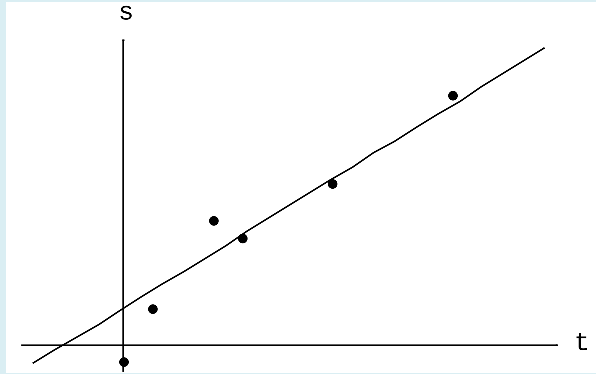
第十七讲 矛盾方程（组）的解---最小二乘法

一、从实验数据处理谈起



设有一组实验数据(t_1，s_1)，(t_2，s_2)，……，(t_n，s_n)，希望由实验数据拟合给定规律，从而测出待测量的有关参数。

假定规律为：$s=c_1t+c_2$，由于存在误差$s_i \neq c_1t_i+c_2((i=1,2,\cdots,n)$

令

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

，则：Ax=b 实际无解，或者说矩阵方程 Ax=b 成为矛盾方程（不自洽、非相容），虽说无解，但在物理上看，我们需要而且也理当有“解”。怎么办？

一般处理是，定义一种目标函数，例如：

$$E(c_1,c_2)=\sum_{i=1}^{n} w_i(s_i-c_1t_i-c_2)^2 \qquad w_i>0\text{为加权系数}$$

使误差 $E(c_1,c_2)$ 最小化。$w_i=1(i=1\sim n)$ 时 $E(c_1,c_2)=\|Ax-b\|_2^2$

二、 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。即求使$\|Ax-b\|_2=\min$的解。

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。

证：1° 方程既然相容，设 X 是其某个解，则

(i) $AXA=AA^{(1,3)}A=A\rightarrow X\in A\{1\}$

(iii) $(AX)^H=(AA^{(1,3)})^H=AA^{(1,3)}=AX\rightarrow X\in A\{3\}$

即方程的解必在 A{1,3}中。

2° 设 X 为 A 的一个{1,3}-逆矩阵，则

$$AX = AA^{(1,3)}AX \overset{\text{iii}}{=} \left(AA^{(1,3)}\right)^H (AX)^H$$

$$= \left(A^{(1,3)}\right)^H A^H X^H A^H$$

$$= \left(A^{(1,3)}\right)^H (AXA)^H$$

$$= \left(AA^{(1,3)}\right)^H = AA^{(1,3)}$$

即，A 的{1,3}-逆矩阵必满足方程 $AX=AA^{(1,3)}$

$$\therefore \quad A\{1,3\}=\left\{\text{方程}AX=AA^{(1,3)}\text{的所有解}\right\}$$

$$=\left\{A^{(1,3)}+(I-A^{(1,3)}A)Z \,\middle|\, Z\in C^{n\times m}\right\}$$

令 $X=A^{(1,3)}+(I-A^{(1,3)}A)Z$，则

$$(i)\ AXA=AA^{(1,3)}A+AZA-AA^{(1,3)}AZA=A \qquad X\in A\{1\}$$

$$(iii)\ AX=AA^{(1,3)}+(A-AA^{(1,3)}A)Z=AA^{(1,3)}=(AX)^H \qquad X\in A\{3\}$$

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 $X\in A\{1,3\}$。

证明：$Ax-b=(Ax-P_{R(A)}b)+(P_{R(A)}b-b)$

$(Ax-P_{R(A)}b)\in R(A),(P_{R(A)}b-b)=-(I-P_{R(A)})b=-P_{R^{\perp}(A)}b\in R^{\perp}(A)$

所以，$\|Ax-b\|_2^2=\|Ax-P_{R(A)}b\|_2^2+\|P_{R(A)}b-b\|_2^2\geq\|b-P_{R(A)}b\|_2^2$，

故$\|Ax-b\|_2^2$取得极小值的条件是 x 为方程 $Ax=P_{R(A)}b$ 的解。任取一个 $A^{(1,3)}\in A\{1,3\}$，我们知道 $AA^{(1,3)}=P_{R(A)}$。而对于 $x=A^{(1,3)}b$，有

$Ax=AA^{(1,3)}b=P_{R(A)}b$（但最小二乘解是否一定具有 $A^{(1,3)}b$ 的形式呢？）

方程 $Ax=AA^{(1,3)}b$ 的通解为

$$x=\left\{A^{(1,3)}AA^{(1,3)}b+y-A^{(1,3)}Ay \middle| y\in C^{n}\right\} \qquad y=A^{(1,3)}b+z$$

$$=\left\{A^{(1,3)}b+(I-A^{(1,3)}A)z \middle| z\in C^{n}\right\}$$

显然最小二乘解并不一定都具有 $A^{(1,3)}b$ 的形式。

反 之 ， 若 对 于 $\forall b\in C^{m}, x=Xb$均使$Ax=P_{R(A)}b=AA^{(1,3)}b$ ， 即

$$\forall b,\text{有}AXb=AA^{(1,3)}b\rightarrow AX=AA^{(1,3)}\rightarrow X\in A\{1,3\}$$

最小二乘解一般不唯一。

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

证：最小二乘解满足 Ax＝AA$^{(1,3)}$b，其极小范数解唯一，且为

$$x = A^{(1,4)}(AA^{(1,3)}b) = A^{+}b$$

反之，$\forall b \in C^{m}, Xb$均成为唯一的极小范数最小二乘解$A^{+}b$，

所以：X＝A^{+}。

定理 3：矩阵方程 AXB＝D 的极小范数最小二乘解唯一，且为

$$X = A^{+}DB^{+}$$

证明略（教材 P86）

作业：P343－344，1，2，5