第六章 广义逆矩阵

6.1 Penrose 广义逆矩阵

1. 定义及存在性

对于满秩方阵 A, A^{-1}存在，且 $AA^{-1}=A^{-1}A=I$ 故，当然有

$$\begin{cases} AA^{-1}A = A \\ A^{-1}AA^{-1} = A^{-1} \\ (AA^{-1})^H = AA^{-1} \\ (A^{-1}A)^H = A^{-1}A \end{cases}$$

•Penrose 广义逆矩阵的定义：

设 $A\in C^{m\times n}$，若 $Z\in C^{n\times m}$ 且使如下四个等式成立，

$$AZA = A,\quad ZAZ = Z,\quad (AZ)^H = AZ,\quad (ZA)^H = ZA$$

则称 Z 为 A 的 Moore-Penrose(广义)逆，记为 A^+。而上述四个等式有依次称为

Penrose 方程(i),(ii) ,(iii) ,(iv)。

•定理： 对任意$A\in C^{m\times n}$，A^+存在且唯一。

证明：存在性. $\forall\ A\in C_r^{m\times n}$，均存在酉矩阵 $U\in C_m^{m\times m}$，$V\in C_n^{n\times n}$ 使

$$U^HAV = D = \begin{bmatrix} \sigma_1 & & & & \vdots & \\ & \sigma_2 & & & \vdots & \\ & & \ddots & & \vdots & 0 \\ & & & \sigma_r & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots \\ & 0 & & & \vdots & 0 \end{bmatrix}_{m\times n} \quad 即\ A = UDV^H$$

此时，令 $Z=V\tilde{D}U\ \in C_r^{n\times m}$ 则 其中，$\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2$是 A^HA 的全部非零特征值。

$$\tilde{D} = \begin{bmatrix} \sigma_1^{-1} & & & & \vdots & \\ & \sigma_2^{-1} & & & \vdots & \\ & & \ddots & & \vdots & 0 \\ & & & \sigma_r^{-1} & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots \\ & 0 & & & \vdots & 0 \end{bmatrix}_{n\times m}$$

(i) $AZA=(UDV^H)(V\tilde{D}U^H)(UDV^H)=UD\tilde{D}DV^H=UDV^H=A$

(ii) $ZAZ=(V\tilde{D}U^H)(UDV^H)(V\tilde{D}U^H)=V\tilde{D}D\tilde{D}U^H=V\tilde{D}U^H=Z$

(iii) $(AZ)^H=[(UDV^H)(V\tilde{D}U^H)]^H=(UD\tilde{D}U^H)^H=UD\tilde{D}U^H=AZ$

(iv) $(ZA)^H=(V\tilde{D}DV^H)^H=V\tilde{D}DV^H=ZA$

即，$Z=A^+$

其中 $D\tilde{D}=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times m}$ $\tilde{D}D=\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times n}$ $AZ=UD\tilde{D}U^H$，$ZA=V\tilde{D}DV^H$

唯一性：设 Z ,Y 均满足四个 Penrose 方程，则

$$\begin{aligned} Z &= ZAZ = Z(AZ)^H = ZZ^HA^H = ZZ^H(AYA)^H = Z(AZ)^H(AY)^H = Z(AZ)(AY) \\ &= ZAY = (ZA)^HY = A^HZ^HY = A^H Z^H(YAY) = A^HZ^H(YA)^HY = A^HZ^HA^HY^HY \\ &= (AZA)^HY^HY = A^HY^HY = (YA)^HY = YAY = Y \end{aligned}$$ 即，满足

四个 Penrose 方程的 Z 是唯一的.

由A^+的唯一性可知：当 A 为满秩矩阵时，$A^+=A^{-1}$

•$\{i,j,\dots,l\}$-逆的定义：

$\forall A\in C^{m\times n}$，若 $Z\in C^{n\times m}$ 且满足 Penrose 方程中的第(i),(j),⋯,(l) 个方程，则称 Z 为 A 的{i, j,⋯,l} -逆，记为 $A^{(i,j,\cdots,l)}$，其全体记为 A{i, j,⋯,l}。

{i, j,⋯,l} -逆共有 $C_4^1+C_4^2+C_4^3+C_4^4=15$ 类，但实际上常用的为如下 5 类：

$A\{1\}, A\{1,2\}, A\{1,3\}, A\{1,4\}, A\{1,2,3,4\}(A^+)$

2. 广义逆矩阵的性质

引理：$rank(AB)\le \min(rankA, rankB)$

•**定理 1**：设 $A\in C^{m\times n}$，$B\in C^{n\times p}$，$\lambda\in C$，$\lambda^+=\begin{cases}\lambda^{-1} & \lambda\ne 0\\ 0 & \lambda=0\end{cases}$，则：

(1) $(A^{(1)})^H\in A^H\{1\}$

(2) $\lambda^+A^{(1)}\in(\lambda A)\{1\}$

(3) S、T 为可逆方阵且与 A 可乘，则

$$T^{-1}A^{(1)}S^{-1}\in(SAT)\{1\}, \quad (S\in C_m^{m\times m}, T\in C_n^{n\times n})$$

(4) $rank(A^{(1)})\ge rankA$

(5) $AA^{(1)}$ 和 $A^{(1)}A$ 均为幂等矩阵且与 A 同秩 $\quad (P^2=P)$

(6) $R(AA^{(1)})=R(A)$, $\quad N(A^{(1)}A)=N(A)$, $\quad R((A^{(1)}A)^H)=R(A^H)$

(7) $A^{(1)}A=I_n \iff rank(A)=n$

$AA^{(1)}=I_m \iff rank(A)=m$

(8) $AB(AB)^{(1)}A=A \iff rank(AB)=rank(A)$

$B(AB)^{(1)}AB=B \iff rank(AB)=rank(B)$

•**定理 2**：设矩阵 $Y,Z\in A\{1\}$，又设 $X=YAZ$，则 $X\in A\{1,2\}$

证：已知 $AYA=AZA=A$，故：

(i) $AXA=A(YAZ)A=(AYA)ZA=AZA=A$

(ii) $XAX=(YAZ)A(YAZ)=Y(AZA)YAZ=Y(AYA)Z=YAZ=X$

•**定理 3**：给定矩阵 A 和 $X\in A\{1\}$，则 $X\in A\{1,2\}$ 的充要条件为：$rankX=rankA$

证：必要性. $Z\in A\{1,2\}$ 则 (i) AZA=A; (ii) ZAZ=Z

$\to A\in Z\{1,2\}$

而由 $rankA^{(1)}\ge rankA$ 可知 $rankZ\ge rankA$, $rankA\ge rankZ$

$\Rightarrow rankZ=rankA$

充分性. 已知 $rankZ=rankA$, $Z\in A\{1\}$

即 AZA=A，$\to rank(ZA)\ge rank(A)$，而 rank（Z）$\ge$ rank(ZA)

即：rank（Z）$\ge$ rank(ZA) $\ge$ rank(A)。再由已知条件，得：

rank(ZA)=rank(A)=rank(Z)

另一方面，$R(ZA)\subseteq R(Z)$，结合 rank(ZA)=rank(Z) $\quad\to R(ZA)=R(Z)$

$\forall e\in C^m$，$\exists u\in C^n$，使 $ZAu=Ze$

$\to \quad ZA[u_1 u_2\cdots u_m]=Z[e_1e_2\cdots e_m]$

令 $[e_1e_2\cdots e_m]=I_m$，$[u_1 u_2\cdots u_m]=U$ (u_i=n 维，e_j=m 维)

$\Rightarrow \exists U$ 使 $Z=ZAU$

故 $\quad ZAZ=ZA(ZAU)=ZAU=Z \quad \to Z$ 满足 Penrose 方程(ii)

可见，$Z\in A\{1,2\}$.

引理：对任意矩阵 A，均有$rank(A^HA) = rankA = rank(AA^H)$

•**定理 4：** 设矩阵 A 给定，则：

$$Y = (A^HA)^{(1)}A^H \in A\{1,2,3\}$$
$$Z = A^H(AA^H)^{(1)} \in A\{1,2,4\}$$

证：显然 $R(A^HA) \subseteq R(A^H)$，又由引理可知 $R(A^HA) = R(A^H)$，

即存在 U 使 $A^H = A^HAU \quad \rightarrow A = U^HA^HA$

$AYA = (U^HA^HA)[(A^HA)^{(1)}A^H]A \overset{(i)}{=} U^HA^HA = A$ 满足(i) $\rightarrow Y \in A\{1\}$

可见 $rankY \geq rankA$

但 $rankY = rank\left(\left(A^HA\right)^{(1)}A^H\right) \leq rankA^H = rankA.$

即 $rankY = rankA. \quad \rightarrow Y \in A\{1,2\}$

$AY = (U^HA^HA)(A^HA)^{(1)}A^H = U^HA^HA\,(A^HA)^{(1)}A^HAU$

$= U^H(A^HA)U = (AY)^{(H)}$

$\Rightarrow Y \in A\{3\}$ 综合之，即 $Y \in A\{1,2,3\}$

同理可证另式。

•**定理 5：** 设矩阵 A 给定，则：

$$A^+ = A^{(1,4)}AA^{(1,3)}$$

证：(1)由定理 1 知，$A^{(1,4)}AA^{(1,3)} \overset{\Delta}{=} X \in A\{1,2\}$

(2) $AX = A\,A^{(1,4)}AA^{(1,3)} \overset{i}{=} AA^{(1,3)} \overset{iii}{=} (AA^{(1,3)})^H = (AX)^H$

(3) $XA = A^{(1,4)}AA^{(1,3)}A \overset{i}{=} A^{(1,4)}A \overset{iv}{=} (A^{(1,4)}A)^H = (XA)^H$

$\Rightarrow X \in A\{1,2,3,4\} = \{A^{\dagger}\}$

•**定理 6：** 设矩阵 A 给定，则：

(1) $rankA^{\dagger} = rankA$

(2) $(A^{\dagger})^{\dagger} = A$

(3) $\left(A^H\right)^{\dagger} = (A^{\dagger})^H$，$\left(A^T\right)^{\dagger} = \left(A^{\dagger}\right)^T$

(4) $\left(A^HA\right)^{\dagger} = A^{\dagger}\left(A^H\right)^{\dagger}$，$\left(AA^H\right)^{\dagger} = \left(A^H\right)^{\dagger}A^{\dagger}$

(5) $A^{\dagger} = \left(A^HA\right)^{\dagger}A^H = A^H\left(AA^H\right)^{\dagger}$

(6) $R(A^{\dagger}) = R(A^H), N(A^{\dagger}) = N(A^H)$

推论 1：

若$A \in C_n^{m\times n}$(列满秩矩阵)，则：$A^+ = (A^HA)^{-1}A^H$

若$A \in C_n^{m\times n}$(行满秩矩阵)，则：$A^+ = A^H(AA^H)^{-1}$

推论 2：设α为 n 维列向量，且$\alpha \neq 0$，则：$\alpha^+ = (\alpha^H\alpha)^{-1}\alpha^H$；

对非零行向量有：$\beta^+ = \beta^H(\beta\beta^H)^{-1}$；

A,B 可逆，则 $(AB)^{-1} = B^{-1}A^{-1}$，但一般$(AB)^{\dagger} \neq B^{\dagger}A^{\dagger}$

如 $A = [1 \quad 0]$，$B = \begin{bmatrix}1\\1\end{bmatrix}$，$AB = [1]$，$BA = \begin{bmatrix}1 & 0\\1 & 0\end{bmatrix}$

$(AB)^{\dagger} = [1]$，$A^{\dagger} = \begin{bmatrix}1\\0\end{bmatrix}$，$B^{\dagger} = [\frac{1}{2} \quad \frac{1}{2}]$，$B^{\dagger}A^{\dagger} = \begin{bmatrix}\frac{1}{2}\end{bmatrix}$

6.2 投影矩阵与 Moore-Penrose 逆

1.投影与投影矩阵

设 L，M 为$\mathbf{C}^n$的子空间并构成直和$\mathbf{L}+\mathbf{M}=\mathbf{L}\oplus\mathbf{M}=\mathbf{C}^n$.即：

$\forall \mathbf{x}\in\mathbf{C}^n$，∃唯一的$\mathbf{y}\in\mathbf{L},\mathbf{z}\in\mathbf{M}$使 x=y+z，称 y 为 x 沿着 M 到 L 的投影。

•定义：

将任意$\mathbf{x}\in\mathbf{C}^n$变为其沿着 M 到 L 的投影的变换称为沿着 M 到 L 的投影算子,记为$\mathbf{P}_{\mathbf{L},\mathbf{M}}$

即$\mathbf{P}_{\mathbf{L},\mathbf{M}}\ \mathbf{x}=\mathbf{y}\in\mathbf{L}$,投影算子是线性变换，其矩阵称为投影矩阵,仍记为$\mathbf{P}_{\mathbf{L},\mathbf{M}}$。

特别地，若$x\in L$，则$P_{L,M}x=x$；若$x\in M$，则$P_{L,M}x=0$，故$P_{L,M}$的值域为$R(P_{L,M})=L$，零空间为$N(P_{L,M})=M$；

$P_{L,M}$为一个线性算子，即对任意向量$x_1,x_2\in C^n$和任意复数λ，μ，恒有$P_{L,M}(\lambda x_1+\mu x_2)=\lambda P_{L,M}x_1+\mu P_{L,M}x_2$.

引理：设 n 阶方阵 A 为幂等矩阵，则$N(A)=R(I-A)$.

•定理： n 阶方阵 P 成为投影矩阵的充要条件是 P 为幂等矩阵。

证：

充分性 $\mathbf{P}^2=\mathbf{P},\forall\mathbf{x}\in\mathbf{C}^n$,令$\mathbf{y}=\mathbf{Px}\in\mathbf{R}(\mathbf{P}),\mathbf{z}=(\mathbf{I}-\mathbf{P})\mathbf{x}\in\mathbf{R}(\mathbf{I}-\mathbf{P})=\mathbf{N}(\mathbf{P})$。

若 $\mathbf{R}(\mathbf{P})\cap\mathbf{N}(\mathbf{P})=\{0\}$,则$\mathbf{P}=\mathbf{P}_{\mathbf{R}(\mathbf{P}),\mathbf{N}(\mathbf{P})}$确为投影矩阵，下面证之

$\forall\mathbf{x}\in\mathbf{R}(\mathbf{P})\cap\mathbf{N}(\mathbf{P})$,

一方面，因$\mathbf{x}\in\mathbf{R}(\mathbf{P})$,存在$\mathbf{u}\in\mathbf{C}^n$使$\mathbf{x}=\mathbf{Pu}$

另一方面$\mathbf{x}\in\mathbf{N}(\mathbf{P})$,即$\mathbf{Px}=0$.但$\mathbf{Px}=\mathbf{P}^2\mathbf{u}=\mathbf{Pu}=\mathbf{x}\to\mathbf{x}=0$

$$\Rightarrow\mathbf{R}(\mathbf{P})\cap\mathbf{N}(\mathbf{P})=\{0\}。$$

必要性 $\mathbf{P}=\mathbf{P}_{\mathbf{L},\mathbf{M}}$ 故 $\forall\mathbf{x}\in\mathbf{C}^n$，∃唯一分解 $\mathbf{y}\in\mathbf{L},\mathbf{z}\in\mathbf{M}$使

$$\mathbf{x}=\mathbf{y}+\mathbf{z}\text{且 }\mathbf{Px}=\mathbf{y}$$

$$\to\mathbf{P}^2\mathbf{x}=\mathbf{Py}=\underset{\substack{\uparrow\\ \mathbf{y}=\mathbf{y}+0}}{\mathbf{y}}=\mathbf{Px}\xrightarrow{\mathbf{x}\text{任意}}\mathbf{P}^2=\mathbf{P}$$

•投影矩阵的构造：

取 L 的一个基$\{\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r\}$(设 L 为 r 维子空间),M 的一个基$\{\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}\}$(则 M 的维数为 n-r)。由直和关系知$\{\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r;\ \mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}\}$即构成$\mathbf{C}^n$的一个基。

故，即令：

$$\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r],\quad\mathbf{Y}=[\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}]$$

则[X Y]为可逆方阵。另一方面

$$\mathbf{x}_i\in\mathbf{L}\to\mathbf{P}_{\mathbf{L},\mathbf{M}}\mathbf{x}_i=\mathbf{x}_i;\mathbf{y}_i\in\mathbf{M}\to\mathbf{P}_{\mathbf{L},\mathbf{M}}\mathbf{y}_i=0$$

即 $\mathbf{P}_{\mathbf{L},\mathbf{M}}[\mathbf{X}\ \mathbf{Y}]=[\mathbf{X}\ \mathbf{O}]\to\mathbf{P}_{\mathbf{L},\mathbf{M}}=[\mathbf{X}\ \mathbf{O}][\mathbf{X}\ \mathbf{Y}]^{-1}$

可见，$\mathbf{P}_{\mathbf{L},\mathbf{M}}$的秩为 r（$\mathbf{rank}(\mathbf{P}_{\mathbf{L},\mathbf{M}})=\dim\mathbf{R}(\mathbf{P}_{\mathbf{L},\mathbf{M}})=\dim\mathbf{L}$）

•例题：

例 6.1 设 L 是由向量$\begin{bmatrix}1\\0\end{bmatrix}$张成的子空间,$M$ 是由向量$\begin{bmatrix}1\\-1\end{bmatrix}$张成的子空间,则由式(6.1.5)可求得,$\mathbf{R}^2$ 上沿着 M 到 L 的投影矩阵为

$$\boldsymbol{P}_{L,M}=\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}1&1\\0&-1\end{bmatrix}^{-1}=\begin{bmatrix}1&0\\0&0\end{bmatrix}\begin{bmatrix}1&1\\0&-1\end{bmatrix}=\begin{bmatrix}1&1\\0&0\end{bmatrix}$$

2.正交投影算子与正交投影矩阵

L 为$\mathbf{C}^n$的子空间，其正交补空间$\mathbf{L}^{\perp}=\left\{\mathbf{x}\,\middle|\,(\mathbf{x},\ \mathbf{y})=0,\mathbf{x}\in\mathbf{C}^n,\mathbf{y}\in\mathbf{L}\right\}$（无特别声明取$\mathbf{x}^H\mathbf{y}$）

•定义：

设 L 是$\mathbf{C}^n$的子空间，则称沿着$\mathbf{L}^{\perp}$到 L 的投影算子$\mathbf{P}_{\mathbf{L},\mathbf{L}^{\perp}}$为正交投影算子，简记为

$\mathbf{P}_L$。正交投影算子的矩阵称为正交投影矩阵，仍记为$\mathbf{P}_L$。

引理：

（1）对 n 阶方阵 A，$\forall\mathbf{x}\in\mathbf{C}^n$均有$\mathbf{x}^H\mathbf{A}\mathbf{x}=\mathbf{0}$则 A=0

（2）N（$\mathbf{P}^H$）$=\mathbf{R}^{\perp}(\mathbf{P})$

•定理： n 阶方阵 P 为正交投影矩阵的充要条件是 P 为幂等的 Hermite 矩阵。

证：

充分性：

$$\mathbf{P}^2=\mathbf{P},\ \mathbf{P}^H=\mathbf{P}\to\mathbf{P}=\mathbf{P}_{R(p),N(p)}=\mathbf{P}_{R(p),N(p^H)}=\mathbf{P}_{R(p),R^{\perp}(p)}=\mathbf{P}_{R(p)}$$

必要性：

$$\mathbf{P}=\mathbf{P}_L$$

$$\forall\mathbf{x}\in\mathbf{C}^n,\text{可唯一地分解成 }\mathbf{y}=\mathbf{P}\mathbf{x}\in\mathbf{L},\mathbf{z}=(\mathbf{I}-\mathbf{P})\mathbf{x}\in\mathbf{L}^{\perp}\text{使 }\mathbf{x}=\mathbf{y}+\mathbf{z}$$

又 $\mathbf{y}\in\mathbf{L},\mathbf{z}\in\mathbf{L}^{\perp}\to\mathbf{y}^H\mathbf{z}=0\to\mathbf{x}^H\mathbf{P}^H(\mathbf{I}-\mathbf{P})\mathbf{x}=0\xrightarrow{x\text{任意}}\mathbf{P}^H(\mathbf{I}-\mathbf{P})=0$

$\mathbf{P}^H=\mathbf{P}^H\mathbf{P}=(\mathbf{P}^H\mathbf{P})^H=(\mathbf{P}^H)^H=\mathbf{P}$， $\mathbf{P}$ 为 Hermite 矩阵。

•正交投影矩阵的构造：

设 L 的一个基为$\{\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r\}$，$\mathbf{L}^{\perp}$的一个基为$\{\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}\}$。则$\mathbf{x}_i^H\mathbf{y}_j=0$

令$\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2\cdots\mathbf{x}_r]$，$\mathbf{Y}=[\mathbf{y}_1,\mathbf{y}_2\cdots\mathbf{y}_{n-r}]$ 则$\mathbf{X}^H\mathbf{Y}=0$

$$(\mathbf{A}^{-1}=(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H)$$

$$\mathbf{P}_L=[\mathbf{X}\ \mathbf{O}][\mathbf{X}\ \mathbf{Y}]^{-1}=[\mathbf{X}\ \mathbf{O}]\left\{[\mathbf{X}\ \mathbf{Y}]^H[\mathbf{X}\ \mathbf{Y}]\right\}^{-1}[\mathbf{X}\ \mathbf{Y}]^H$$

$$=[\mathbf{X}\ \mathbf{O}]\begin{bmatrix}\mathbf{X}^H\mathbf{X} & \mathbf{X}^H\mathbf{Y}\\ \mathbf{Y}^H\mathbf{X} & \mathbf{Y}^H\mathbf{Y}\end{bmatrix}^{-1}[\mathbf{X}\ \mathbf{Y}]^H=[\mathbf{X}\ \mathbf{O}]\begin{bmatrix}\mathbf{X}^H\mathbf{X} & 0\\ 0 & \mathbf{Y}^H\mathbf{Y}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}^H\\ \mathbf{Y}^H\end{bmatrix}$$

$$=\left[\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\ \ \mathbf{O}\right]\begin{bmatrix}\mathbf{X}^H\\ \mathbf{Y}^H\end{bmatrix}=\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$$

•例题：

例 6.2 在 $\mathbf{R}^3$ 中 L 是由向量 $\boldsymbol{\alpha}=(1,2,0)^{\mathrm{T}}$ 和 $\boldsymbol{\beta}=(0,1,1)^{\mathrm{T}}$ 张成的子空间，求正交投影矩阵 $\boldsymbol{P}_L$ 和向量 $\boldsymbol{x}=(1,2,3,)^{\mathrm{T}}$ 沿 $L^{\perp}$ 到 L 的投影.

解 因为

$$\boldsymbol{X}=\begin{bmatrix}1&0\\2&1\\0&1\end{bmatrix},\quad \boldsymbol{X}^{\mathrm{H}}\boldsymbol{X}=\begin{bmatrix}5&2\\2&2\end{bmatrix}$$

$$(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X})^{-1}=\frac{1}{6}\begin{bmatrix}2&-2\\-2&5\end{bmatrix}$$

由式(6.1.6)得

$$\boldsymbol{P}_L=\boldsymbol{X}(\boldsymbol{X}^{\mathrm{H}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{H}}=\frac{1}{6}\begin{bmatrix}2&2&-2\\2&5&1\\-2&1&5\end{bmatrix}$$

$\boldsymbol{x}$ 在 L 上的投影为

$$\boldsymbol{P}_L\boldsymbol{x}=\left(0,\frac{5}{2},\frac{5}{2}\right)^{\mathrm{T}}$$

3.投影矩阵与广义逆矩阵

(i) $\mathbf{AXA}=\mathbf{A}\rightarrow\mathbf{AX}=\mathbf{P}_{\mathbf{R(A),N(AX)}}$

(ii) $\mathbf{XAX}=\mathbf{X}\rightarrow\mathbf{XA}=\mathbf{P}_{\mathbf{R(XA),N(A)}}$

$$\begin{cases}\mathbf{AXA}=\mathbf{A}\\(\mathbf{AX})^{\mathbf{H}}=\mathbf{AX}\end{cases}\rightarrow\mathbf{AX}=\mathbf{P}_{\mathbf{R}\ (\mathbf{A})}$$

$$\begin{cases}\mathbf{XAX}=\mathbf{X}\\(\mathbf{XA})^{\mathbf{H}}=\mathbf{XA}\end{cases}\rightarrow\mathbf{XA}=\mathbf{P}_{\mathbf{R}\ (\mathbf{XA})}$$

Moore 定义：设 $\mathbf{A}\in\mathbf{C}^{\mathbf{m\times n}},\mathbf{X}\in\mathbf{C}^{\mathbf{n\times m}}$ 且 $\mathbf{AX}=\mathbf{P}_{\mathbf{R}\ (\mathbf{A})}$，$\mathbf{XA}=\mathbf{P}_{\mathbf{R}\ (\mathbf{XA})}$

则 X 为 A 的 Moore 广义矩阵。事实上，Moore 广义矩阵正是 $\mathbf{A}^{\dagger}$

6.3 广义逆矩阵的计算方法

1. 利用 Hermite 标准形计算矩阵的{1}-逆和{1，2}-逆

（1）求{1}-逆：

•**定理：**设$A \in C_r^{m\times n}$，$Q \in C_m^{m\times m}$，$P \in C_n^{n\times n}$，使得：$QAP = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}$成立，则对任意$L \in C^{(n-r)\times(m-r)}$，$n\times m$阶矩阵：$X = P\begin{bmatrix} I_r & O \\ O & L \end{bmatrix}Q$是 A 的{1}-逆。（Q 为初等变换矩阵即 QA=B，B 为 Hermite 标准形，P 为置换矩阵）

证：将$QAP = \begin{bmatrix} I_r & K \\ O & O \end{bmatrix}$改写为：$A = Q^{-1}\begin{bmatrix} I_r & K \\ O & O \end{bmatrix}P^{-1}$，容易验证$X$满足$AXA = A$.

例题：

例 6.4 求矩阵

$$A = \begin{bmatrix} 0 & 2\mathrm{j} & \mathrm{j} & 0 & 4+2\mathrm{j} & 1 \\ 0 & 0 & 0 & -3 & -6 & -3-3\mathrm{j} \\ 0 & 2 & 1 & 1 & 4-4\mathrm{j} & 1 \end{bmatrix} \quad (\mathrm{j} = \sqrt{-1})$$

的{1}-逆.

解 因为对$[A \,\vdots\, I]$进行初等行变换可得

$$\left[\begin{array}{cccccc:ccc} 0 & 1 & \frac{1}{2} & 0 & 1-2\mathrm{j} & -\frac{1}{2}\mathrm{j} & -\frac{1}{2}\mathrm{j} & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 1+\mathrm{j} & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathrm{j} & \frac{1}{3} & 1 \end{array}\right]$$

所以 A 的 Hermite 标准形为

$$QA = \begin{bmatrix} 0 & 1 & \frac{1}{2} & 0 & 1-2\mathrm{j} & -\frac{1}{2}\mathrm{j} \\ 0 & 0 & 0 & 1 & 2 & 1+\mathrm{j} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

根据式(6.3.3)，矩阵 A 的{1}-逆为

$$X = P\begin{bmatrix} I_r & O \\ O & L \end{bmatrix} =$$

$$\left[\begin{array}{cc:cccc} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]\left[\begin{array}{cc:c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hdashline 0 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & c \\ 0 & 0 & d \end{array}\right]\begin{bmatrix} -\frac{1}{2}\mathrm{j} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ \mathrm{j} & \frac{1}{3} & 1 \end{bmatrix} =$$

$$\begin{bmatrix} a\mathrm{j} & \frac{1}{3}a & a \\ -\frac{1}{2}\mathrm{j} & 0 & 0 \\ b\mathrm{j} & \frac{1}{3}b & b \\ 0 & -\frac{1}{3} & 0 \\ c\mathrm{j} & \frac{1}{3}c & c \\ d\mathrm{j} & \frac{1}{3}d & d \end{bmatrix}$$

其中 $a,b,c,d \in \mathbf{C}$ 任意.

(2）求{1，2}-逆：

当$X \in A\{1\}$时，对$X = P\begin{bmatrix} I_r & O \\ O & L \end{bmatrix}Q$取$L = O$，即$X = P\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}Q$，故$rankX = rankA = r$，则$X \in A\{1,2\}$

2. 利用满秩分解求广义逆矩阵

给定任意矩阵$A \in C_r^{m\times n}$，它的满秩分解为：$A = FG$，其中$F \in C_r^{m\times r}$，$G \in C_r^{r\times n}$.

•定理：

(1）$G^{(i)}F^{(1)} \in A\{i\} \qquad i = 1,2,4$

(2）$G^{(1)}F^{(i)} \in A\{i\} \qquad i = 1,2,3$

(3）$G^{(1)}F^{+} \in A\{1,2,3\}$，$G^{+}F^{(1)} \in A\{1,2,4\}$

(4）$A^{+} = G^{+}F^{(1,3)} = G^{(1,4)}F^{+}$

(5）$A^{+} = G^{+}F^{+} = G^{H}(GG^{H})^{-1}(F^{H}F)^{-1}F^{H} = G^{H}(F^{H}AG^{H})^{-1}F^{H}$

例题：

例 6.5 求例 6.4 中矩阵 A 的 Moore-Penrose 逆.

解 A 的一个满秩分解式为

$$A = \begin{bmatrix} 2j & 0 \\ 0 & -3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & \frac{1}{2} & 0 & 1-2j & -\frac{1}{2}j \\ 0 & 0 & 0 & 1 & 2 & 1+j \end{bmatrix}$$

于是有

$$F^{H}F = \begin{bmatrix} 8 & 2 \\ 2 & 10 \end{bmatrix},\quad GG^{H} = \begin{bmatrix} \frac{13}{2} & \frac{3}{2}-\frac{9}{2}j \\ \frac{3}{2}+\frac{9}{2}j & 7 \end{bmatrix}$$

从而 $A^{+} = G^{H}(GG^{H})^{-1}(F^{H}F)^{-1}F^{H} =$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ \frac{1}{2} & 0 \\ 0 & 1 \\ 1+2j & 2 \\ \frac{1}{2}j & 1-j \end{bmatrix}\frac{1}{46}\begin{bmatrix} 14 & -3(1-3j) \\ -3(1+3j) & 13 \end{bmatrix}\times$$

$$\frac{1}{38}\begin{bmatrix} 5 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} -2j & 0 & 2 \\ 0 & -3 & 1 \end{bmatrix} =$$

$$\frac{1}{1748}\begin{bmatrix} 0 & 0 & 0 \\ -18-146j & 78-108j & 120+18j \\ -9-73j & 39-54j & 60+9j \\ -90+56j & -165-2j & -1-81j \\ 94-70j & -36-6j & 82+96j \\ 39+137j & -138+177j & -91-20j \end{bmatrix}$$

例题：

⑰ $A=\begin{bmatrix}1&0&0&1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{bmatrix}$

(1) 求 A 的 $\{1\}$-逆和 $\{1,2\}$-逆

(2) 求 A^+

⑰ (1) $(A \vdots I)=\left(\begin{array}{cccc|cccc}1&0&0&1&1&0&0&0\\1&1&0&0&0&1&0&0\\0&1&1&0&0&0&1&0\\0&0&1&1&0&0&0&1\end{array}\right)$

$\xrightarrow{行}\left(\begin{array}{cccc|cccc}1&0&0&1&1&0&0&0\\0&1&0&-1&-1&1&0&0\\0&0&1&1&1&-1&1&0\\0&0&0&0&-1&1&-1&1\end{array}\right)$

故 $Q=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\1&-1&1&0\\-1&1&-1&1\end{pmatrix}$

$\operatorname{rank} A=3$

$X=P\begin{bmatrix}I_r&0\\0&L\end{bmatrix}Q$

$=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\left(\begin{array}{ccc|c}1&0&0&0\\0&1&0&0\\0&0&1&0\\\hline 0&0&0&a\end{array}\right)\begin{pmatrix}1&0&0&0\\-1&1&0&0\\1&-1&1&0\\-1&1&-1&1\end{pmatrix}$

$=\begin{pmatrix}1&0&0&0\\-1&1&0&0\\1&-1&1&0\\-a&a&-a&a\end{pmatrix}=A^{(1)}$

$A^{(1,2)}=P\begin{bmatrix}I_r&0\\0&0\end{bmatrix}Q$

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

(2) 由(1)知 $A=FG$

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

$A^+=G^H(GG^H)^{-1}(F^HF)^{-1}F^H$

$=G^H(F^HAG^H)^{-1}F^H$

$=\frac{1}{8}\begin{pmatrix}3&3&-1&-1\\-1&3&3&-1\\-1&-1&3&3\\3&-1&-1&3\end{pmatrix}$

6.4 广义逆矩阵的应用

对非齐次方程组 Ax=b，其中$A\in C^{m\times n}$，$b\in C^m$给定，而$x\in C^n$为待定向量。如果存在向量 x 使得方程组 Ax=b 成立，则称方程组相容，否则称为不相容或矛盾方程组。

（1）若方程组相容，其解可能有无穷多个，求出具有极小范数的解，即$\min\limits_{Ax=b}||x||$，满足该条件的解是唯一的，称之为极小范数解。

（2）若方程组不相容，则不存在通常意义下的解，需要求出极值问题，即$\min\limits_{x\in C^n}||Ax-b||$的解$x$，称这个极值问题为求矛盾方程组的最小二乘问题，相应的x称为矛盾方程组的最小二乘解。

（3）一般地，矛盾方程组的最小二乘解不唯一，但在最小二乘解的集合中，具有极小范数的解$\min\limits_{\min||Ax=b||}||x||$是唯一的，称之为极小范数最小二乘解。

1. 线性方程组的相容性、通解与广义{1}-逆

•定理：

设$A \in C^{m\times n}, B \in C^{p\times q}, D \in C^{m\times q}$，则矩阵方程$AXB = D$的相容（有解）的充要条件：$AA^{(1)}DB^{(1)}B = D$，在相容情况下矩阵方程的通解为：$\{X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)} | Y \in C^{n\times p}$任意$\}$

证：

充分性：

已知$\mathbf{AA^{(1)}DB^{(1)}B = D}$，显然有解$\mathbf{X = A^{(1)}DB^{(1)}}$

必要性：

已知$\mathbf{AXB = D}$有解，设某个解为$\mathbf{X}$，即

$$\mathbf{D = AXB = AA^{(1)}AXBB^{(1)}B = AA^{(1)}DB^{(1)}B}$$

现在证明通解：

（1）令$\mathbf{X = A^{(1)}DB^{(1)} + Y - A^{(1)}AYBB^{(1)}}$，代入$\mathbf{AXB = D}$

$$\mathbf{AXB = D + AYB - AYB = D}$$

∴ 集合中的元素为方程的解

（2）设$\mathbf{X}$为方程的解，即$\mathbf{AXB = D}$

$$\mathbf{X = A^{(1)}DB^{(1)} + X - A^{(1)}DB^{(1)} = A^{(1)}DB^{(1)} + X - A^{(1)}AXBB^{(1)}}$$

对应于集合中$\mathbf{Y = X}$的情况。

由上述证明可见：

（1）通解中两个$A^{(1)}$及两个$B^{(1)}$完全可以不同。

（2）通解集合中，不同的 Y 完全可能对应同一个解。

•推论：

(1)设$A \in C^{m\times n}$，$A^{(1)} \in A\{1\}$，则$A^{(1)} = \{A^{(1)} + Z - A^{(1)}AZAA^{(1)} | Z \in C^{n\times m}\}$

(2)Ax=b 相容的充要条件是$AA^{(1)}b = b$，且其通解为$x = A^{(1)}b + (I - A^{(1)}A)y$，其中$y \in C^n$任意。注意，该通解中的$A^{(1)}b$正是该方程的特解，而$(I - A^{(1)}A)y \in N(A)$是 Ax=0 的通解。

(3) Ax=b 相容的充要条件是$b \in R(A)$.

•例题：

例 6.9 取例 6.4 的矩阵 $\boldsymbol{A}$ 和

$$\boldsymbol{b} = \begin{bmatrix} 14+5\mathrm{j} \\ -15+3\mathrm{j} \\ 10-15\mathrm{j} \end{bmatrix}$$

求解线性方程组 $\boldsymbol{Ax} = \boldsymbol{b}$.

解 由例 6.4 得 $\boldsymbol{A}$ 的一个{1}- 逆为

$$\boldsymbol{A}^{(1)} = \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{2}\mathrm{j} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

容易验证 $\boldsymbol{AA}^{(1)}\boldsymbol{b} = \boldsymbol{b}$，所以方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 相容，且其通解为

$$\boldsymbol{x} = \boldsymbol{A}^{(1)}\boldsymbol{b} + (\boldsymbol{I} - \boldsymbol{A}^{(1)}\boldsymbol{A})\boldsymbol{y} =$$

$$\begin{bmatrix} 0 \\ \frac{5}{2}-7\mathrm{j} \\ 0 \\ 5-\mathrm{j} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & -1+2\mathrm{j} & \frac{1}{2}\mathrm{j} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & -1-\mathrm{j} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \xi_5 \\ \xi_6 \end{bmatrix}$$

其中，$\xi_1, \xi_2, \cdots, \xi_6 \in \mathrm{C}$ 任意.

式(6.4.10) 表明，利用某个{1}- 逆就可以解决相容方程组的求解问题. 反之，利用相容方程组的解，也可以给出{1}- 逆.

2. 相容线性方程组的极小范数解与广义{1,4}-逆

引理 1：方程组$Ax=b$相容（有解），则必存在唯一的极小范数解$\min\limits_{Ax=b}||x||$（2-范数），且该唯一解在$R(A^H)$中。

证：

设$\mathbf{x}$是方程$\mathbf{AX}=\mathbf{b}$的解，可将其分解为$\mathbf{x}=\mathbf{x}_0+\mathbf{y}$，其中

$$\mathbf{x}_0\in\mathbf{R}(\mathbf{A}^H)=\mathbf{N}^{\perp}(\mathbf{A})\to\mathbf{x}_0\perp\mathbf{N}(\mathbf{A}),\quad \mathbf{y}\in\mathbf{N}(\mathbf{A})$$

$$\|\mathbf{x}\|_2^2=\|\mathbf{x}_0+\mathbf{y}\|_2^2=(\mathbf{x}_0+\mathbf{y})^H(\mathbf{x}_0+\mathbf{y})=\mathbf{x}_0^H\mathbf{x}_0+\mathbf{y}^H\mathbf{y}\ =\|\mathbf{x}_0\|_2^2+\|\mathbf{y}\|_2^2\geq\|\mathbf{x}_0\|_2^2$$

而$\mathbf{A}\mathbf{x}=\mathbf{A}\mathbf{x}_0+\mathbf{A}\mathbf{y}=\mathbf{A}\mathbf{x}_0+0=\mathbf{A}\mathbf{x}_0=\mathbf{b}$

即：$\mathbf{x}_0$也是方程的解，也就是$\mathbf{R}(\mathbf{A}^H)$中存在$\mathbf{AX}=\mathbf{b}$的解。

假设$\mathbf{R}(\mathbf{A}^H)$中存在方程$\mathbf{AX}=\mathbf{b}$的两个解$\mathbf{x}_1$和$\mathbf{x}_2$，即$\mathbf{A}\mathbf{x}_1=\mathbf{A}\mathbf{x}_2=\mathbf{b}$

$\to\mathbf{A}\mathbf{x}_1-\mathbf{A}\mathbf{x}_2=0\to(\mathbf{x}_1-\mathbf{x}_2)\in\mathbf{N}(\mathbf{A})$ 同时$(\mathbf{x}_1-\mathbf{x}_2)\in\mathbf{N}^{\perp}(\mathbf{A})$

$\therefore\quad(\mathbf{x}_1-\mathbf{x}_2)\in\mathbf{N}(\mathbf{A})\bigcap\mathbf{N}^{\perp}(\mathbf{A})=\{0\}$

$\therefore\quad\mathbf{x}_1=\mathbf{x}_2$

也就是说在$\mathbf{R}(\mathbf{A}^H)$中方程$\mathbf{AX}=\mathbf{b}$只有唯一的解（若方程有解）

$\therefore$ 方程的任何其它解的 2-范数均大于$\mathbf{x}_0$的 2-范数

$\therefore$ $\mathbf{x}_0$是极小范数解

[得证]

由证明可知，方程$\mathbf{AX}=\mathbf{b}$在$\mathbf{R}(\mathbf{A}^H)$的解必定是极小范数解。

引理 2：A{1,4}由矩阵方程$XA=A^{(1,4)}A$的所有解 X 组成，其中$A^{(1,4)}$是 A 的某一个{1,4}-逆。
证：

一方面：上述方程的解一定是 A 的某一个{1,4}-逆,设 X 为其解

(i) $\mathbf{AXA}=\mathbf{AA}^{(1,4)}\mathbf{A}=\mathbf{A}$

(iv) $\mathbf{XA}=\mathbf{A}^{(1,4)}\mathbf{A}$ 是厄米矩阵

另一方面：A 的任何{1,4}-逆均满足上述方程，设 X 是 A 的{1,4}-逆，$\mathbf{A}^{(1,4)}$是某个给定的{1,4}-逆，X 满足(i) (iv)Penrose 方程

$$\mathbf{A}^{(1,4)}\mathbf{A}\overset{(i)}{=}\mathbf{A}^{(1,4)}\mathbf{AXA}\overset{(iv)}{=}\left(\mathbf{A}^{(1,4)}\mathbf{A}\right)^H(\mathbf{XA})^H=\left(\mathbf{XAA}^{(1,4)}\mathbf{A}\right)^H=(\mathbf{XA})^H=\mathbf{XA}$$

[得证]

以上引理说明，对于$\mathbf{X}\in\mathbf{A}\{1,4\}$，$\mathbf{XA}$是个不变量。

•**定理 1：** 设$A\in C^{m\times n}$，$A^{(1,4)}\in A\{1,4\}$，则$A\{1,4\}=\{A^{(1,4)}+Z(I-AA^{(1,4)})|Z\in C^{n\times m}\}$
证：

方程组$XA=A^{(1,4)}A$的通解为：

$$X=A^{(1,4)}AA^{(1,4)}+Y-YAA^{(1,4)},\qquad Y\in C^{n\times m}$$

令$Y=A^{(1,4)}+Z$，即得上式。

•**定理 2：** 设方程$Ax=b$相容，则$X=A^{(1,4)}b$是方程的极小范数解；反之，若对任意$b\in R(A)$，存在 X 使得 Xb 成为该方程的极小范数解，则$X\in A\{1,4\}$。

证:

先证前半部分:

推论 1 $\rightarrow \mathbf{A}^{(1)}\mathbf{b}$ 是 $\mathbf{Ax}=\mathbf{b}$ 的解

$$\begin{cases} \mathbf{A}^{(1,4)} \in \mathbf{A}\{1\} \rightarrow \mathbf{x}=\mathbf{A}^{(1,4)}\mathbf{b}\text{是方程的解} \\ \mathbf{A}^{(1,4)} \in \mathbf{A}\{4\} \rightarrow \mathbf{x}=\mathbf{A}^{(1,4)}\mathbf{b}=\mathbf{A}^{(1,4)}\mathbf{A}\mathbf{A}^{(1)}\mathbf{b}=\left(\mathbf{A}^{(1,4)}\mathbf{A}\right)^{H}\mathbf{A}^{(1)}\mathbf{b} \end{cases}$$

$$=\mathbf{A}^{H}(\mathbf{A}^{(1,4)})^{H}\mathbf{A}^{(1)}\mathbf{b} \in \mathbf{R}(\mathbf{A}^{H})$$

由引理 1 知，$\mathbf{A}^{(1,4)}\mathbf{b}$ 是极小范数解。

后半部分：

存在对于任意 $\mathbf{b} \in \mathbf{R}(\mathbf{A})$，均有 $\mathbf{Xb}$ 为 $\mathbf{Ax}=\mathbf{b}$ 的极小范数解，即 $\mathbf{Xb}=\mathbf{A}^{(1,4)}\mathbf{b}$ 为极小范数解。

因为 $\forall \mathbf{b} \in \mathbf{R}(\mathbf{A})$，上式都成立，将 $\mathbf{b}$ 依次取为 $\mathbf{A}$ 的各列，合起来得

$$\mathbf{XA}=\mathbf{A}^{(1,4)}\mathbf{A}$$

由引理 2 知 $\mathbf{X} \in \mathbf{A}\{1,4\}$

3. 矛盾方程组的最小二乘解与广义{1,3}-逆

设有一组实验数据(t_1，s_1)，(t_2，s_2)，……，(t_n，s_n)，希望由实验数据拟合给定规律，从而测出待测量的有关参数。

假定规律为：$s=c_1 t+c_2$，由于存在误差 $s_i \neq c_1 t_i + c_2 \quad (i=1,2,\cdots,n)$，令

$$\mathbf{A}=\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \vdots \\ t_n & 1 \end{bmatrix}, \mathbf{x}=\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \mathbf{b}=\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix},$$

则：Ax=b 实际无解，或者说矩阵方程 Ax=b 成为矛盾方程（不自洽、非相容），虽说无解，但在物理上看，我们需要而且也理当有“解”。怎么办？

一般处理是，定义一种目标函数，例如：

$$\mathbf{E}(c_1,c_2)=\sum_{i=1}^{n} w_i (s_i - c_1 t_i - c_2)^2 \qquad w_i>0\text{为加权系数}$$

使误差 $\mathbf{E}(c_1,c_2)$ 最小化。$w_i=1$($i=1\sim n$)时 $\mathbf{E}(c_1,c_2)=\|\mathbf{Ax}-\mathbf{b}\|_2^2$

对于矛盾方程 Ax=b，最小二乘法是求其“解”的一种方法，即求 $\min\limits_{x\in C^n} ||Ax-b||$，也就是求使 $||Ax-b||_2=min$ 的解。

引理：设$A\in C^{m\times n}$，A{1,3}由矩阵方程$AX=AA^{(1,3)}$的所有解 X 组成，其中$A^{(1,3)}\in A\{1,3\}$.

证：设 X 是其某个解，则

$$\text{(i)}\quad AXA=AA^{(1,3)}A=A\to X\in A\{1\}$$
$$\text{(iii)}\quad (AX)^H=(AA^{(1,3)})^H=AA^{(1,3)}=AX\to X\in A\{3\}$$

即方程的解必在 A{1,3}中。

设 X 为 A 的一个{1,3}-逆矩阵，则

$$\begin{aligned}AX=AA^{(1,3)}AX&\overset{\text{iii}}{=}\left(AA^{(1,3)}\right)^H(AX)^H\\&=\left(A^{(1,3)}\right)^H A^H X^H A^H\\&=\left(A^{(1,3)}\right)^H(AXA)^H\\&=\left(AA^{(1,3)}\right)^H=AA^{(1,3)}\end{aligned}$$

即，A 的{1,3}-逆矩阵必满足方程 $AX=AA^{(1,3)}$

$$\therefore\ A\{1,3\}=\left\{\text{方程}AX=AA^{(1,3)}\text{的所有解}\right\}$$

•**定理 1：** 设$A\in C^{m\times n}$，$A^{(1,3)}\in A\{1,3\}$，则$A\{1,3\}=\{A^{(1,3)}+(I-A^{(1,3)}A)Z|Z\in C^{n\times m}\}$.

证：

方程组$AX=AA^{(1,3)}$的通解为：

$$X=A^{(1,3)}AA^{(1,3)}+Y-A^{(1,3)}AY,\qquad Y\in C^{n\times m}$$

令$Y=A^{(1,3)}+Z$，即得上式。

•**定理 2：** 设$A\in C^{m\times n}$，$b\in C^m,A^{(1,3)}\in A\{1,3\}$，则$x=A^{(1,3)}b$是方程组 Ax=b 的最小二乘解。反之，设$X\in C^{n\times m}$，若对所有$b\in C^m$，$x=Xb$都是方程组 Ax=b 的最小二乘解，则$X\in A\{1,3\}$.

证：

$$Ax-b=(Ax-P_{R(A)}b)+(P_{R(A)}b-b)$$
$$(Ax-P_{R(A)}b)\in R(A),(P_{R(A)}b-b)=-(I-P_{R(A)})b=-P_{R^\perp(A)}b\in R^\perp(A)$$

所以，$\|Ax-b\|_2^2=\|Ax-P_{R(A)}b\|_2^2+\|P_{R(A)}b-b\|_2^2\ge\|b-P_{R(A)}b\|_2^2$，

故$\|Ax-b\|_2^2$取得极小值的条件是 x 为方程$Ax=P_{R(A)}b$ 的解。任取一个$A^{(1,3)}\in A\{1,3\}$，我们知道$AA^{(1,3)}=P_{R(A)}$。而对于$x=A^{(1,3)}b$，有$Ax=AA^{(1,3)}b=P_{R(A)}b$

方程$Ax=AA^{(1,3)}b$的通解为：

$$\begin{aligned}x&=\left\{A^{(1,3)}AA^{(1,3)}b+y-A^{(1,3)}Ay\,\middle|\,y\in C^n\right\}\qquad y=A^{(1,3)}b+z\\&=\left\{A^{(1,3)}b+(I-A^{(1,3)}A)z\,\middle|\,z\in C^n\right\}\end{aligned}$$

显然最小二乘解并不一定都具有 $A^{(1,3)}b$ 的形式。

反之，若对于$\forall b\in C^m,x=Xb$均使$Ax=P_{R(A)}b=AA^{(1,3)}b$，

即$\forall b$,有$AXb=AA^{(1,3)}b\to AX=AA^{(1,3)}\to X\in A\{1,3\}$

推论：x 是方程 Ax＝b 的最小二乘解的充要条件是，x 为方程$A^HAx=A^Hb$的解。该方程组称为矛盾方程组的法方程组（正规方程组）。

证 因为

$$b=P_{R(A)}b+(I-P_{R(A)})b=P_{R(A)}b+P_{N(A^H)}b$$

而由式(6.4.17)知，x 是方程组(6.4.1)的最小二乘解的充要条件是

$$Ax-b=Ax-(P_{R(A)}b+P_{N(A^H)}b)=-P_{N(A^H)}b\in N(A^H)$$

所以 $A^H(Ax-b)=0$. 证毕

4. 矛盾方程组的极小范数最小二乘解与广义逆矩阵A^+

虽然最小二乘解一般不唯一，但是极小范数最小二乘解$\min_{\min\|Ax=b\|}\|x\|$是唯一的，并且它可由 Moore-Penrose 逆A^+表出。

•定理 1：

设$\mathbf{A}\in\mathbf{C}^{m\times n},\mathbf{b}\in\mathbf{C}^{m}$，则 $x=A^{\dagger}b$ 是方程 $Ax=b$ 的极小范数最小二乘解。反之，若存在$\mathbf{X}\in\mathbf{C}^{n\times m}$，若对于所有$\mathbf{b}\in\mathbf{C}^{m}$，$x=Xb$ 均成为方程 $Ax=b$ 的极小范数最小二乘解，则 $X=A^{\dagger}$。

证：

最小二乘解满足 $Ax=AA^{(1,3)}b$，其极小范数解唯一，且为$\mathbf{x}=\mathbf{A}^{(1,4)}(\mathbf{A}\mathbf{A}^{(1,3)}\mathbf{b})=\mathbf{A}^{\dagger}\mathbf{b}$，

反之，$\forall\mathbf{b}\in\mathbf{C}^{m}$，$\mathbf{x}\mathbf{b}$ 均成为唯一的极小范数最小二乘解$\mathbf{A}^{\dagger}\mathbf{b}$，所以：$X=A^{\dagger}$。

•定理 2：矩阵方程 AXB=D 的极小范数最小二乘解唯一，且为 X=A^+DB^+

例:

例 6.10 取例 6.4 的矩阵 $\boldsymbol{A}$ 和

$$\boldsymbol{b}=\begin{bmatrix}-j\\1\\1\end{bmatrix}$$

求方程组(6.4.1)的极小范数最小二乘解.

解 由例 6.5 的结果知，方程组(6.4.1)的极小范数最小二乘解为

$$\boldsymbol{x}=\boldsymbol{A}^{+}\boldsymbol{b}=\frac{1}{874}\begin{bmatrix}0\\26-36j\\13-18j\\-55-9j\\-12-2j\\-46+59j\end{bmatrix}$$

5. 全面最小二乘法

一组测量数据(t_i,s_i),欲拟和直线：$s=c_1t+c_2$

最小二乘法采取目标函数：$E(c_1,c_2)=\sum_{i=1}^{n}|s_i-c_1t_i-c_2|^2=\min$

它隐含了在测量中，t_i是精确测量的，只有s_i才测得不准确，而在实际测量中，t_i，s_i都是无法准确测量的，因此，采用法向回归更有可能。

点(t_i,s_i)到直线$s=c_1t+c_2$的距离为：

$$\frac{1}{\sqrt{1+c_1^2}}|s_i-c_1t_i-c_2|$$

故法向回归的目标函数为：

$$E(c_1,c_2)=\left(\frac{1}{\sqrt{1+c_1^2}}\right)^2\sum_{i=1}^{n}|s_i-c_1t_i-c_2|^2=\min$$

$$\frac{\partial E}{\partial c_2}=\frac{1}{1+c_1^2}\sum_{i=1}^{n}(-2)(s_i-c_1t_i-c_2)=0\rightarrow c_2=\frac{1}{n}\sum_{i=1}^{n}s_i-c_1t_i$$

$$\frac{\partial E}{\partial c_1}=-\frac{2c_1}{\left(1+c_1^2\right)^2}\sum_{i=1}^{n}\left(s_i-c_1t_i-c_2\right)^2+\frac{2}{1+c_1^2}\sum_{i=1}^{n}\left(-t_i\right)\left(s_i-c_1t_i-c_2\right)$$

$$=\frac{2}{1+c_1^2}\sum_{i=1}^{n}\left(c_1c_2-c_1s_i-t_i\right)\left(s_i-c_1t_i-c_2\right)$$

$$=\frac{2}{1+c_1^2}\left\{c_1c\sum_{i=1}^{n}\left(s_i-c_1t_i-c_2\right)-c_1\sum_{i=1}^{n}s_i\left(s_i-c_1t_i-c_2\right)-\sum_{i=1}^{n}t_i\left(s_i-c_1t_i-c_2\right)\right\}$$

$$=\frac{-2}{1+c_1^2}\left\{c_1\sum_{i=1}^{n}s_i\left(s_i-c_1t_i-c_2\right)+\sum_{i=1}^{n}t_i\left(s_i-c_1t_i-c_2\right)\right\}=0$$

将 c_2 代入之，可得

$$\begin{cases} c_1=\dfrac{\left(l_{ss}-l_{tt}\right)+\sqrt{\left(l_{ss}-l_{tt}\right)^2+4l_{st}^2}}{2l_{st}} \\ c_2=\bar{s}-c_1\bar{t} \end{cases}$$

其中

$$\begin{cases} \bar{s}=\dfrac{1}{n}\sum_{i=1}^{n}s_i \\ \bar{t}=\dfrac{1}{n}\sum_{i=1}^{n}t_i \\ l_{ss}=\sum_{i=1}^{n}\left(s_i-\bar{s}\right)^2=\sum_{i=1}^{n}s_i^2-\dfrac{1}{n}\left(\sum_{i=1}^{n}s_i\right)^2, \\ l_{st}=\sum_{i=1}^{n}\left(s_i-\bar{s}\right)\left(t_i-\bar{t}\right)=\sum_{i=1}^{n}s_it_i-\dfrac{1}{n}\left(\sum_{i=1}^{n}s_i\right)\left(\sum_{i=1}^{n}t_i\right) \\ l_{tt}=\sum_{i=1}^{n}\left(t_i-\bar{t}\right)^2=\sum_{i=1}^{n}t_i^2-\dfrac{1}{n}\left(\sum_{i=1}^{n}t_i\right)^2 \end{cases}$$

作为比较，最小二乘法 $\sum_{i=1}^{n}\left|s_i-c_1t_i-c_2\right|^2=\min$ 给出

$$\begin{cases} c_1=l_{st}/l_{tt} \\ c_2=\bar{s}-c_1\bar{t} \end{cases}$$

例 1. 7 点测量 $\left(t_i,s_i\right)=\left(0,3.1\right),\left(0.5,3.9\right),\left(1,5.2\right),\left(1.5,6.0\right),\left(2,6.9\right),\left(2.5,8.0\right),\left(3.0,9.1\right)$

拟合直线 $c_1t+c_2=s$

解：计算结果 $\bar{t}=1.5,\bar{s}\approx 6.02857,l_{tt}=7,l_{ss}=27.8743,l_{st}=13.95$

最小二乘法给出 $c_1=1.99286,c_2=3.03929$

全面最小二乘法（法向回归）给出 $c_1=1.99709,c_2=3.03293$

测量数据误差小，分布合理时，两种方法效果非常接近。

全面最小二乘法：

当方程 $\mathbf{Ax}=\mathbf{b}$ 成为矛盾方程时，采用最小二乘法求解的观点实际上认为 b 存在误差，而 A 不存在误差，故应有 $\boldsymbol{\varepsilon}$，使得 $\mathbf{Ax}=\mathbf{b}+\boldsymbol{\varepsilon}$； $\boldsymbol{\varepsilon}$ 应尽量小以使得不至于严重得破坏方程 $\rightarrow \|\boldsymbol{\varepsilon}\|_2=\mathbf{min}$.

全面最小二乘法采取如下观点解决矛盾方程的问题，不仅 b 存在误差，A 也存在误差，故存在 E 和 $\boldsymbol{\varepsilon}$，使 $(\mathbf{A}+\mathbf{E})\mathbf{x}=\mathbf{b}+\boldsymbol{\varepsilon}$；$\mathbf{E}$、$\boldsymbol{\varepsilon}$ 也应该尽量小，以使得不至于严重偏离原方程 $\rightarrow \|[\mathbf{E}\,|\,\boldsymbol{\varepsilon}]\|_F=\mathbf{min}$.

$$(\mathbf{A}+\mathbf{E})\mathbf{x}=\mathbf{b}+\boldsymbol{\varepsilon} \Leftrightarrow ([\mathbf{A}\,|\,\mathbf{b}]+[\mathbf{E}\,|\,\boldsymbol{\varepsilon}])\begin{bmatrix}\mathbf{x}\\-1\end{bmatrix}=0$$

记 $\mathbf{c}=[\mathbf{A}\,|\,\mathbf{b}],\Delta=[\mathbf{E}\,|\,\boldsymbol{\varepsilon}],\mathbf{v}=\begin{bmatrix}\mathbf{x}\\-1\end{bmatrix}$，则全面最小二乘解即求方程 $(\mathbf{c}+\Delta)\mathbf{v}=0$ 的非零解 v，且 v 的最后分量不能为零，而其中 Δ 应满足 $\|\Delta\|_F=\mathbf{min}$.

引理：设 $\mathbf{z}\in C_r^{m\times n}$，且存在奇异值分解，$\mathbf{Z}=\mathbf{U}\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0\end{bmatrix}_{m\times n}\mathbf{V}^H$，其中 $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0$。又设 $\mathbf{Y}=\mathbf{U}\begin{bmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_s & \\ & & & 0\end{bmatrix}_{m\times n}\mathbf{V}^n \quad (s<r)$，则：

$$\|\mathbf{Z}-\mathbf{Y}\|_F=\min_{\substack{z\in C^{m\times n}\\ \mathrm{rank}z=s}}\|\mathbf{Z}-\mathbf{z}\|_F$$

•定理 1：

设 $\mathbf{A}\in C_n^{m\times n},[\mathbf{A}\,|\,\mathbf{b}]\in C_{n+1}^{m\times(n+1)}$ 具有如下的奇异值分解：

$$\mathbf{C}=\mathbf{U}\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_{n+1}\\ & 0 & \end{bmatrix}\mathbf{V}^H,(\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_{n+1})$$

则使方程 $(\mathbf{C}+\Delta)\mathbf{v}=0$ 具有非零解，且 F 范数最小的 Δ 存在，并且 $\|\Delta\|_F=\sigma_{n+1}$

证明：方程 $(\mathbf{C}+\Delta)\mathbf{v}=0$ 要有非零解，必须 $\mathbf{rank}(\mathbf{C}+\Delta)<\mathbf{n}+1$，故由引理知

$$\begin{aligned}\min\|\Delta\|_F &= \min_{\mathrm{rank}(C+\Delta)<n+1}\|C-(C+\Delta)\|_F\\ &= \min_{\mathrm{rank}(C+\Delta)=n}\|C-(C+\Delta)\|=\sigma_{n+1}\end{aligned}$$

显然满足

$$\Delta=\mathbf{U}\begin{bmatrix}0 & & \\ & \ddots & \\ & 0 & \\ & & \sigma_{n+1}\\ & O & \end{bmatrix}\mathbf{V}^H$$

•**定理 2：**

设 σ_{n+1} 为 C 的 n-k+1 重奇异值，且 $v_{k+1}, v_{k+2}, \cdots v_{n+1}$ 相应的为 $C^H C$ 的属于 $(n-k+1)$ 重特征值 σ_{n+1}^2 的正交归一特征向量，则使方程 $(C+\Delta)v=0$ 具有非零的解且 F 范数最小的 Δ 为

$$\Delta = -Cv_s v_s^H / v_s^H v_s$$

而方程的解则为 $v = v_s$，其中 $v_s \in S_c \overset{\Delta}{=} \mathrm{span}\{v_{k+1}, v_{k+2}, \cdots v_{n+1}\}$

证：(1) 显然 $CC^H v_s = \sigma_{n+1}^2 v_s$

$$\|\Delta\|_F^2 = \|Cv_s v_s^H\|_F^2 \Big/ (v_s^H v_s)^2 = tr(v_s v_s^H C^H C v_s v_s^H) \Big/ (v_s^H v_s)^2$$

$$= \frac{\sigma_{n+1}^2}{(v_s^H v_s)^2} tr(v_s v_s^H v_s v_s^H) = \frac{\sigma_{n+1}^2}{v_s^H v_s} tr(v_s v_s^H) = \frac{\sigma_{n+1}^2}{v_s^H v_s} tr(v_s^H v_s)$$

$$= \sigma_{n+1}^2$$

(2) $$(C+\Delta)v_s = Cv_s - \frac{Cv_s v_s^H v_s}{v_s^H v_s} = 0$$

(3) $\forall v \in C^{n+1}$, 有

$$v = \left(\sum_{i=1}^{n-k+1} v_{k+i} v_{k+i}^H\right) v + \left(I_{n+1} - \sum_{i=1}^{n-k+1} v_{k+i} v_{k+i}^H\right) v \overset{\Delta}{=} v_s + v_T$$

虽然，$\left(C - C\dfrac{vv^H}{v^H v}\right)v = 0$, 但 $\left\|-C\dfrac{vv^H}{v^H v}\right\|_F > \sigma_{n+1}$

$$\because \quad C^H C v = C^H C(v_s + v_T) = \sigma_{n+1}^2 v_s + C^H C v_T > \sigma_{n+1}^2 v_s + \sigma_{n+1}^2 v_T$$

$$= \sigma_{n+1}^2 (v_s + v_T) = \sigma_{n+1}^2 v$$

$$\therefore \quad \left\|C\frac{vv^H}{v^H v}\right\|_F^2 = \frac{1}{(v^H v)^2} tr(vv^H C^H C vv^H) > \frac{\sigma_{n+1}^2}{(v^H v)^2} tr(vv^H vv^H) = \sigma_{n+1}^2$$

•**定理 3：**

在定理 2 的条件下，全面最小二乘解存在的充要条件为：向量 $e_{n=1} = \begin{bmatrix} \underbrace{0 \quad \cdots \quad 0}_{n个} & 1 \end{bmatrix}^T$ 不正交于 S_c。此时，$\forall v \in \left\{ q = \begin{bmatrix} y \\ \alpha \end{bmatrix} \middle| q \in S_c, \alpha \neq 0 \right\}$，则最小二乘解为 $x = \dfrac{1}{\alpha} y$

说明：(1) 最小二乘解一定存在，但全面最小二乘解不一定

(2) 存在全面最小二乘解时，若 σ_{n+1} 为 C 的单重奇异值，全面最小二乘解唯一，否则，解不唯一

例 2. 采用全面最小二乘法重新研究（上例）法向回归的问题

$$\mathbf{Ax}=\mathbf{b}\qquad \mathbf{A}=\begin{bmatrix} t_1 & 1 \\ t_2 & 1 \\ \vdots & \\ t_n & 1 \end{bmatrix}\qquad \mathbf{x}=\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}\qquad \mathbf{b}=\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

$$\mathbf{C}=\begin{bmatrix} t_1 & 1 & s_1 \\ t_2 & 1 & s_2 \\ \vdots & \vdots & \vdots \\ t_n & 1 & s_n \end{bmatrix}$$

$$\mathbf{C}^T\mathbf{C}=\begin{bmatrix} t_1 & t_2 & \cdots & t_n \\ 1 & 1 & \cdots & 1 \\ s_1 & s_2 & \cdots & s_n \end{bmatrix}\begin{bmatrix} t_1 & 1 & s_1 \\ t_2 & 1 & s_2 \\ \vdots & \vdots & \vdots \\ t_n & 1 & s_n \end{bmatrix}=\begin{bmatrix} \sum t_i^2 & \sum t_i & \sum s_i t_i \\ \sum t_i & n & \sum s_i \\ \sum s_i t_i & \sum s_i & \sum s_i^2 \end{bmatrix}$$

$$\mathbf{C}^T\mathbf{C}=\begin{bmatrix} 22.75 & 10.5 & 77.25 \\ 10.5 & 7 & 42.2 \\ 77.25 & 42.2 & 282.28 \end{bmatrix},\ \lambda=309.7754,2.249389,0.0051987257$$

$$v_3=[\underset{\substack{\downarrow \\ c_1}}{1.990944}\quad \underset{\substack{\downarrow \\ c_2}}{3.044416}\quad -1]^T$$

（对应 λ_3）

与法向回归结果并不相同，但亦十分接近。值得注意的是 $\bar{s}\neq c_1\bar{t}+c_2$（全面最小二乘解）