

Pseudoinverse  
and Least  
Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

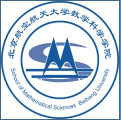
Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

最小二乘法

# 伪逆与最小二乘法

## Pseudoinverse and Least Squares



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

最小二乘法

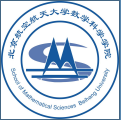
考虑解线性方程组  $A\mathbf{x} = \mathbf{b}$  的问题。我们知道当  $A$  是可逆方阵的时候，该线性方程组有唯一解  $\mathbf{x} = A^{-1}\mathbf{b}$ 。

当  $A$  没有逆的时候，有什么补救办法呢？

希望：定义一个伪逆 (Pseudoinverse)，记作  $A^+$ ，使得当  $A$  可逆的时候， $A$  的伪逆就是它的逆，即

$$A^+ = A^{-1}$$

如果可能的话，还要求当  $A\mathbf{x} = \mathbf{b}$  无解时， $\mathbf{x} = A^+\mathbf{b}$  是“近似程度”最好的“解”。



# 内容提要

## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质

广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质

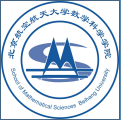
伪逆与线性方程组

最小二乘法

### 1 广义逆

### 2 Moore-Penrose 伪逆

### 3 最小二乘法



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

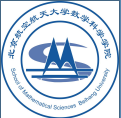
伪逆的定义与性质  
伪逆与线性方程组

### 最小二乘法

## 1 广义逆

## 2 Moore-Penrose 伪逆

## 3 最小二乘法



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质

广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质

伪逆与线性方程组

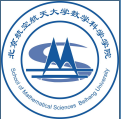
最小二乘法

## 矩阵的广义逆, Generalized Inverse

$m \times n$  矩阵  $A$  的广义逆指的是一个  $n \times m$  的矩阵  $X$ , 满足

$$AXA = A$$

通常把  $A$  的广义逆记作  $A^-$ .



## 矩阵的广义逆, Generalized Inverse

$m \times n$  矩阵  $A$  的广义逆指的是一个  $n \times m$  的矩阵  $X$ , 满足

$$AXA = A$$

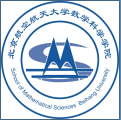
通常把  $A$  的广义逆记作  $A^-$ .

## 广义逆的普遍存在性

任意矩阵  $A$  的广义逆  $A^-$  总存在。设  $P, Q$  分别为  $m$  阶与  $n$  阶可逆阵, 使得  $A = P \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} Q$ , 那么

$$A^- = Q^{-1} \begin{pmatrix} I_r & * \\ *' & *'' \end{pmatrix} P^{-1}$$

其中  $r = r(A)$ ,  $*, *', *''$  分别是大小为  $r \times (n - r)$ ,  $(m - r) \times r$ , 以及  $(m - r) \times (n - r)$  的块, 块中元素可以任取。



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

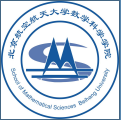
最小二乘法

### 广义逆与齐次线性方程组

齐次线性方程组  $A\mathbf{x} = \mathbf{0}$  的解的全体为

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

其中  $A^-$  取遍  $A$  的所有广义逆,  $\mathbf{z}$  取遍所有的  $n$  维列向量。



## 广义逆与齐次线性方程组

齐次线性方程组  $A\mathbf{x} = \mathbf{0}$  的解的全体为

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

其中  $A^-$  取遍  $A$  的所有广义逆,  $\mathbf{z}$  取遍所有的  $n$  维列向量。

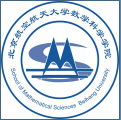
## 广义逆与非齐次线性方程组

非齐次线性方程组  $A\mathbf{x} = \mathbf{b}$  有解, 当且仅当存在  $A$  的广义逆  $A^-$  使得

$$\mathbf{b} = AA^-\mathbf{b}.$$

有解时, 它的解的全体为  $A^-\mathbf{b}$ ,  $A^-$  取遍  $A$  的所有广义逆。





## Pseudoinverse and Least Squares

### 广义逆

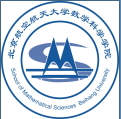
广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

伪逆的定义与性质  
伪逆与线性方程组

### 最小二乘法

广义逆只解决了我们的第一部分要求，即它“像是”逆，而且当矩阵的确可逆的时候它就是矩阵通常意义下的逆。但是它并不唯一，而且仍然没有解决求“最佳近似解”的需求。



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

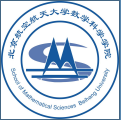
伪逆的定义与性质  
伪逆与线性方程组

### 最小二乘法

## 1 广义逆

## 2 Moore-Penrose 伪逆

## 3 最小二乘法



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

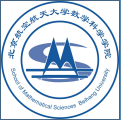
最小二乘法

### Moore-Penrose 伪逆, Pseudoinverse

$m \times n$  矩阵  $A$  的 Moore-Penrose 伪逆, 或简称伪逆, 指的是同时满足如下四个条件的  $n \times m$  矩阵  $X$

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

通常,  $A$  的伪逆被记作  $A^+$ .



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

最小二乘法

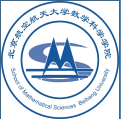
### 伪逆的存在性与唯一性

任意  $m \times n$  矩阵  $A$  的伪逆  $A^+$  都存在, 而且

$$A^+ = V\Sigma^+U^T,$$

其中  $A = U\Sigma V^T$  为矩阵  $A$  的奇异值分解, 奇异值为  $\sigma_1, \dots, \sigma_r$ ,

$$\Sigma^+ = \begin{pmatrix} \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{1}{\sigma_r} & \\ & & & 0 \end{pmatrix}_{n \times m}$$



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

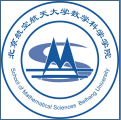
最小二乘法

### 伪逆的另一种表达

$A$  的奇异值分解又可以写作  $A = U\Sigma V^T = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ . 于是我们又有  $A^+$  的另一种表达

$$A^+ = \sum_{i=1}^r \sigma_i^{-1} \mathbf{v}_i \mathbf{u}_i^T$$

其中这些  $\mathbf{v}_i, \mathbf{u}_i$  分别是正交阵  $V$  与  $U$  的列。



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

最小二乘法

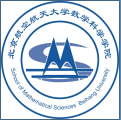
### 伪逆与线性方程组

对于线性方程组  $A\mathbf{x} = \mathbf{b}$  ( $\mathbf{b}$  可以是零向量也可以不是), 它若有解, 则解可以由系数矩阵  $A$  的 Moore-Penrose 伪逆统一给出:

$$A^+\mathbf{b} + (I - A^+A)\mathbf{w}$$

其中  $\mathbf{w}$  是任意一个  $n$  维列向量。当然,  $A\mathbf{x} = \mathbf{b}$  有解的充要条件为

$$AA^+\mathbf{b} = \mathbf{b}$$



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

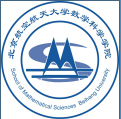
伪逆的定义与性质  
伪逆与线性方程组

### 最小二乘法

## 1 广义逆

## 2 Moore-Penrose 伪逆

## 3 最小二乘法



## 最小二乘法：问题分析

我们来考虑非齐次线性方程组  $A\mathbf{x} = \mathbf{b}$  无解的情况， $A$  是一个  $m \times n$  矩阵。此时，我们有

$$A\mathbf{x} = \mathbf{b} \text{ 无解} \iff \mathbf{b} \notin C(A)$$

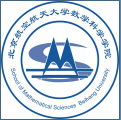
转而：在  $C(A)$  中寻找  $A\hat{\mathbf{x}}$ , 使得它与  $\mathbf{b}$  最接近，即求

$$\operatorname{argmin}_{\hat{\mathbf{x}}} \|A\hat{\mathbf{x}} - \mathbf{b}\|,$$

并称其为一个最小二乘解。

可以将以上结果与之前的广义逆相关的结果对比来看。





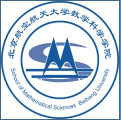
## 最小二乘法：法方程组

**事实：**  $\mathbb{R}^m$  中的一个点，与一个线性子空间中距离最短的点是这个点在这个子空间上的投影。于是

$$\begin{aligned}\|A\hat{\mathbf{x}} - \mathbf{b}\| \text{取得最小值} &\iff A\hat{\mathbf{x}} - \mathbf{b} \perp C(A) \\ &\iff A^T(A\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0} \\ &\iff A^T A\hat{\mathbf{x}} = A^T \mathbf{b}\end{aligned}$$

于是问题转化为了求解非齐次线性方程组

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



## 最小二乘法：法方程组

**事实：**  $\mathbb{R}^m$  中的一个点，与一个线性子空间中距离最短的点是这个点在这个子空间上的投影。于是

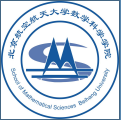
$$\begin{aligned}\|A\hat{\mathbf{x}} - \mathbf{b}\| \text{取得最小值} &\iff A\hat{\mathbf{x}} - \mathbf{b} \perp C(A) \\ &\iff A^T(A\hat{\mathbf{x}} - \mathbf{b}) = \mathbf{0} \\ &\iff A^T A\hat{\mathbf{x}} = A^T \mathbf{b}\end{aligned}$$

于是问题转化为了求解非齐次线性方程组

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

以上方程组（被称作法方程组，Normal Equations）的特点：

- 总有解；
- 即使解  $\hat{\mathbf{x}}$  不唯一，投影  $A\hat{\mathbf{x}}$  是唯一的。



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

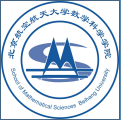
伪逆的定义与性质  
伪逆与线性方程组

### 最小二乘法

## 最小二乘解

- 若  $r(A) = n$ , 那么  $r(A^T) = r(A) = n$ . 此时  $A^T A$  可逆,  $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$  有唯一解。于是我们得到唯一的最小二乘解

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}.$$



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

伪逆的定义与性质  
伪逆与线性方程组

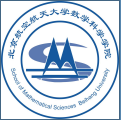
### 最小二乘法

## 最小二乘解

- 若  $r(A) = n$ , 那么  $r(A^T) = r(A) = n$ . 此时  $A^T A$  可逆,  
 $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$  有唯一解。于是我们得到唯一的最小二乘解

$$\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}.$$

- 若  $r(A) < n$ , 那么  $r(A^T) = r(A) < n$ . 此时  $A^T A$  不可逆,  
 $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$  有无穷多组解, 亦即最小二乘解有无穷多组。



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

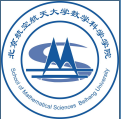
Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

最小二乘法

### 最小二乘解

**重要:**  $\mathbf{x}^+ := A^+ \mathbf{b}$  是长度最小的最小二乘解。其中  $A^+$  为矩阵  $A$  的 Moore-Penrose 伪逆。



## 最小二乘解

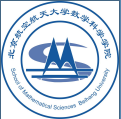
**重要:**  $\mathbf{x}^+ := A^+ \mathbf{b}$  是长度最小的最小二乘解。其中  $A^+$  为矩阵  $A$  的 Moore-Penrose 伪逆。

证明: 以下设  $A$  的奇异值分解为  $A = U\Sigma V^T$ .

(1)  $\mathbf{x}^+$  是最小二乘解: 将其代入法方程组  $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$  左边有

$$\begin{aligned} A^T A \cdot \mathbf{x}^+ &= A^T A A^+ \cdot \mathbf{b} \\ &= (V \Sigma^T U^T)(U \Sigma V^T)(V \Sigma^+ U^T) \mathbf{b} = V \Sigma^T \Sigma \Sigma^+ U^T \cdot \mathbf{b} \\ &= V \begin{pmatrix} \Sigma_0^T & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Sigma_0^{-1} & 0 \\ 0 & 0 \end{pmatrix} U^T \cdot \mathbf{b} \\ &= V \Sigma^T U^T \cdot \mathbf{b} = A^T \cdot \mathbf{b} = \text{法方程组右边} \end{aligned}$$

其中  $\Sigma_0 = \text{diag}(\sigma_1, \dots, \sigma_r)$ . 由上式即知,  $\mathbf{x}^+$  是最小二乘解。



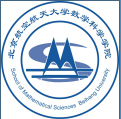
## 最小二乘解

(2)  $\mathbf{x}^+$  是最小二乘解中长度最小者：设  $\hat{\mathbf{x}}_0$  也是一个最小二乘解，即满足  $A^T A \hat{\mathbf{x}}_0 = A^T \mathbf{b}$ . 于是

$$A^T A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) = A^T \mathbf{b} - A^T \mathbf{b} = \mathbf{0}$$

$$\Rightarrow \|A(\hat{\mathbf{x}}_0 - \mathbf{x}^+)\| = (\hat{\mathbf{x}}_0 - \mathbf{x}^+)^T A^T A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) = 0$$

$$\Rightarrow A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) = \mathbf{0}$$



## 最小二乘解

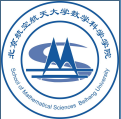
(2)  $\mathbf{x}^+$  是最小二乘解中长度最小者: 设  $\hat{\mathbf{x}}_0$  也是一个最小二乘解, 即满足  $A^T A \hat{\mathbf{x}}_0 = A^T \mathbf{b}$ . 于是

$$\begin{aligned} A^T A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) &= A^T \mathbf{b} - A^T \mathbf{b} = \mathbf{0} \\ \implies \|A(\hat{\mathbf{x}}_0 - \mathbf{x}^+)\| &= (\hat{\mathbf{x}}_0 - \mathbf{x}^+)^T A^T A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) = 0 \\ \implies A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) &= \mathbf{0} \end{aligned}$$

$A$  用其奇异值分解又可写作  $A = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ . 于是从上式我们有

$$\begin{aligned} A(\hat{\mathbf{x}}_0 - \mathbf{x}^+) = \mathbf{0} &\implies \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) = \mathbf{0} \\ &\implies \sum_{i=1}^r (\sigma_i \cdot (\mathbf{v}_i^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+))) \mathbf{u}_i = \mathbf{0} \end{aligned}$$





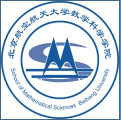
## 最小二乘解

但是, 我们知道  $\mathbf{u}_1, \dots, \mathbf{u}_r$  是线性无关的 (甚至是相互正交的), 而且  $\sigma_1, \dots, \sigma_r$  都是大于 0 的实数。于是

$$\mathbf{v}_1^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) = \dots = \mathbf{v}_r^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) = 0.$$

那么

$$\begin{aligned} \langle \mathbf{x}^+, \hat{\mathbf{x}}_0 - \mathbf{x}^+ \rangle &= (\mathbf{x}^+)^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) = (A^+ \mathbf{b})^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) \\ &= \mathbf{b}^T (A^+)^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) \\ &= \mathbf{b}^T \left( \sum_{i=1}^r \sigma_i^{-1} \mathbf{v}_i \mathbf{u}_i^T \right)^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+) \\ &= \mathbf{b}^T \left( \sum_{i=1}^r \sigma_i^{-1} \mathbf{u}_i \underbrace{\mathbf{v}_i^T (\hat{\mathbf{x}}_0 - \mathbf{x}^+)}_{=0} \right) = 0 \end{aligned}$$



## Pseudoinverse and Least Squares

广义逆

广义逆的定义与性质  
广义逆与线性方程组  
的解

Moore-  
Penrose 伪  
逆

伪逆的定义与性质  
伪逆与线性方程组

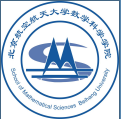
最小二乘法

### 最小二乘解

因此, 我们有

$$\|\hat{\mathbf{x}}_0\|^2 = \|\mathbf{x}^+ + (\hat{\mathbf{x}}_0 - \mathbf{x}^+)\|^2 = \|\mathbf{x}^+\|^2 + \|(\hat{\mathbf{x}}_0 - \mathbf{x}^+)\|^2 \geq \|\mathbf{x}^+\|^2$$

也就是说  $\mathbf{x}^+ = A^+\mathbf{b}$  的确是长度最小的最小二乘解。



## Pseudoinverse and Least Squares

### 广义逆

广义逆的定义与性质

广义逆与线性方程组  
的解

### Moore- Penrose 伪 逆

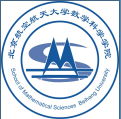
伪逆的定义与性质

伪逆与线性方程组

### 最小二乘法

## 最小二乘解

**注意!** 最小二乘解  $\mathbf{x}^+ = A^+\mathbf{b}$  包含了  $A^T A$  可逆时的唯一最小二乘解  $(A^T A)^{-1} A^T \mathbf{b}$ .



## 最小二乘解

**注意!** 最小二乘解  $\mathbf{x}^+ = A^+ \mathbf{b}$  包含了  $A^T A$  可逆时的唯一最小二乘解  $(A^T A)^{-1} A^T \mathbf{b}$ .

根据矩阵乘积的秩关系

$$A^T A \text{ 可逆} \implies n = r(A^T A) = r(A) = r(\Sigma),$$

于是,  $\Sigma$  就必须形如  $\Sigma = \begin{pmatrix} \Sigma_0 \\ 0 \end{pmatrix}$ , 其中  $\Sigma_0 = \text{diag}(\sigma_1, \dots, \sigma_n)$ . 于是

$$A^+ = V \begin{pmatrix} \Sigma_0^{-1} & 0 \end{pmatrix} U^T$$

另一方面,

$$\begin{aligned} (A^T A)^{-1} A^T &= (V \Sigma^T U^T U \Sigma V^T)^{-1} V \Sigma^T U^T \\ &= V \Sigma_0^{-2} V^T V \Sigma^T U^T = V \begin{pmatrix} \Sigma_0^{-1} & 0 \end{pmatrix} U^T \end{aligned}$$