矩 阵 论

主讲教师：徐乐

2014年12月24日星期三

上讲回顾

- 第19讲 最小二乘法
  - 最小二乘法
  - 极小范数最小二乘解

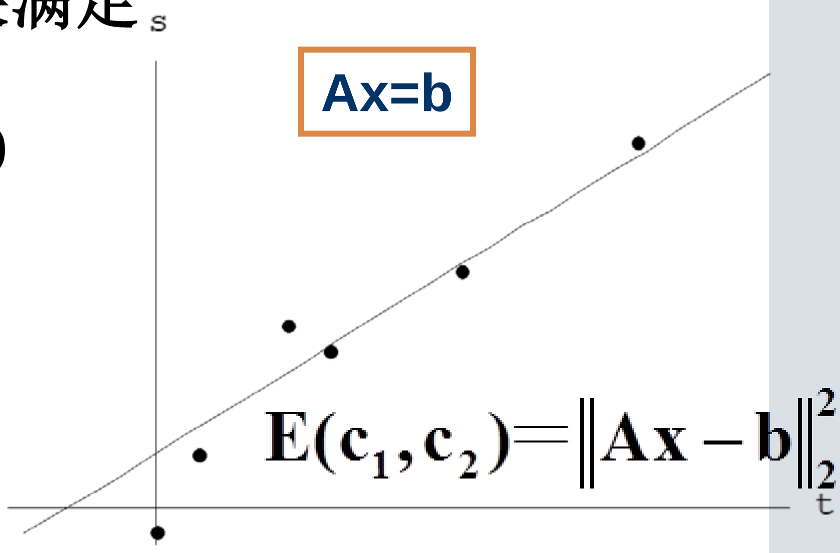
❖引例：实验数据处理

- 由实验数据拟合给定规律，从而测出待测量的有关参数
  - 理论值： $s=c_1 t+c_2$
  - 由于误差存在，实验结果满足

$$s_i \neq c_1 t_i + c_2 \qquad (i=1,2,\cdots,n)$$

  - 令

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



$\boxed{Ax=b}$

$$E(c_1,c_2)=\|Ax-b\|_2^2$$

最小二乘法（解）

- 对于矛盾方程Ax=b，最小二乘法是求其“解”的一种方法
  - 即 $\|Ax-b\|_2 = \min$
- 引理
  - 设$A\in C^{m\times n}$，A{1,3}由如下方程的通解构成：

$$AX = AA^{(1,3)} \rightarrow A\{1,3\} = \{A^{(1,3)} + (I - A^{(1,3)}A)Z \mid Z \in C^{n\times m}\}$$

    - 其中，A(1,3)为A{1,3}中的某个矩阵

最小二乘法（解）

定理

- 矩阵方程Ax=b的最小二乘解为 $x = A^{(1,3)}b$
  - 其中A(1,3)为A的任何一个{1,3}-逆矩阵
  - 反之，存在X，对于任何$b \in C^m$均有Xb成为Ax=b的最小二乘解，则 $X \in A\{1,3\}$

推论

- x是方程Ax＝b的最小二乘解的充要条件是
  - x为方程$A^H AX=A^H b$的解

极小范数最小二乘解

定理2

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

定理3

- 矩阵方程$AXB = D$极小范数最小二乘解唯一
- 且为 $X = A^{+}DB^{+}$

lexu@mail.xidian.edu.cn

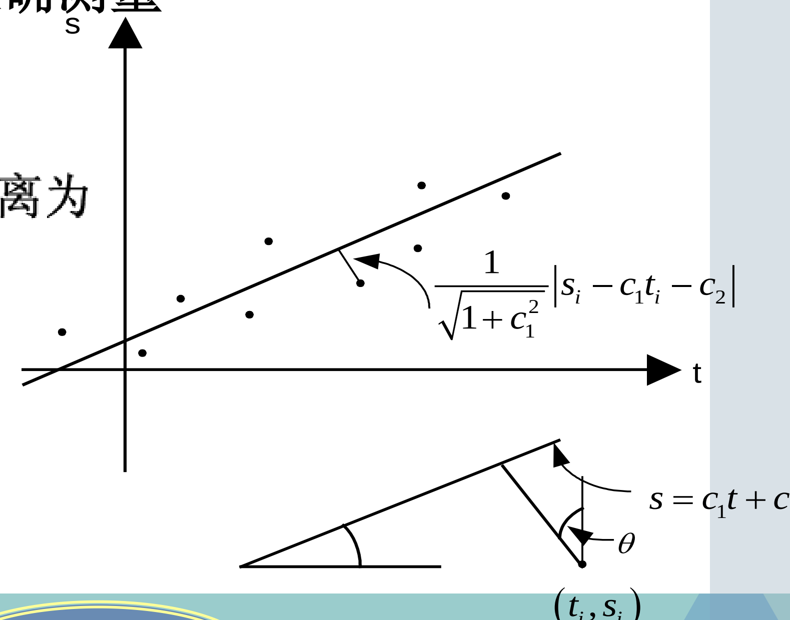
- 法向回归
- 全面最小二乘法

法向回归

再看实验数据处理

- 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 $E(c_1,c_2)=\sum_{i=1}^{n}\left|s_i-c_1t_i-c_2\right|^2=\min$
    - 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都无法准确测量
  - 法向回归
    - 点(t_i,s_i)到直线$s=c_1t+c_2$的距离为

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



❖法向回归的目标函数

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

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

$$\boxed{c_2=\frac{1}{n}\sum_{i=1}^{n}s_i-c_1t_i}$$

法向回归

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

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

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

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

$$=0$$

法向回归

- 令 $\bar{s}=\frac{1}{n}\sum_{i=1}^{n}s_i \qquad \bar{t}=\frac{1}{n}\sum_{i=1}^{n}t_i$

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

- 则

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

❖ 另一种推导方法 $E(c_1,c_2)=\frac{1}{1+c_1^2}\sum_{i=1}^{n}(s_i-c_1t_i-c_2)^2$

$$\frac{\partial E}{\partial c_2}=0$$

$$\boxed{c_2=\frac{1}{n}\sum_{i=1}^{n}(s_i-c_1t_i)=\bar{s}-c_1\bar{t}}$$

$$E(c_1,c_2)=\frac{\sum_{i=1}^{n}\left[(s_i-\bar{s})-c_1(t_i-\bar{t})\right]^2}{1+c_1^2}$$

$$E(c_1,c_2)=\frac{l_{ss}-2c_1l_{st}+l_{tt}c_1^2}{1+c_1^2}$$

$$\frac{\partial E}{\partial c_1}=0$$

$$\boxed{c_1=\frac{(l_{ss}-l_{tt})\pm\sqrt{(l_{ss}-l_{tt})^2+4l_{st}^2}}{2l_{st}}}$$

“—”对应的 E 的最大值

法向回归

❖类似地，可以给出最小二乘法的结果

- 即满足方程的解

$$\sum_{i=1}^{n}\left|s_i - c_1 t_i - c_2\right|^2 = \min$$

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

法向回归

❖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 $c_1 t + c_2 = s$
  - [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

- **Totally Least Square Method**
- 当方程AX=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ε，使得AX= b +ε
  - 因而ε应尽量小以使得不至于严重得破坏方程
  - 即就是 $\|\varepsilon\|_2 = \min$

全面最小二乘法

- 全面最小二乘法采取如下观点解决矛盾方程的问题 $(A+E)x=b+\varepsilon \Leftrightarrow ([A|b]+[E|\varepsilon])\begin{bmatrix} x \\ -1 \end{bmatrix}=0$
  - 不仅b存在误差，A也存在误差
  - 故存在E和ε，使 $(A+E)x=b+\varepsilon$
  - 因而E和ε也应该尽量小，以使得不至于严重偏离原方程，即 $\|[E|\varepsilon]\|_F=\min$
  - 令$C=[A|b], \Delta=[E|\varepsilon], v=\begin{bmatrix} x \\ -1 \end{bmatrix}$
  - 则全面最小二乘解即求如下方程的非零解v

$$(C+\Delta)v=0$$

– v的最后分量不能为零，而其中△应满足$\|\Delta\|_F=\min$

❖引理

- 设$X \in C_r^{m\times n}$
  - 且存在奇异值分解 $X = U\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & 0 \end{bmatrix}_{m\times n} V^H$
  - 其中$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- 又设$Y = U\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_s & \\ & & & 0 \end{bmatrix}_{m\times n} V^H \quad (s < r)$
- 则$\|X - Y\|_F = \min\limits_{\substack{z\in C^{m\times n} \\ \mathrm{rank}z=s}} \|X - Z\|_F$

正交相抵或酉相抵的矩阵与F范数相同

❖[证明]

- 首先考虑F-范数
- 设$P_{m\times n}=UQV^H$
- U、V分别为m阶、n阶酉矩阵。Q为m×n阶矩阵（上式不一定是奇异值分解）。则

$$\|P\|_F^2=\sum_{i,j}|p_{ij}|^2=\sum_{i=1}^{m}\left(\sum_{j=1}^{n}p_{ij}\overline{p_{ij}}\right)=\sum_{i=1}^{m}\left(PP^H\right)_{ii}=tr\left(PP^H\right)=tr\left(P^HP\right)$$

$$=tr\left(UQV^HVQ^HU^H\right)=tr\left(UQQ^HU^H\right)=tr\left(Q^HU^HUQ\right)$$

$$=tr(Q^HQ)=tr\left(QQ^H\right)=\|Q\|_F^2$$

全面最小二乘法

- 因此可得: $\|X - Y\|_F^2 = \sum_{i=s+1}^{r} \sigma_i^2$
- 令 $T = U^H Z V$ ⇒ $Z = UTV^H$
- 则 $\|X - Z\|_F^2 = \sum_{i=1}^{r} |t_{ii} - \sigma_i|^2 + \sum_{\substack{i=1 \\ i\neq j}}^{r} \sum_{j=1}^{n} |t_{ij}|^2 + \sum_{i=r+1}^{m} \sum_{j=1}^{n} |t_{ij}|^2$
  - 对任意Z矩阵而言，各t_{ij}之间完全独立
    - 则$\|X\text{-}Z\|_F$是可能等于零的
  - 但是rank(Z)=r ⇒ $\|X - Z\|_F \neq 0$
  - 详细论证可知$\|X\text{-}Z\|_F$最小的条件为

$$t_{ij} = 0(i \neq j), t_{ii} = 0(i > s), t_{ii} = \sigma_i (i = 1, 2, \cdots, s)$$

全面最小二乘法

- 下面仅考虑在实际应用中非常常见的一种情况
  - $A \in C_n^{m\times n}$　$[A|b] \in C_{n+1}^{m\times(n+1)}$
  - 即A是列满秩的，[A|b]也是列满秩的
  - 这样，系数矩阵与增广矩阵的秩不相等
  - 方程AX=B不相容

❖定理1

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

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

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

- **[证明]**
  - 使方程(C+△)v=0具有非零解，必须

$$\mathrm{rank}(C+\triangle)<n+1$$

  - 由引理知

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

  - 显然有

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

定理2

- 设σ_{n+1}为C的n-k+1重奇异值
- 且v_{k+1},v_{k+2}，…，v_{n+1}相应的为C^HC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+△)v=0具有非零的解且F范数最小的△为 $\Delta = -Cv_s v_s^H / v_s^H v_s$
- 而方程的解则为$v=v_s$
- 其中$v_s \in S_c \overset{\Delta}{=} \mathrm{span}\{v_{k+1}, v_{k+2}, \cdots v_{n+1}\}$

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 $\mathbf{e}_{n+1} = \begin{bmatrix} \underbrace{0 \quad \cdots \quad 0}_{n个} & 1 \end{bmatrix}^T$ 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 $x = \dfrac{1}{\alpha} y$

全面最小二乘法

- 说明
  - 最小二乘解一定存在，但全面最小二乘解不一定
  - 存在全面最小二乘解时，若为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

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

- [解]显然C=[A|b] $\Longrightarrow$ $C=\begin{bmatrix} t_1 & 1 & s_1 \\ t_2 & 1 & s_2 \\ \vdots & \vdots & \vdots \\ t_n & 1 & s_n \end{bmatrix}$

全面最小二乘法

$$C^TC=\begin{bmatrix}22.75 & 10.5 & 77.25\\ 10.5 & 7 & 42.2\\ 77.25 & 42.2 & 282.28\end{bmatrix}$$

$$\boxed{\lambda = 309.7754, 2.249389, 0.0051987257}$$

$$v_3=\begin{bmatrix}1.990944 & 3.044416 & -1\end{bmatrix}^T$$

$$\begin{matrix}\downarrow & \downarrow\\ c_1 & c_2\end{matrix}$$

全面最小二乘解

与法向回归结果并不相同，但亦十分接近，值得注意的是$\bar{s}\neq c_1\bar{t}+c_2$

作业

- **P343－344**
  - **1，2，5**