§6 伪逆

6.1 伪逆

线性代数的中心问题是求解线性方程组$A\mathbf{x} = \mathbf{b}$.

最简单的情形是, 若A为n阶可逆方阵, 则对任意n维向量$\mathbf{b}$, 方程组$A\mathbf{x} = \mathbf{b}$有唯一解$\mathbf{x} = A^{-1}\mathbf{b}$.

问题: 对$m \times n$的矩阵A, 定义其伪逆(pseudoinverse)A^+, 使得当A为n阶可逆方阵时, 有$A^+ = A^{-1}$.

6.1 伪逆

设$m\times n$实矩阵$A=U\Sigma V^T$.　　(SVD)

其中U,V分别为m阶, n阶正交阵; Σ为$m\times n$矩阵,

前$r=r(A)$个“对角元”为A的奇异值$\sigma_1\geq\cdots\geq\sigma_r>0$.

即:

$$A(\overbrace{\underbrace{\mathbf{v}_1\ \cdots\ \mathbf{v}_r}_{C(A^T)}\underbrace{\mathbf{v}_{r+1}\ \cdots\ \mathbf{v}_n}_{N(A)}}^{V_{n\times n}})=(\overbrace{\underbrace{\mathbf{u}_1\ \cdots\ \mathbf{u}_r}_{C(A)}\underbrace{\mathbf{u}_{r+1}\ \cdots\ \mathbf{u}_m}_{N(A^T)}}^{U_{m\times m}})\underbrace{\begin{pmatrix}\sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0\end{pmatrix}}_{\Sigma_{m\times n}}$$

6.1 伪逆

$A\mathbf{v}_j = \sigma_j \mathbf{u}_j \quad (1 \le j \le r),$

$A\mathbf{v}_j = \mathbf{0} \quad (r+1 \le j \le n).$

若A可逆, 则$A^{-1}\mathbf{u}_j = \frac{1}{\sigma_j}\mathbf{v}_j \quad (1 \le j \le r = m = n).$

对$A_{m\times n}$, 令

$$A^+(\mathbf{u}_1 \ \cdots \ \mathbf{u}_r \mathbf{u}_{r+1} \ \cdots \ \mathbf{u}_m) = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_r \mathbf{v}_{r+1} \ \cdots \ \mathbf{v}_n) \underbrace{\begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ & & & 0 \end{pmatrix}_{n\times m}}_{\Sigma^+}$$

6.1 伪逆

即:

$$A^{+}\mathbf{u}_j = \frac{1}{\sigma_j}\mathbf{v}_j \quad (1 \le j \le r),$$
$$A^{+}\mathbf{u}_j = \mathbf{0} \quad (r+1 \le j \le m).$$

$$\boxed{A^{+}_{n\times m} := V\Sigma^{+}U^{T}}$$

6.1 伪逆

(1) 若A可逆, 即$r = m = n$, 则

$$A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T = A^+.$$

(2) $AA^+ = (U\Sigma V^T)(V\Sigma^+U^T) = U\Sigma\Sigma^+U^T = U \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times m} U^T$

- $(AA^+)^T = AA^+$
- $AA^+ = \mathbf{u}_1\mathbf{u}_1^T + \cdots + \mathbf{u}_r\mathbf{u}_r^T$
- $AA^+ = \mathbb{R}^m$到$C(A)$的正交投影矩阵

$(AA^+|_{C(A)} = id, AA^+|_{N(A^T)} = 0)$

6.1 伪逆

(3) $A^+A = (V\Sigma^+U^T)(U\Sigma V^T) = V\Sigma^+\Sigma V^T = V\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times n} V^T$

- $(A^+A)^T = A^+A$
- $A^+A = \mathbf{v}_1\mathbf{v}_1^T + \cdots + \mathbf{v}_r\mathbf{v}_r^T$
- $A^+A = \mathbb{R}^n$到$C(A^T)$的正交投影矩阵

$(A^+A|_{C(A^T)} = id, A^+A|_{N(A)} = 0)$

6.1 伪逆

若$r = n$ (A列满秩), 则

$$AA^+ = U \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times m} U^T = \mathbb{R}^m\text{到}C(A)\text{的正交投影矩阵},$$

$$A^+A = VV^T = I_n.$$

称A^+为A的左逆.

6.1 伪逆

若$r=m$ (A行满秩), 则

$AA^+=UU^T=I_m$,

$A^+A=V\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{n\times n}V^T=\mathbb{R}^n$到$C(A^T)$的正交投影矩阵,

称A^+为A的右逆.

若$r=m=n$ (A满秩), 则

$AA^+=A^+A=I, A^+=A^{-1}$ (双边逆).

6.1 伪逆

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

解： A的奇异值分解为

$$A=U\Sigma V^T=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\0&0&1\end{pmatrix}\begin{pmatrix}2&0\\0&0\\0&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix},$$

故

$$A^+=V\Sigma^+U^T=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}\frac{1}{2}&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{2}}&0\\0&0&1\end{pmatrix}=\frac{1}{4}\begin{pmatrix}1&1&0\\1&1&0\end{pmatrix}$$

6.1 伪逆

例: 若矩阵A的秩为1, 则$A=\sigma\mathbf{u}\mathbf{v}^T, A^+=\frac{1}{\sigma}\mathbf{v}\mathbf{u}^T$.则$AA^+=\mathbf{u}\mathbf{u}^T$, 为到$\mathbf{u}$所在直线的正交投影矩阵; $A^+A=\mathbf{v}\mathbf{v}^T$, 为到$\mathbf{v}$所在直线的正交投影矩阵.

特别, 任意非零向量$\mathbf{x}$的伪逆$\mathbf{x}^+=\frac{\mathbf{x}^T}{\mathbf{x}^T\mathbf{x}}$, 单位向量$\mathbf{x}$的伪逆$\mathbf{x}^+=\mathbf{x}^T$.

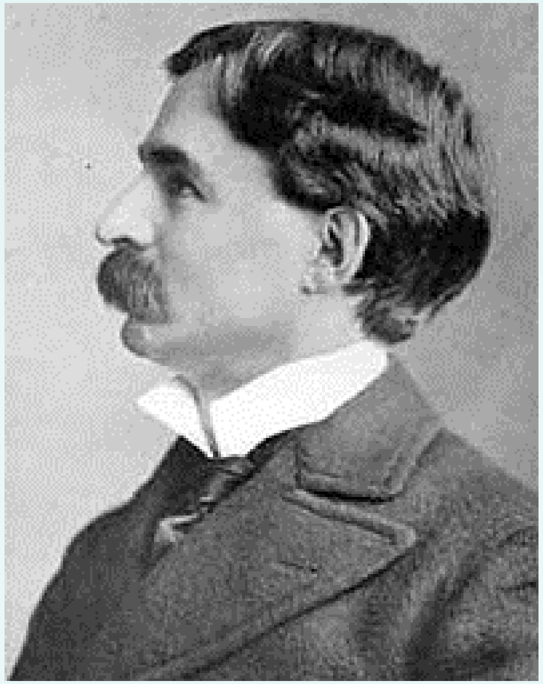
矩阵$A=\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$的伪逆为自身.

矩阵$B=\left(\begin{smallmatrix}0&1&0\\0&0&0\end{smallmatrix}\right)$的伪逆为$B^T$.

6.1 伪逆

例： n阶Jordan矩阵$J_n = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{n\times n}$ 的伪逆为J_n^T.

6.2 Moore – Penrose 伪逆

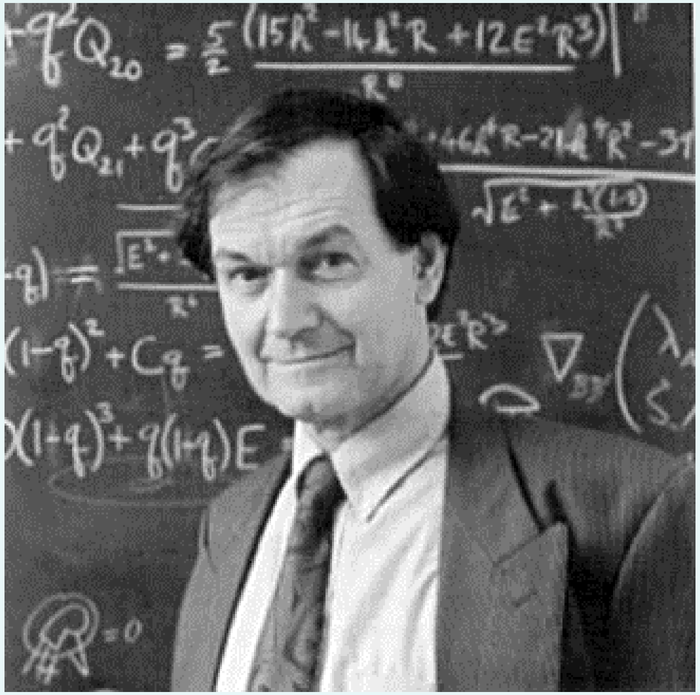


Eliakim Hastings Moore (1862 - 1932),

美国数学家,

是二十世纪初美国数学的奠基人.

6.2 Moore – Penrose 伪逆



Sir Roger Penrose (1931 -),
英国著名数学物理学家,
1988年Wolf奖得主,
与Stephen Hawking合作证明了
广义相对论的奇点存在性.

6.2 Moore – Penrose 伪逆

对$m \times n$矩阵A, Moore意义下的伪逆为满足

$$\boxed{AX = P_{C(A)}, \quad XA = P_{C(X)}}$$

的$n \times m$矩阵X. 这里P_V表示到空间V的正交投影矩阵.

6.2 Moore – Penrose 伪逆

1955年, 英国剑桥大学博士研究生Penrose给出了伪逆的如下定义:

设A为$m \times n$实矩阵, 若$n \times m$矩阵X满足如下方程组:

$$\boxed{\begin{aligned} AXA &= A & (1)\\ XAX &= X & (2)\\ (AX)^T &= AX & (3)\\ (XA)^T &= XA & (4)\end{aligned}}$$

则称X为矩阵A的Penrose伪逆.

6.2 Moore – Penrose 伪逆

命题: 给定任一$m \times n$实矩阵A, A的伪逆A^+是满足Penrose方程组(1) – (4)的唯一$n \times m$矩阵.

证明:

存在性: 由$A^+ = V\Sigma^+U^T, A = U\Sigma V^T$容易验证$A^+$满足Penrose方程组.

唯一性: 若X与Y均是A的Penrose伪逆, 则对X与Y反复利用Penrose方程可得

$$X = XAX = XX^TA^T = XX^TA^TY^TA^T = XAY = XAA^TY^TY$$
$$= A^TY^TY = YAY = Y.$$

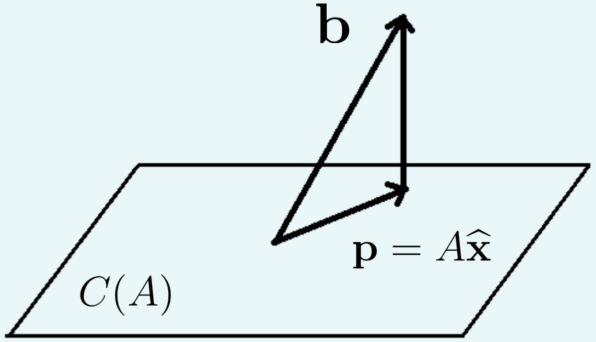
6.3 最小二乘法

- 线性方程组$A\mathbf{x}=\mathbf{b}$ 有解 $\Longleftrightarrow$ $\mathbf{b}\in C(A)$.
  $\mathbf{b}\in C(A)$ 且A列满秩 $\Longrightarrow$ 存在唯一解$\mathbf{x}$.
  特别, A可逆 $\Longrightarrow$ $\mathbf{x}=A^{-1}\mathbf{b}$.
- $A\mathbf{x}=\mathbf{b}$ 无解 $\Longleftrightarrow$ $\mathbf{b}\notin C(A)$.

6.3 最小二乘法

- $A\mathbf{x} = \mathbf{b}$ 无解 $\Longleftrightarrow \mathbf{b} \notin C(A)$.

求近似解$\widehat{\mathbf{x}}$, 使得$||\mathbf{b} - A\widehat{\mathbf{x}}||$最小 (最小二乘解)



$$\Updownarrow$$

$$\mathbf{e} = \mathbf{b} - A\widehat{\mathbf{x}} \perp C(A)$$

$$\Updownarrow \quad C(A) \oplus N(A^T) = \mathbb{R}^m$$

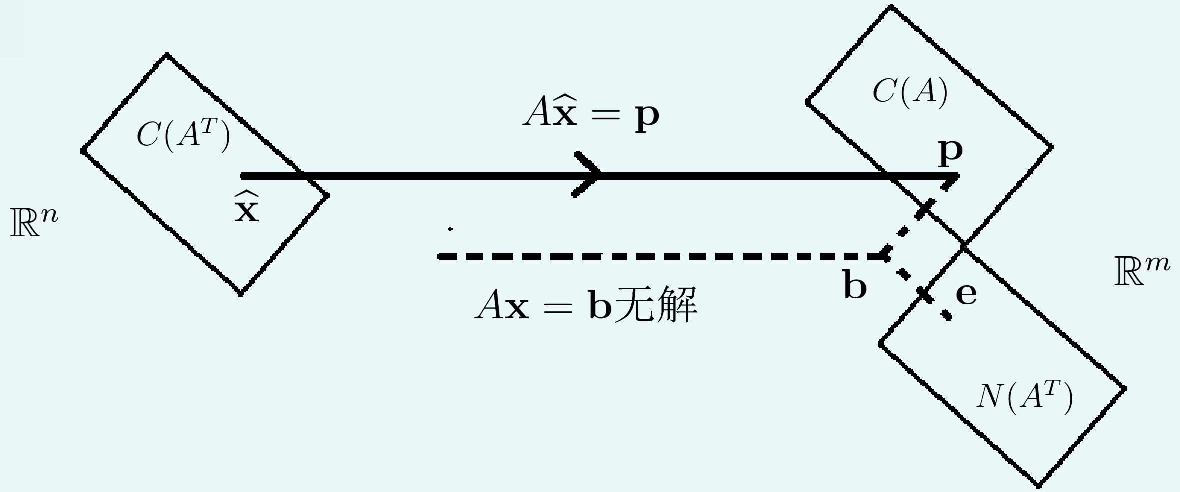
$$\mathbf{e} = \mathbf{b} - A\widehat{\mathbf{x}} \in N(A^T)$$

$$\Updownarrow$$

$$A^T A\widehat{\mathbf{x}} = A^T\mathbf{b}$$

6.3 最小二乘法

(a) 若$r(A)=n$ (A列满秩), 则$r(A^TA)=r(A)=n$. 即A^TA可逆, 于是有唯一最小二乘解: $\widehat{\mathbf{x}}=(A^TA)^{-1}A^T\mathbf{b}$.



(b) 若$r(A) < n$ (A列相关), 则$r(A^TA) = r(A) < n$. 正规方程 $A^TA\widehat{\mathbf{x}} = A^T\mathbf{b}$解不唯一, 即最小二乘解不唯一.

6.3 最小二乘法

命题: $\mathbf{x}^+ := A^+\mathbf{b}$为一个最小二乘解.

证明:

$A^T\mathbf{b} - A^TA\mathbf{x}^+ = A^T(\mathbf{b} - A\mathbf{x}^+) = A^T(\mathbf{b} - AA^+\mathbf{b})$

由于AA^+为$\mathbb{R}^m$到$C(A)$的正交投影矩阵且$\mathbb{R}^m = C(A) \oplus N(A^T)$, 故

$\mathbf{b} - AA^+\mathbf{b} \in N(A^T)$,

$A^T\mathbf{b} - A^TA\mathbf{x}^+ = A^T(\mathbf{b} - AA^+\mathbf{b}) = \mathbf{0}.$

6.3 最小二乘法

命题: 在$A\mathbf{x}=\mathbf{b}$的所有最小二乘解中, $\mathbf{x}^+$的长度最小.

称$\mathbf{x}^+=A^+\mathbf{b}$为$A\mathbf{x}=\mathbf{b}$的最佳最小二乘解.

证明: 设$\hat{\mathbf{x}}$也是$A^TA\hat{\mathbf{x}}=A^T\mathbf{b}$的一个解, 即一个最小二乘解. 于是,

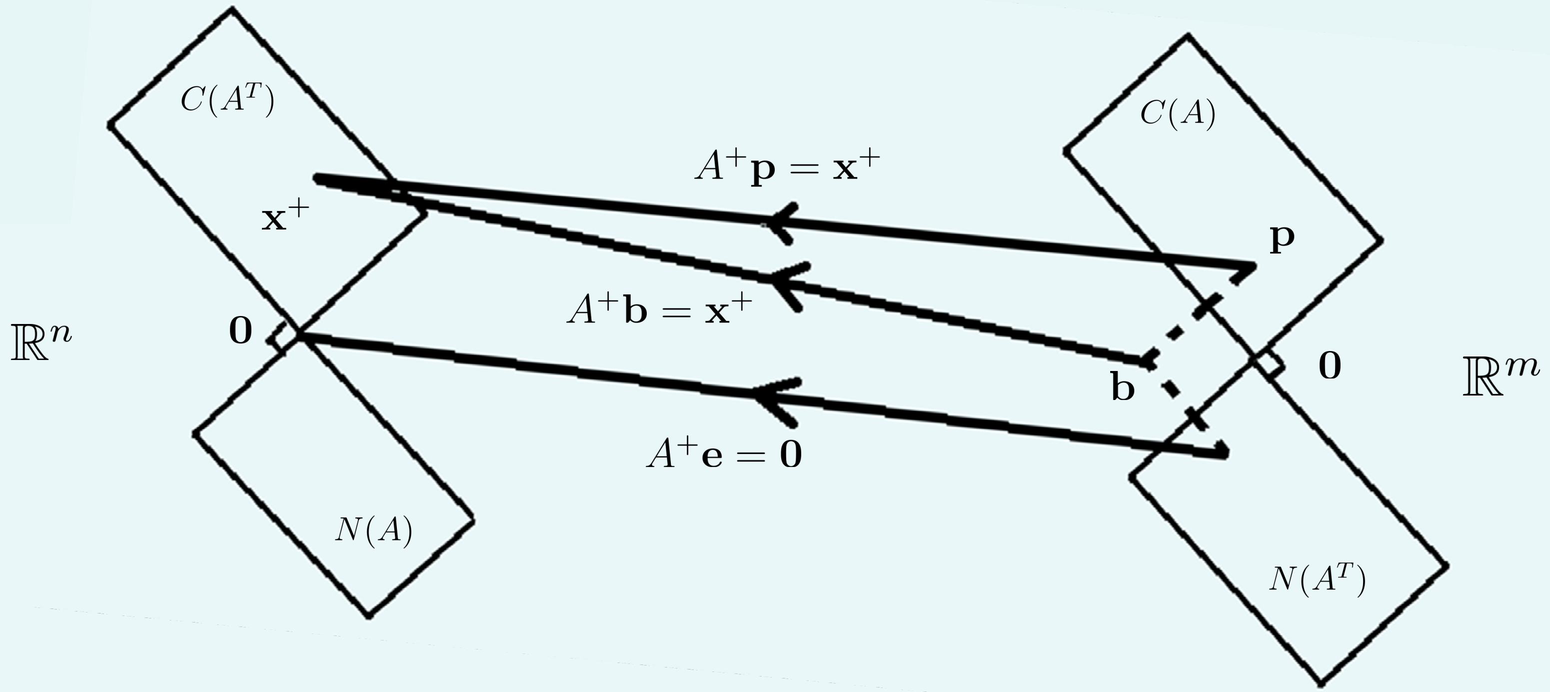
$$\left.\begin{array}{r} A^TA\hat{\mathbf{x}}=A^T\mathbf{b} \\ A^TA\mathbf{x}^+=A^T\mathbf{b} \end{array}\right. \Rightarrow A^TA(\hat{\mathbf{x}}-\mathbf{x}^+)=\mathbf{0} \Rightarrow \hat{\mathbf{x}}-\mathbf{x}^+\in N(A^TA)=N(A)$$

而$\mathbf{x}^+=A^+\mathbf{b}\in C(A^T)$, 故$\mathbf{x}^+\perp\hat{\mathbf{x}}-\mathbf{x}^+$,

$$\Rightarrow ||\hat{\mathbf{x}}||^2=||(\hat{\mathbf{x}}-\mathbf{x}^+)+\mathbf{x}^+||^2=||\hat{\mathbf{x}}-\mathbf{x}^+||^2+||\mathbf{x}^+||^2\geq||\mathbf{x}^+||^2$$

即$\mathbf{x}^+$是长度最小的最小二乘解.

6.3 最小二乘法


$C(A^T)$
$\mathbf{x}^+$
$\mathbb{R}^n$
$\mathbf{0}$
$N(A)$
$A^+\mathbf{p} = \mathbf{x}^+$
$A^+\mathbf{b} = \mathbf{x}^+$
$A^+\mathbf{e} = \mathbf{0}$
$C(A)$
$\mathbf{p}$
$\mathbf{b}$
$\mathbf{0}$
$\mathbb{R}^m$
$N(A^T)$