第六章　广义逆矩阵

引　言　不可逆矩阵的逆矩阵

我们知道, 如果线性方程组 $Ax = b$ 的系数矩阵 $A_{m\times n}$ 可逆 (此时 $m = n$), 则该方程组有唯一解 $x = A^{-1}b$. 实际上, 如果 A 是列满秩的矩阵, 则该方程组的唯一解或者最小二乘解也可以由 A 的正交三角分解 $A = UR$ 得到:

$$x = R^{-1}U^*b \tag{6.0.1}$$

(或者由原方程的正规化方程 $A^*Ax = A^*b$ 求得, 因为此时系数矩阵 A^*A 可逆.) 如果记 $A^{\ominus} = R^{-1}U^*$, 则有 $A^{\ominus}A = I_n$, 因此, 如果 A 是方阵, 则 $A^{\ominus}$ 确为 A 的逆矩阵. 但若 $n < m$, 则有

$$AA^{\ominus} = \begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix}_{m\times m} \tag{6.0.2}$$

由于 A 的秩为 $n < m$, 自然不能要求 $AA^{\ominus} = I_m$, 因此, 矩阵 $A^{\ominus} = R^{-1}U^*$ 仍可以看作是 A 的某种意义下的逆矩阵. 能否对任意的矩阵定义这样的"逆矩阵"$A^{\ominus}$ 呢? 如果 A 不是行或列满秩的, 则对任意矩阵 B, 矩阵 AB 与 BA 均不是满秩矩阵, 因此需要对"逆矩阵"的概念作一些拓展. 为了得到有意义的拓展, 我们将任意矩阵 $A_{m\times n}$ 看作是 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性变换. 此时线性变换 A 的"逆" – $A^{\ominus}$ 是从 $\mathbb{F}^m$ 到 $\mathbb{F}^n$ 的线性变换. 因此乘积 (即线性变换的复合)$A^{\ominus}A$(或 $AA^{\ominus}$) 是从 $\mathbb{F}^n$ 到 $\mathbb{F}^n$(或 $\mathbb{F}^m$ 到 $\mathbb{F}^m$) 的线性变换. 自然应该要求 $A^{\ominus}A$ 尽可能接近 $\mathbb{F}^n$ 的恒等变换, 因此 $A^{\ominus}A$ 应该将尽可能多的向量 $x \in \mathbb{F}^n$ 固定, 但 $\mathrm{Im}(A^{\ominus}A) \subseteq R(A^{\ominus})$. 因此最多要求 $A^{\ominus}A$ 在 $A^{\ominus}$ 的列空间 $R(A^{\ominus})$ 上是恒等变换. 同理可知最多要求 $AA^{\ominus}$ 在 A 的列空间 $R(A)$ 上是恒等变换. 这样的线性变换 $A^{\ominus}A$ 的确存在 (一般还有无穷多个, 为什么?), 其中最简单者当属正交投影变换. 基于上述理由, Moore[55] 于1920年提出了 $m \times n$ 矩阵的**广义逆矩阵** 的概念. 若 $A \in \mathbb{C}^{m\times n}$, Moore意义下的广义逆为满足

$$AX = P_{R(A)} \quad \text{与} \quad XA = P_{R(X)} \tag{6.0.3}$$

的矩阵 $X \in \mathbb{C}^{n\times m}$, 这里 P_L 表示在子空间 L 上的正交投影矩阵. 但公式 (6.0.3) 的含义不容易理解和应用, 因此Moore给出的广义逆矩阵一直未被重视. 1955年, 英国剑桥大学的博士研究生 Penrose[56] 给出了广义逆矩阵的下述定义: 如果矩阵 $X \in \mathbb{C}^{n\times m}$ 满足下述方程组

$$\begin{array}{ll} (1) & AXA = A \\ (2) & XAX = X \\ (3) & (AX)^* = AX \\ (4) & (XA)^* = XA \end{array} \tag{6.0.4}$$

则称 X 是矩阵 $A \in \mathbb{C}^{m\times n}$ 的广义逆矩阵.

方程组 (6.0.3) 与 (6.0.4) 分别称为**Moore方程组** 与**Penrose方程组**. 注意, Moore方程组虽然只有两个方程, 却涉及了四个矩阵, 其中除了 A 之外, 其余三个均是未知的 (尽管矩阵 $P_{R(A)}$ 仅与 A 有关), 而Penrose方程组尽管有四个方程, 但却仅涉及两个矩阵! 因此Penrose方程组更易于研究和应用. 历史的进展正是如此, 自Penrose的广义逆矩阵的论文发

[55]Eliakim Hastings Moore(1862-1932), 美国数学家, 是二十世纪初美国数学的奠基人, 曾任美国数学会主席.

[56]Sir Roger Penrose(1931-), 著名英国数学家, 物理学家, 哲学家. 1988年Wolf奖得主. 与Stephen Hawking (霍金) 合作证明了广义相对论的奇点存在性.

表以来, 广义逆矩阵迅速成为矩阵理论的研究热点, 并开始广泛地应用于数理统计, 多元分析, 最优化理论, 控制论, 网络理论等众多学科.

通常, 将满足Penrose方程组中等式 $i_1, \cdots, i_j$ 的矩阵 X 称为矩阵 A 的 $\{i_1, \cdots, i_j\}$- 逆, 比如满足第一及第三个等式的矩阵 X 称为 A 的 $\{1,3\}$- 逆, 记为 $A^{(1,3)}$; 而矩阵 A 的**Moore-Penrose广义逆**$A^\dagger = A^{(1,2,3,4)}$. 因此, 一个矩阵 A 共有 (至少) 15种广义逆, 其中研究最深, 应用最广的当属Moore-Penrose广义逆 $A^\dagger$ 以及 $\{1\}$- 逆 $A^{(1)}$, 通常记为 A^-. 本章主要介绍 $m \times n$ 矩阵的Moore-Penrose广义逆 $A^\dagger$, $\{1\}$- 逆 A^-, $\{1,3\}$- 逆与 $\{1,4\}$- 逆, 以及这几类广义逆矩阵的一些应用.

第一节 投影矩阵与Moore-Penrose广义逆矩阵

本节我们将证明Moore方程组与Penrose方程组是等价的, 因此矩阵的Moore广义逆与Penrose广义逆实际上是相同的. 为此需要研究Moore方程组中的投影矩阵 $P_{R(A)}$ 与 $P_{R(X)}$. 回顾第二章, 若 $\mathbb{C}$ 上 n 阶方阵 A 满足 $A^2 = A$, 则有

$$\mathbb{C}^n = R(A) \oplus N(A),$$

因此 A 表示 $\mathbb{C}^n$ 中的投影算子, 即 A 在 $R(A)$ 上的限制 (即诱导变换) 为恒等变换, 而 A 在 $N(A)$ 上的限制为零变换. 这时称 A 为 $\mathbb{C}^n$ 沿 $N(A)$ 到 $R(A)$ 上的投影矩阵. 进一步, 若 $R(A)^\perp = N(A)$, 即 A 的核空间与像空间正交, 则称 A 是正交投影矩阵, 因为 A 对应的线性变换 $x \mapsto Ax$ 是正交投影变换. 由于 $R(A)^\perp = N(A^*)$, 故若 $A^2 = A$ 且 $A^* = A$, 则 A 是正交投影矩阵. 容易证明, 这两个条件也是充分的, 即有下述结论 (对照第二章定理2.3.2)

定理 6.1.1 矩阵 A 为正交投影矩阵 $\iff$ $A^2 = A$, $A^* = A$.

一般地, 设 L 与 M 是 $\mathbb{C}^n$ 的两个互补的子空间, 即 $\mathbb{C}^n = L \oplus M$, 则将线性空间 $\mathbb{C}^n$ 的沿子空间 M 在子空间 L 上的投影变换记为 $P_{L,M}$. 如果还有 $M = L^\perp$, 则将 $P_{L,M}$ 简记为 P_L. 投影变换 $P_{L,M}$ 在标准基下的矩阵也记为 $P_{L,M}$(这不会引起混淆).

例 6.1.1 设 $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$, 容易验证这是一个投影矩阵, 它所对应的投影变换 $P_{L,M}$ 是

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x+y \\ 0 \end{pmatrix}.$$

因此 $L = \mathrm{Im}(P_{L,M}) = R(A) = \{(x,y)^T \in \mathbb{C}^2 \,|\, y = 0\}$, 而 $M = \mathrm{Ker}(P_{L,M}) = N(A) = \{(x,y)^T \in \mathbb{C}^2 \,|\, x + y = 0\}$.

投影矩阵 $P_{L,M}$ 的计算方法 如下.

设 $\dim L = r$, 则 $\dim M = n - r$. 设 $\{\alpha_1, \alpha_2, ..., \alpha_r\}$ 与 $\{\alpha_{r+1}, \alpha_{r+2}, ..., \alpha_n\}$ 分别是子空间 L 与 M 的一组基. 则

$$P_{L,M}\alpha_i = \alpha_i, \quad 1 \le i \le r,$$
$$P_{L,M}\alpha_j = 0, \quad r+1 \le j \le n.$$

令

$$X = (\alpha_1, \alpha_2, ..., \alpha_r), \quad Y = (\alpha_{r+1}, \alpha_{r+2}, ..., \alpha_n),$$

则 X, Y 分别为 $n\times r$ 和 $n\times(n-r)$ 矩阵, 且

$$P_{L,M}X=X,\quad P_{L,M}Y=0.$$

因而

$$P_{L,M}(X,Y)=(X,0).$$

因 (X,Y) 为 n 阶满秩矩阵, 所以

$$P_{L,M}=(X,0)(X,Y)^{-1}. \tag{6.1.1}$$

例 6.1.2 设 L 是由向量 $(1,-1)^T$ 张成的 $\mathbb{R}^2$ 的子空间, M 是由向量 $(1,0)^T$ 张成的 $\mathbb{R}^2$ 的子空间, 则 $\mathbb{R}^2$ 上沿着 M 到 L 上的投影矩阵为

$$P_{L,M}=\begin{pmatrix}1&0\\-1&0\end{pmatrix}\begin{pmatrix}1&1\\-1&0\end{pmatrix}^{-1}=\begin{pmatrix}1&0\\-1&0\end{pmatrix}\begin{pmatrix}0&-1\\1&1\end{pmatrix}=\begin{pmatrix}0&-1\\0&1\end{pmatrix}.$$

例 6.1.2 中的投影矩阵 $\begin{pmatrix}0&-1\\0&1\end{pmatrix}$ 不是正交投影矩阵, 因为它不是对称的.

由 定理 6.1.1 立即可以得到正交投影矩阵的下述性质, 证明见习题 4.

例 6.1.3 设 A 是正交投影矩阵, 则方程组 $Ax=b$ 的解或最小二乘解均为 Ab. 等价地, 向量 b 在子空间 $R(A)$ 中的正交投影向量为 Ab. 实际上, 我们知道, 方程组的解应该是 $A^*Ax=A^*b$ 的解, 而这正是方程 $Ax=Ab$, 所以 $x=Ab$ 是解.

正交投影矩阵 P_L 的计算方法 如下.

设 L 的一组基为 $\alpha_1,\alpha_2,\cdots,\alpha_r$, 任取 $L^\perp$ 的一组基 $\{\alpha_{r+1},\alpha_{r+2},\cdots,\alpha_n\}$. 令

$$X=(\alpha_1,\alpha_2,\cdots,\alpha_r),\ Y=(\alpha_{r+1},\alpha_{r+2},\cdots,\alpha_n).$$

则 X, Y 分别为 $n\times r$ 和 $n\times(n-r)$ 矩阵且 $X^*Y=Y^*X=0$, 利用公式 (6.1.1) 可得

$$\begin{aligned}P_L&=(X,0)(X,Y)^{-1}=(X,0)(X,Y)^{-1}((X,Y)^*)^{-1}(X,Y)^*\\&=(X,0)\left(\begin{pmatrix}X^*\\Y^*\end{pmatrix}(X,Y)\right)^{-1}(X,Y)^*\\&=(X,0)\begin{pmatrix}X^*X&0\\0&Y^*Y\end{pmatrix}^{-1}\begin{pmatrix}X^*\\Y^*\end{pmatrix}=X(X^*X)^{-1}X^*.\end{aligned}$$

故有公式

$$P_L=X(X^*X)^{-1}X^* \tag{6.1.2}$$

其中, X 的列是子空间 L 的任意一组基. 特别, 若 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是 L 的一组标准正交基, 则 $X^*X=I$, 于是上面的公式变为

$$P_L=XX^* \tag{6.1.3}$$

例 6.1.4 在 $\mathbb{R}^3$ 中 L 为由 $\alpha_1=(1,1,0)^T$ 和 $\alpha_2=(0,1,0)^T$ 生成的子空间, 求正交投影矩阵 P_L 和向量 $\alpha=(1,0,1)^T$ 在 L 上的正交投影向量.

解 利用公式 (6.1.2) 可得

$$X=(\alpha_1,\alpha_2)=\begin{pmatrix}1&0\\1&1\\0&0\end{pmatrix},$$

$$X^*X=\begin{pmatrix}1&1&0\\0&1&0\end{pmatrix}\begin{pmatrix}1&0\\1&1\\0&1\end{pmatrix}=\begin{pmatrix}2&1\\1&1\end{pmatrix},$$

$$(X^*X)^{-1}=\begin{pmatrix}1&-1\\-1&2\end{pmatrix}.$$

所以

$$\begin{aligned}P_L&=X(X^*X)^{-1}X^*\\&=\begin{pmatrix}1&0\\1&1\\0&0\end{pmatrix}\begin{pmatrix}1&-1\\-1&2\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&0\end{pmatrix}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}.\end{aligned}$$

由 例6.1.3 可知, α 在 L 上的正交投影向量为

$$P_L\alpha=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

(实际上 $P_L\alpha$ 无需计算即可“猜”到, 为什么?)

定义 6.1.1 设矩阵 $A\in\mathbb{C}^{m\times n}$, 若矩阵 $X\in\mathbb{C}^{n\times m}$ 满足Penrose方程组 (6.0.4), 则称 X 为 A 的一个Penrose广义逆 (矩阵).

显然, 若 A 为非奇异矩阵, 则 A^{-1} 是 A 的一个Penrose广义逆. 任意 $m\times n$ 阶零矩阵0的一个Penrose广义逆是 $n\times m$ 阶0矩阵.

例 6.1.5 (Penrose方程的几何意义)将矩阵 $A\in\mathbb{C}^{m\times n}$ 看作是从 $\mathbb{C}^n$ 到 $\mathbb{C}^m$ 的线性变换, 则矩阵 $X\in\mathbb{C}^{n\times m}$ 是从 $\mathbb{C}^m$ 到 $\mathbb{C}^n$ 的线性变换. 因此Penrose方程 (1) 与 (2) 隐含线性变换 AX 与 XA 均是幂等变换, 从而分别是 ($\mathbb{C}^m$ 的子空间) 列空间 $R(A)$ 与 ($\mathbb{C}^n$ 的子空间)$R(X)$ 上的恒等变换. 因此, 它们分别是列空间 $R(A)$ 与 $R(X)$ 上的投影变换. Penrose方程 (3) 与 (4) 表示线性变换 AX 与 XA 均是Hermite变换, 因此Penrose方程组表示线性变换 AX 与 XA 分别是列空间 $R(A)$ 与 $R(X)$ 上的正交投影变换, 这正是Moore方程组 (6.0.3) 的含义!

定义 6.1.2 设矩阵 $A\in\mathbb{C}^{m\times n}$. 若矩阵 $X\in\mathbb{C}^{n\times m}$ 满足Moore方程组 (6.0.3), 则称 X 为 A 的一个Moore广义逆矩阵.

注意, 单位矩阵 I_n 是整个空间 $\mathbb{C}^n$ 上的正交投影变换, 于是在 A 可逆时, 我们知道Moore广义逆矩阵与Penrose广义逆矩阵是相同的. 为了一般地证明这个结论, 我们先讨论Penrose广义逆矩阵的基本性质.

定理 6.1.2 对任意 $A\in\mathbb{C}^{m\times n}$, A 的Penrose广义逆存在并且唯一, 记为 $A^\dagger$.

证 若 A 为零矩阵, 可取 X 也为零矩阵. 现设 $A \neq 0$. 则 A 有奇异值分解:

$$A = U \begin{pmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}_{m\times n} V^*, \tag{6.1.4}$$

其中 U, V 分别为 n 阶和 m 阶酉矩阵, r 为 A 的秩. 令

$$X = V \begin{pmatrix} \sigma_1^{-1} & & & & & \\ & \ddots & & & & \\ & & \sigma_r^{-1} & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}_{n\times m} U^*, \tag{6.1.5}$$

则 X 显然满足Penrose方程. 所以 $A^\dagger$ 总是存在的.

设 X 与 Y 均是 A 的Penrose广义逆, 则对 X 与 Y 重复利用Penrose方程可得

$$X = XAX = XX^*A^* = XX^*A^*Y^*A^* = XAY = XAA^*Y^*Y = A^*Y^*Y = YAY = Y.$$

因此Penrose广义逆是唯一的. □

注. Penrose广义逆 $A^\dagger$ 是Penrose的原始记号, 也是国际通行的记号, 国内常将 $A^\dagger$ 记为 A^+ 并称为“加号逆”.

定理6.1.2 的证明实际上给出了利用奇异值分解计算Penrose广义逆的一种方法, 见本章第二节.

例 6.1.6 任意非零向量 x 的Penrose广义逆为 $\frac{x^*}{x^*x}$. 特别地, 单位向量 x 的Penrose广义逆为 x^*.

例 6.1.7 矩阵 $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ 的Penrose广义逆为自身. 而矩阵 $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ 的Penrose广义逆为 B^T.

例 6.1.8 n 阶幂零Jordan块 J_n 的Penrose广义逆为 J_n^T.

容易证明Penrose广义逆具有下述性质, 见习题6.

命题 6.1.1 (Penrose广义逆的性质) 矩阵 $A^\dagger A$ 与 $AA^\dagger$ 均为正交投影矩阵, 且

$$R(A^\dagger) = R(A^*), \qquad N(A^\dagger) = N(A^*).$$

定理 6.1.3 (Moore广义逆等于Penrose广义逆) 任意矩阵 A 的Moore广义逆矩阵与Penrose广义逆矩阵相等. 因此Moore广义逆矩阵的定义与Penrose广义逆矩阵的定义等价. 通常将 $A^\dagger$ 称为Moore-Penrose广义逆或**伪逆**.

证 若 X 是 A 的Moore广义逆矩阵, 则 AX 满足 (6.0.4), 即对任意 n 维向量 α, 有 $AX(A\alpha) = A\alpha$, 故 $AXA = A$. 同样, $XAX = X$. 再由 定理6.1.1 知, $(AX)^* = AX$, $(XA)^* = XA$. 所以 X 是 A 的Penrose广义逆矩阵.

反之, 若 X 是 A 的Penrose广义逆矩阵, 则由 $AXA = A$ 知, AX 在 $R(A)$ 上为恒等变换. 因 $R(A)^\perp = N(A^*)$ 而由 命题6.1.1, $N(A^*) = N(X)$, 所以对 $\forall\,\alpha \in R(A)^\perp$, 有 $X\alpha = 0$. 因此 $AX\alpha = 0$, 即 AX 是 $R(A)^\perp$ 上的零变换. 所以 $AX = P_{R(A)}$. 同理可证 $XA = P_{R(X)}$. 由此推出 X 满足 (6.0.3), 即 X 是 A 的Moore广义逆矩阵. 所以 定义6.1.1 与 定义6.1.2 是等价的. □

现在可以回答我们在第二章第三节提出的一个问题了, 即如何表示投影向量 $\mathrm{Proj}_{R(A)}b$(见第二章, 命题 2.3.2 之后)? 由 定理6.1.3 以及Moore方程可知矩阵 A 的列空间 $R(A)$ 上的正交投影矩阵是 $AA^\dagger$, 于是有

$$\mathrm{Proj}_{R(A)}b = AA^\dagger b \tag{6.1.6}$$

注. 由于普通逆矩阵只是Moore-Penrose广义逆矩阵的一种特例, 故Moore-Penrose 广义逆矩阵可能不具备普通逆矩阵的一些性质, 如下例.

例 6.1.9 设 $A = (1, 0), B = (1, 1)^T$, 则 $(AB)^\dagger = 1$ 而 $B^\dagger A^\dagger = 1/2$, 因此 $(AB)^\dagger \neq B^\dagger A^\dagger$.

为了进一步讨论 $A^\dagger$ 的性质, 我们引入以下符号

$$\lambda^\dagger = \begin{cases} \lambda^{-1}, & \lambda \neq 0 \\ 0, & \lambda = 0. \end{cases} \tag{6.1.7}$$

下面的命题汇总了Moore-Penrose广义逆矩阵 $A^\dagger$ 的一些性质, 证明均较为直接, 见习题12. 请读者比较这些性质与普通逆矩阵的类似性质.

命题 6.1.2 对任意矩阵 A, 有
(1) $(A^\dagger)^\dagger = A$;
(2) $(A^*)^\dagger = (A^\dagger)^*$; $(A^T)^\dagger = (A^\dagger)^T$;
(3) $(\lambda A)^\dagger = \lambda^\dagger A^\dagger$;
(4) $\mathrm{diag}\,(\lambda_1, \cdots, \lambda_n)^\dagger = \mathrm{diag}\,(\lambda_1^\dagger, \cdots, \lambda_n^\dagger)$;
(5) $A^\dagger AA^* = A^*$, $A^*AA^\dagger = A^*$, $AA^\dagger A^* = A^*$, $A^*A^\dagger A = A^*$;
(6) $(A^*A)^\dagger = A^\dagger(A^\dagger)^*$;
(7) $A^\dagger = (A^*A)^\dagger A^*$;
(8) 设 $A = B + C, B^*C = BC^* = 0$, 则 $A^\dagger = B^\dagger + C^\dagger$;
(9) $r(A) = r(A^\dagger) = r(A^\dagger A) = \mathrm{tr}(A^\dagger A)$.

按照矩阵与线性变换的对应关系, 我们可以讨论线性变换的Moore-Penrose广义逆变换, 此只需将Penrose方程组与Moore方程组中的矩阵理解成线性变换即可. 于是 $0 \in \mathrm{Hom}(U, V)$ 的Moore-Penrose广义逆变换为 $0 \in \mathrm{Hom}(V, U)$.

例 6.1.10 平面 $\mathbb{R}^2$ 上的移位变换 $\sigma : (x,y)^T \mapsto (y,0)^T$ 是不可逆变换, 其Moore-Penrose广义逆变换为 $\sigma^\dagger : (x,y)^T \mapsto (0,x)^T$, 这是因为 $\sigma\sigma^\dagger : (x,y)^T \mapsto (x,0)^T$ 正是 $\mathrm{Im}(\sigma)$ 上的正交投影变换, 而 $\sigma^\dagger\sigma : (x,y)^T \mapsto (0,y)^T$ 正是 $\mathrm{Im}(\sigma^\dagger)$ 上的正交投影变换.

例 6.1.11 设 $V = \mathbb{F}[x]_n$, 考虑 V 上的求导变换 $\partial : f(x) \mapsto f'(x)$, 我们知道 ∂ 不是可逆线性变换, 但只要定义 V 上的内积, 则其Moore-Penrose广义逆变换就唯一地存在. 比如, 设 $1, x, x^2, \cdots, x^{n-1}$ 是 V 的一组标准正交基, 则 ∂ 的Moore-Penrose广义逆变换为 (见习题13)

$$\partial^\dagger(x^{i-1}) = \begin{cases} x^i/i, & 1 \le i \le n-1 \\ 0, & i = n. \end{cases}$$

思考题

1. 2×1 矩阵与 1×2 矩阵的广义逆矩阵的几何意义是什么?
2. 两个 n 阶矩阵 A 与 B 何时满足条件 $AB = BA = 0$?
3. 设 P, Q 是两个可逆矩阵, 等式 $(PAQ)^\dagger = Q^{-1}A^\dagger P^{-1}$ 成立吗?

第二节 Moore-Penrose广义逆矩阵的计算

本节我们讨论Moore-Penrose广义逆矩阵 $A^\dagger$ 的计算. 首先, 由 定理6.1.2 的证明可知 (见公式 (6.1.4) 与 (6.1.5)), 利用奇异值分解可以计算 $A^\dagger$.

定理 6.2.1 ($A^\dagger$ 的SVD算法) 设 A 的奇异值分解为 $A = UDV^*$, 则 $A^\dagger = VD^\dagger U^*$.

例 6.2.1 用奇异值分解求 $A^\dagger$, 其中

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

解 A 的奇异值分解为

$$A = UDV^* = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

所以

$$A^\dagger = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

利用矩阵的满秩分解也可以计算 $A^\dagger$, 即有下述定理, 证明见习题17.

定理 6.2.2 ($A^\dagger$ 的满秩算法) (1) 设 A 为列满秩矩阵, 则 $A^\dagger = (A^*A)^{-1}A^*$;
(2) 设 A 为行满秩矩阵, 则 $A^\dagger = A^*(AA^*)^{-1}$;
(3) 设 $A = LR \in \mathbb{C}^{m\times n}$ 的秩为 r, 其中 L 为列满秩矩阵, R 为行满秩矩阵. 则

$$A^\dagger = R^\dagger L^\dagger = R^*(RR^*)^{-1}(L^*L)^{-1}L^* \tag{6.2.1}$$

例 6.2.2 设 $A=\alpha\beta^*$ 是秩为 1 的矩阵, 则

$$A^\dagger = \frac{A^*}{\|\alpha\|_2^2\,\|\beta\|_2^2} \tag{6.2.2}$$

重新考察 例 6.2.1, 因为 $A=(1,1,0)^T(1,1)$, 由公式 (6.2.2) 可得 $A^\dagger=(1/4)A^T$.

例 6.2.3 求 $A^\dagger$, 其中

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

解 A 的满秩分解为

$$A=LR=\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

则

$$\begin{aligned}
RR^* &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \\
(RR^*)^{-1} &= \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \\
L^*L &= \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix}, \\
(L^*L)^{-1} &= \frac{1}{6}\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}.
\end{aligned}$$

所以

$$\begin{aligned}
A^\dagger &= R^*(RR^*)^{-1}(L^*L)^{-1}L^* \\
&= \frac{1}{18}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \\
&= \frac{1}{6}\begin{pmatrix} 2 & 1 & -3 \\ -2 & 0 & 4 \\ 0 & 1 & 1 \end{pmatrix}.
\end{aligned}$$

如果矩阵 A 是列满秩或行满秩矩阵, 则 $A^\dagger$ 也可由 A 的正交三角分解求出.

定理 6.2.3 ($A^\dagger$ 的 QR 算法) 设列满秩矩阵 A 有正交三角分解 $A=QR$, 其中 Q 的列向量为单位正交向量组, R 为非奇异的上三角矩阵. 则

$$A^\dagger = R^{-1}Q^*. \tag{6.2.3}$$

由 命题6.1.2(7) 可知, $A^\dagger$ 可由 $(A^*A)^\dagger$ 与 A^* 的乘积得到. 因此我们讨论一般Hermite矩阵的Moore-Penrose广义逆. 设 A 的互不相同的非零奇异值为 $\sigma_1, \cdots, \sigma_s, s > 1$, 则Hermite矩阵 A^*A 的谱分解为

$$A^*A = \sigma_1^2 P_1 + \cdots + \sigma_s^2 P_s,$$

由正规矩阵的谱分解定理以及 命题6.1.2(8) 可得

$$(A^*A)^\dagger = \sigma_1^{-2} P_1 + \cdots + \sigma_s^{-2} P_s \tag{6.2.4}$$

设

$$\phi(x) = \prod_{i=1}^{s}(x - \sigma_i^2), \qquad \phi_i(x) = \prod_{j\neq i}(x - \sigma_j^2) \tag{6.2.5}$$

则由于 $P_i^* = P_i, P_iP_j = \delta_{ij}P_i$ 以及 $\sum\limits_{i=1}^{s} P_i = I$, 可得 (细节见习题17)

$$\phi_j(A^*A) = \phi_j(\sigma_j^2)P_j, 1 \leq j \leq s \tag{6.2.6}$$

因此

$$P_j = \frac{\phi_j(A^*A)}{\phi_j(\sigma_j^2)}, 1 \leq j \leq s \tag{6.2.7}$$

将上式代入公式 (6.2.4) 可得下述计算 $A^\dagger$ 的 **Lagrange-Sylvester 展开公式** .

定理 6.2.4 (Lagrange-Sylvester 展开公式)

$$A^\dagger = \sum_{i=1}^{s} \sigma_i^{-2} \frac{\prod\limits_{j\neq i}(A^*A - \sigma_j^2 I)}{\prod\limits_{j\neq i}(\sigma_i^2 - \sigma_j^2)} A^* \tag{6.2.8}$$

例 6.2.4 设 $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$, 求 $A^\dagger$.

解 由第四章, 例 4.2.4 可知

$$A^TA = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}, |\lambda I - A^TA| = \lambda(\lambda-1)(\lambda-3).$$

故 A 的非零奇异值为 $\sigma_1 = \sqrt{3}, \sigma_2 = 1$. 故由公式 (6.2.8) 可得

$$\begin{aligned} A^\dagger &= [\tfrac{1}{3}(\tfrac{A^TA-3I}{1-3}) + \tfrac{A^TA-I}{3-1}]A^T \\ &= (\tfrac{4}{3}I - \tfrac{1}{3}A^TA)A^T = \tfrac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \\ -1 & 2 \end{pmatrix}. \end{aligned}$$

注. 实际计算 例6.2.4 中的 $A^\dagger$ 时, 解线性矩阵方程 $AX = I_2$ 较为简捷, 因为 A 行满秩, 故该矩阵方程有唯一解, 即为所需.

思考题

1. 公式 (6.2.1) 的几何意义是什么?

2. 列满秩矩阵与行满秩矩阵的 Moore-Penrose 广义逆的几何意义是什么?

3. 利用谱分解计算 Moore-Penrose 广义逆的几何意义是什么?

第三节 矩阵的 $\{1\}$- 广义逆

Moore 方程组实际上是用特殊的线性变换来定义矩阵的广义逆. 如果从线性方程组的角度讨论矩阵的广义逆, 则可得到下面的定义.

定义 6.3.1 设 A 是 $m\times n$ 矩阵. 一个 $n\times m$ 矩阵 G 称为 A 的一个 **{1}- 广义逆矩阵** , 记为 A^-, 若对任意给定的 m 维向量 b, 只要方程组 $Ax=b$ 有解, 则 $x=Gb$ 也一定是解.

注. 国内常将 **{1}- 广义逆矩阵**称为“减号”逆.

例 6.3.1 设 $A=\begin{pmatrix}1\\0\end{pmatrix}$, 则 $\begin{pmatrix}1\\0\end{pmatrix}^\dagger=(1,0)$. 由 定义6.3.1, A^- 应满足条件 $AA^-b=b$, 于是可取 $A^-=(1,y)$, 其中 y 是任意常数.

下面的定理用 Penrose 方程组中的一个方程来刻画 $\{1\}$- 广义逆矩阵.

定理 6.3.1 $n\times m$ 矩阵 G 是 $m\times n$ 矩阵 A 的一个 $\{1\}$- 广义逆矩阵 $\Longleftrightarrow$ $AGA=A$.

证 “$\Rightarrow$”: 对任意的 $z\in\mathbb{C}^n$, $b=Az$ 是 m 维向量, 且 z 为 $Ax=b$ 的解. 因而 $x=Gb$ 也是一个解, 即 $AGb=b$. 于是
$$AGAz=Az.$$
由于 z 的任意性, 可知 $AGA=A$.

“$\Leftarrow$”: 若 $AGA=A$, 设 $Ax=b$ 有解, 则
$$AGb=AGAx=Ax=b.$$
所以 $x=Gb$ 也是解. 由定义, G 是 A 的一个 $\{1\}$- 广义逆矩阵. □

显然 $A^\dagger$ 是 A 的一个 $\{1\}$- 广义逆矩阵. 由 例6.3.1 可知, 一般情况下, A^- 不唯一. 因此我们用符号 $A\{1\}$ 表示矩阵 A 的所有 $\{1\}$- 广义逆矩阵, 即
$$A\{1\}=\{X\mid AXA=A\} \tag{6.3.1}$$
例如, $0_{m\times n}\{1\}=\mathbb{C}^{n\times m}$.

例 6.3.2 若 A 为可逆矩阵, 则方程 $AXA=A$ 只有唯一解 A^{-1}, 故此时 A 的 $\{1\}$- 广义逆矩阵唯一, 即 $A\{1\}=\{A^{-1}=A^\dagger\}$.

例 6.3.3 由 定理6.3.1 可知, 任何一个 $n\times m$ 阶矩阵都是 $0_{m\times n}$ 的一个 $\{1\}$- 广义逆矩阵! (请对照, $0_{m\times n}^\dagger=0_{n\times m}$.) 故知 $\{1\}$- 广义逆矩阵不是对称的, 即若 G 是矩阵 A 的一个 $\{1\}$- 广义逆矩阵, 则 A 未必是 G 的一个 $\{1\}$- 广义逆矩阵. 因此公式 $(A^-)^-=A$ 一般不成立.

例 6.3.4 设 $A \in \mathbb{C}^{m\times n}$. 矩阵方程 $AX = I$ 的解称为矩阵 A 的**右逆** . 类似地, 矩阵方程 $XA = I$ 的解称为矩阵 A 的**左逆**. 显然, 左逆与右逆可能均不存在, 也可能只存在一个. 但由 定理6.3.1 可知, 左逆与右逆均是 A 的 $\{1\}$- 广义逆矩阵. 左逆与右逆的其它性质见习题19.

为了得到 $\{1\}$- 广义逆矩阵的一般形式, 需要 $\{1\}$- 广义逆矩阵的以下性质, 证明见习题20.

命题 6.3.1 设 P, Q 为非奇异矩阵, 则 $Q^{-1}A^{-}P^{-1} \in (PAQ)\{1\}$, 即 $Q^{-1}A^{-}P^{-1}$ 是 PAQ 的一个 $\{1\}$- 广义逆矩阵.

命题6.3.1 可以粗略地解释为 $(PAQ)^{-} = Q^{-1}A^{-}P^{-1}$, 请注意此等式不清晰, 因为左右两端既可以表示一个矩阵, 也可以表示一个集合. 但这个记法确实简单且不致引起混乱, 所以常用 A^{-} 或 $A^{(1)}, A^{(1,2)}$ 等表示一个 $\{1\}$- 逆或一个 $\{1,2\}$- 逆. 需要指出, 尽管 命题6.3.1 成立, 但正如 $A^{\dagger}$ 的情形, $(AB)^{-} = B^{-}A^{-}$ 一般不成立, 见 例6.1.9.

定理 6.3.2 设 $A \in \mathbb{C}^{m\times n}, P \in \mathbb{C}^{m\times m}$ 与 $Q \in \mathbb{C}^{n\times n}$ 可逆且满足

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

则

$$M = Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}P, \tag{6.3.2}$$

是 A 的一个 $\{1\}$- 广义逆矩阵, 且 A 的任意一个 $\{1\}$- 广义逆矩阵都可以写成 (6.3.2) 的形式, 其中 $X \in \mathbb{C}^{r\times(m-r)}$, $Y \in \mathbb{C}^{(n-r)\times r}$, $Z \in \mathbb{C}^{(n-r)\times(m-r)}$ 是任意的.

证 直接计算可得

$$\begin{aligned} AMA &= P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1}Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}PP^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} \\ &= P^{-1}\begin{pmatrix} I_r & X \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} \\ &= P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} = A. \end{aligned}$$

所以 M 为 A 的一个 $\{1\}$- 广义逆矩阵. 反之, 设 A^{-} 为 A 的任意一个 $\{1\}$- 广义逆矩阵. 由于 P, Q 可逆, 因此可令

$$A^{-} = Q\begin{pmatrix} W & X \\ Y & Z \end{pmatrix}P.$$

由于 $A = AA^{-}A$, 即

$$\begin{aligned} P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} &= P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1}Q\begin{pmatrix} W & X \\ Y & Z \end{pmatrix}PP^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} \\ &= P^{-1}\begin{pmatrix} W & X \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} \\ &= P^{-1}\begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix}Q^{-1} = P^{-1}\begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix}Q^{-1}. \end{aligned}$$

因此 $W=I_r$, 即 A^- 具有公式 (6.3.2) 的形式. $\square$

下面的命题罗列了 $\{1\}$- 广义逆矩阵的一些基本性质, 证明均较简单, 见习题 21.

命题 6.3.2 若 $A\in\mathbb{C}^{m\times n}$, 则
(1) $(A^-)^*\in A^*\{1\}$, 即 $(A^-)^*$ 是 A^* 的一个 $\{1\}$- 广义逆矩阵;
(2) $r(A)\le r(A^-)$;
(3) A 可逆 $\iff$ $A\{1\}$ 是一元集合, 即 $A\{1\}=\{A^{-1}\}$;
(4) $\lambda^\dagger A^-\in(\lambda A)\{1\}$, 即 $\lambda^\dagger A^-$ 是 λA 的一个 $\{1\}$- 广义逆矩阵;
(5) AA^- 与 A^-A 都是幂等矩阵;
(6) $R(AA^-)=R(A),\ N(A^-A)=N(A)$.

注意, 命题 6.3.2 的 (2) 仅给出了矩阵 A 的 $\{1\}$- 广义逆矩阵的秩的一个下限 $r(A)$, 读者可尝试由 定理 6.3.2 得出一个上限.

例 6.3.5 试求表示成公式 (6.3.2) 形式的广义逆矩阵 A^-, 其中

$$A=\begin{pmatrix}0&1\\-1&-4\\1&5\end{pmatrix}.$$

解

$$\begin{pmatrix}A&I_3\\I_2&0\end{pmatrix}=\begin{pmatrix}0&1&1&0&0\\-1&-4&0&1&0\\1&5&0&0&1\\1&0&0&0&0\\0&1&0&0&0\end{pmatrix}\to\begin{pmatrix}0&1&1&0&0\\-1&0&4&1&0\\1&0&-5&0&1\\1&0&0&0&0\\0&1&0&0&0\end{pmatrix}$$

$$\to\begin{pmatrix}0&1&1&0&0\\1&0&-4&-1&0\\0&0&-1&0&1\\1&0&0&0&0\\0&1&0&0&0\end{pmatrix}\to\begin{pmatrix}1&0&1&0&0\\0&1&-4&-1&0\\0&0&-1&0&1\\0&1&0&0&0\\1&0&0&0&0\end{pmatrix}$$

所以

$$P=\begin{pmatrix}1&0&0\\-4&-1&0\\-1&0&1\end{pmatrix},\quad Q=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

于是 A 的 $\{1\}$- 广义逆矩阵为

$$A^-=Q\begin{pmatrix}I_2&X\end{pmatrix}P,$$

其中 $X\in\mathbb{C}^{2\times1}$ 是任意矩阵.

以下我们从矩阵方程的观点讨论矩阵的 $\{1\}$- 广义逆矩阵. 由 定理 6.3.1 和公式 (6.3.1), 矩阵 A 的全体 $\{1\}$- 广义逆矩阵恰好是矩阵方程 $AXA=A$ 的解集, 而由线性方程组的一般理论, 该方程的通解是相应的线性齐次矩阵方程 $AXA=0$ 的通解与它自身的一个特解的和. 以下, 我们将矩阵方程 $AXA=A$ 称为**对称线性矩阵方程**, 而 $AXA=0$ 称为**对称线性齐次矩阵方程** .

定理 6.3.3 (对称线性齐次矩阵方程的解) 设 $A\in\mathbb{C}^{m\times n}, r(A)=r$. 则
(1) 齐次矩阵方程 $AXA=0$ 的解空间是 $\mathbb{C}^{n\times m}$ 的一个 $mn-r^2$ 维子空间;
(2) 齐次矩阵方程 $AXA=0$ 的通解为

$$X=Y-A^{\dagger}AYAA^{\dagger},\forall Y\in\mathbb{C}^{n\times m} \tag{6.3.3}$$

证 (1) 考察 $\mathbb{C}^{n\times m}$ 到 $\mathbb{C}^{m\times n}$ 的如下线性变换

$$\sigma: X\mapsto AXA, \forall X\in\mathbb{C}^{n\times m}$$

则齐次方程 $AXA=0$ 的解空间恰好是 $\mathrm{Ker}(\sigma)$. 根据矩阵的张量积的性质, σ 在按列顺序的标准基下的矩阵为 $A^T\otimes A$, 因此

$$\dim\mathrm{Ker}(\sigma)=\dim N(A^T\otimes A)=mn-r^2,$$

此处用到 $r(A\otimes B)=r(A)r(B)$.

(2) 公式 (6.3.3) 中的矩阵显然是方程 $AXA=0$ 的解. 现设 Y 是方程 $AXA=0$ 的一个解. 则 $AYA=0$, 故 $A^{\dagger}AYAA^{\dagger}=0$. 于是 $Y=Y-0=Y-A^{\dagger}AYAA^{\dagger}$ 即为公式 (6.3.3) 中的形式, 所以公式 (6.3.3) 确是方程 $AXA=0$ 的通解. □

从 定理6.3.3 的证明可知, 若 A^- 是 A 的任意一个 $\{1\}$- 广义逆矩阵, 则 $AXA=0$ 的通解为

$$X=Y-A^-AYAA^-,\forall Y\in\mathbb{C}^{n\times m} \tag{6.3.4}$$

例 6.3.6 设 $A=\begin{pmatrix}1&1\\0&0\end{pmatrix}$, 则 $A^{\dagger}=(1/2)\begin{pmatrix}1&0\\1&0\end{pmatrix}$, 故 $AXA=0$ 的通解为

$$X=Y-A^{\dagger}AYAA^{\dagger}=\begin{pmatrix}a&b\\-a&c\end{pmatrix},$$

其中 a,b,c 是任意常数. 因此 $AXA=0$ 的解空间是 3 维的.

由对称线性齐次矩阵方程的通解结构 (公式 (6.3.4)), 立即可得下述

定理 6.3.4 (Rao[57] 定理) 设 $A\in\mathbb{C}^{m\times n}$, A^- 是 A 的任意一个 $\{1\}$- 广义逆矩阵, 则

$$A\{1\}=\{A^-+Y-A^-AYAA^-\,|\,\forall Y\in\mathbb{C}^{n\times m}\} \tag{6.3.5}$$

例 6.3.7 设 $A=\begin{pmatrix}1&1\\0&0\end{pmatrix}$, 则由 例6.3.6 及 定理6.3.4 可知, 对称矩阵方程 $AXA=A$ 的通解为

$$X=A^{\dagger}+\begin{pmatrix}x&y\\-x&z\end{pmatrix}=\begin{pmatrix}1&0\\0&0\end{pmatrix}+\begin{pmatrix}a&b\\-a&c\end{pmatrix},$$

其中 a,b,c 是任意常数.

[57] Calyampudi Radhakrishna Rao(1920-), 印度著名统计学家, 被誉为印度历史上最著名的十位科学家之一, 第三世界科学院创始人之一, 现为美国 Pennsylvania 州立大学教授.

对称矩阵方程 $AXA = A$ 还可以写成 $A(XA - I_n) = 0$ 与 $(AX - I_m)A = 0$ 两种等价形式. 分别令

$$Y = XA - I_n, \quad Z = AX - I_m,$$

可得两个齐次矩阵方程

$$AY = 0, \qquad ZA = 0.$$

由上述矩阵方程组可以得出对称线性矩阵方程 $AXA = A$ 通解的如下形式, 证明见习题22.

定理 6.3.5 设 $A \in \mathbb{C}^{m\times n}$, $A^- \in \mathbb{C}^{n\times m}$ 为 A 的一个 $\{1\}$- 广义逆矩阵, 则方程 $AXA = A$ 的通解为

$$X = A^- + Y(I_m - AA^-) + (I_n - A^-A)Z, \forall Y, Z \in \mathbb{C}^{n\times m} \tag{6.3.6}$$

下面我们介绍左逆与 $\{1\}$- 广义逆矩阵之间的一个关系.

定理 6.3.6 设 $A \in \mathbb{C}^{m\times n}$, 则
(1) A 列满秩 $\iff$ $A^-A = I_n$; 即矩阵 A 的 $\{1\}$- 广义逆矩阵是左逆 $\iff$ A 列满秩;
(2) A 行满秩 $\iff$ $AA^- = I_m$; 即矩阵 A 的 $\{1\}$- 广义逆矩阵是右逆 $\iff$ A 行满秩.

证 我们只证 (1), (2) 的证明类似.
"$\Rightarrow$": 设 A 列满秩, 则存在非奇异矩阵 $P \in \mathbb{C}^{m\times m}$, $Q \in \mathbb{C}^{n\times n}$ 使

$$PAQ = \begin{pmatrix} I_n \\ 0 \end{pmatrix}.$$

由 定理6.3.1, A^- 具有形式 $A^- = Q[I_n, X]P$. 所以

$$A^-A = Q[I_n, X]PP^{-1}\begin{pmatrix} I_n \\ 0 \end{pmatrix}Q^{-1} = QI_nQ^{-1} = I_n.$$

"$\Rightarrow$": 若 $A^-A = I_n$, 则显然 A 是列满秩的. □

本节最后, 我们讨论一个重要的问题, 即 A^-A 与单位矩阵究竟有多大差别?

定理 6.3.7 设 $A \in \mathbb{C}^{m\times n}, r(A) = r$, A^- 是 A 的一个 $\{1\}$- 广义逆矩阵, 则 $r(I_n - A^-A) = n - r$.

证 设

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n},$$

其中 P, Q 分别为 m 阶和 n 阶非奇异矩阵, 则

$$A^- = Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}_{n\times m}P.$$

所以

$$\begin{aligned} I_n - A^-A &= I_n - Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}_{n\times m}PP^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n}Q^{-1} \\ &= I_n - Q\begin{pmatrix} I_r & 0 \\ Y & 0 \end{pmatrix}_{n\times n}Q^{-1} = Q\begin{pmatrix} 0 & 0 \\ -Y & I_{n-r} \end{pmatrix}Q^{-1}. \end{aligned}$$

因此 $r(I_n - A^- A) = n - r$. □

定理 6.3.7 表明, A 的秩越大, 则 $A^- A$ 与单位矩阵的差距就越小, 特别, 当 A 可逆时, $A^- A$ 就是单位矩阵.

思考题

1. 零矩阵的 $\{1\}$- 广义逆矩阵是所有矩阵, 是否还有别的矩阵的 $\{1\}$- 广义逆矩阵是所有矩阵?
2. 不可逆的方阵可否有可逆的 $\{1\}$- 广义逆矩阵?
3. $A^- A$ 与 AA^- 的几何意义是什么?
4. 试给出矩阵 A 的 $\{1\}$- 广义逆矩阵的秩的一个上限?

第四节 矩阵的 $\{13\}$- 逆与 $\{14\}$- 逆

本节我们简要介绍矩阵的其它几种广义逆.

定义 6.4.1 如果矩阵 X 满足 Penrose 方程组的前两个方程, 即

$$AXA = A, \qquad XAX = X$$

则称 X 是矩阵 A 的一个 **$\{1,2\}$- 逆或自反逆**.

例 6.4.1 零矩阵的 $\{1,2\}$- 逆是唯一的, 就是其转置. 任意可逆矩阵的 $\{1,2\}$- 逆是唯一的, 就是其逆矩阵. 矩阵 $(1,0)$ 的 $\{1,2\}$- 逆是 $(1,\, x)^T$, 其中 x 为任意常数.

因此矩阵的 $\{1,2\}$- 逆一般不唯一, 我们用符号 $A\{1,2\}$ 表示矩阵 A 的全体 $\{1,2\}$- 逆构成的集合. 显然,

$$A\{1,2\} \subseteq A\{1\}.$$

例 6.4.2 设 $A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix}$, 则下列矩阵

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, D = (1/6)A^T$$

依次是 A 的 $\{1\}$- 逆, $\{1,2\}$- 逆和 Moore-Penrose 逆, 但 B, C 均不是 A 的 Moore-Penrose 逆, 而 B 也不是 A 的 $\{1,2\}$- 逆. 因此 $A\{1,2\} \subseteq A\{1\}$ 通常是真包含.

下面的定理给出了任意矩阵在初等变换下的 $\{1,2\}$- 逆的一般形式, 证明见习题 23.

定理 6.4.1 设 $A \in \mathbb{C}^{m\times n}, r(A) = r$, $P \in \mathbb{C}^{m\times m}$, $Q \in \mathbb{C}^{n\times n}$ 是可逆矩阵使得

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix},$$

则

$$A\{1,2\} = \{Q\begin{pmatrix} I_r & B \\ C & CB \end{pmatrix}P \mid \forall B \in \mathbb{C}^{r\times(n-r)}, C \in \mathbb{C}^{(m-r)\times r}\} \tag{6.4.1}$$

下面的命题罗列了矩阵的 $\{1,2\}$- 逆的几个简单性质, 证明见习题25.

命题 6.4.1 设 $A \in \mathbb{C}^{m\times n}$, 则
(1) $A\{1,2\} = \{X_1AX_2 \,|\, X_1, X_2 \in A\{1\}\}$;
(2) $\forall X \in A\{1,2\}, r(X) = r(A)$;
(3) 设 P, Q 为适当阶数的可逆矩阵, 则 $(PAQ)^{(1,2)} = Q^{-1}A^{(1,2)}P^{-1}$.

定义 6.4.2 如果矩阵 X 满足Penrose方程组的第一及第三个方程, 即

$$AXA = A, \qquad (AX)^* = AX$$

则称 X 是矩阵 A 的一个 $\{1,3\}$**- 逆**.

我们以符号 $A\{1,3\}$ 表示矩阵 A 的全体 $\{1,3\}$- 逆构成的集合. 显然

$$A\{1,3\} \subseteq A\{1\}.$$

例 6.4.3 任何矩阵 $A \in \mathbb{C}^{n\times m}$ 都是 $m\times n$ 阶零矩阵的 $\{1,3\}$- 逆. 任意可逆矩阵的 $\{1,3\}$- 逆是唯一的, 就是其逆矩阵. 矩阵 $(1,0)$ 的 $\{1,3\}$- 逆是 $(1,x)^T$, 其中 x 为任意常数, 即此时有 $A\{1,3\} = A\{1,2\}$. 但矩阵 $(1,0)^T$ 的 $\{1,3\}$- 逆是唯一的 (为什么?), 等于 $(1 \;\; 0) = (1,0)^{T\dagger}$.

定理 6.4.2 设 B 是 $A \in \mathbb{C}^{m\times n}$ 的一个 $\{1,3\}$- 逆, 则

$$A\{1,3\} = \{B + (I_n - BA)Y \,|\, \forall Y \in \mathbb{C}^{n\times m}\} \tag{6.4.2}$$

证 由于 $A\{1,3\} \subseteq A\{1\}$, 而由 定理6.3.4(对称线性矩阵方程的解) 我们知道 $X = B + (I_n - BA)Y$ 满足方程 $AXA = A$. 再由 $A[B + (I_n - BA)Y] = AB$ 以及 $B \in A\{1,3\}$ 可知 AX 还是Hermite的. □

例 6.4.4 设 $A = \alpha\beta^*$ 是秩为1的矩阵, 则

$$A\{1,3\} = \{\frac{A^*}{\alpha^*\alpha\,\beta^*\beta} + (I - \frac{\beta\beta^*}{\beta^*\beta})Y \,|\, \forall Y\} \tag{6.4.3}$$

下面的命题列出了 $\{1,3\}$- 逆的一些性质, 证明见习题28.

命题 6.4.2 设 $A^{(1,3)}$ 是 $A \in \mathbb{C}^{m\times n}$ 的任意一个 $\{1,3\}$- 逆, 则
(1) $A\{1,3\} = \{X \,|\, AX = AA^{(1,3)}\}$;
(2) $AA^{(1,3)} = AA^\dagger = P_{R(A)}$; 特别地, $AA^{(1,3)}$ 是幂等矩阵;
(3) $A^{(1,3)}A$ 是幂等矩阵;
(4) $I_m \otimes A^{(1,3)} \in (I_m \otimes A)\{1,3\}$.

定义 6.4.3 如果矩阵 X 满足Penrose方程组的第一及第四个方程, 即

$$AXA = A, \qquad (XA)^* = XA$$

则称 X 是矩阵 A 的一个 $\{1,4\}$**- 逆**.

类似于前面几种情形, 我们以符号 $A\{1,4\}$ 表示矩阵 A 的全体 $\{1,4\}$- 逆构成的集合. 显然

$$A\{1,4\} \subseteq A\{1\}.$$

由定义可知, 矩阵 A 的 $\{1,4\}$- 逆与其 $\{1,3\}$- 逆的差别是前者满足方程 $(XA)^* = XA$ 而后者满足方程 $(AX)^* = AX$, 因此它们具有较为相近的性质. 我们仅将矩阵 A 的 $\{1,4\}$- 逆的有关结论列出, 所有的证明均见习题 30-32, 请读者参照 $\{1,3\}$- 逆的相关结论与证明.

例 6.4.5 任何矩阵 $A \in \mathbb{C}^{n\times m}$ 都是 $m \times n$ 阶零矩阵的 $\{1,4\}$- 逆. 任意可逆矩阵的 $\{1,4\}$- 逆是唯一的, 就是其逆矩阵. 矩阵 $(1,0)$ 的 $\{1,4\}$- 逆是唯一的 (为什么?), 等于 $(1,0)^T$, 即此时有 $A^{(1,4)} = A^\dagger$. 矩阵 $(1,0)^T$ 的 $\{1,4\}$- 逆是 $(1,x)^T$, 其中 x 为任意常数, 即此时有 $A\{1,4\} = A\{1,2\}$. 请读者比较 例 6.4.3.

定理 6.4.3 设 B 是 $A \in \mathbb{C}^{m\times n}$ 的一个 $\{1,4\}$- 逆, 则

$$A\{1,4\} = \{B + Y(I_m - AB) \,|\, \forall Y \in \mathbb{C}^{n\times m}\} \tag{6.4.4}$$

例 6.4.6 设 $A = \alpha\beta^*$ 是秩为 1 的矩阵, 则

$$A\{1,4\} = \{\frac{A^*}{\alpha^*\alpha\,\beta^*\beta} + Y(I - \frac{\alpha\alpha^*}{\alpha^*\alpha}) \,|\, \forall Y\} \tag{6.4.5}$$

下面的命题列出了 $\{1,4\}$- 逆的一些性质, 证明见习题 32.

命题 6.4.3 设 G 是 $A \in \mathbb{C}^{m\times n}$ 的任意一个 $\{1,4\}$- 逆, 则
(1) $A\{1,4\} = \{X \,|\, XA = GA\}$;
(2) $GA = A^\dagger A = P_{R(A^*)}$; 特别地, GA 是幂等矩阵;
(3) AG 是幂等矩阵;
(4) $G \otimes I_n \in (A \otimes I_n)\{1,4\}$.

思考题
1. 除了零矩阵与可逆矩阵外, 是否还有别的矩阵的 $\{1,2\}$- 逆是唯一的?
2. Hermite 矩阵的 $\{1,2\}$- 逆一定是 Hermite 的吗?
3. 不可逆矩阵的 $\{1,3\}$- 逆与 $\{1,4\}$- 逆一定是不可逆的吗?
4. 矩阵的 $\{1,2\}$- 逆, $\{1,3\}$- 逆, $\{1,4\}$- 逆的几何意义是什么?
5. 何时 $A\{1,i\} = A\{1,j\}, 1 \le i \ne j \le 4$?

第五节 应用：线性方程组, 流量矩阵估计

本节我们将利用广义逆矩阵的理论来统一线性方程组与矩阵方程 $AXB = C$ 的解的问题.
回忆线性方程组 $Ax = b$ 称为相容的或一致的如果该方程组至少存在一个解, 而称为不相容的或矛盾的, 如果该方程组没有解. 相容方程组又分为确定方程组 (如果恰好有一组解) 和超定方程组 (如果有无穷多组解). 对于超定方程组常常需要求出最小范数解, 而对于矛盾方程组则需要求出最小二乘解 (正如第二章讨论的). 一般情况下, 最小二乘解也可能是无限的, 则还需要进一步求出范数最小的最小二乘解.
我们先利用 $\{1\}$- 广义逆矩阵给出一般齐次线性方程组的通解.

定理 6.5.1 设 A^- 为 $m\times n$ 矩阵 A 的一个 $\{1\}$- 广义逆矩阵, 则齐次线性方程组 $Ax=0$ 的通解为

$$x=(I_n-A^-A)z, \tag{6.5.1}$$

其中 z 是任意的 n 维列向量.

证 由 $AA^-A=A$ 知 $(I_n-A^-A)z$ 为 $Ax=0$ 的解. 设 A 的秩为 r, 则 $Ax=0$ 的解空间为 $n-r$ 维的. 而 $L=\{(I_n-A^-A)z\,|\,z \text{ 任意}\}$ 是解空间的子空间, 且 L 就是矩阵 (I_n-A^-A) 的列空间, 故由 定理6.3.7 知其维数为 $r(I_n-A^-A)=n-r$. 定理得证. □

现若线性方程组 $Ax=b$ 是相容的, 则由 1- 逆的定义知 A^-b 是 $Ax=b$ 的一个特解. 故由线性方程组的解的结构和 定理6.5.1 立即可得下面的

定理 6.5.2 设 A^- 为 $m\times n$ 矩阵 A 的一个 $\{1\}$- 广义逆矩阵, 则当方程组 $Ax=b$ 有解时, 其通解可表示为

$$x=A^-b+(I_n-A^-A)z, \tag{6.5.2}$$

其中 z 是任意的 n 维列向量.

定义 6.5.1 设线性方程组 $Ax=b$ 有解, 在所有解中, $\|x\|_2=\sqrt{x^*x}$ 取最小值的解 x 称为方程组的最小范数解.

显然, 为求最小范数解只需使解的范数平方 x^*x 最小即可.

定理 6.5.3 设 $Ax=b$ 为相容方程组.

(1) 设 $G\in A\{1,4\}$, 方程 $Ax=b$ 的通解 为

$$x=Gb+(I-GA)z,\forall\, z \tag{6.5.3}$$

(2) 设 $G\in A\{1\}$, 则 $x=Gb$ 是方程 $Ax=b$ 的最小范数解 (即 $\|Gb\|_2\le\|x\|_2,\forall x, Ax=b$) $\Longleftrightarrow G\in A\{1,4\}$. 因此, 矩阵的 $\{1,4\}$- 逆也称为**最小范数逆**.

证 显然 $y=Gb$ 为方程组的解. 设 x_0 为 $Ax=b$ 的一个特解. 由 定理6.5.2, $Ax=b$ 的通解为

$$x=Gb+(I_n-A^-A)z.$$

我们有

$$\begin{aligned}\|x\|_2^2=x^*x&=(A^-b+(I_n-A^-A)z)^*(A^-b+(I_n-A^-A)z)\\&=(A^-b)^*A^-b+((I_n-A^-A)z)^*(I_n-A^-A)z\\&\quad+(A^-b)^*(I_n-A^-A)z+z^*(I_n-A^-A)^*A^-b.\end{aligned}$$

因

$$\begin{aligned}(A^-b)^*(I_n-A^-A)z&=(A^-Ax_0)^*(I_n-A^-A)z=x_0^*(A^-A)^*(I_n-A^-A)z\\&=x_0^*A^-A(I_n-A^-A)z=x_0^*(A^-A-A^-AA^-A)z=0.\end{aligned}$$

$$z^*(I_n-A^-A)^*A^-b\ =z^*(I_n-A^-A)A^-Ax_0=z^*(A^-A-A^-AA^-A)x_0=0.$$

所以

$$\|x\|_2^2=(A^-b)^*A^-b+((I_n-A^-A)z)^*(I_n-A^-A)z\ge(A^-b)^*A^-b.$$

由此便知 $y = A^{-}b$ 为方程组 $Ax = b$ 的最小范数解. $\square$

因 A 的 Moore-Penrose 逆 $A^{\dagger}$ 一定是 A 的一个 $\{1\}$- 广义逆矩阵A^{-}, 所以若方程组 $Ax = b$ 有解, 则 $y = A^{\dagger}b$ 一定是方程组的最小范数解.

由 定理 6.5.3 并结合矩阵方程的张量积表示, 即可证明下面的 (见习题 43)

推论 6.5.1 设 $A \in \mathbb{C}^{m\times n}, G \in \mathbb{C}^{n\times m}$, 则 G 是矩阵 A 的一个 $\{1,4\}$- 逆 $\iff$ G 是矩阵方程 $XA = I_n$ 的最小二乘解, 即

$$\|GA - I_n\|_F = \min_{X\in\mathbb{C}^{n\times m}} \|XA - I_n\|_F.$$

定理 6.5.4 设 $Ax = b$ 为一矛盾方程. 则 $x = Gb$ 是方程 $Ax = b$ 的最小二乘解 $\iff$ $G \in A\{1,3\}$. 特别地, 方程 $Ax = b$ 的最小二乘解 为

$$x = Gb + (I - GA)y, \forall\, y \tag{6.5.4}$$

因此矩阵的 $\{1,3\}$- 逆也称为**最小二乘逆**.

证 由第二章的讨论我们知道, $Ax = b$ 的最小二乘解为方程组

$$A^*Ax = A^*b$$

的解. 因此需证明 $A^*A(A^{-}b) = A^*b$. 设

$$PAQ = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n},$$

其中 P, Q 分别为 m 阶和 n 阶非奇异矩阵. 则

$$\begin{aligned}
A^{-} &= Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}_{n\times m} P, \\
AA^{-} &= P^{-1}\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n} Q^{-1}Q\begin{pmatrix} I_r & X \\ Y & Z \end{pmatrix}_{n\times m} P = P^{-1}\begin{pmatrix} I_r & X \\ 0 & 0 \end{pmatrix}_{m\times m} P.
\end{aligned}$$

因 $(AA^{-})^* = AA^{-}$, 所以

$$P^*\begin{pmatrix} I_r & 0 \\ X^* & 0 \end{pmatrix}(P^{-1})^* = P^{-1}\begin{pmatrix} I_r & X \\ 0 & 0 \end{pmatrix}P.$$

即

$$\begin{pmatrix} I_r & 0 \\ X^* & 0 \end{pmatrix}(P^{-1})^*P^{-1} = (P^{-1})^*P^{-1}\begin{pmatrix} I_r & X \\ 0 & 0 \end{pmatrix}.$$

因此

$$\begin{aligned}
A^*Ay = A^*AA^{-}b &= (Q^{-1})^*\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times m}(P^{-1})^*P^{-1}\begin{pmatrix} I_r & X \\ 0 & 0 \end{pmatrix}_{m\times m} Pb \\
&= (Q^{-1})^*\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times m}\begin{pmatrix} I_r & 0 \\ X^* & 0 \end{pmatrix}_{m\times m}(P^{-1})^*P^{-1}Pb \\
&= (Q^{-1})^*\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times m}(P^{-1})^*b = A^*b.
\end{aligned}$$

由此推出 $y = A^{-}b$ 为 $Ax = b$ 的最小二乘解. $\square$

推论 6.5.2 设 $A\in\mathbb{C}^{m\times n}, G\in\mathbb{C}^{n\times m}$, 则 G 是矩阵 A 的一个 $\{1,3\}$- 逆 $\iff$ G 是矩阵方程 $AX=I_m$ 的最小二乘解, 即

$$\|AG-I_m\|_F=\min_{X\in\mathbb{C}^{n\times m}}\|AX-I_m\|_F.$$

证 由矩阵的张量积的性质可知, 矩阵方程 $AX=I_m$ 可化为

$$(I_m\otimes A)\mathrm{vec}(X)=\mathrm{vec}(I_m) \tag{6.5.5}$$

由于矩阵 X 的 F- 范数显然等于其列展开 $\mathrm{vec}(X)$ 的 l_2 范数, 故 X 是矩阵方程 $AX=I_m$ 的最小二乘解 $\iff$ $\mathrm{vec}(X)$ 是方程组 (6.5.5) 的最小二乘解.

由 定理6.5.4, $\mathrm{vec}(X)$ 是上述方程组的最小二乘解 $\iff$

$$\mathrm{vec}(X)=(I_m\otimes A)^{(1,3)}\mathrm{vec}(I_m)+(I_{mn}-(I_m\otimes A)^{(1,3)}(I_m\otimes A))y,\forall\, y$$

因此

$$\mathrm{vec}(X)=(I_m\otimes A^{(1,3)})\mathrm{vec}(I_m)+(I_{mn}-(I_m\otimes A^{(1,3)}A)y,\forall\, y$$

此即

$$X=A^{(1,3)}+(I_n-A^{(1,3)}A)Y,\forall\, Y\in\mathbb{C}^{n\times m},$$

$\iff$ X 是 A 的 $\{1,3\}$- 逆. □

定理 6.5.5 设 $A\in\mathbb{C}^{m\times n}$, 且 $Ax=b$ 为不相容方程组, 则 $Ax=b$ 必有唯一的最小范数的最小二乘解, 这个解即为 $y=A^\dagger b$.

证 由 $A^\dagger$ 的定义及 定理6.5.4 我们知道, $y=A^\dagger b$ 为 $Ax=b$ 的一个最小二乘解. 设 x 为 $Ax=b$ 的任意一个最小二乘解, 则 x 为 $A^*Ax=A^*b$ 的解, 因而 $x=A^\dagger b+z$, 其中 z 为 $A^*Ax=0$ 的解. 因 $A^*Ax=0$ 与 $Ax=0$ 的解空间是一样的, 所以 z 也是 $Ax=0$ 的解, 即 $Az=0$. 由于

$$x^*x=(A^\dagger b+z)^*(A^\dagger b+z)=(A^\dagger b)^*(A^\dagger b)+z^*z+z^*A^\dagger b+(A^\dagger b)^*z,$$

及

$$\begin{aligned}&z^*A^\dagger b=z^*A^\dagger AA^\dagger b=z^*(A^\dagger A)^*A^\dagger b=(A^\dagger Az)^*A^\dagger b=0,\\&(A^\dagger b)^*z=(A^\dagger AA^\dagger b)^*z=(A^\dagger b)^*(A^\dagger A)^*z=(A^\dagger b)^*(A^\dagger Az)=0,\end{aligned}$$

所以

$$x^*x=(A^\dagger b)^*(A^\dagger b)+z^*z\geq(A^\dagger b)^*(A^\dagger b). \tag{6.5.6}$$

由此推出 $y=A^\dagger b$ 为 $Ax=b$ 的最小二乘解中范数最小的.

另外, 由式 (6.5.6) 知, $x^*x=(A^\dagger b)^*(A^\dagger b)$ $\iff$ $z=0$, 即 $x=A^\dagger b$, 所以最小范数的最小二乘解是唯一的. □

关于矩阵方程 $AXB=C$ 的解, 有下述结论 (证明见习题 44)

定理 6.5.6 (Penrose 定理) 矩阵方程 $AXB=C$ 有解 $\iff$ $AA^\dagger CB^\dagger B=C$; 且此时的全体解为 $X=A^\dagger CB^\dagger+Y-A^\dagger AYBB^\dagger$.

由于向量是矩阵的特例, 本节关于线性方程组的几个结论实际上均是 定理6.5.6 的推论 (Penrose 的原始论文正是这样做的).

根据本节的上述结论, 利用广义逆矩阵可以方便地求解线性方程组 (无论是对未知向量还是对未知矩阵), 因此归结为线性模型的应用问题 (可以有非线性的约束) 均可利用广义逆矩阵来解决. 下面的例子是广义逆矩阵在网络流量矩阵估算方面的一个应用.

例 6.5.1 (网络流量矩阵估计)

设 $\mathcal{N}$ 为一具有有限节点的网络系统. 一个**源节点**(origin node) 与一个**目的节点**(destination node) 称为一个**OD 对**. 设 $\mathcal{N}$ 共有 n 个 OD 对. 设第 k 个 OD 对的流量为 x_k(设该 OD 对的源节点为 i, 目的节点为 j, 则 x_k 即为从 i 到 j 的流量). 向量 $X=(x_1,\cdots,x_n)^T$ 称为 $\mathcal{N}$ 的**流量矩阵**(实际上是一个向量). 设 $\mathcal{N}$ 共有 m 个链接, 第 i 个链接的链接数为 y_i. 向量 $Y=(y_1,\cdots,y_m)^T$ 称为 $\mathcal{N}$ 的**链接向量** 或链接负载. 令

$$a_{ij}=\begin{cases}1, & \text{如果第 } j \text{ 个 OD 对的流量经过第 } i \text{ 个链接,}\\ 0, & \text{其它.}\end{cases}$$

则矩阵 $A=(a_{ij})_{m\times n}$ 称为网络 $\mathcal{N}$ 的**路由 (信息) 矩阵**. 路由矩阵, 流量矩阵和链接向量之间的关系如下:

$$AX=Y \tag{6.5.7}$$

通常网络 $\mathcal{N}$ 的路由矩阵和链接向量较易得到, 因此主要的问题是估计流量矩阵 X. 一般而言, 网络中的 OD 对数量 n 远大于链接总数 m, 因此方程 (6.5.7) 是超定方程, 故有无穷多组解. 所以需要寻求满足一定条件的解 $\hat{X}$. 比如可以使用方程 (6.5.7) 的最小范数解 $\hat{X}$ 来估计网络的最小流量 X, 此即求解约束问题

$$\begin{cases}\min X^TX\\ AX=Y\end{cases} \tag{6.5.8}$$

由 定理6.5.3(2), 约束问题 (6.5.8) 的解为 $\hat{X}=A^{\dagger}Y$ 或 $\hat{X}=A^{(1,4)}Y$.

思考题

1. 利用广义逆矩阵如何刻画方程组 $Ax=b$ 的相容性?
2. 方程 $Ax=b$ 的最小范数解是否唯一? 几何意义是什么?
3. 利用矩阵的张量积 (第二章定理 2.6.3) 与广义逆 (定理6.5.6) 求解矩阵方程 $AXB=C$ 有何异同?

习 题 六

1. 证明 定理6.1.1.

2. 设 P_1, P_2 均为投影矩阵, 证明:
(1) $P=P_1+P_2$ 是投影矩阵 $\iff$ $P_1P_2=P_2P_1=0$;
(2) $P=P_1-P_2$ 是投影矩阵 $\iff$ $P_1P_2=P_2P_1=P_2$;
(3) $P_1^*, I-P_1, T^{-1}P_1T$(T 为任意一个非奇异矩阵) 均为投影矩阵.

3. 设 $\mathbb{R}^3$ 的子空间 L 由向量 $e=(1,0,0)^T$ 生成.
(1) 若子空间 M 由 $\alpha=(1,1,0)^T$ 和 $\beta=(1,1,1)^T$ 生成, 求投影矩阵 $P_{L,M}$ 和向量 $x=(2,3,1)^T$ 沿着 M 到 L 上的投影;
(2) 求正交投影矩阵 P_L 和向量 $x=(2,3,1)^T$ 在 L 上的正交投影.

4. 证明 例6.1.3.

5. 证明 $\begin{pmatrix} A \\ 0 \end{pmatrix}^{\dagger} = (A^{\dagger}, 0)$.

6. 证明 命题6.1.1.

7. 设 $A \in \mathbb{C}^{m\times n}$, 又 $U \in \mathbb{C}^{m\times m}$ 和 $V \in \mathbb{C}^{n\times n}$ 均为酉矩阵. 证明 $(UAV)^{\dagger} = V^{*}A^{\dagger}U^{*}$.

8. 设 H 为幂等Hermite矩阵, 证明 $H^{\dagger} = H$.

9. 证明 $A^{\dagger} = A \iff A^2$ 为幂等Hermite矩阵且 $r(A^2) = r(A)$.

10. 证明: 若 A 是正规矩阵, 则 $A^{\dagger}A = AA^{\dagger}$, 且 $(A^n)^{\dagger} = (A^{\dagger})^n$, 其中 n 为正整数.

11. 计算基本矩阵 E_{ij} 的Moore-Penrose广义逆和 $\{1\}$- 广义逆矩阵.

12. 证明 命题6.1.2.

13. 验证 例6.1.11. 如果 $\mathbb{F}[x]_3$ 中的内积定义为 $(f,g) = \int\limits_0^1 f(x)g(x)\,\mathrm{d}\,x$, 计算求导变换 ∂ 的Moore-Penrose广义逆 $\partial^{\dagger}$.

14. (1) 设 $r(BC) = r(B)$. 证明存在矩阵 D 使 $B = BCD$, 且 $C(BC)^{-}$ 是 B 的一个 $\{1\}$- 广义逆矩阵.
(2) 设 $r(BC) = r(C)$. 证明存在矩阵 D 使 $C = DBC$, 且 $(BC)^{-}B$ 是 C 的一个 $\{1\}$- 广义逆矩阵.

15. (1) 设 A 是 $m \times n$ 矩阵, B 是 $m \times r$ 矩阵, 则等式 $AA^{-}B = B \iff$ 存在矩阵 D 使 $B = AD$;
(2) 设 A 是 $m \times n$ 矩阵, B 是 $r \times m$ 矩阵, 则等式 $BA^{-}A = B \iff$ 存在矩阵 D 使 $B = DA$.

16. 证明 定理6.2.2.

17. 详细证明 定理6.2.4.

18. 计算下列矩阵 A 的Moore-Penrose广义逆和 $\{1\}$- 广义逆矩阵, 并验证所得的结果.

$$(1) \begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix}; \quad (2) \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 0 & 2 & 1 \end{pmatrix};$$

19. 证明:(1) 如果矩阵 A 的左逆唯一, 则 A 必是可逆矩阵, 于是左逆等于右逆;
(2) 设矩阵 A 存在左逆但不唯一, 则 A 有无穷多个左逆. 类似地, 如果存在两个右逆, 则必存在无穷多个右逆.

20. 证明 命题6.3.1.

21. 证明 命题6.3.2.

22. 证明 定理6.3.5.

23. 证明: $(AB)^{\dagger} = B^{\dagger}A^{\dagger} \iff A^{\dagger}ABB^{*}A^{*} = BB^{*}A^{*}$ 与 $BB^{\dagger}A^{*}AB = A^{*}AB$ 同时成立.

24. 证明 定理6.4.1.

25. 证明 命题6.4.1.

26. 计算下列矩阵的 $\{1,2\}$- 逆:

$$(1)\quad A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}; \qquad (2)\quad A = \begin{pmatrix} \mathrm{i} & 0 \\ 1 & \mathrm{i} \\ 0 & 1 \end{pmatrix}.$$

27. 计算下列矩阵的 $\{1,3\}$- 逆:

$$(1)\quad A = \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & 2 & 2 \\ -1 & 1 & 0 & 1 \end{pmatrix}; \qquad (2)\quad A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & -2 \\ -1 & 1 & 1 \\ 3 & -1 & -3 \end{pmatrix}.$$

28. 证明 命题6.4.2.

29. 计算下列矩阵的 $\{1,4\}$- 逆:

$$(1)\begin{pmatrix} 1 & 1 \\ 2 & 2 \\ -1 & -1 \end{pmatrix};\quad (2)\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}.$$

30. 证明 定理6.4.3.

31. 证明 命题6.4.3.

32. (1) 哪些矩阵的 $\{1,2\}$- 逆等于它的转置矩阵?
(2) 哪些矩阵的 $\{1,4\}$- 逆等于它的转置矩阵?

33. 试求一个与书中公式形式不同的计算秩为1的矩阵的各种广义逆的公式.

34. 不可逆的方阵可否有可逆的 $\{1,2\}$- 逆或 $\{1,3\}$- 逆或 $\{1,4\}$- 逆?

35. 哪些不可逆的方阵有唯一的 $\{1,2\}$- 逆或 $\{1,3\}$- 逆或 $\{1,4\}$- 逆?

36. 是否存在矩阵其 $\{1,2\}$- 逆或 $\{1,3\}$- 逆或 $\{1,4\}$- 逆不唯一但只有有限个?

37. 设正规矩阵 A 仅有一个非零特征值 λ.
(1) 证明 $A^\dagger = \lambda^{-2}A$;
(2) 试求 A 的 $\{1,2\}$- 逆, $\{1,3\}$- 逆及 $\{1,4\}$- 逆的表达式;
(3) 根据 (1) 与 (2) 计算矩阵 $\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$ 的各种广义逆.

38. 设 L, M 是 $\mathbb{C}^n$ 的子空间. 证明:
(1) $P_{L+M} = (P_L + P_M)(P_L + P_M)^\dagger = (P_L + P_M)^\dagger(P_L + P_M)$;
(2) $P_{L\cap M} = 2P_L(P_L + P_M)^\dagger P_M = 2P_M(P_L + P_M)^\dagger P_L$.

39. 证明: $A^\dagger = A^{(1,4)}AA^{(1,3)}$.

40. 取 A_1, A_2 分别为第18题的 (1) 和 (2), 并设 $b_1 = (1,1,0,1)^T$, $b_2 = (1,1,2)^T$. 分别求出方程组 $A_1x = b_1$ 和 $A_2x = b_2$ 的通解.

41. 设 $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$. 求 $Ax = b$ 的最小范数解.

42. 已知 $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & -2 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. 求矛盾方程组 $Ax = b$ 的最小二乘解.

43. 证明 推论6.5.1.

44. 确定矩阵方程矩阵方程 $AXB = 0$ 的通解, 并以此证明 定理6.5.6.

45. 设 $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.

(1) 当 $b = (1,1,1,1)^T$ 时, 方程组 $Ax = b$ 是否相容?

(2) 当 $b = (1,0,1,0)^T$ 时, 方程组 $Ax = b$ 是否相容?

若方程组相容, 求其通解和最小范数解; 若方程组不相容, 求其最小范数的最小二乘解.

46. 证明线性方程组 $Ax = b$ 有解 $\Longleftrightarrow$ $AA^\dagger b = b$. 这里 $A \in \mathbb{C}^{m\times n}$, $b \in \mathbb{C}^m$.

47. 判断矩阵方程 $AXB = C$ 是否有解, 有解时求其解, 其中

(1) $A=\begin{pmatrix}1&1\\1&0\\0&1\end{pmatrix}$; $B=\begin{pmatrix}1&2&-1\\-1&1&1\end{pmatrix}$; $C=\begin{pmatrix}1&-1&1\\2&-2&2\\-1&1&-1\end{pmatrix}$;

(2) $A=\begin{pmatrix}1&-1&0&-1\\0&0&2&2\\-1&1&0&1\end{pmatrix}$; $B=\begin{pmatrix}1&0&2\\1&0&2\\1&0&2\end{pmatrix}$; $C=\begin{pmatrix}1&0&1\\1&0&1\\0&2&1\end{pmatrix}$.

48. 相容方程组 $Ax=a$ 的通解 $x=A^\dagger a+(I-A^\dagger A)y(\forall y)$ 还可以表示为 $A^\dagger a+N(A)$ 的陪集形式. 证明:
(1) 这个表示是**正交表示**, 即向量 $A^\dagger b$ 与向量 $(I-A^\dagger A)y$ 正交, $\forall y$;
(2) 方程组 $Ax=a$ 与 $Bx=b$ 有公共解 $\iff A^\dagger a-B^\dagger b\in N(A)+N(B)$;
(3) 设方程组 $Ax=a$ 与 $Bx=b$ 有公共解. 试用陪集形式表示其解.

49. 设 A,B,C,D 均为 n 阶矩阵, 且矩阵方程 $AX=B$ 与 $XC=D$ 均有解. 证明:
(1) 两个方程有公共解 $\iff AD=BC$;
(2) 设两个方程有公共解. 试利用广义逆矩阵表示它们的公共通解.(提示: 可先研究齐次方程.)

50. 证明约束优化问题 $\min\{x^Tx\}, Ax=b$ 具有唯一解, 并求该解.

51. 证明约束优化问题 $\min\{\mathrm{tr}(X^TX)-2\mathrm{tr}(X)\}, XA=0$ 的解为 $\hat{X}=I-AA^\dagger$.

52. 设 U 与 W 是线性空间 V 的两个子空间, $\alpha,\beta\in V$. 设 $(\alpha+U)\cap(\beta+W)\neq\emptyset$. 证明:
(1) $(\alpha+U)\cap(\beta+W)=\alpha+P_U(P_U+P_W)^\dagger(\beta-\alpha)+(U\cap W)$;
(2) $(\alpha+U)\cap(\beta+W)=\alpha+(P_{U^\perp}+P_{W^\perp})^\dagger P_{W^\perp}(\beta-\alpha)+(U\cap W)$;
(3) $(\alpha+U)\cap(\beta+W)=\alpha+(I-P_WP_U)^\dagger P_{W^\perp}(\beta-\alpha)+(U\cap W)$.
(提示: 参考第二章习题 75.)