

第 6 章

複迴歸之一

Multiple Regression I

6.1 複迴歸模型 Multiple Regression Models

- 需要多個預測變數
- 在第1章迴歸分析時，單一預測變數可能無法對反應變數提供適當完整的描述，因為存在多個對於反應變數有重要或是關鍵性的影響的變數，同時單一預測變數對反應變數所進行的預測並不够精確以致於不實用。
- 在第2章的 Toluca 公司案例中，單一預測變數所進行的預測不够精確的情形，增加新預測變數雖然會複雜化模型，但是經常有助於提供反應變數足够精確的預測。
- 前述的案例其分析資料多是依據觀察記錄得到的，而沒有對預測變數進行控制，這是因為預測變數不易進行控制，如果進行可以控制預測變數的實驗，則複迴歸分析在此情形下將是非常有用的。複迴歸分析模型可以使用於依據觀察記錄得到的案例資料，或是在完全隨機設計下的實驗資料。

- 雙預測變數的第一階模型
- 當有兩個預測變數 X_1 與 X_2 時，迴歸模型為：
- (6.1)
$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$$
- 上面的迴歸模型稱為雙預測變數的第一階模型，第一階模型是指模型中所有的預測變數都是線性的， X_{i1} 與 X_{i2} 分別表示兩項預測變數第 i 次實驗下的水準， β_0 、 β_1 與 β_2 均為模型參數， ε_i 為隨機誤差項。
- 假設 $E\{\varepsilon_i\} = 0$ ，則模型(6.1)迴歸函數為：
- (6.2)
$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$
- 圖6.1所示為以下反應平面的部份圖示：
- (6.3)
$$E\{Y\} = 10 + 2X_1 + 5X_2$$
- 反應平面(6.3)上面的每一個點的高度正是每一對 X_1 與 X_2 的平面反應 $E\{Y\}$ 。

- 圖 6.1 同時也標示出了兩個預測變數水準(X_{i1}, X_{i2})的觀測值 Y_i ，在圖6.1中 Y_i 與反應平面的差距，代表給定(X_{i1}, X_{i2})時， Y_i 與 Y 之機率分配下期望值的差，也就是誤差項 ε_i ，亦即 $\varepsilon_i = Y_i - E\{Y_i\}$ 。複迴歸的迴歸函數常稱為迴歸曲面(regression surface)或是反應曲面(response surface)。

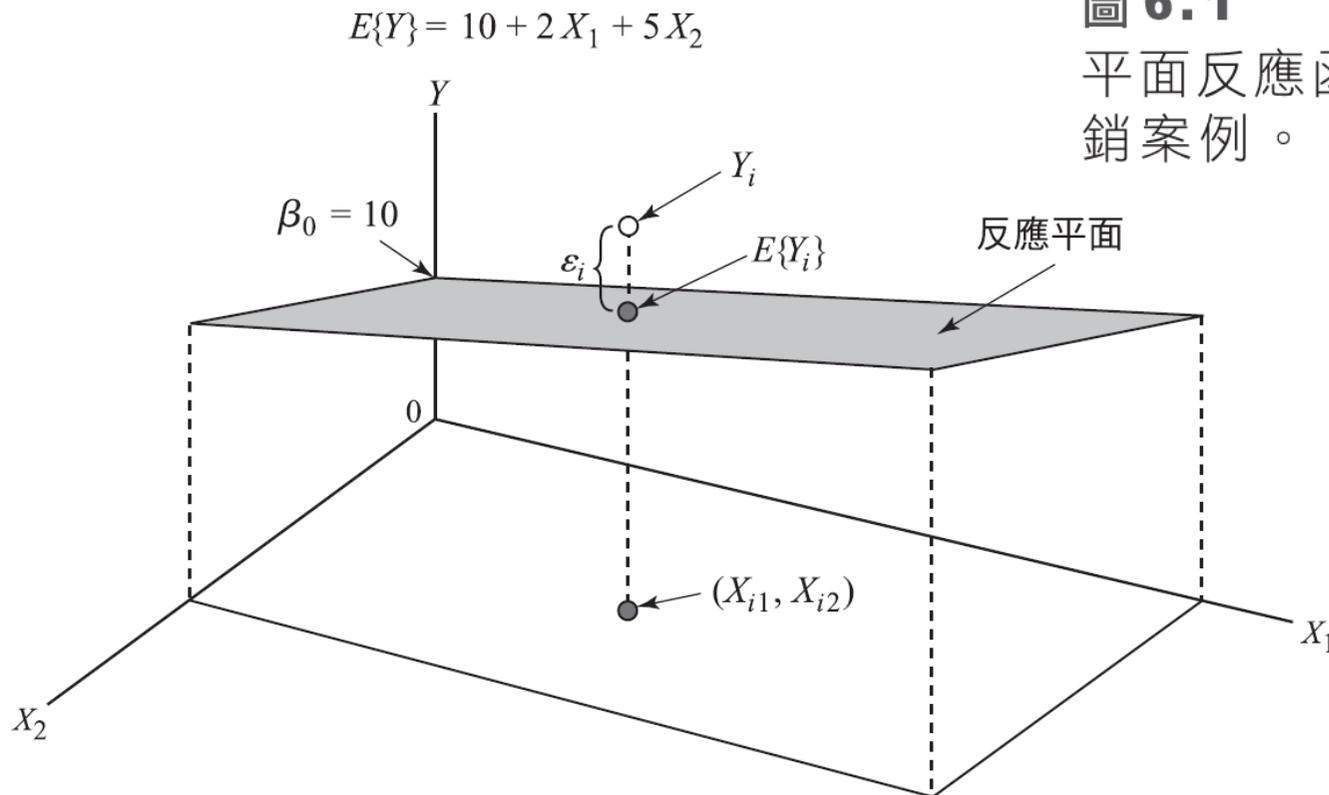


圖 6.1
平面反應函數－促銷案例。

迴歸係數的意義 $E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

- 參數 β_0 是迴歸平面的Y截距，當模型範圍涵蓋到 $X_1=0$ 與 $X_2=0$ ，則 $\beta_0 = 10$ 表示在 $X_1=0$ 與 $X_2=0$ 時的平均反應 $E\{Y\}$ ，否則單獨的參數在迴歸模型中並無特殊意義。
- 參數 β_1 表示當 X_2 固定時，平均反應 $E\{Y\}$ 隨著 X_1 每增加一個單位而改變的量，同樣地，參數 β_2 表示當 X_1 固定時，平均反應隨著 X_2 每增加一個單位而改變的量。
- 用前面的案例說明，當 X_2 固定在水準 $X_2 = 2$ 時，迴歸函數 (6.3) 成為：

$$E\{Y\} = 10 + 2X_1 + 5(2) = 20 + 2X_1 \quad X_2 = 2 \quad (6.4)$$

此迴歸函數是一條斜率為 $\beta_1 = 2$ 的直線，即使在 X_2 等於其他值時，此一結果仍然不變，只是可能會改變反應函數的截距量。

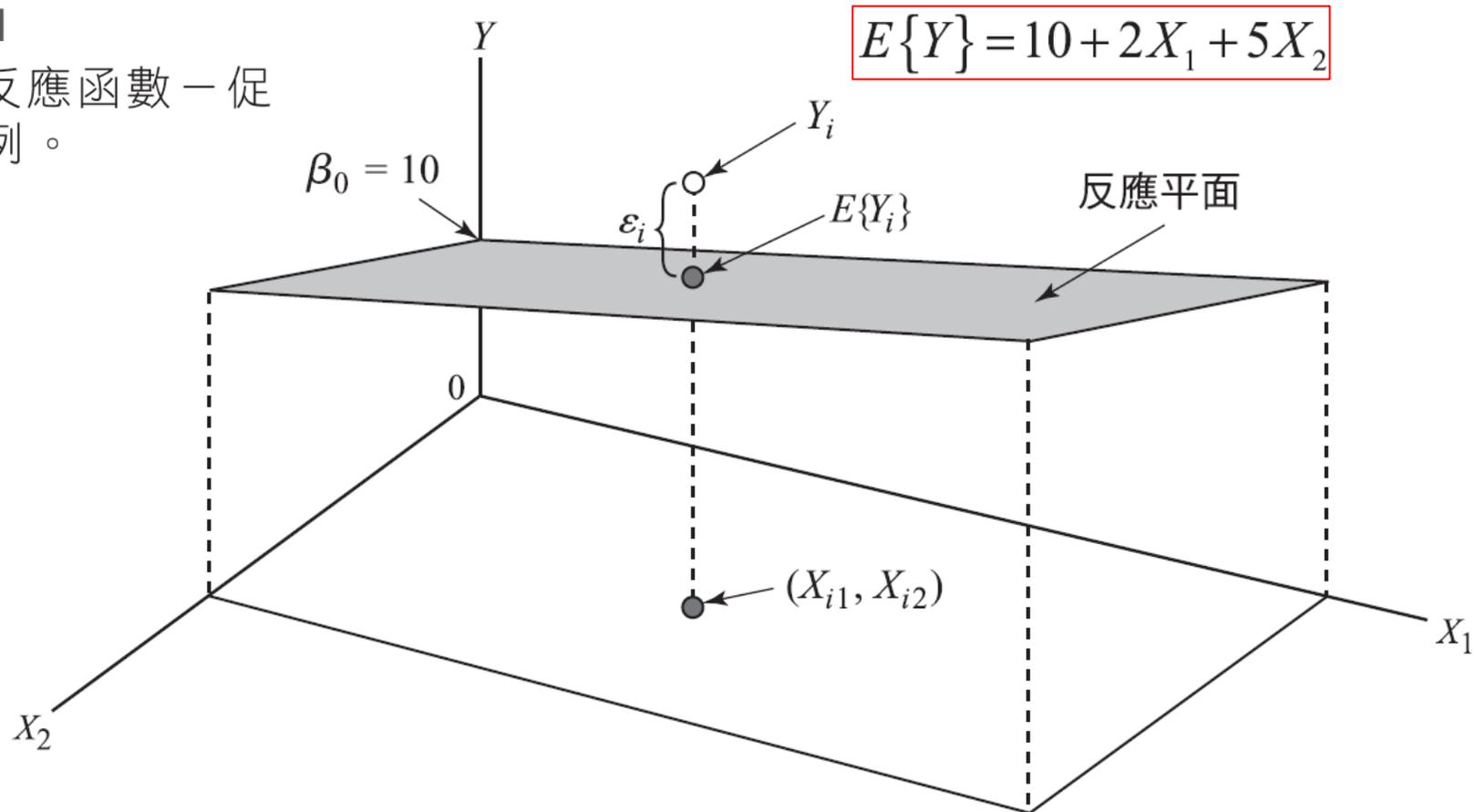
- 在迴歸函數(6.3)中 $\beta_2 = 5$ 表示當固定時，隨著每增加一個單位，平均反應 $E = \{Y\}$ 將增加 5 個單位。
- 若 X_1 對平均反應的效應，不受 X_2 的影響，亦即與 X_2 的水準高低無關，而 X_2 對平均反應的效應，也與 X_1 的水準高低無關，則此時兩個預測變數具有相加效應 (additive effect) 或稱為無交互作用 (not to interact)，所以第一階迴歸模型 (6.1) 的設定是預測變數對於反應變數的效應是相加的或無交互作用的。
- 參數 β_1 與 β_2 有時也稱為偏迴歸係數 (partial regression coefficient)，因為該參數主要在反應所對應的預測變數，在固定其他預測變數的情形下，於模型中的部分效應或是偏效應。

例題

圖 6.1 所示的反應平面(6.3)，是來自一個關於市場銷售量 (Y ，萬元) 對於銷售點的促銷支出 (X_1 ，千元) 與電視廣告支出 (X_2 ，千元) 關係的迴歸模型，因為 $\beta_1 = 2$ ，所以在電視廣告支出 X_2 不變的情形下，該地區銷售點的促銷支出 X_1 每增加一千元，期望的市場銷售量可增加兩萬元。類似的情形，因為 $\beta_2 = 5$ ，所以在地區銷售點的促銷支出 X_1 不變的情形下，電視廣告支出 X_2 每增加一千元，期望的市場銷售量可增加五萬元。

圖 6.1

平面反應函數－促銷案例。



超過兩個預測變數的第一階模型

現在考慮一個配適有 $p - 1$ 個預測變數 X_1, \dots, X_{p-1} 的迴歸模型：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (6.5)$$

上面的迴歸模型稱為「具 $p - 1$ 個預測變數的第一階模型」，此模型也可以寫成：

$$Y_i = \beta_0 + \sum_{k=1}^{p-1} \beta_k X_{ik} + \varepsilon_i \quad (6.5a)$$

或令 $X_{i0} = 1$ ，則可以寫成：

$$Y_i = \sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i \quad \text{其中 } X_{i0} \equiv 1 \quad (6.5b)$$

假設 $E\{\varepsilon_i\} = 0$ ，則迴歸模型(6.5)的反應函數為：

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{p-1} X_{p-1} \quad (6.6)$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (6.5)$$

- 此反應函數為一個 **超平面 (hyperplane)**，亦即它是一個超過二維的平面，這裡所有的參數與兩個預測變數模型參數類似：參數 β_k 表示當其他預測變數固定其值時，預測變數每增加一個單位，平均反應 $E\{Y\}$ 變動的程度，迴歸模型(6.5)中，任何一個預測變數對反應變數影響的程度，都不會受到其他預測變數的干擾而不同，使得預測變數對平均反應的效應是**相加的**，而**沒有交互作用存在**。

說明

- 當 $p-1=1$ 時，迴歸模型(6.5)可以簡化為：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i$$

此模型即為**簡單線性迴歸模型**

一般線性迴歸模型

在迴歸模型中變數 X_1, \dots, X_{p-1} ，所代表的意義不一定是完全不同或沒有關係的預測變數（例如 X_k 可以是 $X_1 \cdot X_2$ ），因此定義下面的一般線性迴歸模型：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (6.7)$$

其中， $\beta_0, \beta_1, \dots, \beta_{p-1}$ 為參數，

$X_{i1}, \dots, X_{i(p-1)}$ 為已知之常數，

ε_i 獨立服從 $N(0, \sigma^2)$ $i = 1, \dots, n$

如果令 $X_{i0} \equiv 1$ ，則迴歸模型(6.7)可以寫成：

$$Y_i = \beta_0 X_{i0} + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i \quad (6.7a)$$

或

$$Y_i = \sum_{k=0}^{p-1} \beta_k X_{ik} + \varepsilon_i \quad \text{其中 } X_{i0} \equiv 1 \quad (6.7b)$$

因為 $E\{\varepsilon_i\} = 0$ ，所以迴歸模型(6.7)的反應函數為

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_{p-1} X_{p-1} \quad (6.8)$$

所以具有常態誤差項的一般線性迴歸模型隱含著
觀測值 Y_i 為相互獨立、服從平均數 $E\{Y_i\}$ 、變異數 σ^2
之常態分配。

$P - 1$ 個預測變數

當 X_1, \dots, X_{p-1} 代表 $p-1$ 個完全無關的預測變數時，一般線性迴歸模型(6.7)就是前面所提的第一階模型，此時預測變數間不存在交互作用，圖 6.1 就是一個雙預測變數下的第一階模型。

質性的預測變數

一般線性迴歸模型(6.7)不只包含了計量的預測變數，也可以是質性的預測變數，例如性別（男、女），或殘障程度（正常、部分、完全），可以用指標變數 (indicator variable) 來標識質性的預測變數的每一種分類如 X_2 。考慮利用病患年齡(X_1)與性別(X_2)預測在醫院中的住院天數(Y)，定義 X_2 如下：

$$X_2 = \begin{cases} 1 & \text{當病患是女性時} \\ 0 & \text{當病患是男性時} \end{cases}$$

則第一階迴歸模型為： $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$ (6.9)

X_{i1} = 病患年齡

其中

$X_{i2} = \begin{cases} 1 & \text{當病患是女性時} \\ 0 & \text{當病患是男性時} \end{cases}$

迴歸模型(6.9)的反應函數為：

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \quad (6.10)$$

對男性病患而言， $X_2 = 0$ ，所以反應函數(6.10)成為：

$$E\{Y\} = \beta_0 + \beta_1 X_1 \quad \text{男性病患} \quad (6.10a)$$

對女性病患而言， $X_2 = 1$ ，所以反應函數(6.10)成為：

$$E\{Y\} = (\beta_0 + \beta_2) + \beta_1 X_1 \quad \text{女性病患} \quad (6.10b)$$

這兩個反應函數(6.10a)與(6.10b)代表著兩條不同截距的平行直線。

- 一般而言，當質性變數有 c 種分類情形時，模型中需要 $c-1$ 個指標變數來描述，在住院天數的案例中，殘障程度是一個質性變數（殘障程度（正常、部分、完全）），它可以用下面的兩個指標變數來描述：

$$X_3 = \begin{cases} 1 & \text{當病患是未殘障時} \\ 0 & \text{其他} \end{cases}$$

$$X_4 = \begin{cases} 1 & \text{當病患是部分殘障時} \\ 0 & \text{其他} \end{cases}$$

- 於是透過病患年齡、性別、殘障程度所構成的第一階迴歸模型為：
$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \varepsilon_i \quad (6.11)$$

其中，

$$X_{i1} = \text{病患年齡}$$

$$X_{i2} = \begin{cases} 1 & \text{當病患是女性時} \\ 0 & \text{當病患是男性時} \end{cases}$$

$$X_{i3} = \begin{cases} 1 & \text{當病患是未殘障時} \\ 0 & \text{其他} \end{cases}$$

$$X_{i4} = \begin{cases} 1 & \text{當病患是部份殘障時} \\ 0 & \text{其他} \end{cases}$$

多項式迴歸

多項式迴歸模型是一般線性迴歸模型的一種特例，包含了預測變數的二次式或更高項次，所以此時反應函數為曲線型式，下面單一預測變數多項式迴歸模型第一種情形：

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i \quad (6.12)$$

當我們令 $X_{i1} = X_i$ 、 $X_{i2} = X_i^2$ ，則(6.12)可以表示為：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$$

而上式正式是一般線性迴歸模型(6.7)的形式，(6.12)的反應函數為一個二次曲線迴歸模型，類似的情形，對於更高項次的多項式反應函數仍然是一般線性迴歸模型的特例。

經轉換後之變數

變數經過轉換後的模型可能含有複雜的曲線型式反應函數，不過它仍是一般線性迴歸模型的特例，考慮下面這個變數 Y 經轉換後的模型：

$$\log Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i \quad (6.13)$$

此模型的反應曲面雖然複雜，不過還是可以依照一般線性迴歸模型來處理。令 $Y'_i = \log Y_i$ ，則迴歸模型(6.13)成為：

$$Y'_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$$

很多模型可以轉換成一般線性迴歸模型，例如：

$$Y_i = \frac{1}{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i} \quad (6.14)$$

進行 $Y'_i = 1/Y_i$ 的變數轉換，它就是一個一般線性迴歸模型：

$$Y'_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$$

交互作用

- 當預測變數對反應變數的影響不是單純相加性的，表示某一或某些預測變數對反應變數的影響，會受到其他預測變數的干擾，一般線性迴歸模型(6.7)也包含了不可相加性或具有交互作用效應的，例如下面的雙預測變數模型：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + \varepsilon_i \quad (6.15)$$

此時由於交互作用項 $\beta_3 X_{i1} X_{i2}$ 存在模型中，造成反應函數複雜化，但是迴歸模型(6.15)仍然是一個一般線性迴歸模型，當令 $X_{i3} = X_{i1} X_{i2}$ ，則迴歸模型(6.15)可以寫成：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$$

還是一般線性迴歸模型 (6.15) 的形式，

組合情形

- 一個迴歸模型可能同時結合了上述幾種情形，但仍是一般線性迴歸模型，考慮下面的迴歸模型：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i2} + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \varepsilon_i \quad (6.16)$$

該模型中同時有預測變數的線性項、二次項與代表交互作用效應的交叉乘積項，定義：

$$Z_{i1} = X_{i1} \quad Z_{i2} = X_{i1}^2 \quad Z_{i3} = X_{i2} \quad Z_{i4} = X_{i2}^2 \quad Z_{i5} = X_{i1} X_{i2}$$

則迴歸模型(6.16)可以寫成：

$$Y_i = \beta_0 + \beta_1 Z_{i1} + \beta_2 Z_{i2} + \beta_3 Z_{i3} + \beta_4 Z_{i4} + \beta_5 Z_{i5} + \varepsilon_i$$

它顯然還是一般線性迴歸模型(6.7)的形式。

- 一般線性迴歸模型(6.7)可以包含許多複雜模型，其中有些具有高度的複雜性，圖6.2為兩個雙預測變數模型下的複雜反應曲面，它們均能用一般線性迴歸模型(6.7)來表示。

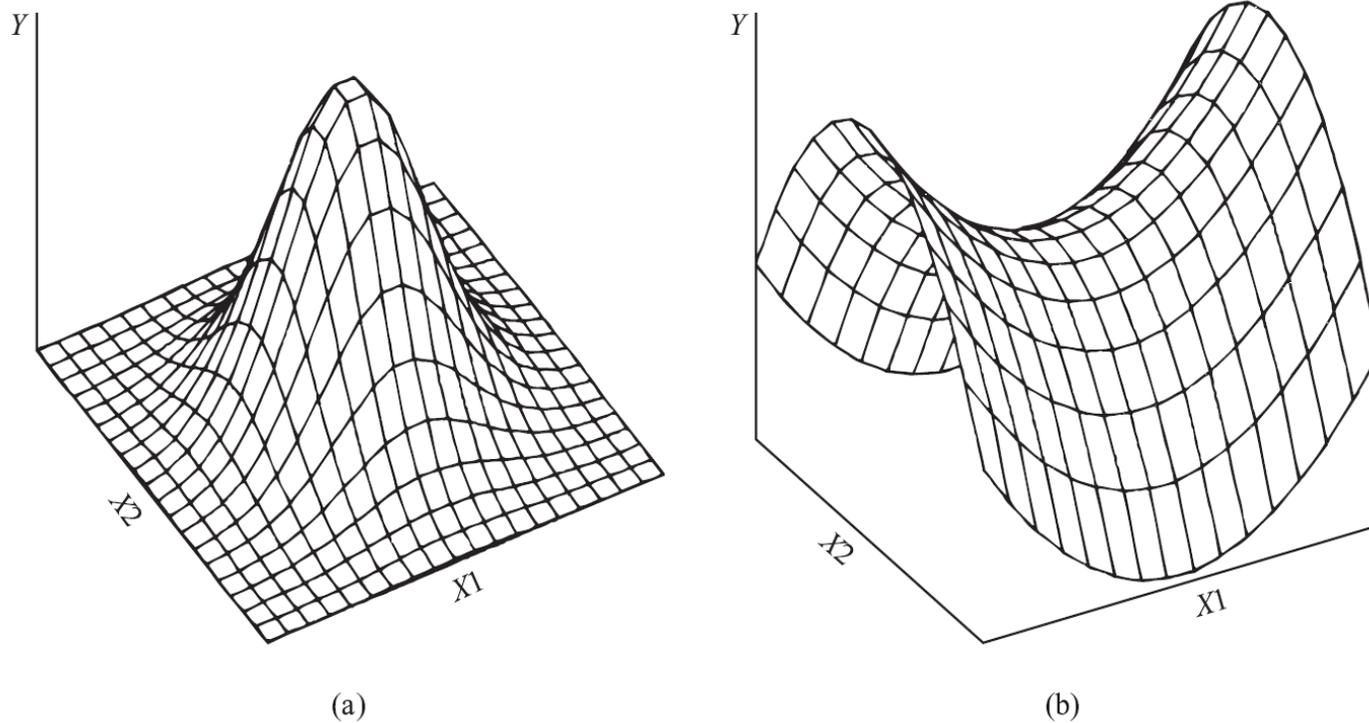


圖 6.2
反應函數的其他例子。

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i2} + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \varepsilon_i$$

一般線性迴歸模型中的線性意義

- 一般線性迴歸模型(6.7) 之為**線性模型(linear model)**是因為**模型(6.7) 對參數，是線性的**，而不是針對反應曲面的形狀，所謂對參數線性是指模型可以寫成：

$$Y_i = c_{i0}\beta_0 + c_{i1}\beta_1 + c_{i2}\beta_2 + \cdots + c_{i,p-1}\beta_{p-1} + \varepsilon_i \quad (6.17)$$

其中 $c_{i0}, c_{i1}, \dots, c_{i(p-1)}$ 是由預測變數的水準所決定出的值

- 舉一個**非線性迴歸模型**的例子如下：

$$Y = \beta_0 \exp(\beta_1 + X_{i1}) + \varepsilon_i$$

因為上面這個模型無法表示成 (6.17) 的形式，所以它是一個**非線性迴歸模型**，本書在第三部分將探討非線性迴歸模。

6.2 矩陣形式下的一般線性迴歸模型

General Linear Regression Model in Matrix Terms

- 利用矩陣的形式來表示一般線性迴歸模型(6.7)的結果，僅有的一些差別如自由度以及預測變數之個數，還有與矩陣階數有關的常數，因此結果將可以被更為簡潔地表示出來。

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

- 由於矩陣代數經常隱藏著龐大而複雜之運算，例如要計算一個 10×10 的矩陣A之反矩陣，只需用符號 A^{-1} 表示即可。強調矩陣代數的原因，是因為它可以簡單地指出一些觀念性的結果，而不需要實際的計算工作
- 對於複雜的計算工作只需交由程式計算與電腦執行即可，因此不需要花費心思在於 $(X'X)^{-1}$ 是一個 2×2 的矩陣，還是一個 10×10 的矩陣，重點應該在於 $(X'X)^{-1}$ 所代表的意義。

- 在將一般線性迴歸模型(6.7)用矩陣的形式來表示前，先定義下列矩陣：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

其中

(6.18a)

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

(6.18c)

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}$$

(6.18b)

$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1,p-1} \\ 1 & X_{21} & X_{22} & \cdots & X_{2,p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{n,p-1} \end{bmatrix}$$

(6.18d)

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

(6.18)

\mathbf{Y} 與 $\boldsymbol{\varepsilon}$ 的定義與簡單線性迴歸模型都相同， $\boldsymbol{\beta}$ 向量新增了一些迴歸參數， \mathbf{X} 矩陣包含第一行為1、另外有**p-1**行分別代表**p-1**個預測變數，在 \mathbf{X} 矩陣中的任一元素 X_{ik} ，其列下標*i*表示第*i*次試驗或是觀察個案，行下標*k*表示第*k*個預測變數。

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

- 用矩陣的形式來表示一般線性迴歸模型(6.7)如下：

$$\underset{n \times 1}{\mathbf{Y}} = \underset{n \times p}{\mathbf{X}} \underset{n \times p}{\boldsymbol{\beta}} + \underset{n \times 1}{\boldsymbol{\varepsilon}} \quad (6.19)$$

其中 \mathbf{Y} 為反應向量，

$\boldsymbol{\beta}$ 為參數向量，

\mathbf{X} 為常數矩陣，

$\boldsymbol{\varepsilon}$ 為獨立常態隨機變數之向量

$\mathbf{E}\{\boldsymbol{\varepsilon}\} = \mathbf{0}$ ，共變異矩陣為：

$$\underset{n \times n}{\sigma^2 \{\boldsymbol{\varepsilon}\}} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} = \sigma^2 \mathbf{I}$$

因此隨機向量 \mathbf{Y} 之期望值 $\mathbf{E}\{\mathbf{Y}\}$ 為：

$$\mathbf{E}\{\mathbf{Y}\}_{n \times 1} = \mathbf{X}\boldsymbol{\beta} \quad (6.20)$$

而 \mathbf{Y} 之共變異矩陣與 $\boldsymbol{\varepsilon}$ 相同為：

$$\boldsymbol{\sigma}^2\{\mathbf{Y}\}_{n \times n} = \sigma^2\mathbf{I} \quad (6.21)$$

6.3 迴歸係數的估計

Estimation of Regression Coefficients

- 將(1.8)中的最小平方準則推廣至一般線性迴歸模型(6.7)中，成為：

$$Q = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \dots - \beta_{p-1} X_{i,p-1})^2 \quad (6.22)$$

最小平方估計量就是在滿足能使 Q 最小化的 $\beta_0, \beta_1, \dots, \beta_{p-1}$ ，用向量 \mathbf{b} 表示所得到的最小平方估計量 b_0, b_1, \dots, b_{p-1}

$$\mathbf{b}_{p \times 1} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{bmatrix} \quad (6.23)$$

一般線性迴歸模型(6.19)的最小平方標準方程式為：

$$\begin{matrix} \mathbf{Y} & = & \mathbf{X} & \boldsymbol{\beta} & + & \boldsymbol{\varepsilon} \\ n \times 1 & & n \times p & n \times p & & n \times 1 \end{matrix} \quad \mathbf{X}'\mathbf{X}\mathbf{b} = \mathbf{X}'\mathbf{Y} \quad (6.24)$$

而最小平方估計量 \mathbf{b} 為：

$$\mathbf{b}_{p \times 1} = (\mathbf{X}'\mathbf{X})^{-1} (\mathbf{X}'\mathbf{Y})_{p \times 1} \quad (6.25)$$

- 根據常態誤差迴歸模型(6.19)所得到的最大概似估計量與(6.25)的最小平方估計量**b**相同，可以先將(1.26)的概似函數推廣至複迴歸模型，如下：

$$L(\boldsymbol{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_{p-1} X_{i,p-1})^2\right] \quad (6.26)$$

最大概似估計量就是在滿足能使(6.26)最大化的，而向量**b**正能滿足能使(6.26)最大化，且具有第1章中所介紹的不偏、最小變異、一致性與充分性等性質。

6.4 配適值及殘差

Fitted Values and Residuals

- 將配適值 \hat{Y}_i 組成向量 $\hat{\mathbf{Y}}$ ，殘差項 $e_i = Y_i - \hat{Y}_i$ 組成殘差向量 \mathbf{e} ：

$$(6.27a) \quad \hat{\mathbf{Y}}_{n \times 1} = \begin{bmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_n \end{bmatrix} \quad (6.27b) \quad \mathbf{e}_{n \times 1} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \quad (6.27)$$

可以將配適值向量 $\hat{\mathbf{Y}}$ 與殘差向量 \mathbf{e} 分別表示成：

$$\hat{\mathbf{Y}}_{n \times 1} = \mathbf{X}\mathbf{b} \quad (6.28)$$

$$\mathbf{e}_{n \times 1} = \mathbf{Y} - \hat{\mathbf{Y}} = \mathbf{Y} - \mathbf{X}\mathbf{b} \quad (6.29)$$

將配適值 $\hat{\mathbf{Y}}$ 用 \mathbf{H} 矩陣表示為：

$$\hat{\mathbf{Y}}_{n \times 1} = \mathbf{H}\mathbf{Y} \quad (6.30)$$

其中，

$$\mathbf{H}_{n \times n} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \quad (6.30a)$$

同樣地，殘差向量 \mathbf{e} 可以表示為：

$$\mathbf{e}_{n \times n} = (\mathbf{I} - \mathbf{H})\mathbf{Y} \quad (6.31)$$

而共變異矩陣則可以表示為：

$$\sigma^2 \{\mathbf{e}\}_{n \times n} = \sigma^2 (\mathbf{I} - \mathbf{H}) \quad (6.32)$$

其估計式為

$$\mathbf{s}^2 \{\mathbf{e}\}_{n \times n} = \text{MSE}(\mathbf{I} - \mathbf{H}) \quad (6.33)$$

6.5 變異數分析的結果

Analysis of Variance Results

- 平方和與均方

變異數分析的平方和透過(5.89)可以用矩陣表示為：

$$SSTO = \mathbf{Y}'\mathbf{Y} - \left(\frac{1}{n}\right)\mathbf{Y}'\mathbf{J}\mathbf{Y} = \mathbf{Y}'\left[\mathbf{I} - \left(\frac{1}{n}\right)\mathbf{J}\right]\mathbf{Y} \quad (6.34)$$

$$SSE = \mathbf{e}'\mathbf{e} = (\mathbf{Y} - \mathbf{X}\mathbf{b})'(\mathbf{Y} - \mathbf{X}\mathbf{b}) = \mathbf{Y}'\mathbf{Y} - \mathbf{b}'\mathbf{X}'\mathbf{Y} = \mathbf{Y}'(\mathbf{I} - \mathbf{H})\mathbf{Y} \quad (6.35)$$

$$SSR = \mathbf{b}'\mathbf{X}'\mathbf{Y} - \left(\frac{1}{n}\right)\mathbf{Y}'\mathbf{J}\mathbf{Y} = \mathbf{Y}'\left[\mathbf{H} - \left(\frac{1}{n}\right)\mathbf{J}\right]\mathbf{Y} \quad (6.36)$$

其中矩陣 \mathbf{J} 與(5.18)之定義完全相同，是一個元素全為1的 $n \times n$ 方陣，而矩陣 \mathbf{H} 則是定義於(6.30a)中的帽子矩陣。

- $SSTO$ 的自由度仍然為 $(n-1)$ ，因為模型(6.19)有 p 個參數需要被估計，所以 SSE 的自由度為 $(n-p)$ ，而 SSR 的自由度則為 $(p-1)$ ，分別代表著 $(p-1)$ 個變數 X_1, \dots, X_{p-1} 。
$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\begin{matrix} n \times 1 & n \times p & n \times p & n \times 1 \end{matrix}$$

在表 6.1 中列出了上面的變異數分析結果以及均方 MSR 與 MSE :

$$MSR = \frac{SSR}{p-1} \quad (6.37)$$

$$MSE = \frac{SSE}{n-p} \quad (6.38)$$

如同簡單線性迴歸的結果， MSE 的期望值為 σ^2 ， MSR 的期望值為 σ^2 再加上一個非負的量，例如當 $p-1=2$ 時：

$$E\{MSR\} = \sigma^2 + \frac{1}{2} \left[\beta_1^2 \sum (X_{i1} - \bar{X}_1)^2 + \beta_2^2 \sum (X_{i2} - \bar{X}_2)^2 + 2\beta_1\beta_2 \sum (X_{i1} - \bar{X}_1)(X_{i2} - \bar{X}_2) \right]$$

除非 $\beta_1 = \beta_2 = 0$ ，則 MSR 的期望值為 σ^2 ，否則 MSR 的期望值將大於 σ^2 。

表 6.1

一般線性迴歸模型 (6.19) 的 ANOVA 表。

變異來源	SS	df	MS
迴歸	$SSR = \mathbf{b}'\mathbf{X}'\mathbf{Y} - \left(\frac{1}{n}\right)\mathbf{Y}'\mathbf{J}\mathbf{Y}$	$p - 1$	$MSR = \frac{SSR}{p - 1}$
誤差	$SSE = \mathbf{Y}'\mathbf{Y} - \mathbf{b}'\mathbf{X}'\mathbf{Y}$	$n - p$	$MSE = \frac{SSE}{n - p}$
總合	$SSTO = \mathbf{Y}'\mathbf{Y} - \left(\frac{1}{n}\right)\mathbf{Y}'\mathbf{J}\mathbf{Y}$	$n - 1$	

• 迴歸關係的F檢定

現在先考慮反應變數 Y 與預測變數 X_1, \dots, X_{p-1} 間是否存在迴歸關係之檢定，虛無假設與對立假設分別為：

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{p-1} = 0$$

$$H_a : \beta_k (k = 1, \dots, p-1) \text{ 非全為 } 0 \quad (6.39a)$$

採用檢定統計量：

$$F^* = \frac{MSR}{MSE} \quad (6.39b)$$

控制型一錯誤的機率不大於 α ，決策法則為：

$$\begin{aligned} &\text{若 } F^* \leq F(1-\alpha; p-1, n-p), \text{ 則結論為 } H_0 \\ &\text{若 } F^* > F(1-\alpha; p-1, n-p), \text{ 則結論為 } H_a \end{aligned} \quad (6.39c)$$

複判定係數 (coefficient of multiple determination)

- 用 R^2 表示，其定義為：
$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \quad (6.40)$$

它是用來表示 Y 的總變異中與預測變數 X_1, \dots, X_{p-1} 有關的部份，當 $p-1=1$ 時，則複判定係數 R^2 成為(2.72)的簡單判定係數 r^2 ，複判定係數 R^2 與簡單判定係數 r^2 有相同的範圍限制：

$$0 \leq R^2 \leq 1 \quad (6.41)$$

- 當所有 b_k 為 0 時($k=1, \dots, p-1$)， R^2 之值為零，當所有的觀測值 Y 均落於所配適的迴歸曲面上時，亦即 $Y_i = \hat{Y}_i$ ，此時 $R^2 = 1$ 。
- 在迴歸模型中增加預測變數「只」會提高 R^2 之值，因為更多的預測變數只可能繼續降低 SSE (如式6.40)，而給定一組樣本觀測值後的 $SSTO$ 將會固定。

- 由於 R^2 之值會因為採用大量的預測變數而增大，有些學者建議採用調整的複判定係數(*adjusted coefficient of multiple determination*)，用符號 R_a^2 表示，其定義為：

$$R_a^2 = 1 - \frac{\frac{SSE}{n-p}}{\frac{SSTO}{n-1}} = 1 - \left(\frac{n-1}{n-p} \right) \frac{SSE}{SSTO} \quad (6.42)$$

此一調整的複判定係數當模型中增加預測變數時，可能會因為 SSE 降低的程度比不上自由度 $(n-p)$ 損失所抵消之效果，而導致 R_a^2 反而降低。

說明：複判定係數很高不代表配適的模型有用，例如預測變數只「取範圍很小」的水準數，此時即使有很高的複判定係數 R^2 ，當所需預測之值總是落在觀測範圍之外，則配適的模型沒有用。同時，就算有很高的複判定係數 R^2 ，其 MSE 還是可能會太大而造成精確過低，而不符合實際上的需求使用。

• 複相關係數

複相關係數(*coefficient of multiple correlation*) 是複判定係數 R^2 的正平方根 R :

$$R = \sqrt{R^2} \quad (6.43)$$

當迴歸模型(6.19)中只有一個預測變數 X ，也就是當 $p - 1 = 1$ 時，複相關係數 R 等於(2.73)中相關係數 r 的絕對值。

6.6 迴歸參數的推論

Inference about Regression Parameter

- 最小平方估計量與最大概似估計量 \mathbf{b} 為不偏之估計量：

$$\mathbf{E}\{\mathbf{b}\} = \boldsymbol{\beta} \quad (6.44)$$

其共變異矩陣：

$$\boldsymbol{\sigma}^2\{\mathbf{b}\} = \begin{bmatrix} \sigma^2\{b_0\} & \sigma\{b_0, b_1\} & \cdots & \sigma\{b_0, b_{p-1}\} \\ \sigma\{b_1, b_0\} & \sigma^2\{b_1\} & \cdots & \sigma\{b_1, b_{p-1}\} \\ \vdots & \vdots & & \vdots \\ \sigma\{b_{p-1}, b_0\} & \sigma\{b_{p-1}, b_1\} & \cdots & \sigma^2\{b_{p-1}\} \end{bmatrix} \quad (6.45)$$

可以藉由下面的公式算出：

$$\boldsymbol{\sigma}^2\{\mathbf{b}\} = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \quad (6.46)$$

而估計的共變異矩陣 $\mathbf{s}^2\{\mathbf{b}\}$ ：

$$\mathbf{s}^2\{\mathbf{b}\} = \begin{bmatrix} s^2\{b_0\} & s\{b_0, b_1\} & \cdots & s\{b_0, b_{p-1}\} \\ s\{b_1, b_0\} & s^2\{b_1\} & \cdots & s\{b_1, b_{p-1}\} \\ \vdots & \vdots & & \vdots \\ s\{b_{p-1}, b_0\} & s\{b_{p-1}, b_1\} & \cdots & s^2\{b_{p-1}\} \end{bmatrix} \quad (6.47)$$

可以藉由下面的公式算出：

$$\mathbf{s}^2 \{\mathbf{b}\} = \text{MSE} (\mathbf{X}'\mathbf{X})^{-1} \quad (6.48)$$

$p \times p$

透過 $\mathbf{s}^2\{\mathbf{b}\}$ ，可以求出 $s^2\{b_0\}$ 、 $s^2\{b_1\}$ 以及其他所需的變異數或共變異數。

• β_k 的區間估計

對於常態誤差迴歸模型(6.19)，有下面的推論：

$$\frac{b_k - \beta_k}{s\{b_k\}} \sim t(n-p) \quad k = 0, 1, \dots, p-1 \quad (6.49)$$

因此關於信賴係數 $(1-\alpha)$ 下 β_k 的信賴區間為：

$$b_k \pm t(1-\alpha/2; n-p) s\{b_k\} \quad (6.50)$$

• β_k 的檢定

關於 β_k 的檢定：

$$\begin{aligned} H_0 : \beta_k &= 0 \\ H_a : \beta_k &\neq 0 \end{aligned} \quad (6.51a)$$

可以採用檢定統計量：

$$t^* = \frac{b_k}{s\{b_k\}}$$

(6.51b)

檢定的法則為：

若 $|t^*| \leq t(1-\alpha/2; n-p)$, 則結論為 H_0
否則結論為 H_a

(6.51c)

6.7 平均反應的估計與新觀測值的預測

Estimation of Mean Response and Prediction of New Observation

- $E\{Y\}$ 的區間估計

給定 X_1, \dots, X_{p-1} 的值為 $X_{h1}, \dots, X_{h(p-1)}$ ，其平均反應為 $E\{Y_h\}$ ，定義向量 \mathbf{X}_h 為：

$$\mathbf{X}_h = \begin{bmatrix} 1 \\ X_{h1} \\ \vdots \\ X_{h,p-1} \end{bmatrix} \quad (6.53)$$

則所要估計的平均反應 $E\{Y_h\}$ 為：

$$E\{Y_h\} = \mathbf{X}'_h \boldsymbol{\beta} \quad (6.54)$$

用 \hat{Y}_h 表示對應向量 \mathbf{X}_h 的平均反應之估計：

$$\hat{Y}_h = \mathbf{X}'_h \mathbf{b} \quad (6.55)$$

上面的估計量也是一個不偏的估計量：

$$E\{\hat{Y}_h\} = \mathbf{X}'_h \boldsymbol{\beta} = E\{Y_h\} \quad (6.56)$$

其變異數為：

$$\sigma^2 \{\hat{Y}_h\} = \sigma^2 \mathbf{X}'_h (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h \quad (6.57)$$

可以將上面的變異數轉換為估計係數的共變異矩陣之函數：

$$\sigma^2 \{\hat{Y}_h\} = \mathbf{X}'_h \boldsymbol{\sigma}^2 \{\mathbf{b}\} \mathbf{X}_h \quad (6.57a)$$

從(6.57a)中， $\sigma^2 \{\hat{Y}_h\}$ 是迴歸係數的估計量其變異數 $\sigma^2 \{\mathbf{b}_k\}$ 與共變異數 $\sigma \{\mathbf{b}_k, \mathbf{b}_k\}$ 之函數，估計的變異數 $s^2 \{\hat{Y}_h\}$ 如下：

$$s^2 \{\hat{Y}_h\} = MSE \left(\mathbf{X}'_h (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}_h \right) = \mathbf{X}'_h \mathbf{s}^2 \{\mathbf{b}\} \mathbf{X}_h \quad (6.58)$$

$E\{Y\}$ 的 $(1-\alpha)$ 信賴界線為：

$$\hat{Y}_h \pm t(1-\alpha/2; n-p) s \{\hat{Y}_h\} \quad (6.59)$$

- 迴歸曲面的信賴域(confidence region)

在 \mathbf{X}_h 的水準下 $(1-\alpha)$ 的信賴域之邊界點，可以根據下面的公式得到：

$$\hat{Y}_h \pm W_s \{ \hat{Y}_h \} \quad (6.60)$$

其中，

$$W^2 = pF(1-\alpha; p, n-p) \quad (6.60a)$$

- 多個平均反應的同步信賴區間

在全族信賴係數 $(1-\alpha)$ 下，要估計在不同的 \mathbf{X}_h 水準下，多個平均反應 $E\{Y_h\}$ 的同步信賴區間可以採用下面的兩種方法：

- 對於所關心的 \mathbf{X}_h ，利用(6.60)的Working-Hotelling信賴域界線：

$$\hat{Y}_h \pm W_s \{ \hat{Y}_h \} \quad (6.61)$$

- 採用Bonferroni同步信賴區間，如果要同時進行 g 個區間估計，則Bonferroni信賴界線為：

$$\hat{Y}_h \pm B_s \{ \hat{Y}_h \} \quad (6.62)$$

其中，

$$B = t(1-\alpha/2g; n-p) \quad (6.62a)$$

- 新觀測值的預測

對應於一個特定水準 \mathbf{X}_h ，一個新觀測值 $Y_{h(\text{new})}$ 的 $(1-\alpha)$ 預測界線為：

$$\hat{Y}_h \pm t(1-\alpha/2; n-p) s\{\text{pred}\} \quad (6.63)$$

其中

$$s^2\{\text{pred}\} = MSE + s^2\{\hat{Y}_h\} = MSE \left(1 + \mathbf{X}'_h (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_h\right) \quad (6.63a)$$

- 預測 \mathbf{X}_h 處下的 m 個觀測值之平均

如果在 \mathbf{X}_h 處下抽選 m 個新觀測值，則可預測其平均數 $\bar{Y}_{h(\text{new})}$ 的 $(1-\alpha)$ 預測界線：

$$\hat{Y}_h \pm t(1-\alpha/2; n-p) s\{\text{predmean}\} \quad (6.64)$$

其中，

$$s^2\{\text{predmean}\} = \frac{MSE}{m} + s^2\{\hat{Y}_h\} = MSE \left(\frac{1}{m} + \mathbf{X}'_h (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_h\right) \quad (6.64a)$$

- g 個新觀測值的預測

在全族信賴係數 $(1-\alpha)$ 下， g 個不同 \mathbf{X}_h 水準下的新觀測值之預測問題，根據Scheffé同步預測程序，其預測界線為：

$$\hat{Y}_h \pm Ss\{\text{pred}\} \quad (6.65)$$

其中，
$$S^2 = gF(1-\alpha; g, n-p) \quad (6.65a)$$

而 $s^2\{\text{pred}\}$ 如(6.63a)所示。

當然，我們也可以利用Bonferroni同步預測程序，在全族信賴係數 $(1-\alpha)$ 下，對 g 個不同 \mathbf{X}_h 水準下進行新觀測值之預測：

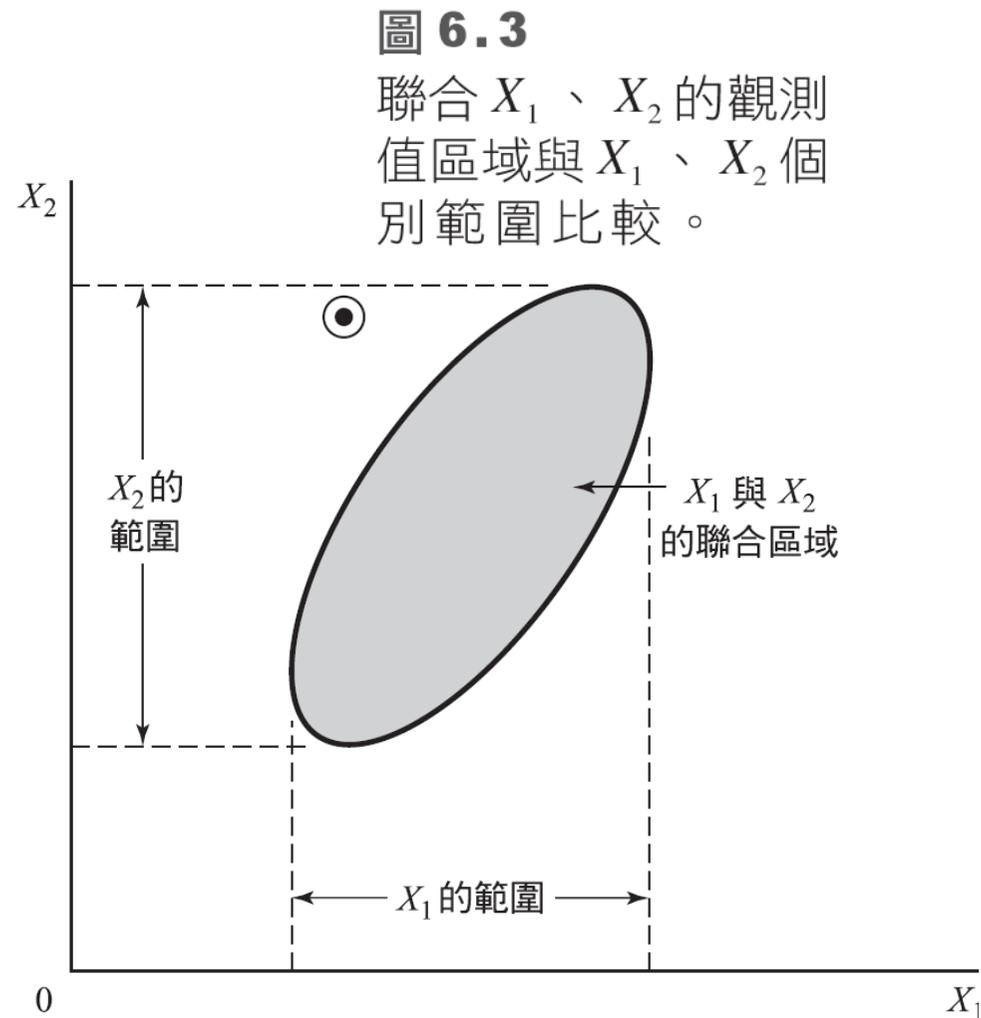
$$\hat{Y}_h \pm Bs\{\text{pred}\} \quad (6.66)$$

其中，
$$B = t(1-\alpha/2g; n-p) \quad (6.66a)$$

在實務界的應用中，經常先比較乘數 S 與 B ，再決定較窄區間的程序

隱藏的外插之注意事項

- 在複迴歸中估計一個平均反應或是預測一個新觀測值，需特別小心是否超出了模型的適用範圍，由於超出範圍時此模型可能並不適當
- 在圖6.3中的陰影區域是一個雙預測變數複迴歸模型，其觀測值的範圍，如果要預測的X水準是針對圖中加圓圈點(X_{h1}, X_{h2})，雖然分開考慮預測變數 X_1 與 X_2 的範圍，都符合模型觀測範圍，但是點(X_{h1}, X_{h2})卻是超出模型適用範圍



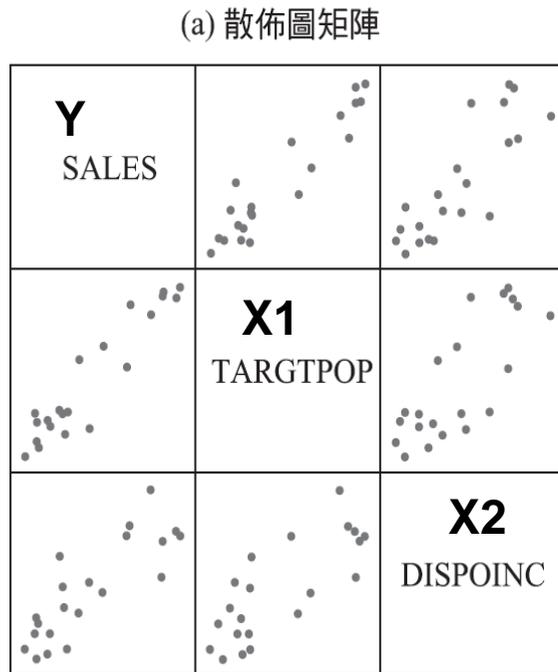
6.8 診斷與矯正測量

Diagnostics and Remedial Measures

散佈圖矩陣

- 各預測變數與反應變數的盒型圖、順序圖、莖葉圖以及點圖，都可以提供這些預測變數與反應變數初步的單變量訊息。
- 如果是對反應變數與各預測變數所做的散佈圖，則可以提供研究人員判斷關於各預測變數與反應變數間關係的性質與強度，並且有助於辨認資料點之間隙與離群資料，
- 而每一個預測變數與其他預測變數所做的散佈圖，可以作為探討兩兩預測變數間的關係，同時也可以辨認資料點之間隙與離群資料。

- 將上述這些散佈圖組成如圖6.4a 的**散佈圖矩陣(scatter plot matrix)**，將更方便進行資料分析，在圖6.4中，散佈圖中的Y變數如該**列**的位置所示，而自變數則該**行**的位置所示
- 散佈圖矩陣主要用於變數間關係的探討。



(b) 相關矩陣

	SALES	TARGTPOP	DISPOINC
SALES	1.000	.945	.836
TARGTPOP		1.000	.781
DISPOINC			1.000

圖 6.4

SYGRAPH 的散佈圖矩陣與相關矩陣－Dwaine 工作室案例。

- 相關矩陣

相關矩陣(correlation matrix)是散佈圖矩陣的一個有用的輔助工具如圖6.4b中，此矩陣包含了Y與其他預測變數間，兩兩的簡單相關係數，其排列形式如同散佈圖矩陣：

$$\begin{bmatrix} 1 & r_{Y1} & r_{Y2} & \cdots & r_{Y,p-1} \\ r_{Y1} & 1 & r_{12} & \cdots & r_{1,p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ r_{Y,p-1} & r_{1,p-1} & r_{2,p-1} & \cdots & 1 \end{bmatrix} \quad (6.67)$$

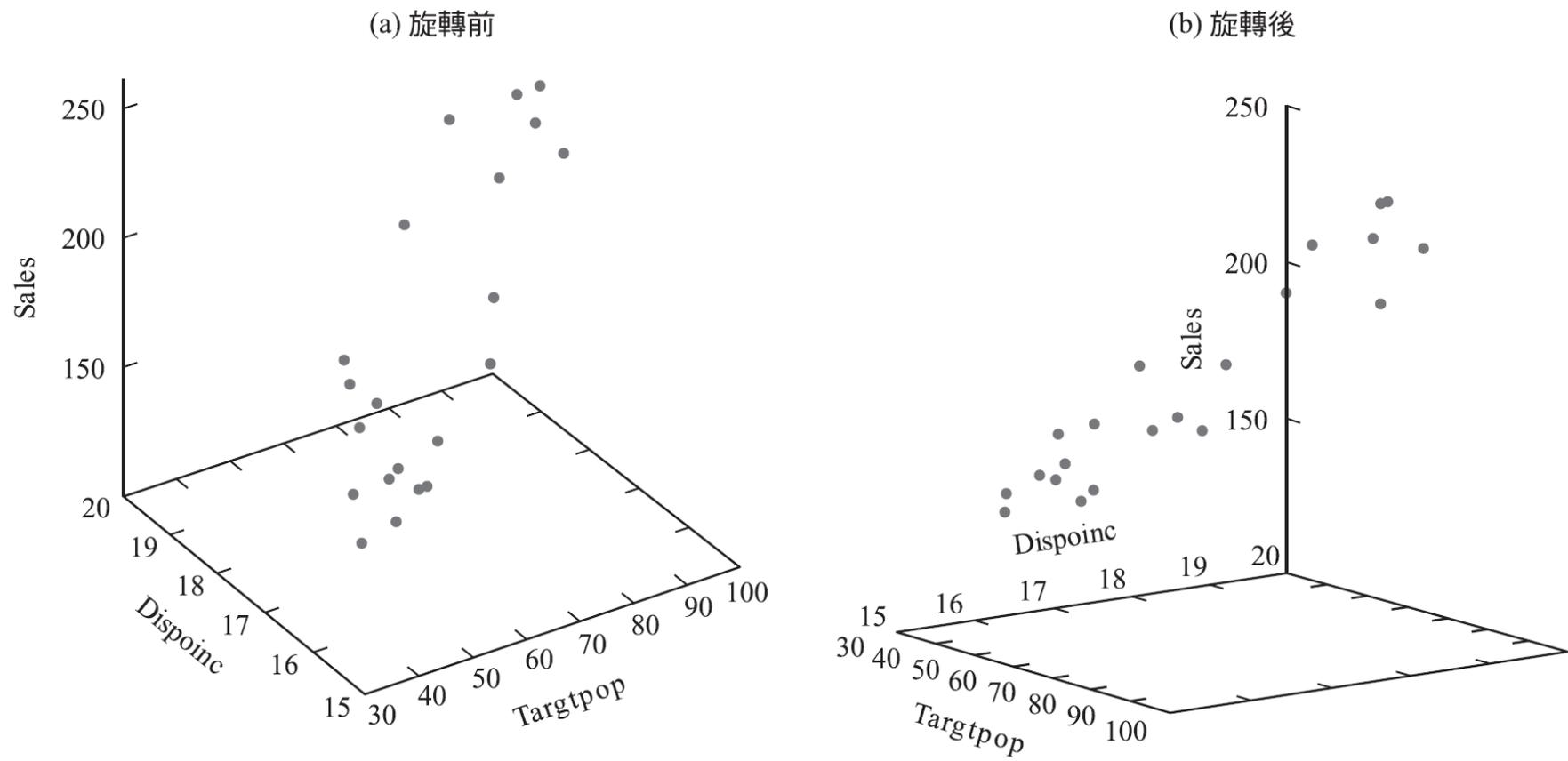
因為變數與自身的簡單相關係數必然為1，而且變數間的簡單相關係數計算不會受到順序影響，所以相關矩陣是一個主對角元素都是1的對稱矩陣

- 三維散佈圖

三維散佈圖(three-dimensional scatter)或是雲點圖(point clouds)，並且允許旋轉，圖6.6例示了一個三維散佈圖及其旋轉功能。

圖 6.6

SYGRAPH 旋轉前後之雲點圖－Dwaine 工作室案例。



殘差圖

- 對配適值的殘差圖可以用於評估複迴歸函數的適當性，以及常數誤差項變異數，並提供離群質的訊息
- 殘差的時序圖或順序圖提供了複迴歸中誤差項間的相關性診斷訊息，殘差的盒型圖與常態機率圖則可用於檢查誤差項是否合理地接近常態分配。
- 如果對每一個預測變數作出殘差圖，將有助於提供迴歸函數對預測變數適當與否的進一步訊息(如是否須考慮變數的曲線效果)，以及是否發生了變異程度不均一的現象。
- 殘差也應該對未被納入模型中的重要預測變數畫圖，觀察這些重要的預測變數是否對反應變數有重大的影響，另外可將迴歸模型中未被考慮的潛在交互作用項如 X_1X_2 、 X_1X_3 、 X_2X_3 作出殘差圖，以觀察模型中是否須加入這些交互作用項。
- 配適值的絕對殘差圖或是平方殘差圖有助於檢查誤差項變異數的均等性，

常態性的相關檢定

- 可以直接套用在複迴歸模型中，將殘差排序後，在常態性假設條件下的期望值依照(3.6)計算，然後求出計算結果與殘差間之相關係數，最後利用表B.6評估常態性假設的合理性。

常數變異數的 *Brown-Forsythe* 檢定

- 當誤差變異隨著某預測變數起伏，則(3.9)的 Breusch-Forsythe 檢定有助於我們判斷常數變異數假設的合理性，其作法首先將資料分成兩組，一組對應的預測變數值較低，而另一組較高，然後進行與簡單迴歸分析時一樣的檢定程序

常數變異數的 *Breusch-Pagan* 檢定

- 當誤差變異隨著某預測變數起伏，則(3.11)的 Breusch-Pagan 檢定也有助於我們判斷常數變異數假設的合理性，其作法首先將殘差平方後對預測變數進行簡單迴歸，計算其迴歸平方和 SSR^* ，並使用原複迴歸模型中的誤差平方和 SSE 進行簡單線迴歸時一樣的檢定程序。

配適不良的 F 檢定

- 在複迴歸模型中，檢定以下反應函數是否為一個適當的反應曲面：

$$E\{Y\} = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1}$$

在複迴歸中，重複觀測值的取法是固定每一個預測變數值，重複觀察數次反應變數之結果。

- 當表6.1的ANOVA表完成後，將SSE分解成(純誤差SSPE)與配適不良成分(SSLE)，(配適不良平方和 $SSLE = SSE - SSPE$)並計算純誤差平方和SSPE，

變異來源	SS	df	MS
迴歸	$SSR = \mathbf{b}'\mathbf{X}'\mathbf{Y} - \left(\frac{1}{n}\right) \mathbf{Y}'\mathbf{J}\mathbf{Y}$	$p - 1$	$MSR = \frac{SSR}{p - 1}$
誤差	$SSE = \mathbf{Y}'\mathbf{Y} - \mathbf{b}'\mathbf{X}'\mathbf{Y}$	$n - p$	$MSE = \frac{SSE}{n - p}$
總合	$SSTO = \mathbf{Y}'\mathbf{Y} - \left(\frac{1}{n}\right) \mathbf{Y}'\mathbf{J}\mathbf{Y}$	$n - 1$	

表 6.1

一般線性迴歸模型 (6.19) 的 ANOVA 表。

- 首先計算每一個重複組內Y觀測值對組內平均的離差平方和SSE， $\sum \sum (Y_{ij} - \hat{Y}_{ij})^2 = SSE$ 假設有c組重複觀測值，表示固定過c次測變數的值，分別進行重複反應變數之觀測值，而第j組的Y觀測值平均數用符號 \bar{Y}_j 表示，所有的c組平方和相加 $\sum \sum (Y_{ij} - \bar{Y}_j)^2 = SSPE$ 即為(3.16)的純誤差平方和SSPE，(3.24)的配適不良平方和 $SSLE = SSE - SSPE$ 。
- $SSPE$ 的自由度為 $(n - c)$ ， $SSLF$ 的自由度為 $(n - p) - (n - c) = (c - p)$ ，所以針對下面的檢定：

$$\begin{aligned} H_0 : E\{Y\} &= \beta_0 + \beta_1 X_1 + \cdots + \beta_{p-1} X_{p-1} \\ H_a : E\{Y\} &\neq \beta_0 + \beta_1 X_1 + \cdots + \beta_{p-1} X_{p-1} \end{aligned} \quad (6.68a)$$

檢定統計量為：

$$F^* = \frac{SSLF}{c - p} \div \frac{SSPE}{n - c} = \frac{MSLF}{MSPE} \quad (6.68b)$$

其中 $SSLF$ 與 $SSPE$ 如(3.24)與(3.16)所示，最後有決策法則：

$$\begin{aligned} \text{若 } F^* &\leq F(1 - \alpha; c - p, n - c), \text{ 則結論為 } H_0 \\ \text{若 } F^* &> F(1 - \alpha; c - p, n - c), \text{ 則結論為 } H_a \end{aligned} \quad (6.68c)$$

矯正測量

- 在複迴歸模型中，有時需要一個較為複雜的模型以描述曲線關係或是交互作用，例如將 X_2^2 視為一個變數，來描述 X_2 的曲線效果，或將 X_1X_3 視為一個變數，來描述 X_1 與 X_3 對反應變數的交互作用
- 對預測變數或反應變數進行變數轉換，以矯正模型的缺失。當誤差項的分配呈現明顯偏態情形，或是違背常數變異的情形，可以對反應變數進行變數轉換，經過對預測變數或反應變數轉換後，有時可以不必考慮交互作用的效應。
- 變數轉換的有效性需要再次經過殘差圖或是其它診斷工具，來檢驗轉換後的複迴歸模型是否適當。

6.9 雙預測變數的複迴歸案例

An Example-Multiple Regression with Two Predictor Variables

案例介紹

Dwaine工作室在21個一般規模的城市中經營兒童肖像業務工作，此時公司考慮在其他一般規模的城市中擴充營業據點，所以想研究社區之營業額(Y)是否可以經由社區內十六歲以下之人口數(X_1)與平均每人可支配所得(X_2)預測，目前該公司所經營的21個城市進一年來的資料如圖6.5(b)所示，其中營業額(Y)的單位是千元，輸出標記為SALES，十六歲以下之人口數(X_1)的單位是千人，輸出標記為TARGTPOP，平均每人可支配所得(X_2)的單位是千元，輸出標記為DISPOINC。

圖 6.5 (b) 基本資料

CASE	X1	X2	Y	FITTED	RESIDUAL
1	68.5	16.7	174.4	187.184	-12.7841
2	45.2	16.8	164.4	154.229	10.1706
3	91.3	18.2	244.2	234.396	9.8037
4	47.8	16.3	154.6	153.329	1.2715
5	46.9	17.3	181.6	161.385	20.2151
6	66.1	18.2	207.5	197.741	9.7586
7	49.5	15.9	152.8	152.055	0.7449
8	52.0	17.2	163.2	167.867	-4.6666
9	48.9	16.6	145.4	157.738	-12.3382
10	38.4	16.0	137.2	136.846	0.3540
11	87.9	18.3	241.9	230.387	11.5126
12	72.8	17.1	191.1	197.185	-6.0849
13	88.4	17.4	232.0	222.686	9.3143
14	42.9	15.8	145.3	141.518	3.7816
15	52.5	17.8	161.1	174.213	-13.1132
16	85.7	18.4	209.7	228.124	-18.4239
17	41.3	16.5	146.4	145.747	0.6530
18	51.7	16.3	144.0	159.001	-15.0013
19	89.6	18.1	232.6	230.987	1.6130
20	82.7	19.1	224.1	230.316	-6.2160
21	52.3	16.0	166.5	157.064	9.4356

圖 6.5

SYSTAT 複迴歸輸出與基本資料－Dwayne 工作室案

(a) 複迴歸輸出

DEP VAR: SALES N: 21 MULTIPLE R: 0.957 SQUARED MULTIPLE R:
0.917
ADJUSTED SQUARED MULTIPLE R: .907 STANDARD ERROR OF ESTIMATE:
11.0074

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-68.8571	60.0170	0.0000	.	-1.1473	0.2663
TARGTPOP	1.4546	0.2118	0.7484	0.3896	6.8682	0.0000
DISPOINC	9.3655	4.0640	0.2511	0.3896	2.3045	0.0333

圖 6.5

SYSTAT 複迴歸輸出與基本資料 - Dwaine 工作室案

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	24015.2821	2	12007.6411	99.1035	0.0000
RESIDUAL	2180.9274	18	121.1626		

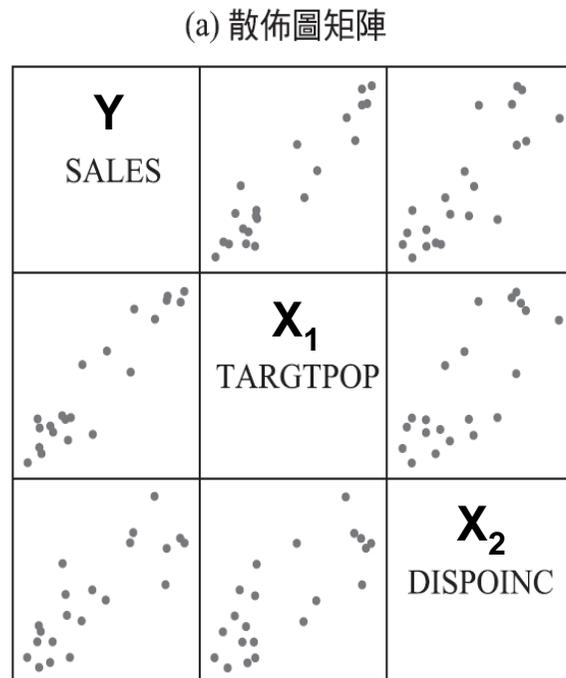
INVERSE ($X'X$)

	1	2	3
1	29.7289		
2	0.0722	0.00037	
3	-1.9926	-0.0056	0.1363

從圖6.4a中的SYGRAPH散佈圖矩陣來看，似乎可以預期為具有常態誤差的第一階迴歸模型：

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i \quad (6.69)$$

應注意的是 X_1 與 Y 、 X_2 與 Y 間的線性關係，而後者的關係似乎較為鬆散，同時 X_1 與 X_2 好像也有線性關係，從圖6.4b中的數據可以支持圖6.4a中視覺判斷。



(b) 相關矩陣

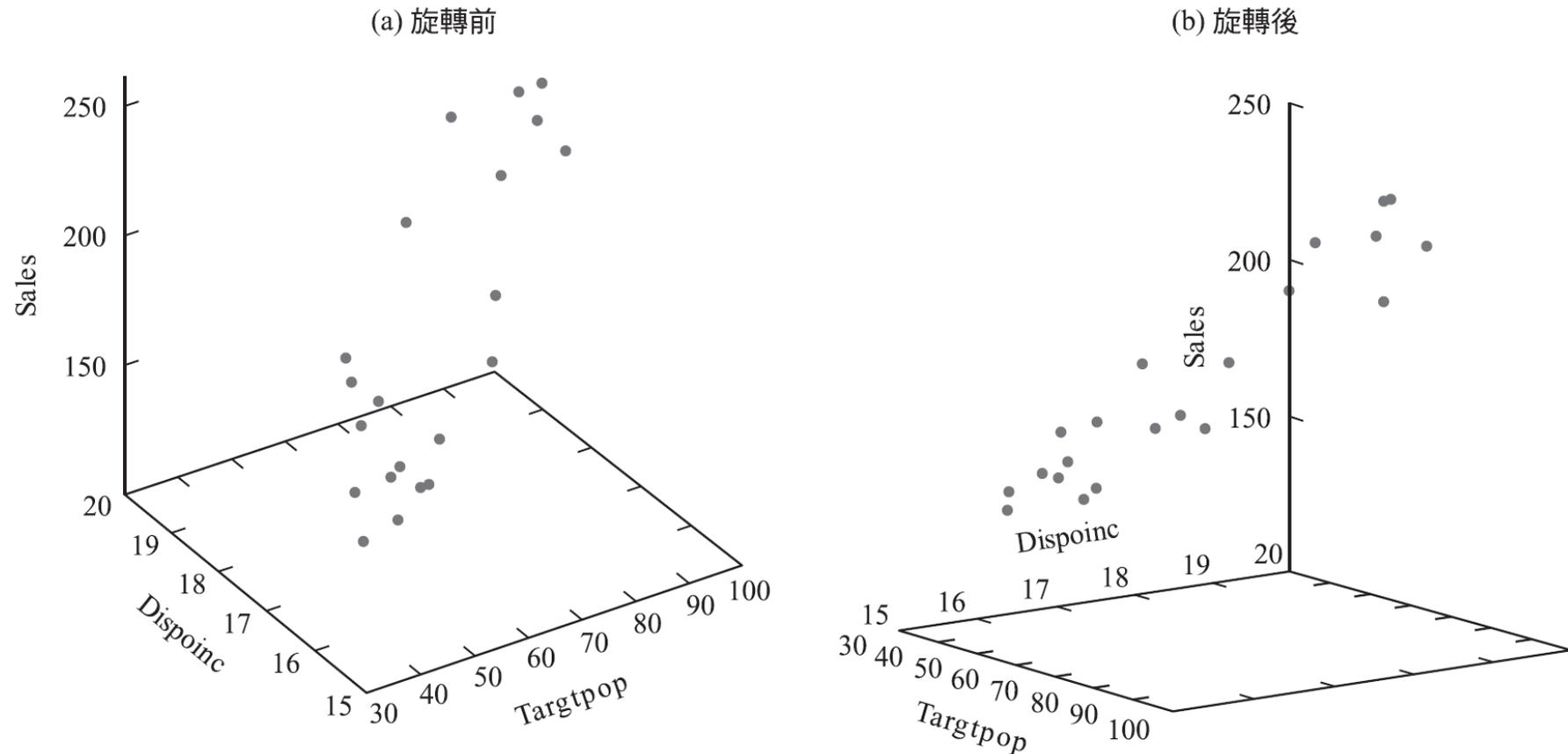
	SALES	TARGTPOP	DISPOINC
SALES	1.000	.945	.836
TARGTPOP		1.000	.781
DISPOINC			1.000

圖 6.4
SYGRAPH 的散佈圖矩陣與相關矩陣 — Dwaine 工作室 案例。

- 透過SYGRAPH所畫出的雲點圖6.6a中，將座標軸旋轉將出現圖6.6b的結果此一結果似乎也暫時支持迴歸函數為反應平面的假設性判斷

圖 6.6

SYGRAPH 旋轉前後之雲點圖－Dwaine 工作室案例。



- 基本計算

Dwaine工作室案例之 \mathbf{X} 與 \mathbf{Y} 如下：

$$\mathbf{X} = \begin{bmatrix} 1 & 68.5 & 16.7 \\ 1 & 45.2 & 16.8 \\ \vdots & \vdots & \vdots \\ 1 & 52.3 & 16.0 \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} 174.4 \\ 164.4 \\ \vdots \\ 166.5 \end{bmatrix} \quad (6.70)$$

先計算：

1.

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 68.5 & 45.2 & \dots & 52.3 \\ 16.7 & 16.8 & \dots & 16.0 \end{bmatrix} \begin{bmatrix} 1 & 68.5 & 16.7 \\ 1 & 45.2 & 16.8 \\ \vdots & \vdots & \vdots \\ 1 & 52.3 & 16.0 \end{bmatrix}$$

結果為：

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 21.0 & 1,302.4 & 360.0 \\ 1,302.4 & 87,707.9 & 22,609.2 \\ 360.0 & 22,609.2 & 6,190.3 \end{bmatrix} \quad (6.71)$$

2.

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 68.5 & 45.2 & \dots & 52.3 \\ 16.7 & 16.8 & \dots & 16.0 \end{bmatrix} \begin{bmatrix} 174.4 \\ 164.4 \\ \vdots \\ 166.5 \end{bmatrix}$$

結果為：

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 3,820 \\ 249,643 \\ 66,073 \end{bmatrix}$$

(6.72)

3.

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 21.0 & 1,302.4 & 360.0 \\ 1,302.4 & 87,707.9 & 22,609.2 \\ 360.0 & 22,609.2 & 6,190.3 \end{bmatrix}^{-1}$$

應用(5.23)可以得到結果：

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} 29.7289 & .0722 & -1.9926 \\ .0722 & .00037 & -.0056 \\ -1.9926 & -.0056 & .1363 \end{bmatrix}$$

(6.73)

- 代數等式 在第一階迴歸模型(6.69)中的 $\mathbf{X}'\mathbf{X}$ 為：

或是：

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & X_{12} \\ 1 & X_{21} & X_{22} \\ \vdots & \vdots & \vdots \\ 1 & X_{n1} & X_{n2} \end{bmatrix}$$

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1}X_{i2} \\ \sum X_{i2} & \sum X_{i2}X_{i1} & \sum X_{i2}^2 \end{bmatrix} \quad (6.74) \quad \mathbf{X}'\mathbf{X} = \begin{bmatrix} 21.0 & 1,302.4 & 360.0 \\ 1,302.4 & 87,707.9 & 22,609.2 \\ 360.0 & 22,609.2 & 6,190.3 \end{bmatrix}$$

- 在Dwaine工作室案例中可以計算出：

$$n=21$$

$$\sum X_{i1} = 68.5 + 45.2 + \dots = 1,302.4$$

$$\sum X_{i1}X_{i2} = 68.5(16.7) + 45.2(16.8) + \dots = 22,609.2$$

$$\sum X_{i2} = 360.0 \text{ 等等}$$

這些結果在(6.71)中都可以找到。

關於雙預測變數下的第一階迴歸模型(6.69)， $\mathbf{X}'\mathbf{Y}$ 為：

$$\mathbf{X}'\mathbf{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{21} & \cdots & X_{n1} \\ X_{12} & X_{22} & \cdots & X_{n2} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1}Y_i \\ \sum X_{i2}Y_i \end{bmatrix} \quad (6.75) \quad \mathbf{X}'\mathbf{Y} = \begin{bmatrix} 3,820 \\ 249,643 \\ 66,073 \end{bmatrix}$$

- 在Dwayne工作室案例中，可以計算出：

$$\sum Y_i = 174.4 + 164.4 + \dots = 3,820$$

$$\sum X_{i1}Y_i = 68.5(174.4) + 45.2(64.4) + \dots = 249,643$$

$$\sum X_{i2}Y_i = 16.7(174.4) + 16.8(64.4) + \dots = 66,073$$

這些結果在(6.72)中都可以找到。

- 迴歸函數的估計

從(6.25)中可以立即利用(6.72)與(6.73)計算最小平方估計量 \mathbf{b} ：

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} = \begin{bmatrix} 29.7289 & .0722 & -1.9926 \\ .0722 & .00037 & -.0056 \\ -1.9926 & -.0056 & .1363 \end{bmatrix} \begin{bmatrix} 3,820 \\ 249,643 \\ 66,073 \end{bmatrix}$$

結果為：

$$\mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} -68.857 \\ 1.455 \\ 9.366 \end{bmatrix} \quad (6.76)$$

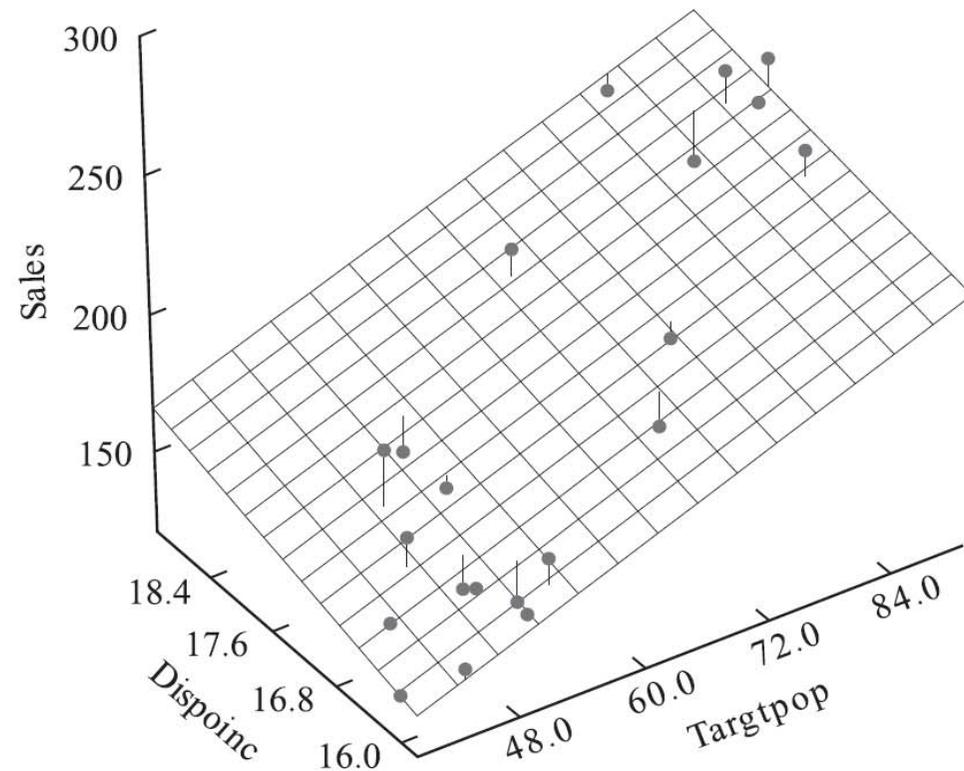
所以所估計的迴歸函數為：

$$\hat{Y} = -68.857 + 1.455 X_1 + 9.366 X_2$$

圖 6.7

S-Plus 估計的迴歸
函數平面圖 —
Dwaine 工作室案
例。

- 畫出了反應變數疊印在所估計的迴歸函數平面上之情形，連接在迴歸函數平面上的許多小段垂直細線，表示對應的觀測值之殘差。



- 圖 6.5a 為 Dwaine 工作室案例中的 SYSTAT 複迴歸輸出，所估計的迴歸係數在欄位 COEFFICIENT 下。
- SYSTAT 也同時輸出 $(X'X)^{-1}$ ，由於它是一個對稱矩陣，所以只有輸出該矩陣的左下角，結果自然與(6.73)相同。

圖6.5a Dwaine工作室案例

VARIABLE	COEFFICIENT	STD ERROR	STD COEF	TOLERANCE	T	P(2 TAIL)
CONSTANT	-68.8571	60.0170	0.0000	.	-1.1473	0.2663
TARGTPOP	1.4546	0.2118	0.7484	0.3896	6.8682	0.0000
DISPOINC	9.3655	4.0640	0.2511	0.3896	2.3045	0.0333

INVERSE (X'X)

	1	2	3
1	29.7289		
2	0.0722	0.00037	
3	-1.9926	-0.0056	0.1363

- 標準方程式的代數版本 雙預測變數下的標準方程式之代數式可以透過(6.74)與(6.75)得出：

$$(\mathbf{X}'\mathbf{X})\mathbf{b} = \mathbf{X}'\mathbf{Y}$$

$$\begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1}X_{i2} \\ \sum X_{i2} & \sum X_{i2}X_{i1} & \sum X_{i2}^2 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum Y_i \\ \sum X_{i1}Y_i \\ \sum X_{i2}Y_i \end{bmatrix}$$

因此有下面的標準方程式：

$$\begin{aligned} \sum Y_i &= nb_0 + b_1 \sum X_{i1} + b_2 \sum X_{i2} \\ \sum X_{i1}Y_i &= b_0 \sum X_{i1} + b_1 \sum X_{i1}^2 + b_2 \sum X_{i1}X_{i2} \\ \sum X_{i2}Y_i &= b_0 \sum X_{i2} + b_1 \sum X_{i2}X_{i1} + b_2 \sum X_{i2}^2 \end{aligned} \tag{6.77}$$

配適值及殘差

- 有關迴歸模型(6.69)對於資料的適切性的檢查工作，先計算配適值 \hat{Y}_i 與殘差 $e_i=Y_i-\hat{Y}_i$ ，根據(6.28)可以得到： $\hat{\mathbf{Y}}=\mathbf{Xb}$

$$\begin{pmatrix} \hat{Y}_1 \\ \hat{Y}_2 \\ \vdots \\ \hat{Y}_{21} \end{pmatrix} = \begin{pmatrix} 1 & 68.5 & 16.7 \\ 1 & 45.2 & 16.8 \\ \vdots & \vdots & \vdots \\ 1 & 52.3 & 16.0 \end{pmatrix} \begin{pmatrix} -68.857 \\ 1.455 \\ 9.366 \end{pmatrix} = \begin{pmatrix} 187.2 \\ 154.2 \\ \vdots \\ 157.1 \end{pmatrix}$$

再利用(6.29)，可以得到 $\mathbf{e}=\mathbf{Y}-\hat{\mathbf{Y}}$

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_{21} \end{pmatrix} = \begin{pmatrix} 174.4 \\ 164.4 \\ \vdots \\ 166.5 \end{pmatrix} - \begin{pmatrix} 187.2 \\ 154.2 \\ \vdots \\ 157.1 \end{pmatrix} = \begin{pmatrix} -12.8 \\ 10.2 \\ \vdots \\ 9.4 \end{pmatrix}$$

(b) 基本資料

- 圖 6.5b 第五欄及第六欄為配適值與殘差的電腦計算結果輸出，再計算之結果較為精細模型配適度分析

CASE	X1	X2	Y	FITTED	RESIDUAL
1	68.5	16.7	174.4	187.184	-12.7841
2	45.2	16.8	164.4	154.229	10.1706
3	91.3	18.2	244.2	234.396	9.8037
4	47.8	16.3	154.6	153.329	1.2715
5	46.9	17.3	181.6	161.385	20.2151
6	66.1	18.2	207.5	197.741	9.7586
7	49.5	15.9	152.8	152.055	0.7449
8	52.0	17.2	163.2	167.867	-4.6666
9	48.9	16.6	145.4	157.738	-12.3382
10	38.4	16.0	137.2	136.846	0.3540
11	87.9	18.3	241.9	230.387	11.5126
12	72.8	17.1	191.1	197.185	-6.0849
13	88.4	17.4	232.0	222.686	9.3143
14	42.9	15.8	145.3	141.518	3.7816
15	52.5	17.8	161.1	174.213	-13.1132
16	85.7	18.4	209.7	228.124	-18.4239
17	41.3	16.5	146.4	145.747	0.6530
18	51.7	16.3	144.0	159.001	-15.0013
19	89.6	18.1	232.6	230.987	1.6130
20	82.7	19.1	224.1	230.316	-6.2160
21	52.3	16.0	166.5	157.064	9.4356

模型配適度分析

- 圖6.8a中的配適值殘差圖來看，圖中似乎沒有出現任何對於反應平面的系統性偏離，也有看不出誤差項變異數與 \hat{Y} 之高低有相關性，圖6.8b與圖6.8c是殘差 e 分別對於 X_1 與 X_2 的關係圖，圖示反應函數的良好配適與常數的誤差變異數。

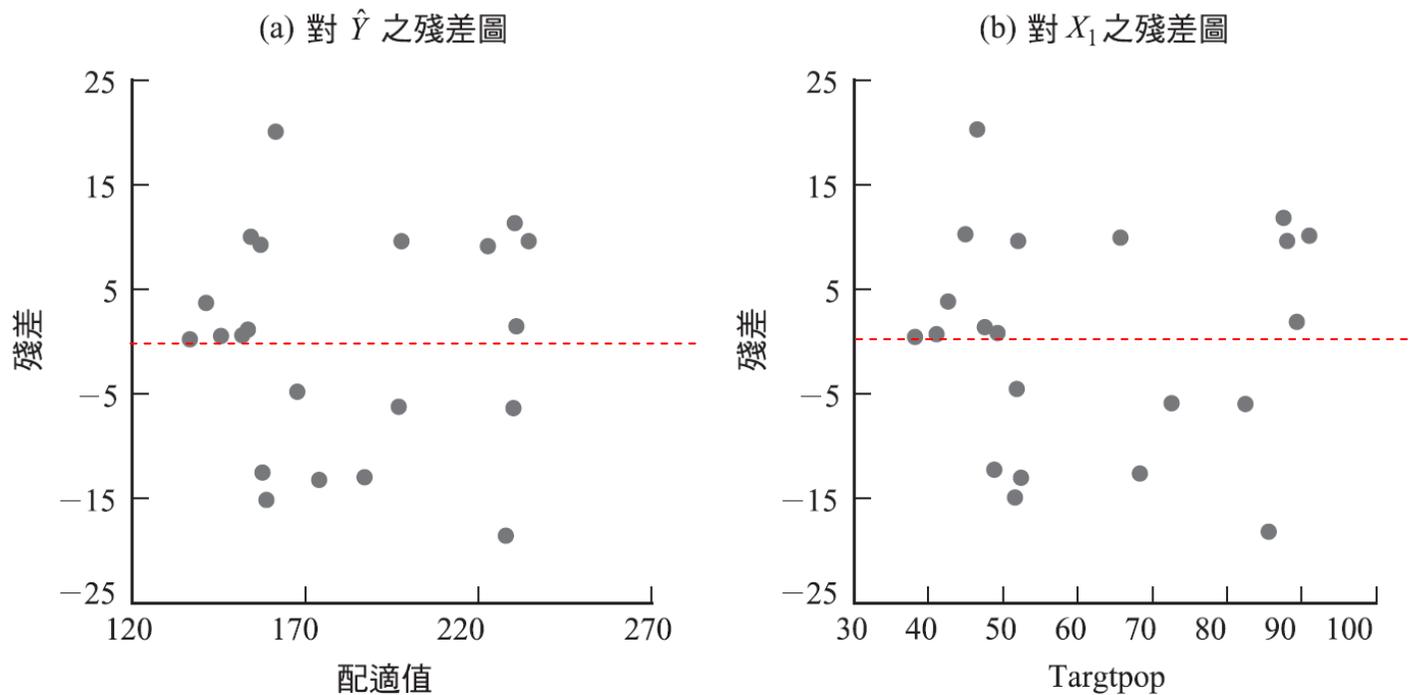
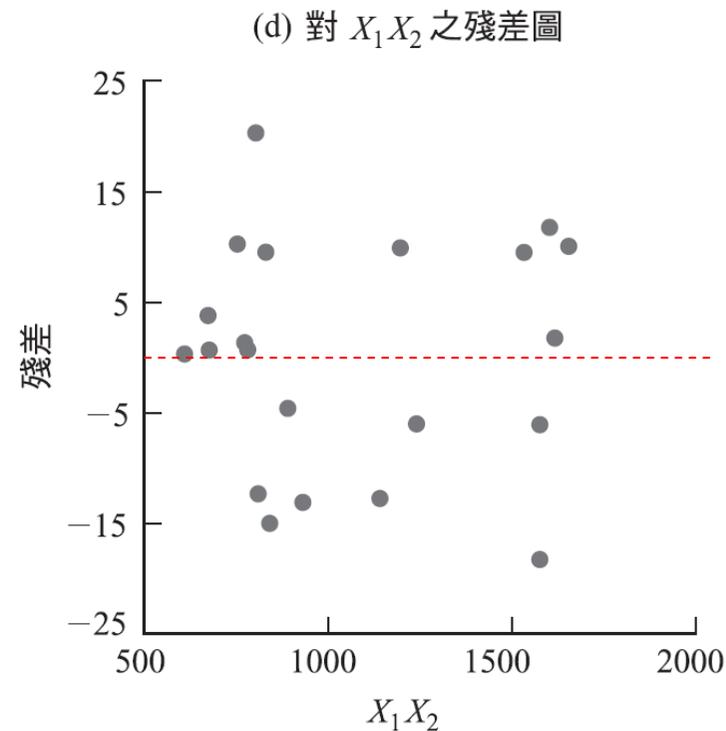
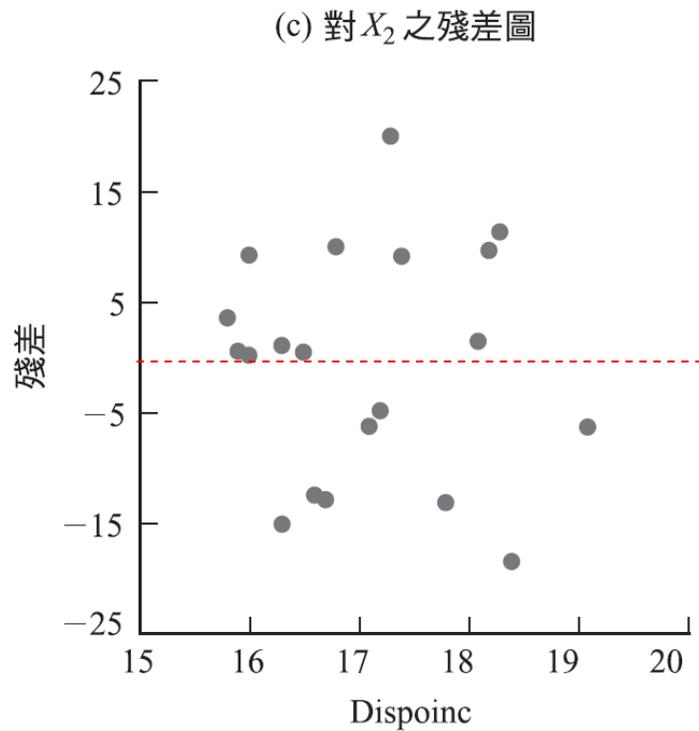


圖 6.8
SYGRAPH 診斷圖
— Dwaine 工作室
案例。

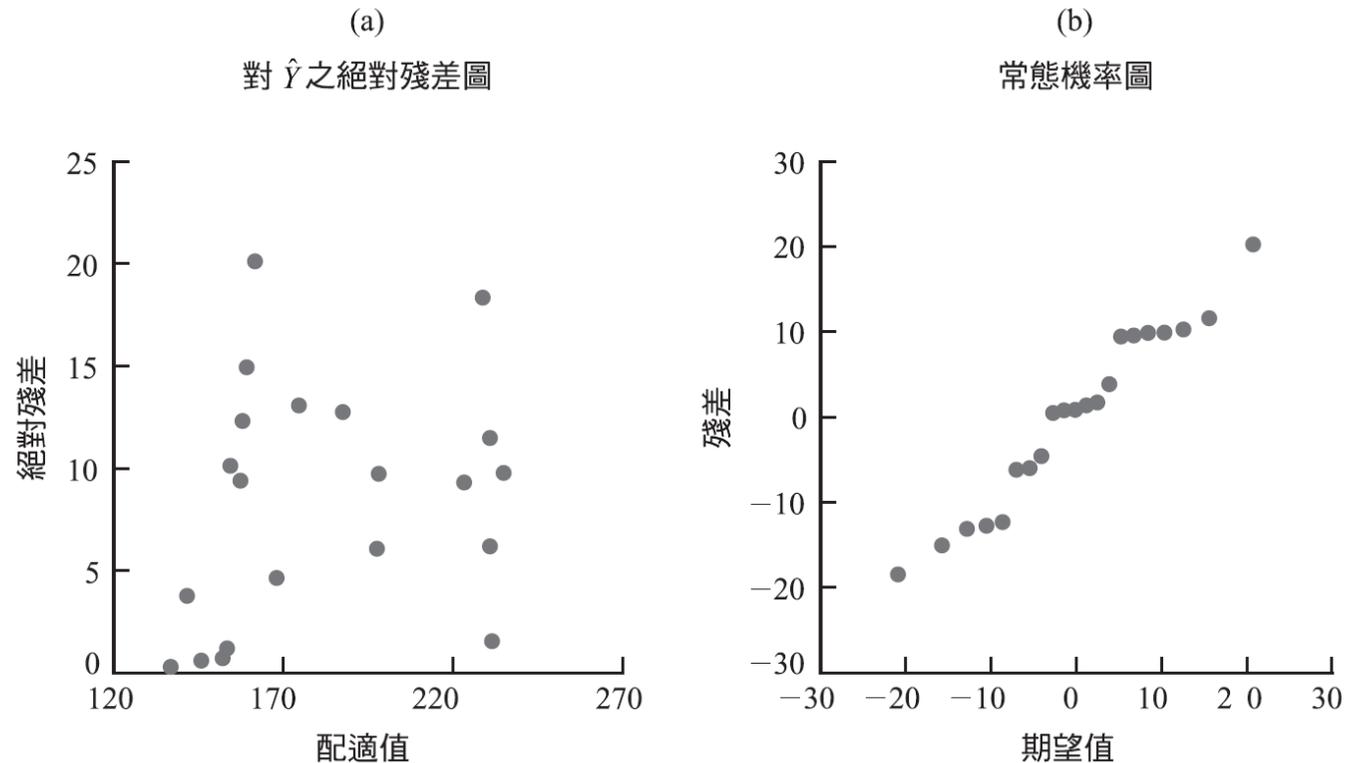


- 在複迴歸模型的應用上，可以透過對 $X_1 X_2$ 畫出殘差圖來檢查 X_1 與 X_2 是否存在交互作用，在對 $X_1 X_2$ 殘差圖在圖 6.8d，當圖中顯示某些型態時，則可能需考慮交互作用效應，此時反應函數可能為：
$$E\{Y\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$
- 不過，在圖 6.8d 中沒有顯示任何系統性的型態，因此不考慮交互作用效應。

- 圖6.9是另外兩個診斷圖形，在圖6.9a中所畫為配適值對絕對殘差圖，圖中看不出來變異數有不穩定大小的情形，圖6.9b是殘差的常態機率圖，散佈的型態略為呈現直線，將殘差排序後與常態期望值得相關係數為 0.980 (根據表B.6，在 $n=21$ ， $\alpha=0.05$ 經過內插結果得到臨界值為.9525)，高的相關係數有助於確認誤差項沒有明顯背離常態分配之假設。

圖 6.9

額外診斷圖 —
Dwaine 工作室案
例。



變異數分析

- 檢定營業額是否真與十六歲以下人口數及平均每人可支配所得有迴歸關係，可以製作ANOVA表來分析，表中所需數據包括：

$$\mathbf{Y}'\mathbf{Y} = [174.4 \ 164.4 \dots 166.5] \begin{pmatrix} 174.4 \\ 164.4 \\ \vdots \\ 166.5 \end{pmatrix} = 721.072.40$$

$$(1/n) \mathbf{Y}'\mathbf{J}\mathbf{Y} = [174.4 \ 164.4 \dots 166.5] \begin{pmatrix} 1 & 1 \dots 1 \\ 1 & 1 \dots 1 \\ \vdots & \vdots \vdots \\ 1 & 1 \dots 1 \end{pmatrix} \begin{pmatrix} 174.4 \\ 164.4 \\ \vdots \\ 166.5 \end{pmatrix}$$

$$\mathbf{SSTO} = \mathbf{Y}'\mathbf{Y} - (1/n)\mathbf{Y}'\mathbf{J}\mathbf{Y}$$

$$\mathbf{SSE} = \mathbf{Y}'\mathbf{Y} - \mathbf{b}'\mathbf{X}'\mathbf{Y}$$

$$= (3820.0/21)^2 = 694.876.1$$

因此， $SSTO = Y'Y - (1/n)Y'JY = 721.072.40 - 694.876.19 = 26,196.21$

另一方面，從(6.72)與(6.76)的結果，我們有

$$\begin{aligned}
 SSE &= Y'Y - \mathbf{b}'X'Y \\
 &= 721,072.40 - \begin{bmatrix} -68.857 & 1.455 & 9.366 \end{bmatrix} \begin{pmatrix} 3,820 \\ 249,643 \\ 66,073 \end{pmatrix} \\
 &= 721,072.40 - 718,891.47 = 2,180.93
 \end{aligned}$$

將上面兩個結果相減：

$$SSR = SSTO - SSE = 26,196.21 - 2,180.93 = 24,015.28$$

圖 6.5a 中ANOVA

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
REGRESSION	24015.2821	2	MSR 12007.6411	99.1035	0.0000
RESIDUAL	2180.9274	18	MSE 121.1626		

迴歸關係的檢定

- 要檢定營業額是否真與十六歲以下人口數及平均每人可支配所得有迴歸關係，可以考慮：

$$H_0 : \beta_1 = 0 \text{ 與 } \beta_2 = 0$$

$$H_a : \beta_1 \text{ 與 } \beta_2 \text{ 非全為 } 0$$

採用檢定統計量(6.39b)：

$$F^* = \frac{MSR}{MSE} = \frac{12,007.64}{121.1626} = 99.1$$

- 上面的檢定統計量在SYSTAT輸出位置在標視為 F-RATIO 的地方，在顯著水準 $\alpha = 0.05$ 下，臨界值 $F(0.95; 2, 18) = 3.55 < F^* = 99.1$ ，所以結論為拒絕 H_0 之虛無假設（接受 H_a 之對立假設），也就是營業額的確與十六歲以下人口數及平均每人可支配所得有迴歸關係，此一檢定的p-值在SYSTAT的輸出位置標視為 p 的地方，SYSTAT 輸出為 0.0000。

(a) 複迴歸輸出

```
DEP VAR: SALES N: 21 MULTIPLE R: 0.957 SQUARED MULTIPLE R:  
                                0.917  
ADJUSTED SQUARED MULTIPLE R: .907 STANDARD ERROR OF ESTIMATE:  
                                11.0074
```

複判定係數 透過(6.40)可以得到：

$$R^2 = \frac{SSR}{SSTO} = \frac{24,015.28}{26,196.21} = .917$$

- 只考慮十六歲以下人口數及平均每人可支配所得這兩個預測變數，則可以解釋了營業額資料變異原因的 91.7% (0.917)，複判定係數 R_a 在SYSTAT的輸出位置在標示為SQUARED MULTIPLE R的地方，SYSTAT的輸出還包括了複相關係數 $R = 0.957$ 以及(6.42)的調整的複判定係數 $R_a^2 = 0.907$ ，SYSTAT的輸出位置在標示為 ADJUSTED SQUARED MULTIPLE R的地方，在本案例中變數的個數對 R^2 影響很小。

迴歸參數的估計

- 案例中由於截距項落在模型範圍外很遠，所以參數 β_0 並不感興趣，現在要進行的是關於全族信賴係數 0.90 下， β_1 與 β_2 的聯合估計，將採用 Bonferroni 同步信賴界限的程序。
- 首先需先計算共變異矩陣的估計量 $s^2\{\mathbf{b}\}$:

$$s^2\{\mathbf{b}\} = MSE(\mathbf{X}'\mathbf{X})^{-1}$$

- 在圖6.5a中已經輸出了 MSE ，而 $(\mathbf{X}'\mathbf{X})^{-1}$ 則在(6.73)中計算過，所以：

$$\begin{aligned} s^2\{\mathbf{b}\} &= 121.1626 \begin{bmatrix} 29.7289 & .0722 & -1.9926 \\ .0722 & .00037 & -.0056 \\ -1.9926 & -.0056 & .1363 \end{bmatrix} \\ &= \begin{bmatrix} 3,602.0 & 8.748 & -241.43 \\ 8.748 & .0448 & -.679 \\ -241.43 & -.679 & 16.514 \end{bmatrix} \end{aligned} \quad (6.78)$$

- 所需的兩個變異數或是標準差之估計量為：

$$s^2\{b_1\} = 0.448 \quad \text{或} \quad s\{b_1\} = 0.212$$

$$s^2\{b_2\} = 16.514 \quad \text{或} \quad s\{b_2\} = 4.06 \quad n-p=21-3$$

- 其次需要 $g = 2$ 個聯合估計數：

$$B = t(1 - \alpha/2g; n - p)$$

$$B = t[1 - 0.10 / 2(2); 18] = t(0.975; 18) = 2.101$$

- 所以同步信賴界限的結果為 $1.455 \pm 2.101(0.212)$ 與 $9.366 \pm 2.101(4.06)$ ，而信賴區間為：

$$b_k \pm t(1 - \alpha/2; n - p) s\{b_k\}$$

$$1.01 \leq \beta_1 \leq 1.90$$

$$0.84 \leq \beta_2 \leq 17.9$$

在全族信賴係數 0.90 ，可以合理推論 β_1 落在 1.01 與 1.09 之間， β_2 落在 0.84 與 17.9 之間。

- 上面關於 β_1 與 β_2 的聯合信賴區間指出了範圍都是正值，這一點也正式當其他條件不變時，隨著十六歲以下人口數增加，以及平均每人可支配所得增加，符合預期營業額增加的理論

平均反應的估計

Dwaine工作室想估計，當十六歲以下人口數 $X_{h1} = 65.4$ 千人，平均每人可支配所得 $X_{h2} = 17.6$ 千元時，一城市的期望營業額之95%信賴區間，先定義：

$$\mathbf{X}_h = \begin{pmatrix} 1 \\ 6.54 \\ 17.6 \end{pmatrix}$$

平均營業額可以從(6.55)計算出：

$$\hat{\mathbf{Y}}_h = \mathbf{X}'_h \mathbf{b} = [1 \ 6.54 \ 17.6] \begin{pmatrix} -68.875 \\ 1.455 \\ 9.366 \end{pmatrix} = 191.10$$

- 根據(6.58)，同時應用(6.78)的結果，其變異數之估計量為：

$$\begin{aligned}
 s^2\{\hat{Y}_h\} &= \mathbf{X}'_h \mathbf{s}^2\{\mathbf{b}\} \mathbf{X}_h \\
 &= [1 \quad 65.4 \quad 17.6] \begin{pmatrix} 3602.0 & 8.748 & -241.43 \\ 8.748 & 0.448 & -0.679 \\ -241.43 & -0.679 & 16.514 \end{pmatrix} \begin{pmatrix} 1 \\ 65.4 \\ 17.6 \end{pmatrix} \\
 &= 7.656
 \end{aligned}$$

或是 $s\{\hat{Y}_h\} = 2.77$

- 要符合信賴係數 0.95，我們需要 $t(0.975; 18) = 2.101$ ，利用(6.59)可以得到信賴界限為 $191.10 \pm 2.101(2.77)$ ，所以 $E\{\hat{Y}_h\}$ 之 95% 信賴區間為：
 $185.3 \leq E\{\hat{Y}_h\} \leq 196.9$
- 亦即在信賴係數.95 下，當十六歲以下人口數 $X_{h1} = 65.4$ 千人，平均每人可支配所得 $X_{h2} = 17.6$ 千元時，一城市的期望營業額將介於 183.5 至 196.9 之間

變異數估計量 $s^2\{\hat{Y}_h\}$ 的代數版本 利用(6.58)

$$s^2\{\hat{Y}_h\} = \mathbf{X}'_h \mathbf{s}^2\{\mathbf{b}\} \mathbf{X}_h$$

在雙預測變數的第一階模型中，我們有：

$$\begin{aligned} s^2\{\hat{Y}_h\} &= s^2\{b_0\} + X_{h1}^2 s^2\{b_1\} + X_{h2}^2 s^2\{b_2\} + 2X_{h1}s\{b_0, b_1\} \\ &\quad + 2X_{h2}s\{b_0, b_2\} + 2X_{h1}X_{h2}s\{b_1, b_2\} \end{aligned} \quad (6.79)$$

新觀測值的預測界限

在Dwayne工作室的可能擴充方案中，需要預測兩個新城市的營業額，有如下的資料：

	A城市	B城市
X_{h1}	65.4	53.1
X_{h2}	17.6	17.7

預測區間需要有90%的全族信賴係數，注意這兩個城市的資料並沒有在原先21各城市的資料範圍內。

- 為了評估哪一個同步預測區間較佳，先依照(6.65a)與(6.66a)分別計算S與B，其中 $g = 2$ ， $1 - \alpha = 0.90$ ，結果為：

$$S^2 = 2F(.90; 2, 18) = 2(2.62) = 5.24, \quad S = 2.29$$

$$\text{且 } B = t [1 - .10/2(2); 18] = t (.975; 18) = 2.101$$

所以在本案例中 Bonferroni 程序的界限效率較佳（較小）。

- 對於A城市而言，因為預測變數的值跟前面估計平均營業額相同，所以可以引用：

$$\hat{Y}_h = 191.10 \quad s^2 \{ \hat{Y}_h \} = 7.656 \quad \text{MSE} = 121.1626$$

透過(6.63a)可以得到：

$$s^2 \{ \text{pred} \} = \text{MSE} + s^2 \{ \hat{Y}_h \} = 121.1626 + 7.656 = 128.82$$

$$\text{或是 } s \{ \text{pred} \} = 11.35$$

$$S^2 = gF(1 - \alpha; g, n - p)$$

類似的過程在B城市上(計算省略)：

$$B = t(1 - \alpha / 2g; n - p)$$

$$\hat{Y}_h = 174.15 \quad s \{ \text{pred} \} = 11.93$$

- Bonferroni 乘數 $B=2.101$ ，所以利用(6.66)可以得到全族信賴係數 0.90 的同步預測界限為 $191.10 \pm 2.101(11.35)$ 與 $174.15 \pm 2.101(11.93)$ ，所以同步預測區間為 $\hat{Y}_h \pm Bs \{\text{pred}\}$

$$A \text{ 城市} : 167.3 \leq Y_{h(\text{new})} \leq 214.9$$

$$B \text{ 城市} : 149.1 \leq Y_{h(\text{new})} \leq 199.2$$

- 在全族信賴係數 0.90 下，預測這兩個城市之營業額將會落在上面的界限內，Dwayne 工作室認為這樣的預測界限對計畫有些幫助，但是希望能針對特定城市之營業額有更窄的預測區間，因此聘請顧問研究是否尚有其他預測變數可以幫助 Dwayne 工作室做出更窄的預測區間。
- 在上例中雖然複判定係數 $R^2 = 0.917$ 相當高，但是預測界限卻未能使人滿意，這提醒了就算是高的 R^2 值，也並不代表必然有高精確度的預測。