



矩阵论及其应用

主讲老师：徐跃东

2022-2023学年第二学期

广义逆矩阵

□ 广义逆矩阵:

- 左逆与右逆

- 减号广义逆

- 加号广义逆

□ 投影变换

□ 最小二乘解

投影变换

□ **定义** 设 $C^n = L \oplus M$ ，向量 $x \in C^n$, $x = y + z$, $y \in L$, $z \in M$ ，如果线性变换 $\sigma: C^n \rightarrow C^n$, $\sigma(x) = y$ ，则称 σ 为从 C^n 沿子空间 M 到子空间 L 的投影变换。

➤ $R(\sigma) = L$; $N(\sigma) = M$, $\Rightarrow C^n = R(\sigma) \oplus N(\sigma)$

➤ L 和 M 是 σ 的不变子空间; $\sigma|_L = I$; $\sigma|_M = 0$

投影变换

□ 投影变换的矩阵

$$\sigma \xleftrightarrow{\exists \{ \alpha_1, \alpha_2, \dots, \alpha_n \}} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

投影变换

□ 定理: σ 是投影 $\Leftrightarrow \sigma$ 是幂等变换

□ 推论: σ 为投影变换的充要条件是变换矩阵是幂等矩阵

证明 必要性: 设 σ 是 $\mathbf{C}^n$ 空间沿 M 到 L 上的投影变换, 则 $\forall x \in \mathbf{C}^n$, 存在 $y \in L, z \in M$, 使得

$$x = y + z, \quad \sigma(x) = y,$$

于是

$$\sigma^2(x) = \sigma[\sigma(x)] = \sigma(y) = y = \sigma(x),$$

故

$$\sigma^2 = \sigma.$$

充分性: 首先证明

$$\mathbf{C}^n = R(\sigma) + N(\sigma),$$

$\forall x \in \mathbf{C}^n$, 有 $x = \sigma(x) + [x - \sigma(x)]$, 注意到 $\sigma^2 = \sigma$, 有 $\sigma[x - \sigma(x)] = \sigma(x) - \sigma^2(x) = 0$, 于是 $\sigma(x) \in R(\sigma), x - \sigma(x) \in N(\sigma)$, 故

$$\mathbf{C}^n = R(\sigma) + N(\sigma).$$

其次, $\forall x \in R(\sigma) \cap N(\sigma)$, 因 $x \in R(\sigma)$, 故存在 $y \in \mathbf{C}^n$, 使得 $x = \sigma(y)$. 又因 $x \in N(\sigma)$, 故 $\sigma(x) = 0$, 于是

$$0 = \sigma(x) = \sigma^2(y) = \sigma(y) = x,$$

故

$$R(\sigma) \cap N(\sigma) = \{0\}.$$

于是

$$\mathbf{C}^n = R(\sigma) \oplus N(\sigma).$$

此时, $\forall x \in \mathbf{C}^n$, 存在 $y \in R(\sigma), z \in N(\sigma)$, 使得 $x = y + z, y = \sigma(x), \sigma(z) = 0$, 故 $\sigma(x) = \sigma(y) = \sigma^2(x) = \sigma(x) = y$. 这便证明了 σ 是 $\mathbf{C}^n$ 空间沿 $N(\sigma)$ 到 $R(\sigma)$ 上的投影变换. $\square$

投影变换

□ 投影变换矩阵求法

➤ 投影矩阵A, 分块矩阵B,C;

➤ $A(B|C) = (B, 0) \rightarrow A = (B|0)(B|C)^{-1}$

$C^n = L \oplus M$ L 的一组基: $y_1, y_2, \dots, y_r$

M : $z_1, z_2, \dots, z_{n-r}$

C^n 的基 $\{y_1, \dots, y_r, z_1, \dots, z_{n-r}\}$

$$Ay_1 = y_1$$

$$Ay_2 = y_2$$

$$Ay_r = y_r$$

$$Az_1 = 0$$

$$Az_2 = 0$$

$$\vdots$$

$$Az_{n-r} = 0$$

$$A \left[\underbrace{y_1 \dots y_r}_B, \underbrace{z_1 \dots z_{n-r}}_C \right] = \left[\underbrace{y_1 \dots y_r}_B, \underbrace{0 \dots 0}_0 \right]$$

$$A[B|C] = [B|0]$$

$$A = [B|0][B|C]^{-1}$$

投影变换

□例：设 L 是由向量 $(1, 0)^T$ 所生成的子空间， M 是由向量 $(-1, 1)^T$ 所生成的子空间，则 R^2 沿子空间 M 到子空间 L 上的投影矩阵为

$$A = \begin{bmatrix} \underbrace{1}_{B} & \underbrace{0}_0 \\ \underbrace{0} & \underbrace{0}_B \end{bmatrix} \begin{bmatrix} \underbrace{1}_B & \underbrace{1}_C \\ \underbrace{0} & \underbrace{-1}_C \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

正交投影变换

□ 正交投影：定义 设 $\sigma: C^n \rightarrow C^n$ 是投影变换， $C^n = R(\sigma) \oplus N(\sigma)$ ，如果 $R^\perp(\sigma) = N(\sigma)$ ，则称为正交投影。

□ 定理 σ 是正交投影 $\Leftrightarrow$ 投影矩阵 A 满足： $A^2 = A$ 和 $A^H = A$

证明 必要性: 设 A 是线性变换 σ 在某组基下的矩阵, 由定理 4.13 只需证 $A^H = A$.

由

$$\begin{cases} \mathbf{C}^n = R(\sigma) \oplus N(\sigma), \\ R(\sigma)^\perp = N(\sigma), \end{cases}$$

有

$$\begin{cases} \mathbf{C}^n = R(A) \oplus N(A), \\ R(A)^\perp = N(A), \end{cases}$$

又令 $x \in R(A), y \in N(A^H)$, 由 $A^2 = A$ 有 $x = Ax$, 于是

$$(x, y) = (Ax, y) = (x, A^H y) = (x, 0) = 0,$$

所以

$$N(A^H) \subseteq R(A)^\perp.$$

另一方面, 对任意 $y \in R(A)^\perp$, 有

$$(A^H y, A^H y) = (y, AA^H y) = 0.$$

从而 $A^H y = 0$, 即 $y \in N(A^H)$, 故 $R(A)^\perp = N(A^H)$. 于是

$$\begin{cases} \mathbf{C}^n = R(A) \oplus N(A^H), \\ R(A)^\perp = N(A^H). \end{cases}$$

由正交补的惟一性推得 $N(A) = N(A^H)$, 同理, 由 $(A^H)^2 = A^H$, 推得

$$\begin{cases} \mathbf{C}^n = R(A^H) \oplus N(A), \\ R(A^H)^\perp = N(A), \end{cases}$$

故

$$R(A) = R(A^H).$$

$\forall x \in \mathbf{C}^n, x = y + z, y \in R(A) = R(A^H), z \in N(A) = N(A^H)$, 由

$$\begin{aligned} Ax &= Ay + \cancel{Az} = Ay = y, \\ A^H x &= A^H y + \cancel{A^H z} = A^H y = y, \\ A^H &= A. \end{aligned}$$

证得

充分性: 因 $A^2 = A$, 所以 σ 是 C^n 空间的投影变换, 且

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

$\forall x \in R(A), y \in N(A)$, 有

$$(x, y) = (Ax, y) = (A^H x, y) = (x, Ay) = (x, 0) = 0,$$

证得 $N(A) \subseteq R(A)^\perp$. 又 $\forall x \in R(A)^\perp$,

$$\underline{(Ax, Ax)} = (x, A^H Ax) = (x, Ax) = 0,$$

所以 $Ax=0$, 即 $x \in N(A)$, 故 $R(A)^\perp = N(A)$.

$$x \in R(A)^\perp$$

$$y \in R(A) \text{ 则 } (x, y) = 0$$

$$Ax \in R(A) \Rightarrow (Ax, Ax) \text{ 未必是 } 0$$

(

正交投影变换

□例1 设 W 是 C^n 的子空间，证明 存在到 W 的投影变换，使 $R(\sigma) = W$ 。

□在内积空间 C^n 中，存在到 W 的正交投影变换，使 $R(\sigma) = W$ 。

正交投影变换

□定理 设 W 是 C^n 的子空间, $x_0 \in C^n$, $x_0 \notin W$, 如果 σ 是空间 C^n 向空间 W 的正交投影, 则

$$\|\sigma(x_0) - x_0\| \leq \|y - x_0\| \quad \forall y \in W$$

证明 因 σ 是 C^n 空间向 W 的正交投影变换, 所以

$$C^n = W \oplus W^\perp, \quad W = R(\sigma), \quad W^\perp = N(\sigma),$$

$\forall y \in W$, 由于 $[y - \sigma(x_0)] \in W$, $[\sigma(x_0) - x_0] \in W^\perp$, 因此

$$\begin{aligned} \|y - x_0\|^2 &= \|[y - \sigma(x_0)] + [\sigma(x_0) - x_0]\|^2 \\ &= \|y - \sigma(x_0)\|^2 + \|\sigma(x_0) - x_0\|^2 \\ &\geq \|\sigma(x_0) - x_0\|^2, \end{aligned}$$

故

$$\|\sigma(x_0) - x_0\| \leq \|y - x_0\|, \quad \forall y \in W.$$

□

含义: 点 $\sigma(x_0)$ 是空间 W 中与点 x_0 距离最近的点。

正交投影变换

□ 正交投影的求法

$$\begin{aligned} A &= (B \mid 0)(B \mid C)^{-1} \\ &= (B \mid 0)((B \mid C)^H(B \mid C))^{-1} (B \mid C)^H \\ &= B(B^H B)^{-1} B \end{aligned}$$

□ 例：考虑 R^3 中由向量 $\alpha = (1, 2, 0)^T$ 和 $\beta = (0, 1, 1)^T$ 所生成的子空间 L , 求正交投影矩阵 A 和向量 $x = (1, 2, 3)^T$ 沿 $L^\perp$ 到 L 的投影。

正交投影变换

□ 定理 $A^+ A$ 与 AA^+ 的性质

➤ $A^+ A$ 的性质:

$$(A^+ A)^2 = A^+ A, \quad (A^+ A)^H = A^+ A$$

$$C^n = R(A^+) \oplus N(A)$$

$$R^\perp(A^+) = N(A)$$

➤ AA^+ 的性质:

$$(AA^+)^2 = AA^+, \quad (AA^+)^H = AA^+$$

$$C^m = R(A) \oplus N(A^+)$$

$$R^\perp(A) = N(A^+)$$

$A^+ A$ 是正交投影，将向量 x 投影到空间 $R(A^+)$ 中。

AA^+ 是正交投影，将向量 x 投影到空间 $R(A)$ 中。

正交投影变换

证明 由 A^+ 的定义可直接证明(1)和(3)成立. 下面证明(2)成立.

由定理 4.14 及(1)知, $A^+ A$ 是正交投影矩阵, 故

$$\begin{cases} \mathbf{C}^n = R(A^+ A) \oplus N(A^+ A), \\ R(A^+ A)^\perp = N(A^+ A). \end{cases}$$

下证 $R(A^+ A) = R(A^+)$, $N(A^+ A) = N(A)$.

显然, $R(A^+ A) \subseteq R(A^+)$, $N(A) \subseteq N(A^+ A)$, 只需证 $R(A^+) \subseteq R(A^+ A)$, $N(A^+ A) \subseteq N(A)$.

$\forall y \in R(A^+)$, 存在 x 使 $y = A^+ x = A^+ A(A^+ x)$, 故 $R(A^+) \subseteq R(A^+ A)$.

$\forall x \in N(A^+ A)$, $A^+ Ax = 0$, $Ax = AA^+ Ax = 0$, 故 $N(A^+ A) \subseteq N(A)$.

同理可证(4)成立.

□

广义逆矩阵

□ 广义逆矩阵:

- 左逆与右逆

- 减号广义逆

- 加号广义逆

□ 投影变换

□ 最小二乘解

线性最小二乘问题

一、最小二乘问题的一般提法

已知函数 $f(x)$ 在 m 个点上的数据表，寻求其近似函数。

x_i	x_1	x_2	$\dots$	x_m
$f(x_i)$	f_1	f_2	$\dots$	f_m

设 $f(x)$ 的近似函数为

$$F(x) = \alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) + \dots + \alpha_n \varphi_n(x)$$

其中 $\{\varphi_i(x)\}_{i=1}^n$ 是某函数族中的已知线性无关函数。

寻求一组常数 $\alpha_i (i=1, 2, \dots, n)$, 要求

$$\mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} f(x_1) - F(x_1) \\ f(x_2) - F(x_2) \\ \vdots \\ f(x_m) - F(x_m) \end{pmatrix}$$

称为
残向量

的2-范数达到最小:

$$\|\mathbf{r}\|_2 = \min$$

如果 $m=n$, $\varphi_i(x) = x^{i-1}$ 以及 $\mathbf{r} = 0$ 即多项式插值。
且

记 $A = \begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{pmatrix} \quad b = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}$

$x = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T$ 则得到最小二乘问题:

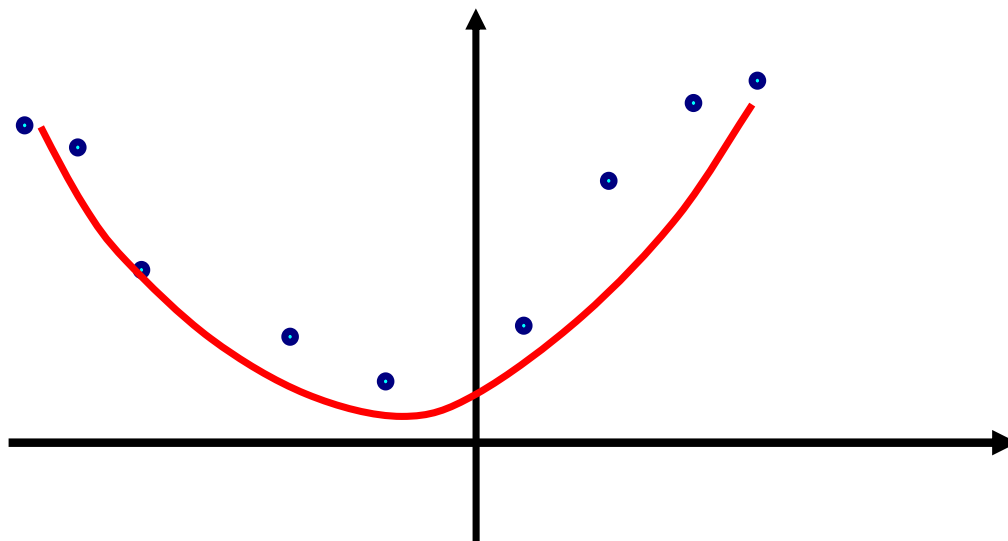
$$\|r\|_2^2 = \|b - Ax\|_2^2 = \min \quad A \in R^{m \times n}$$

上述问题的解也称为方程组 $Ax = b$ 的最小二乘解。

当 $m > n$ 时称之为超定（或矛盾）方程组。

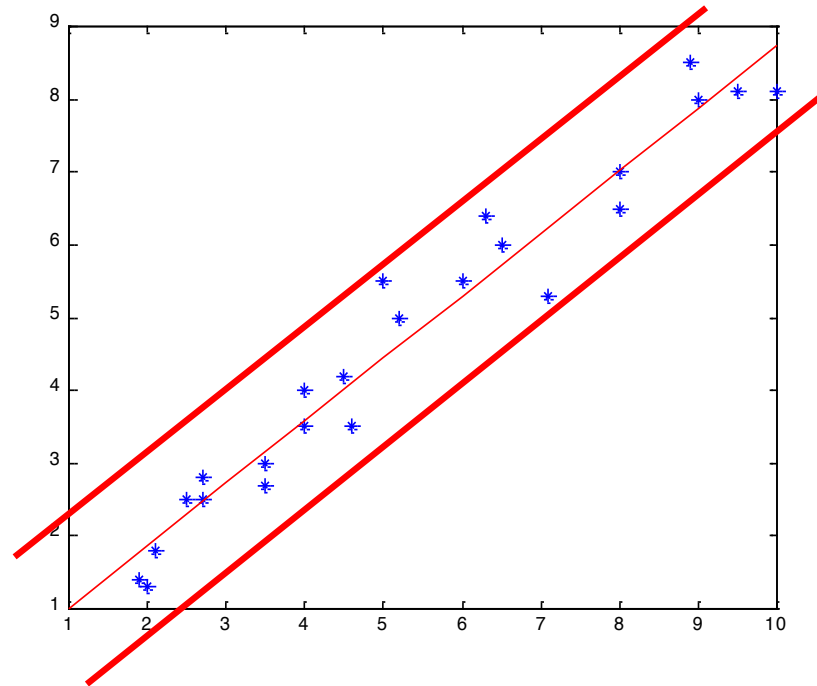
二、最小二乘多项式拟合

“**曲线拟合**”，是指根据给定的数据表，寻找一个简单的表达式来“**拟合**”该组数据，此处的“**拟合**”的含义为：不要求该表达式对应的近似曲线**完全通过**所有的数据点，只要求该**近似曲线**能够反映数据的基本变化趋势。



例1：考察某种纤维的强度与其拉伸倍数的关系。

编号	拉伸倍数	强 度	编号	拉伸倍数	强 度
1	1.9 x_i	1.4 y_i	13	5 x_i	5.5 y_i
2	2	1.3	14	5.2	5
3	2.1	1.8	15	6	5.5
4	2.5	2.5	16	6.3	6.4
5	2.7	2.8	17	6.5	6
6	2.7	2.5	18	7.1	5.3
7	3.5	3	19	8	6.5
8	3.5	2.7	20	8	7
9	4	4	21	8.9	8.5
10	4	3.5	22	9	8
11	4.5	4.2	23	9.5	8.1
12	4.6	3.5	24	10	8.1



可认为强度与拉伸倍数之间的主要关系是线性关系

$$y \approx \varphi(x) = a + bx$$

该直线称为这一问题的数学模型。

怎样确定 a, b ? 采用最小二乘的思想!

$$\text{令 } S(a, b) = \sum_{i=1}^m (a + bx_i - y_i)^2$$

问题转化为求参数 a, b 使 $S(a, b)$ 达到最小值。

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0 \quad \longrightarrow \quad \begin{bmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{bmatrix}$$

$$\begin{pmatrix} 24 & 127.5 \\ 127.5 & 829.61 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 113.1 \\ 731.6 \end{pmatrix} \quad \begin{aligned} a &\approx 0.1505 \\ b &\approx 0.8587 \end{aligned}$$

$$\varphi(x) = 0.1505 + 0.8587x$$

这种求线性函数 $y=a+bx$ 的过程称为线性拟合。

设 $f(x)$ 的近似函数为 $p(x) = \sum_{j=0}^n a_j x^j$

寻求 $a_i (0 \leq i \leq n)$, 使得

$$S(a_0, a_1, a_2, \dots, a_n) = \sum_{i=1}^m (p(x_i) - y_i)^2 = \min$$

则称 $p(x)$ 为函数 $f(x)$ 的多项式拟合。

$a_i (0 \leq i \leq n)$ 满足下列法方程组:

$$\sum_{j=0}^n \sum_{i=1}^m x_i^{j+l} a_j = \sum_{i=1}^m x_i^l y_i, \quad l = 0, 1, 2, \dots, n$$

即 $A^T A x = A^T b$

三、最小二乘问题解的存在性、唯一性

定义：设 $A \in R^{m \times n}$ ，若存在 $x \in R^n$ 精确地满足 $Ax = b$ ，则称该方程组是相容的。

定理：方程组 $Ax = b$ 相容的充要条件是

$$\text{rank}(A) = \text{rank}([A, b])$$

引理：设 $A \in R^{m \times n}$ ，且 $\text{rank}(A) = r > 0$

则总存在分解 $A = FG$

满秩分解

其中 $F \in R^{m \times r}$ ， $G \in R^{r \times n}$ ， $\text{rank}(F) = \text{rank}(G) = r$

定理： $\eta \in R^n$ 是方程组 $Ax = b; A \in R^{m \times n}$ 的最小二乘解
的充要条件是 η 为方程组 $A^T Ax = A^T b$ 的解。

证明：  充分性 设 η 是 $A^T Ax = A^T b$ 的解

对 $\forall y \in R^n$ ， 令 $y = \eta + z$

$$(x, x) = \|x\|_2^2$$

$$\|Ay - b\|_2^2 = \|A(\eta + z) - b\|_2^2$$

$$= \|A\eta - b\|_2^2 + \|Az\|_2^2 + 2(Az)^T (A\eta - b)$$

$$= \|A\eta - b\|_2^2 + \|Az\|_2^2 + 2z^T A^T (A\eta - b)$$

$$= \|A\eta - b\|_2^2 + \|Az\|_2^2 \geq \|A\eta - b\|_2^2$$

必要性

设 $\eta = [\eta_1, \eta_2, \dots, \eta_n]^T$ 是方程组的最小二乘解

记 $r = Ax - b$, 则 η 必使 $\|r\|_2^2$ 达到极小.

由极值的必要条件知: $\nabla \|r\|_2^2 \Big|_{\eta} = 0$

$$\text{即 } \left. \frac{\partial \|r\|_2^2}{\partial x_i} \right|_{\eta} = 0; i = 1, 2, \dots, n$$

$$A^T (A\eta - b) = 0$$

$$\frac{\partial \|r\|_2^2}{\partial x_i} = \frac{\partial}{\partial x_i} \left(\sum_{k=1}^m \left(\sum_{j=1}^n a_{kj} x_j - b_k \right)^2 \right)$$

$$= 2 \sum_{k=1}^m a_{ki} \left(\sum_{j=1}^n a_{kj} x_j - b_k \right) \Rightarrow 2A^T (Ax - b)$$

称 $A^T A x = A^T b$ 为方程组 $Ax = b$ 的法方程组

推论：若 $\text{rank} A = n (n \leq m)$ ，则方程组有

唯一的最小二乘解 $\eta = (A^T A)^{-1} A^T b$ 。

定理：方程组 $Ax = b; A \in R^{m \times n}$ 必存在最小二乘解。

证明：记 $\text{rank}(A) = r > 0$ 则存在满秩分解 $A = FG$

法方程组可写成： $G^T F^T FGx = G^T F^T b$

可以验证 $\tilde{x} = G^T (GG^T)^{-1} (F^T F)^{-1} F^T b$

是法方程组的一个解，故是原方程组的一个最小二乘解



推论 若 $\text{rank}A = r < n$, 则方程组 $Ax = b$

有无穷多个最小二乘解。

定义：方程组 $Ax = b$ 的所有最小二乘解中2-范数最小者称为方程组的极小最小二乘解。

定理：方程组 $Ax = b$ 存在唯一的极小最小二乘解，

且可以表示为 $\tilde{x} = G^T (GG^T)^{-1} (F^T F)^{-1} F^T b$

其中 $A = FG$ 为满秩分解。

证明：由前述定理可知， $\tilde{x}$ 是一个最小二乘解。

设 η 是方程组的任一最小二乘解，下证： $\|\tilde{x}\|_2 \leq \|\eta\|_2$

$$\left. \begin{aligned} A^T A \tilde{x} &= A^T b \\ A^T A \eta &= A^T b \end{aligned} \right\}$$

$$\Rightarrow A^T A(\eta - \tilde{x}) = 0$$

$$G^T F^T F G(\eta - \tilde{x}) = 0$$

$$G G^T F^T F G(\eta - \tilde{x}) = 0 \Rightarrow G(\eta - \tilde{x}) = 0$$

$$\|\eta\|_2^2 = \|\tilde{x} + (\eta - \tilde{x})\|_2^2$$

$$= \|\tilde{x}\|_2^2 + \|\eta - \tilde{x}\|_2^2 + 2(\eta - \tilde{x}, \tilde{x}) = 0$$

$$= \|\tilde{x}\|_2^2 + \|\eta - \tilde{x}\|_2^2 \geq \|\tilde{x}\|_2^2$$

例：求下列方程组的最小二乘解

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 2 & -4 & 3 \\ 4 & -7 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ 1 \\ -6 \end{bmatrix} \quad x^* = \begin{bmatrix} \frac{75}{7} \\ \frac{88}{7} \\ \frac{69}{7} \end{bmatrix}$$

解： $\text{rank}(A) = 3 \neq 4 = \text{rank}(A, b)$

$$A^T A = \begin{bmatrix} 21 & -38 & 23 \\ -38 & 70 & -43 \\ 23 & -43 & 27 \end{bmatrix} \quad A^T b = \begin{bmatrix} -26 \\ 49 \\ -28 \end{bmatrix}$$

广义逆矩阵与最小二乘解

若 $A \in R^{m \times n}$, 则 Moore-Penrose 方程变为

$$\begin{cases} AXA = A \\ XAX = X \\ (AX)^T = AX \\ (XA)^T = XA \end{cases}$$

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

$$Ax = b \text{ 的极小最小二乘解} \\ x = A^+ b$$

$$\text{若 } m=n, \text{ 且 } A \text{ 非奇异, 则 } A^+ = A^{-1}$$

最小二乘解

□ 定理： 设 $A \in C_{m \times n}$, $b \in C_m$, 则 $x_0 = A_+ b$ 是线性方程组 $Ax = b$ 的最佳的最小二乘解.



证明 由定理 4.15 知道, AA^+ 是 C^m 空间向 $R(A)$ 上的一个正交投影变换所对应的矩阵, 再由定理 4.16, 有

$$\|A(A^+b) - b\| \leq \|Ax - b\|, x \in C^n,$$

故 $x_0 = A^+b$ 是 $Ax=b$ 的最小二乘解.

又由定理 4.15, 有

$$\begin{cases} C^m = R(A) \oplus N(A^+), \\ R(A)^\perp = N(A^+), \end{cases}$$

每个 $b \in C^m$ 可惟一地分解为

$$b = b_1 + b_2, \quad b_1 \in R(A), \quad b_2 \in N(A^+),$$

$\forall x \in C^n$, 有

$$\begin{aligned} \|Ax - b\|^2 &= \|(Ax - b_1) + (-b_2)\|^2 \\ &= \|Ax - b_1\|^2 + \|b_2\|^2, \end{aligned}$$

因此, u 是 $Ax=b$ 的最小二乘解当且仅当 u 是 $Ax=b_1$ 的解.

设 u 是 $Ax=b$ 的任一个最小二乘解, 因 x_0 也是 $Ax=b$ 的最小二乘解, 故 $A(x_0 - u) = Ax_0 - Au = b_1 - b_1 = 0$, 从而 $(x_0 - u) \in N(A)$, 又因 $x_0 = A^+b \in R(A^+)$, $R(A^+)^\perp = N(A)$, 故 x_0 与 $(x_0 - u)$ 正交. 因而

$$\|u\|^2 = \|x_0 + (u - x_0)\|^2 = \|x_0\|^2 + \|u - x_0\|^2 \geq \|x_0\|^2,$$

证得 $x_0 = A^+b$ 是 $Ax=b$ 的最佳的最小二乘解. $\square$

推论 设矩阵方程 $AX=B$, 其中 $A \in C^{m \times n}$, $B \in C^{m \times k}$, 则 $X_0 = A^+B$ 是 $AX=B$ 的最佳的最小二乘解.



Thanks