算術平均數、最小二乘方及方差

高中組數學第三名

省立豐原高級中學　　製作學生：李　瑞　吉
指導老師：林　春　景

一、前言：

人類爲了要了解或描寫自然現象，常採用一種「逼近」的過程來處理問題。比如某位同學使用一尺來量取一線段長幾次，結果產生每次的讀數都不相同，此時產生逼近解的問題。又比如有四支50公分左右的尺，每次使用兩支來量取一公尺左右的線段長，亦產生類似的問題。然而這些「逼近」解的可靠性到底如何？我們則用方差來權衡之。

以下是自然界中常發生的幾個問題：

二、算術平均數的問題：

當觀測一量至若干次，則所有觀測值之算術平均數爲此量之逼近解。

證明：設 x 爲此量之逼近解，ℓ_i爲第 i 次之觀測值：

ε_i爲第 i 次之改正值，則

$$\varepsilon_i = x - \ell_i, \quad i = 1, 2, \cdots\cdots, n \tag{1}$$

依最小二乘方原理，此改正數之平方和應爲最小值，即

$$\sum_{i=1}^{n} \varepsilon_i^2 = \text{最小值}$$

(2)式依 x 而微分之而令之爲零得

$$\frac{d}{dx}\left(\sum_{i=1}^{n} \varepsilon_i^2\right) = \frac{d}{dx}\sum_{i=1}^{n}(x-\ell_i)^2$$

$$= 2\sum_{i=1}^{n}(x-\ell_i) = 0$$

即 $(x-\ell_1)+(x-\ell_2)+\cdots\cdots+(x-\ell_n)=0$

$\Rightarrow n\,x=\ell_1+\ell_2+\cdots\cdots+\ell_n$

$$\Rightarrow \quad x=\frac{\ell_1+\ell_2+\cdots\cdots+\ell_n}{n} \qquad (3)$$

例一：測量一角10次，其結果為：

觀　測　值　ℓ_i	改　正　數 $\varepsilon_i = x-\ell_i$
25°　33′　07.8″	−0.79″
07.6″	−0.59″
08.5″	−1.49″
07.8″	−0.79″
05.1″	1.91″
06.4″	0.61″
06.8″	0.21″
06.3″	0.71″
05.8″	1.21″
08.0″	−0.99″
平均值：25°　33′　07.01″	$\sum_{i=1}^{n}\varepsilon_i=$　0

$$\sum_{i=1}^{n}\varepsilon_i=n\,x-\sum_{i=1}^{n}\ell_i$$

$$=\sum_{i=1}^{n}\ell_i-\sum_{i=1}^{n}\ell_i$$

$$=0$$

此式可做為驗算之用

三、方程式的個數多於未知數的個數問題：

設 $x_1, x_2, \cdots x_r$ 爲 r 元一次方程組之逼近解，$L_1, L_2, \cdots L_n$ 爲觀測值，$\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n$ 分別爲其改正數，$a_i, b_i, \cdots k_i$ 爲已知之係數及常數（ $i = 1, 2, \cdots\cdots, n$）且爲實數，於是得一方程組爲

$$\left.\begin{aligned} L_1+\varepsilon_1 &= a_1x_1+b_1x_2+\cdots\cdots+k_1 \\ L_2+\varepsilon_2 &= a_2x_1+b_2x_2+\cdots\cdots+k_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ L_n+\varepsilon_n &\underbrace{= a_nx_1+b_nx_2+}_{r\text{ 個未知數}}\cdots\cdots+k_n \end{aligned}\right\} \text{共 } n \text{ 個方程式} \quad (4)$$

(4)式中令 $L_i - k_i = \ell_i$（ $i = 1, 2, \cdots\cdots, n$ ）得

$$\left.\begin{aligned} \varepsilon_1 &= a_1x_1+b_1x_2+\cdots\cdots-\ell_1 \\ \varepsilon_2 &= a_2x_1+b_2x_2+\cdots\cdots-\ell_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \varepsilon_n &\underbrace{= a_nx_1+b_nx_2+}_{r\text{ 個未知數}}\cdots\cdots-\ell_n \end{aligned}\right\} \text{共 } n \text{ 個方程式} \quad (5)$$

(5)式中若 $n > r$ 時，可由最小二乘方之理論知 $\sum\limits_{i=1}^{n} \varepsilon_i^{2}$ = 最小值，然後依 $x_1, x_2, \cdots\cdots x_r$ 微分之而令之爲零得

$$\left.\begin{aligned} &\frac{d}{dx_1}\Big(\sum_{i=1}^{n}\varepsilon_i^{2}\Big)=0 \Rightarrow 2\sum_{i=1}^{n}\varepsilon_i a_i \\ &\qquad\qquad\qquad = 0 \Rightarrow \sum_{i=1}^{n}\varepsilon_i b_i = 0 \\ &\text{同理可求得} \qquad \sum_{i=1}^{n}\varepsilon_i b_i = 0 \\ &\qquad\qquad\qquad \cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}\right\} \text{共 } r \text{ 個方程式} \quad (6)$$

將(5)式代入(6)式得

$$\left.\begin{array}{l} \displaystyle\sum_{i=1}^{n}\varepsilon_i a_i=\sum_{i=1}^{n}a_i^2x_1+\sum_{i=1}^{n}a_ib_ix_2+\cdots \\ \qquad\qquad \displaystyle-\sum_{i=1}^{n}a_i\ell_i=0 \\ \displaystyle\sum_{i=1}^{n}\varepsilon_i b_i=\sum_{i=1}^{n}a_ib_ix_1+\sum_{i=1}^{n}b_i^2x_2+\cdots\cdots \\ \qquad\qquad \displaystyle-\sum_{i=1}^{n}b_i\ell_i=0 \\ \underbrace{\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots}_{\text{共 r 個方程式}} \end{array}\right\}\text{共 r 個方程式}\quad(7)$$

此時方程式之個數與未知數之個數相等，方程式可解矣！

討論：

(4)式中任一方程式之係數 $a_i, b_i, \cdots\cdots k_i$ 皆不與其他一組方程式之係數成比例且不能同時爲零。設若成比例，可調整其觀測（或實驗）使其不成比例。此時可用克萊瑪法則（Cramer's rule）求得其解。

例二：某化學實驗如下：

實　　驗	因		果
	甲成分重 a	乙成分重 b	主 產 品 c
第　1　次	2	3	1
2	1	4	1
3	3	4	2
4	3	2	1

根據實驗，我們初步想獲得一次關係式

$$ax+by=c \tag{8}$$

依實驗列出方程式如下：

$$\begin{cases} 2x+3y=1 & L_1 \\ x+4y=1 & L_2 \\ 3x+4y=2 & L_3 \\ 3x+2y=1 & L_4 \end{cases} \qquad (8')$$

(8′)式轉化成幾何之關係式意即四直線相交於一點之問題，然而而四線交於一點的機會太小 ，但我們可求其迫近解x_0，y_0如下：

$$\begin{cases} 2x_0+3y_0=1+\varepsilon_1 \\ n_0+4x_0=1+\varepsilon_2 \\ 3x_0+4y_0=2+\varepsilon_3 \\ 3x_0+2y_0=1+\varepsilon_4 \end{cases} \Rightarrow \begin{cases} \varepsilon_1=2x_0+3y_0-1 \\ \varepsilon_2=x_0+4y_0-1 \\ \varepsilon_3=3x_0\omega 4y_0-2 \\ \varepsilon_4=3x_0+2y_L-1 \end{cases} \qquad (9)$$

解(9)式先計算方程式之係數如下：

$$\sum_{i=1}^{4} a_i^2=23 \qquad \sum_{i=1}^{4} b_i^2=28 \qquad \sum_{i=1}^{n} a_i\ell_i=-12$$

$$\sum_{i=1}^{4} b_i^2=45 \qquad \sum_{i=1}^{4} b_i\ell_i=-17$$

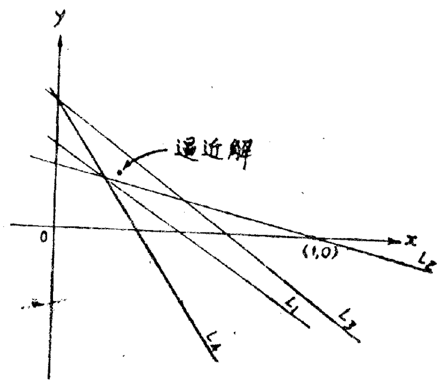
故得方程式

$$\begin{cases} 23x_0+28y_0-12=0 \\ 28x_0+45y_0-17=0 \end{cases} \Rightarrow x_0=\frac{64}{251},\ y_0=\frac{55}{251} \qquad (10)$$

(10)代入(8)得

$$\frac{64}{251}a+\frac{55}{251}b=c \qquad (11)$$

(11)式可以用來推測未做實驗的結果。



例三：物理實驗中，想檢定四支500mm左右之尺，每次使用兩支放在一長1030.20mm（檢定過）尺上檢定，其結果如下：

尺　　號	讀　　　數
I II	27.94mm
I III	27.11mm
I IV	27.91mm
II III	27.87mm
II IV	28.22mm
III IV	27.58mm

設 x，y，z，t，分別為 I，II，III，IV號尺之逼近解，並設 $x = x_0+500$，　$y = y_0+500$，　$z = z_0+500$，$t = t_0+500$，由上實驗得方程組如下：

$$\left.\begin{array}{l}
27.94+\varepsilon_1=-x_0-y_0+30.20 \\
27.11+\varepsilon_2=-x_0-z_0+30.20 \\
27.91+\varepsilon_3=-x_0-t_0+30.20 \\
27.87+\varepsilon_4=-y_0-z_0+30.20 \\
28.22+\varepsilon_5=-y_0-t_0+30.20 \\
27.58+\varepsilon_6=-z_0-t_0+30.20
\end{array}\right\} \Rightarrow$$

$$\left.\begin{array}{l}
\varepsilon_1=-x_0-y_0+2.26 \\
\varepsilon_2=-x_0-z_0+3.09 \\
\varepsilon_3=-x_0-t_0+2.29 \\
\varepsilon_4=-y_0-z_0+2.33 \\
\varepsilon_5=-y_0-t_0+1.98 \\
\varepsilon_6=-z_0-t_0+2.62
\end{array}\right\} \qquad (12)$$

欲解(12)式，先計算方程式之係數如下：

$$\sum_{i=1}^{4} a_i^2 = 3 \qquad \sum_{i=1}^{4} a_i b_i = 1 \qquad \sum_{i=1}^{4} a_i c_i = 1$$

$$\sum_{i=1}^{4} a_i d_i = 1 \qquad \sum_{i=1}^{4} a_i \ell_i = 7.64 \qquad \sum_{i=1}^{4} b_i^2 = 3$$

$$\sum_{i=1}^{4} b_i c_i = 1 \qquad \sum_{i=1}^{4} b_i d_i = 1 \qquad \sum_{i=1}^{4} b_i \ell_i = 6.57$$

$$\sum_{i=1}^{4} c_i^2 = 3 \qquad \sum_{i=1}^{4} c_i d_i = 1 \qquad \sum_{i=1}^{4} c_i \ell_i = 8.04$$

$$\sum_{i=1}^{4} d_i^2 = 3 \qquad \sum_{i=1}^{4} d_i \ell_i = 6.89 \qquad \sum_{i=1}^{4} \ell_i = 36.11$$

方程式如下：

$$\left.\begin{array}{l} 3x_0 + y_0 + z_0 + t_0 - 7.64 = 0 \\ x_0 + 3y_0 + z_0 + t_0 - 6.57 = 0 \\ x_0 + y_0 + 3z_0 + t_0 - 8.04 = 0 \\ x_0 + y_0 + z_0 + 3t_0 - 6.89 = 0 \end{array}\right\} \qquad (12')$$

解(12′)得　$x_0 = 1.39$，　$y_0 = 0.86$，　$z_0 = 1.59$，　$t_0 = 1.02$

故其逼近解爲　$x = 501.39\text{mm}$　　$y = 500.86\text{mm}$

$z = 501.59\text{mm}$　　$t = 501.02\text{mm}$

四、誤差絕對值之平均數與方差問題：

以上所求之逼近解是否可靠，可由誤差絕對值之平均數或誤差平方之平均數（即方差）做爲其精確度之權衡。但我們寧可取方差做爲精度之標準。其理由如下：

設 ε_1，ε_2，……ε_n 爲誤差，m爲誤差絕對值平均數。

σ^2為方差，則

$$m=\frac{\sum_{i=1}^{n}\left|\varepsilon_i\right|}{n} \qquad \sigma^2=\frac{\sum_{i=1}^{n}\varepsilon_i^2}{n}=\frac{n\sum_{i=1}^{n}\varepsilon_i^2}{n^2}$$

$$m^2=\frac{\varepsilon_1^2+\varepsilon_2^2+\cdots\varepsilon_n^2+2(|\varepsilon_1\varepsilon_2|+|\varepsilon_1\varepsilon_3+\cdots+|\varepsilon_{n-1}\varepsilon_n|}{n^2}$$

$$\sigma^2-m^2=\frac{(n-1)(\varepsilon_1^2+\varepsilon_2^2+\cdots+\varepsilon_n^2)}{n^2}$$

$$\frac{-2(|\varepsilon_1\varepsilon_2|+|\varepsilon_1\varepsilon_3|+\cdots+|\varepsilon_{n-1}\varepsilon_n|)}{n^2}$$

$$=\frac{(\varepsilon_1-\varepsilon_2)^2+(\varepsilon_1-\varepsilon_3)^2+\cdots}{n^2}$$

$$\frac{+(\varepsilon_2-\varepsilon_3)^2+\cdots+(\varepsilon_{n-1}-\varepsilon_n)^2}{n^2}>0 \tag{13}$$

由此可知$\sigma>m$，意卽誤差大的經平方後較顯著增加，而誤差絕對值之平均數變化較緩慢。因此取方差有較好之效果。

例三：如例一中　$m=0.93$，　$\sigma=1.04$

討論：

(13)式中若$\varepsilon_1=\varepsilon_2$，　$\varepsilon_1=\varepsilon_3$，……$\varepsilon_{n-1}=\varepsilon_n$　卽各誤差皆相等時$\sigma=m$。但在觀測一量，或做某一實驗時，誤差不可能皆相同，因此$\sigma>m$。如果有二組觀測組，一組方差大，另一組方差小，我們當然要取方差小的一組做爲成果較準確。

五、結論：

數學必須理論與實際問題互相配合，才能引人入勝，不乏味，以上所論都是實驗本中第三册曾約略提過，如「四直線交於一點的機會不大，但可用最小二乘方原理求其逈近解」，但也只是「只聽樓梯響，不見人下來」。因此爲文申論之。