\beta$$

$$= \underbrace{Y^H A^H}(A^+)^H A^+\beta$$

$$= \beta^H (A^+)^H A^H (A^+)^H A^+\beta$$

$$= \beta^H (A^+)^H \cdot A^+\beta = \|A^+\beta\|^2$$

故 $\|Y\|^2 = \|Y - A^+\beta\|^2 + \|A^+\beta\|^2 \geq \|A^+\beta\|^2$

故 $A^+\beta$ 模最小.

(4) 代入验证即可. □

注：(3) 的简易做法.

设 $A = P\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}Q$ 为奇异值分解.

则 $\|Ax-\beta\| = \left\|\begin{pmatrix}\Sigma & \\ & 0\end{pmatrix}Qx - P^{-1}b\right\|$ (利用 $\|P\vec{y}\| = \|\vec{y}\|$)

则最小二乘解为 $x = Q^{-1}\begin{pmatrix}\Sigma^{-1}c_1 \\ t\end{pmatrix}$

其中 $P^{-1}b = \begin{pmatrix}c_1 \\ c_2\end{pmatrix}$,

t 为自由参数.

则 $\|x\|$ 最小的为 $t=0$ 的 $x_0 = Q^{-1}\begin{pmatrix}\Sigma^{-1}c_1 \\ 0\end{pmatrix}$

$$= A^+\beta.$$

思考：(1) 在得到 A_1 时，A_1 并不依赖 U_1, U_2,

但是 $B = U_2^H B_1 U_1^H$, "似乎"依赖于 U_1, U_2.

我们在哪一步实质证明了不依赖 U_1, U_2?

(Steinitz 替换定理)

若 $S=\{\alpha_1,\cdots,\alpha_s\}$ 线性无关，且可被 $T=\{\beta_1,\cdots,\beta_t\}$ 线性表出

则可用 S 替换 $\beta_{i_1},\cdots,\beta_{i_s}$

s.t. $\{\alpha_1,\cdots,\alpha_s,\beta_{i_{s+1}},\cdots,\beta_{i_t}\}$ 与 T 等价

证明. 矩阵证法

设 $(\alpha_1\cdots\alpha_s)=(\beta_1\cdots\beta_t)\ C$

$C\in F^{t\times s}$

由于 $\operatorname{rank}(\alpha_1\cdots\alpha_s)=s$

故 C 列满秩. 即

$\exists\ i_1<\cdots<i_s.\quad \det C\begin{pmatrix} i_1\cdots i_s\\ 1\cdots s\end{pmatrix}\neq 0$

不妨设 $i_j=j$

(否则考虑 $(\alpha_1\cdots\alpha_s)=(\beta_{i_1},\cdots,\beta_{i_s},\beta_{i_{s+1}},\cdots,\beta_t)\hat{C}$)

则 $(\alpha_1\cdots\alpha_s)=(\beta_1\cdots\beta_s)C_1+(\beta_{s+1}\cdots\beta_t)C_2$

$C=\begin{pmatrix}C_1\\C_2\end{pmatrix}.\quad \underbrace{C_1=C\begin{pmatrix}1\cdots s\\1\cdots s\end{pmatrix}}_{\text{可逆}}$

则 $(\beta_1\cdots\beta_s)=\left[(\alpha_1\cdots\alpha_s)-(\beta_{s+1}\cdots\beta_t)C_2\right]C_1^{-1}$

即 $\beta_1\cdots\beta_s$ 可被 $\alpha_1\cdots\alpha_s$ 及 $\beta_{s+1},\cdots,\beta_t$ 线性表出

$\tilde{T}=\{\alpha_1,\cdots\alpha_s,\beta_{s+1},\cdots,\beta_t\}\Longleftrightarrow T$

□

6. 设向量组 $\alpha_1,\cdots,\alpha_s$ 的秩为 r，在其中任取 m 个向量 $\alpha_{i_1},\cdots,\alpha_{i_m}$ 组成向量组 S. 求证：S 的秩 $\geqslant r+m-s$.

不妨设 $i_j = j$

I_m

则 $(\alpha_1 \cdots \alpha_m) = (\alpha_1 \cdots \alpha_s)\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1 \\ & 0 & \end{pmatrix}$

则 $\operatorname{rank}(\alpha_1 \cdots \alpha_m) = \operatorname{rank}\left((\alpha_1 \cdots \alpha_s)\begin{pmatrix} I_m \\ 0 \end{pmatrix}\right)$

由以下不等式

例 4.9. 对于任意 $A \in \mathbb{F}^{m\times n}$，$B \in \mathbb{F}^{n\times p}$，$C \in \mathbb{F}^{p\times q}$，有 **Frobenius**[5]秩不等式

$$\operatorname{rank}(AB) + \operatorname{rank}(BC) \leqslant \operatorname{rank}(ABC) + \operatorname{rank}(B).$$

特别，当 $B = I_n$ 时，上式成为 **Sylvester** 秩不等式

$$\operatorname{rank}(A) + \operatorname{rank}(C) \leqslant \operatorname{rank}(AC) + n.$$

则 $\operatorname{rank}(\alpha_1 \cdots \alpha_m) \geqslant \operatorname{rank}(\alpha_1 \cdots \alpha_s) + \operatorname{rank}\begin{pmatrix} I_m \\ 0 \end{pmatrix} - s$
$= r + m - s$

$\square$

3. 设 V 是由复数组成的无穷数列 $\{a_n\} = \{a_1, a_2, \cdots, a_n, \cdots\}$ 的全体组成的集合，定义 V 中任意两个数列的加法 $\{a_n\} + \{b_n\} = \{a_n + b_n\}$ 及任意数列与任意复数的乘法 $\lambda\{a_n\} = \{\lambda a_n\}$ 之后成为复数域 $\mathbf{C}$ 上线性空间.

(1) 求证：V 中满足条件 $a_n = a_{n-1} + a_{n-2}$ ($\forall n \geqslant 3$) 的全体数列 $\{a_n\}$ 组成 V 的子空间 W. W 的维数是多少？

(2) 对任意 $(a_1, a_2) \in \mathbf{C}^2$，定义 $\sigma(a_1, a_2) = \{a_1, a_2, \cdots, a_n, \cdots\} \in W$. 求证：$\sigma$ 是 $\mathbf{C}^2$ 到 W 的同构映射.

(3) 求证：W 中存在一组由等比数列组成的基 M.

(4) 设数列 $\{F_n\}$ 满足条件 $F_1 = F_2 = 1$ 且 $F_n = F_{n-1} + F_{n-2}$. 求 $\{F_n\}$ 在基 M 下的坐标，并由此求出 $\{F_n\}$ 的通项公式.

(1). 取 $v_1 = (1, 0, 1, 1, \cdots)$

$v_2 = (0, 1, 1, 2, \cdots)$

$\forall\ a = (a_1, a_2, \cdots) \in W$

用归纳法易证 $a = a_1 v_1 + a_2 v_2$

而 v_1, v_2 线性无关，则 $\dim W = 2$.

另证

$\tau: W \to \mathbb{C}^2$

$(a_1, a_2, \cdots) \mapsto (a_1, a_2)$

则易证

$\sigma\tau = id_W$

$\tau\sigma = id_{\mathbb{C}^2}$

$\Rightarrow \sigma$ 为同构.

(2) $\sigma: \mathbb{C}^2 \to W$ 　　是线性映射.

$(a_1, a_2) \mapsto (a_1, a_2, \cdots)$

单: 若 $\sigma(a_1, a_2) = (0, 0, \cdots)$

则 $a_1 = a_2 = 0$

满: $\forall\ v = (a_1, a_2, \cdots) \in W$

$\sigma(a_1, a_2) = (a_1, a_2, \cdots)$ 　故为同构.

(3) 设 $(1, q, q^2, \cdots) \in W$

Ex: 若 $\sigma: V \to W$

$\tau: W \to V$

线性映射

且 $\sigma \circ \tau = id_W$

则 σ 满, τ 单

则 $q^{n+1} = q^n + q^{n-1}$ 　　$n \geq 1$

q 显然不为 0, 故 $q^2 - q - 1 = 0$

$\Rightarrow q_{1,2} = \dfrac{1 \pm \sqrt{5}}{2}$

而 $(1, q_1)$, $(1, q_2)$ 为 $\mathbb{C}^2$ 的基

故由 σ 同构. $\sigma(1, q_1)$, $\sigma(1, q_2)$ 为基.

(3) $(F_n)_n \in W$

设 $F_n = x\,\sigma(1, q_1) + y\,\sigma(1, q_2)$

即 $1 = x + y$

$1 = x q_1 + y q_2$

$\Rightarrow F_n = \dfrac{(\sqrt{5}+1)^n - (1-\sqrt{5})^n}{2^n \sqrt{5}}$

□

4. 设 $\mathbf{R}^+$ 是所有的正实数组成的集合. 对任意 $a, b \in \mathbf{R}^+$ 定义 $a \oplus b = ab$ (实数 a, b 按通常乘法的乘积), 对任意 $a \in \mathbf{R}^+$ 和 $\lambda \in \mathbf{R}$ 定义 $\lambda \circ a = a^{\lambda}$. 求证:

(1) $\mathbf{R}^+$ 按上述定义的加法 $a \oplus b$ 和数乘 $\lambda \circ a$ 成为实数域 $\mathbf{R}$ 上的线性空间.

(2) 实数集合 $\mathbf{R}$ 按通常方式定义加法和乘法看成 $\mathbf{R}$ 上的线性空间, 求证: 通常的这个线性空间 $\mathbf{R}$ 与按上述方式定义的线性空间 $\mathbf{R}^+$ 同构. 并给出这两个空间之间的全部同构映射.

(2) $\sigma: \mathbb{R} \to \mathbb{R}^+$

$a \mapsto e^a$

线性映射: 　$\sigma(\lambda a) = e^{\lambda a}$

$\lambda \circ \sigma(a) = \lambda \circ e^a = e^{\lambda a}$

$\sigma(a+b) = e^{a+b} = e^a \cdot e^b = \sigma(a) \oplus \sigma(b)$

单：　若 $\sigma(a) = 1$ 则 $e^a = 1 \Rightarrow a = 0$

满．　$\forall\ b \in \mathbb{R}^+$　$\sigma(\log b) = b$

$\forall\ \sigma: \mathbb{R} \to \mathbb{R}^+$

设 $\sigma(1) = a$

则 $\sigma(\lambda) = \sigma(\lambda \cdot 1) = \lambda \circ \sigma(1) = a^\lambda$

故 σ 由 $\sigma(1) = a \in \mathbb{R}^+$ 唯一决定

由于 σ 同构．故 $a \neq 1$

故 同构映射全体 $= \left\{ \sigma_a: \begin{array}{l} \mathbb{R} \to \mathbb{R}^+ \\ \lambda \mapsto a^\lambda \end{array} \;\middle|\; a \in \mathbb{R}^+ .\ a \neq 1 \right\}$

$\square$

小测题　$V_1, V_2, V_3 \in V$ 子空间

证明 $V_1 \cap (V_2 + V_1 \cap V_3) = V_1 \cap V_2 + V_1 \cap V_3$

并举例 $V_1 \cap (V_2 + V_3) = V_1 \cap V_2 + V_1 \cap V_3$

不一定成立．

证．　"$\subseteq$"　设 $v \in V_1 \cap (V_2 + V_1 \cap V_3)$

即 $v \in V_1$，$v = v' + v''$

其中 $v' \in V_2$，$v'' \in V_1 \cap V_3$

则 $v' = \underset{\in V_1}{v} - \underset{\in V_1}{v''} \Rightarrow v' \in V_1$

$\Rightarrow v' \in V_1 \cap V_2 \Rightarrow v = v' + v''$

$\in$ 右侧

"$\supseteq$" $v = v' + v'' \in V_1 \cap V_2 + V_1 \cap V_3$

其中 $v' \in V_1 \cap V_2$, $v'' \in V_1 \cap V_3$

则 $v \in V_1$

且 $v \in V_2 + V_1 \cap V_3$ □

反例. 你发现 "$\supseteq$" 显然,

故要找 "$\subseteq$" 的反例

找 $v \in V_1 \cap (V_2 + V_3)$ ①

$v \notin V_1 \cap V_2 + V_1 \cap V_3$ ②

① $\Rightarrow$ $v = v' + v''$. $v' \in V_2$, $v'' \in V_3$

只需 $v' \notin V_1$, $v'' \notin V_1$

但 $v' + v'' \in V_1$ 即可

故要找 $V_2 + V_3 \supseteq V_1$ 的例子

例. $V_2 = \langle (0,1) \rangle$ $V = \mathbb{R}^2$

$V_3 = \langle (1,0) \rangle$

$V_1 = \langle (1,1) \rangle$

则 $(1,1) \in V_1$ 且 $(1,1) = (0,1) + (1,0)$

但 $V_1 \cap V_2 = \{\vec{0}\}$, $V_1 \cap V_3 = \{\vec{0}\}$

故 $V_1 \cap V_2 + V_1 \cap V_3 = \{\vec{0}\}$

$(1,1) \notin \{\vec{0}\}$ □