最小平方擬合

本節將使用內積空間上的正交投影，求得在平面中對一組實驗數據點以直線或多項式曲線擬合的技巧。

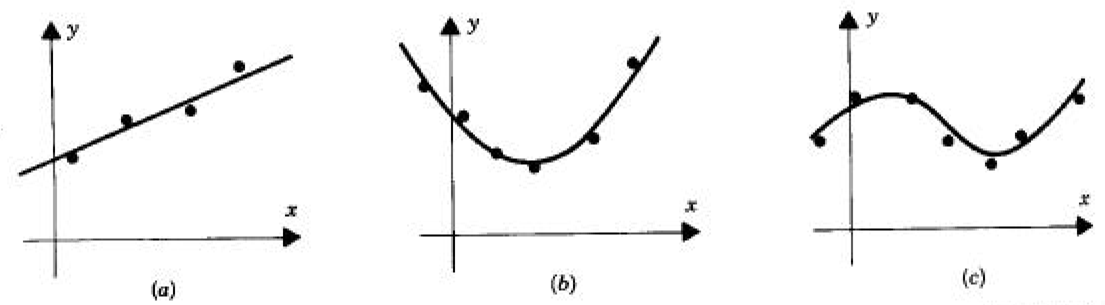
以曲線擬合實驗數據

實驗工作的共同問題是如何利用曲線"擬合"(fitting) 平面中對應於由不同實驗所得各點的 x,y 值，比如說

$$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$$

以獲得二變數 x 與 y 之間的數學關係式 $y=f(x)$。

無論是以理論上之考慮作基礎或僅依據點之分佈離型，都能決定用進行擬合之曲線的一般型式 $y=f(x)$。一些可能之曲線爲（參考圖 1）:



$y=a+bx$　　$y=a+bx+cx^2$　　$y=a+bx+cx^2+dx^3$

圖 1

(a) 直線: $y=a+bx$
(b) 二次多項式: $y=a+bx+cx^2$
(c) 三次多項式: $y=a+bx+cx^2+dx^3$

由於各點皆是實驗所得，"誤差" 自是難免，因而求得通過所有各點之曲線幾乎不可能。因此，應選取"最" 能滿足數據的曲線（藉由決定其係數）。首先考慮最簡單的情形: 對數據點進行直線擬合。

直線之最小平方擬合

假設想以直線

$$y=a+bx$$

擬合實驗所得的各點

$$(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$$

若數據點皆位於同一直線上，則此直線將通過所有 n 點且未知係數 a 及 b 將滿足

$$\begin{aligned} y_1 &= a+bx_1 \\ y_2 &= a+bx_2 \\ &\vdots \\ y_n &= a+bx_n \end{aligned}$$

此方程式組可用矩陣表示爲

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

或更簡潔地表示爲

$$M\mathbf{v}=\mathbf{y} \tag{1}$$

其中

$$\mathbf{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\quad M=\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},\quad \mathbf{v}=\begin{bmatrix} a \\ b \end{bmatrix} \tag{2}$$

假若各數據點並不位於同一直線上，則無法求得完全滿足 (1) 式的係數 a 與 b; 也就是方程式組是矛盾的，在此情況下應該尋求最小平方解。

$$\mathbf{v}=\mathbf{v}^*=\begin{bmatrix} a^* \\ b^* \end{bmatrix}$$

係數取自最小平方解的直線 $y=a^*+b^*x$，稱爲對數據的**最小平方直線擬合** (least squares straight line fit)。爲說明這個術語，回顧(1) 式的最小平方解爲最小化

$$\|\mathbf{y} - M\mathbf{v}\| \tag{3}$$

若以分量來表示 (3) 式的平方，可得

$$\|\mathbf{y} - M\mathbf{v}\|^2 = (y_1 - a - bx_1)^2 + (y_2 - a - bx_2)^2 + \ldots + (y_n - a - bx_n)^2 \tag{4}$$

現在若令

$$d_1 = |y_1 - a - bx_1|, d_2 = |y_2 - a - bx_2|, \ldots, d_n = |y_n - a - bx_n|$$

則 (4) 式能以

$$\|\mathbf{y} - M\mathbf{v}\|^2 = d_1^2 + d_2^2 + \cdots + d_n^2 \tag{5}$$

表示。如圖 2 所示，d_i 可以解釋為直線 $y = a + bx$ 及數據點(x_i, y_i) 間的垂直距離。

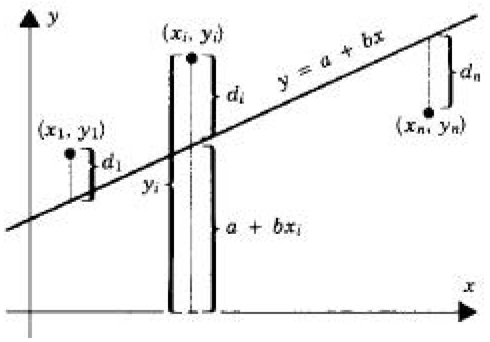


圖 2 d_i 為最小平方直線擬合的垂直誤差量度

此距離即為得自 $y = a + bx$ 對數據點的不完全擬合在點 $(x_i,\ y_i)$ 處的誤差量度。因為 (3) 式及 (5) 式均以同一向量 $\mathbf{v}^*$ 最小化，最小平方直線將這些誤差的平方和最小化，因此稱之為最小平方直線擬合。

法線方程式

回顧定理 6.4.2 可知(1) 式的最小平方解能以求解伴隨法線方程式組

$$M^T M\mathbf{v} = M^T\mathbf{y}$$

求得，此方程式組稱為**法線方程式組** (normal equations)。

在習題中將證明 M 的行向量為線性獨立的向量，若且唯若 n 個數據點不在與 xy- 平面垂直的同一直線上。在此情況下由定理 6.4.4 可知最小平方解是唯一的，而且可由

$$\mathbf{v}^* = (M^T M)^{-1} M^T\mathbf{y}$$

求得。總而言之，可得到下列的定理

定理 9.3.1：令 $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ 為一個、二個或更多數據點的集合，且不全在一垂直線上，且令

$$M = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{及} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

則對所有的數據點有唯一的最小平方直線擬合

$$y = a^* + b^* x$$

存在。此外，

$$\mathbf{v}^* = \begin{bmatrix} a^* \\ b^* \end{bmatrix}$$

可由公式

$$\mathbf{v}^* = (M^T M)^{-1} M^T\mathbf{y} \tag{6}$$

求得。這表示，$\mathbf{v} = \mathbf{v}^*$ 為法線方程式

$$M^T M\mathbf{v} = M^T\mathbf{y} \tag{7}$$

的唯一解。

例題 1

求對四點 (0,1)，(1,3)，(2,4) 及(3,4) 作最小平方直線擬合（參考圖 3）。

解：由題意知

$$M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$M^T M = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

$$(M^T M)^{-1} = \frac{1}{10}\begin{bmatrix} 7 & -3 \\ -3 & 2 \end{bmatrix}$$

$$\mathbf{v}^* = (M^TM)^{-1}M^T\mathbf{y} = \frac{1}{10}\begin{bmatrix} 7 & -3 \\ -3 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$$

故所求直線爲 $y = 1.5 + x$。 ▲

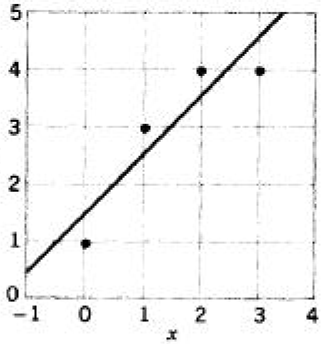
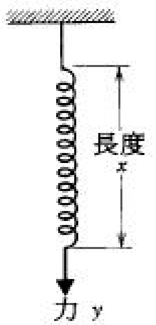


圖 3

例題 2

物理學中，虎克定律說明“均勻彈簧的長度 x 乃作用力 y 的線性函數”。若將虎克定律以數學式 $y = a + bx$ 表示，則係數 b 稱爲彈簧常數。假設有某自然長度爲 6.1 吋 (即 $y = 0$ 時 $x = 6.1$) 的彈簧，分別以 2 磅力，4 磅力及 6 磅力作用其上而將此彈簧拉長，若各力分別將此彈簧拉長至 7.6 吋，8.7 吋及 10.4 吋 (參考圖 5)，試求此彈簧的彈簧常數。



x_i	6.1	7.6	8.7	10.4
y_i	0	2	4	6

圖 4

解：由題意可知

$$M = \begin{bmatrix} 1 & 6.1 \\ 1 & 7.6 \\ 1 & 8.7 \\ 1 & 10.4 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 0 \\ 2 \\ 4 \\ 6 \end{bmatrix},$$

而

$$\mathbf{v}^* = \begin{bmatrix} a^* \\ b^* \end{bmatrix} = (M^TM)^{-1}M^T \quad \mathbf{y} \simeq \begin{bmatrix} -8.6 \\ 1.4 \end{bmatrix}$$

其中的數值已捨入成 1 位小數。因此，彈簧常數的估計值爲 $b^* \simeq 1.4$ lb/in。▲

多項式的最小平方擬合

對以直線擬合各數據點的技巧描述，使得以任意規定次數的多項式擬合各數據點顯得容易。現在企圖以固定爲 m 次的多項式

$$y = a_0 + a_1x + \cdots + a_mx^m \tag{8}$$

擬合 m 個數據點

$$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$$

將此 n 個 x 及 y 值分別代入 (8) 式，則可得 n 個方程式

$$\begin{aligned} y_1 &= a_0 + a_1x_1 + \cdots + a_mx_1^m \\ y_2 &= a_0 + a_1x_2 + \cdots + a_mx_2^m \\ \vdots \quad & \quad \vdots \quad \quad \vdots \quad \quad \quad \quad \vdots \\ y_n &= a_0 + a_1x_n + \cdots + a_mx_n^m \end{aligned}$$

或以矩陣式表示爲

$$M\mathbf{v} = \mathbf{y} \tag{9}$$

其中

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad M = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^m \\ 1 & x_2 & x_2^2 & \cdots & x_2^m \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^n & \cdots & x_n^m \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}$$

與先前一樣，法線方程式

$$M^TM\mathbf{v} = M^T\mathbf{y}$$

的解決定極小化

$$\|\mathbf{y} - M\mathbf{v}\|$$

的多項式係數。保證 M^TM 之可逆性的條件將在習題裡討論。若 M^TM 爲可逆的，則法線方程式有唯一解 $\mathbf{v} = \mathbf{v}^*$，且此解爲

$$\mathbf{v}^* = (M^TM)^{-1}M^T\mathbf{y}$$

例題 3

依據牛頓第二運動定律，接近地球表面之物體墜落距離係依方程式

$$s = s_0 + v_0t + \frac{1}{2}gt^2 \tag{10}$$

計算，此處

s ＝ 相對於某固定點落下之垂直位移

s_0 ＝ 初位移（ $t = 0$ 時）

v_0 ＝ 初速度（ $t = 0$ 時）

g ＝ 地球表面之重力加速度

假設某實驗室使用此方程式執行 g 值的評估。該試驗是這麼做的：某重錘在未知初位移及初速度的情況下釋放，並測量該物體由某固定參考點之下降距離數次，假設在時間 $t = 0.1, 0.2, 0.3, 0.4$ 及 0.5 秒時，分別測得物體由參考點下降之位移爲 $s = -0.18, 0.31, 1.03, 2.48$ 及 3.73 呎。試利用這些數據求 g 的近似值。

解：本題可使用二次曲線

$$s = a_0 + a_1t + a_2t^2 \tag{11}$$

擬合五個數據點：

$$(.1, -0.18),\ (.2, 0.31),\ (.3, 1.03),\ (.4, 2.48),\ (.5, 3.73)$$

來解。所需的計算爲

$$M = \begin{bmatrix} 1 & t_1 & t_1^2 \\ 1 & t_2 & t_2^2 \\ 1 & t_3 & t_3^2 \\ 1 & t_4 & t_4^2 \\ 1 & t_5 & t_5^2 \end{bmatrix} = \begin{bmatrix} 1 & .1 & 1.0 \\ 1 & .2 & .04 \\ 1 & .3 & .09 \\ 1 & .4 & .16 \\ 1 & .5 & .25 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{bmatrix} = \begin{bmatrix} -0.18 \\ 0.31 \\ 1.03 \\ 2.48 \\ 3.73 \end{bmatrix}$$

且

$$\mathbf{v}^* = \begin{bmatrix} a_0^* \\ a_1^* \\ a_2^* \end{bmatrix} = (M^TM)^{-1}M^T \quad \mathbf{y} = \begin{bmatrix} -0.40 \\ 0.35 \\ 16.1 \end{bmatrix}$$

由 (10) 及 (11) 式可得 $a_2 = \frac{1}{2}g$，因此

$$g = 2a_2^* = 2(16.1) = 32.2 \text{ 呎/秒}^2$$

若有需要，也能算出物體的初位移及初速度：

$$s_0 = a_0^* = -0.40 \text{ 呎}$$

$$v_0 = a_1^* = 0.35 \text{ 呎/秒}$$

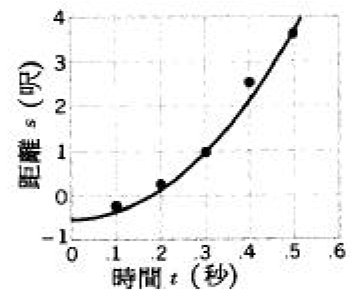


圖 5

將五個數據點及近似多項式繪於圖 5 中。▲

習題集 9.3

1.試對三點 (0,0), (1,2) 及 (2,7) 進行最小平方直線擬合。

2.試對四點 (0,1), (2,0), (3,1) 及 (3,2) 進行最小平方直線擬合。

3.試求擬合四點 (2,0), (3,−10), (5,−48) 及(6,−76) 的最佳二次多項式。

4.試求擬合五點 (−1,−14), (0,−5), (1,−4), (2,1) 及 (3,22) 的最佳三次多項式。

5.試證明 (2) 式中之矩陣 M 具有線性獨立的行若且唯若在 $x_1, x_2, \cdots, x_n$ 各數中至少有兩個數是相異的。

6.試證明 (9) 式中 $n \times (m+1)$ 階矩陣 M 之各行皆爲線性獨立的，若 $n > m$ 且在 $x_1, x_2, \cdots, x_n$ 各數中至少有 $m+1$ 個數是相異的。

7.令 M 爲 (9) 式中的矩陣，試使用習題 6 的結果，證明(9) 式中之矩陣 M^TM 爲可逆的，若 $n > m$ 且在 $x_1, x_2, \cdots, x_n$ 各數中至少有 $m+1$ 個數是相異的。

8.某快速擴展之商店的業主發現本年度前五個月的銷售額（以千計）分別爲$4.0, 4.4, 5.2, 6.4 及 8.0。他將這些數據標示於圖紙上，並臆測本年度其餘月份之銷售曲線能以二次多項式來近似。試對銷售曲線進行最小平方二次多項式擬合，並依據所得之擬合曲線推測十二月份的銷售額。