第三章　變數刪除過程

3.1 M.W.G.變數刪除之理論與步驟

本節係介紹 Mansfield et al.(1977)文中刪除過多自變數理論的詳細過程。

考慮模型(2.2.1)，並將正交矩陣V分割成$V=[V_s : V_{k-s}]$，其中$V_s=[v_1,\cdots,v_s]$是那些對應於矩陣$X'X$中較大的s個特徵值之特徵向量。很明顯地，W可對應分割成$W=[W_s : W_{k-s}]=X[V_s : V_{k-s}]$。於是模型(2.2.1)可表示成

$$Y=\beta_0 1+[W_s : W_{k-s}]\begin{bmatrix}\gamma_s \\ \gamma_{k-s}\end{bmatrix}+\varepsilon$$

令Λ_s=diag($\lambda_1,\cdots,\lambda_s$)，則$\gamma_s=(\gamma_1,\cdots,\gamma_s)'$的最小平方估計為

$$\hat{\gamma}_s=(\hat{\gamma}_1,\cdots,\hat{\gamma}_s)'=\Lambda_s^{-1}W_s'Y$$

而這正是(2.2.3)式中主成份迴歸模型的係數。

假設要刪除r個自變數。因為自變數的順序無關緊要，為了說明方便起見，可令要刪除的是最後r個自變數。將V_s分割成$V_s=\begin{bmatrix}V_1 \\ V_2\end{bmatrix}$，其中$V_1:(k-r)\times s$，$V_2$：$r\times s$，而$V_2$中的列向量就是對應於即將被刪除的最後$r$個自變數。

Mansfield et al. (1977)考慮了其他的估計量

$$\tilde{\gamma}_s = A_s\hat{\gamma}_s = A_s\Lambda_s^{-1}W_s'Y \qquad (3.1.1)$$

來取代原來的$\hat{\gamma}_s$，此處的矩陣A_s為一$s\times s$的對角矩陣。當A_s為單位矩陣I_s時，$\tilde{\gamma}_s=\hat{\gamma}_s$。

在採用$\tilde{\gamma}_s$取代$\hat{\gamma}_s$後，參數β的估計則變為

$$\tilde{\beta} \equiv V_s\tilde{\gamma}_s = \begin{pmatrix} V_1A_s\hat{\gamma}_s \\ V_2A_s\hat{\gamma}_s \end{pmatrix} \equiv \begin{pmatrix} \tilde{\beta}_{k-r} \\ \tilde{\beta}_r \end{pmatrix} \qquad (3.1.2)$$

所以，欲刪除最後r個自變數，也就是希望$\tilde{\beta}$的最後r個元素$\tilde{\beta}_r$為零。另一方面，採用主成份估計並刪掉最後r個自變數後，預測方程式會變成

$$\tilde{Y} = \bar{Y} + \sum_{j=1}^{s} u_j\tilde{\gamma}_j = \sum_{j=1}^{k-r} z_j\tilde{\beta}_j \qquad (3.1.3)$$

此時$\tilde{Y}$的殘差平方和為

$$\widetilde{\mathrm{SSE}} = \sum_{i=1}^{n}\left(Y_i - \tilde{Y}_i\right)^2 = SSE^* + \sum_{j=1}^{s}(1-a_j)^2\lambda_j\hat{\gamma}_j^2$$

$$= SSE^* + u_r \qquad (3.1.4)$$

其中$u_r = \sum_{j=1}^{s}(1-a_j)^2\lambda_j\hat{\gamma}_j^2$ 為刪除最後r個自變數後殘差平方和的增量。

於是，考慮是否要刪除最後r個自變數，便須先處理一個如後的數學問題：在滿足$\tilde{\beta}_r = V_2A_s\hat{\gamma}_s = 0$ 的限制條件下，求矩陣A_s使得u_r的值最小。在本文中我們稱滿足被刪除自變數係數為零且使對應的殘差平方和增量u_r為最小的(3.1.2)式估計$\tilde{\beta}$為 Mansfield,Webster, and Gunst 估計，簡稱 M.W.G.估計。

Mansfield et al.(1977)證明了滿足$\tilde{\beta}_r$=0 且使u_r最小的對角矩陣A_s，其對角線上的元素為

$$a_j = 1 - v_{2j}^{'}(V_2\Lambda_s^{-1}V_2^{'})^{-1}V_2\hat{\gamma}_s / \lambda_j\hat{\gamma}_j \quad , \quad j = 1,2,\cdots,s \tag{3.1.5}$$

且

$$\hat{\beta}_{k-r} = V_1(I - \Lambda_s^{-1}V_2^{'}(V_2\Lambda_s^{-1}V_2^{'})^{-1}V_2)\hat{\gamma}_s \tag{3.1.6}$$

$$u_r = \hat{\gamma}_s^{'}V_2^{'}(V_2\Lambda_s^{-1}V_2^{'})^{-1}V_2\hat{\gamma}_s \tag{3.1.7}$$

其中$V_2 = (v_{21},\cdots,v_{2s})$。

當u_r夠大時，我們認為該r個自變數在模型中所能解釋的總變異夠多，所以應保留這些自變數；反之，若u_r不夠大時，表示這些自變數在模型中所能解釋的總變異不夠多，可視為多餘的自變數並將之刪除。

Mansfield,Webster, and Gunst 變數刪除法(以下簡稱 M.W.G.變數刪除法)係依下列步驟進行：

步驟一：當我們要刪除第一個自變數時，必須先計算刪除每個自變數後殘差平方和的增量，稱為u_1，取這k個u_1中的最小值，記為u_{1m}，隨後進行檢定。若仿F-型檢定統計量$F = u_{1m}/MSE$的值夠大時，表示該變數能夠解釋的總變異很大，理當保留在模型中不被刪除，此時本刪除法便告結束。反之，若檢定統計量的值不顯著，表示該自變數能夠解釋的總變異不大，因此可去掉此多餘的變數。檢定統計量F中的$MSE = \sum_i (Y_i - \hat{Y}_i)^2 \big/ (n-k-1)$是完全模型

(2.2.1)下的均方誤(mean squared error)。若步驟一的結果係應刪除某一自變數，則隨即進行步驟二。

步驟二:將剩餘的$k-1$個自變數分別與被刪除的第一個自變數配對計算殘差平方和的增量。亦即在此階段共需計算$k-1$個u_2值，取他們的最小值，記為u_{2m}。此時的仿F-型檢定統計量為

$$F=(u_{2m}-u_{1m})/MSE。$$

當檢定統計量的值夠大時，保留該殘差平方和增量為u_{2m}的自變數並結束刪除法。若檢定統計量F值很小，則刪除該自變數，並隨即進行步驟三。

步驟三:將剩餘的$k-2$個自變數分別與被刪除的第一、二個自變數搭配計算u_3。此階段共需計算$k-2$次u_3值，取他們的最小值，記為u_{3m}。當$F=(u_{3m}-u_{2m})/MSE$的值夠大時，不刪除殘差平方和增量為u_{3m}的自變數並結束刪除法；若仿F-型的檢定統計量F值很小，則刪除該自變數，並隨即進行步驟四。

第$j+1$個步驟時衡量殘差平方和增量顯著與否的仿F-型檢定統計量為：

$$F=(u_{(j+1)m}-u_{jm})/MSE \qquad j=0,1,2,\cdots,l,$$

其中$u_{0m}=0$, l是步驟中止時被刪除自變數的個數。

然而，Mansfield et al.(1977)另外提出了一個改良的方法，該改良法與原方法不同的地方在於每刪除一個變數後，矩陣$X'X$及

$X'Y$ 中對應後的行與列也要隨著被刪除，也就是說以降階後的 $X'X$ 矩陣、$X'Y$ 矩陣取代原來的矩陣。於是在每個步驟中對剩下的變數分別計算殘差平方和增量時，該增量僅與該變數有關而與先前己遭刪除的變數無關。在第 j 個步驟時，仿 F-型的檢定統計量 F 的分子即為殘差平方和增量的最小值。若仍以 u_{jm} 表示此最小值為，則仿 F-型的檢定統計量 F 的計算公式為 $F = u_{jm}/MSE$ 。

在分析幾個實例後，Mansfield et al.發現改良法的效果能令人更滿意。

3.2 M.W.G.估計與廣義壓縮最小平方估計之關連

在本節中我們先探討(3.1.2)式的估計 $\tilde{\beta}$ 和(2.3.1)式廣義壓縮最小平方估計 $\hat{\beta}_{GS}$ 有何種的關連，並由此導引出為何我們會考慮將(3.1.1)或(3.1.2)式中對角矩陣 A_s 的對角線上元素設限的原因。

定義模型(2.1.1)中參數 β 的最廣義脊估計(most generalized ridge estimator)如下:

$$\hat{\beta}_{MGR} \equiv VAV'\hat{\beta} \tag{3.2.1}$$

其中 $\hat{\beta}$ 是 β 的最小平方估計，V 是正交矩陣滿足 $V'(X'X)V = \Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_k)$ ，$A = diag(a_1, a_2, \cdots, a_k)$ 是對角線矩陣，$-\infty < a_j < \infty$ ，$j = 1,2,\cdots,k$ 。

當我們把(3.2.1)式中 A 的對角線元素 a_i 限制在區間[0,1]的範圍內時，即令 a_j 滿足 $0 \le a_j \le 1$，$j=1,2,\cdots,k$ 時，$\hat{\beta}_{MGR}$ 即為王力群(1990)的廣義壓縮最小平方估計 $\hat{\beta}_{GS}$。若我們根據 M.W.G.刪除自變數的理論要保留對應於最大的 s 個特徵值之主成份時，可令對角線矩陣 $A \equiv \begin{pmatrix} \mathrm{A}_s & 0 \\ 0 & A_{k-s} \end{pmatrix}$ 中的 $\mathrm{A}_{k-s}=0$，即可得到(3.1.2)式的估計 $\tilde{\beta}$，其證明過程如下：

$$
\begin{aligned}
\hat{\beta}_{MGR} &= VAV'\hat{\beta} \\
&= \left(V_{s,}V_{k-s}\right)\begin{pmatrix} \mathrm{A}_s & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} V_s' \\ V_{k-s}' \end{pmatrix}\hat{\beta} \\
&= V_s A_s V_s'\hat{\beta} \\
&= V_s A_s \hat{\gamma}_s \\
&= V_s \tilde{\gamma}_s \\
&= \tilde{\beta}
\end{aligned}
$$

而所要刪除的自變數就分別對應了 V_s 中的列向量，也就是 V 中的列向量。例如要刪除最後 r 個自變數，則所對應的列向量即為 V_s 的最後 r 列或 V 的最後 r 列。故求 β 的 M.W.G.估計就是在 $\mathrm{A}_{k-s}=0$ 時，求對角矩陣 A 中的 A_s，使

$$
\tilde{\beta} = \begin{pmatrix} \tilde{\beta}_{k-r} \\ \tilde{\beta}_r \end{pmatrix}
$$

$$
= \begin{pmatrix} V_1 A_s \hat{\gamma}_s \\ V_2 A_s \hat{\gamma}_s \end{pmatrix}
$$

中的 $\tilde{\beta}_r = V_2 A_s \hat{\gamma}_s = 0$，並使對應的殘差平方和增量 u_r 最小。

由此可知，估計 $\tilde{\beta}$ 與 $\hat{\beta}_{GS}$ 都是 $\hat{\beta}_{MGR}$ 的特例。若要增加估計 $\tilde{\beta}$ 的精確度，可將 A_s 對角線上的元素 a_j 限制在[0,1]的範圍內，此時 $\tilde{\beta}$ 就變成廣義壓縮最小平方估計 $\hat{\beta}_{GS}$ 的特例。而本文主要目標就是探討矩陣 A_s 的對角線上元素，在設限與未設限的情況下，對 M.W.G. 刪除自變數的結果會產生怎樣的影響？

3.3 M.W.G.估計之性質

在本節中，我們要探討 M.W.G.估計的一些性質。由 Mansfield et al.(1977)文中的附錄，可求得

$$\tilde{\gamma}_s = A\hat{\gamma}_s = [I - \Lambda_s^{-1}V_2'(V_2\Lambda_s^{-1}V_2')^{-1}V]\hat{\gamma}_s \tag{3.3.1}$$

於是我們可導出 $\tilde{\gamma}_s$ 的期望值及與共變異矩陣分別為

$$E(\tilde{\gamma}_s) = [I - \Lambda_s^{-1}V_2'(V_2\Lambda_s^{-1}V_2')^{-1}V_2]\gamma_s \tag{3.3.2}$$

$$V(\tilde{\gamma}_s) = \sigma^2(\Lambda_s^{-1} - \Lambda_s^{-1}V_2'(V_2\Lambda_s^{-1}V_2')^{-1}V_2\Lambda_s^{-1}) \tag{3.3.3}$$

故 $\tilde{\gamma}_s$ 的偏誤(bias) $E(\tilde{\gamma}_s) - \gamma_s = -\Lambda_s^{-1}V_2'(V_2\Lambda_s^{-1}V_2')^{-1}V_2\gamma_s$。又 $\hat{\gamma}_s$ 的共變異矩陣為 $V(\hat{\gamma}_s) = \sigma^2\Lambda_s^{-1}$，且 $\Lambda_s^{-1}V_2'(V_2\Lambda_s^{-1}V_2')^{-1}V_2\Lambda_s^{-1}$ 係一正定矩陣，因此 $V(\tilde{\gamma}_s) \le V(\hat{\gamma}_s)$，即 M.W.G.估計的精確度確是優於最小平方估計的精確度，式(3.3.1)至(3.3.3)的證明請參閱附錄 A。

在一般線性迴歸模型中，最小平方估計具備下面兩個性質：

性質 1：殘差和等於零。

性質 2：殘差向量與配適值向量正交

至於 M.W.G.估計是否也具備了上述的兩個性質呢?在本節中我們將證實 M.W.G.估計的確也滿足了這兩個性質。

假設依據 M.W.G.刪除變數法所建立起的模型其殘差向量為 $\widetilde{e} \equiv \mathrm{Y} - \overline{\mathrm{Y}} \cdot 1 - X\widetilde{\beta}$ 。由於矩陣 X 的行向量已標準化，利用行和為零的條件，可容易證得 $1'\widetilde{e} = 1'(\mathrm{Y} - \overline{\mathrm{Y}} \cdot 1 - X\widetilde{\beta}) = 0$，故 M.W.G.估計滿足性質 1。

接下來我們不直接而從另一角度證明 M.W.G.估計也滿足性質 2。前面曾提到，若欲刪除的是最後 r 個自變數，則求 M.W.G.估計的問題就是要找對角矩陣 A_s 使得(3.1.2)式中的 $\widetilde{\beta}_r = V_2 A_s \hat{\gamma}_s = 0$，並使(3.1.4)式中的 $u_r = \sum_{j=1}^{s} (1 - a_j)^2 \lambda_j \hat{\gamma}_j^2$ 部份最小。當我們進一步希望 M.W.G.估計也具備性質 2 時，則(3.1.2)式中的 $\widetilde{\beta}$ 除了要滿足 $\widetilde{\beta}_r = 0$ 外，也要滿足

$$\hat{\gamma}_s^{'}(A_s\Lambda_s - A_s\Lambda_s A_s)\hat{\gamma}_s = 0 \tag{3.3.4}$$

(3.3.4)式的推導請參閱附錄 B。

現考慮下列的最佳化問題：

$$\begin{aligned} \text{Min} \quad & u_r = \sum_{j=1}^{s} (1 - a_j)^2 \lambda_j \hat{\gamma}_j^2 \\ \text{s.t.} \quad & V_2 A_s \hat{\gamma}_s = 0 \\ & \hat{\gamma}_s^{'}(A_s\Lambda_s - A_s\Lambda_s A_s)\hat{\gamma}_s = 0 \end{aligned} \tag{3.3.5}$$

令我們驚訝的是(3.3.5)式的最適解 a_j 正好與(3.1.5)式一樣，其

推導過程有些複雜，詳見附錄 C。因此，我們便證實了 M.W.G. 估計也具備性質 2 這個結論。

在只刪除一個自變數的情形下，即 $r=1$ 時，我們可以直接證明(3.1.5)式中的 a_j 確實滿足(3.3.4)式，其過程如下：首先(3.3.4)式的左端可改寫成 $\sum_{j=1}^{k} a_j(1-a_j)\lambda_j\hat{\gamma}_j^2$ 。又當 $r=1$ 時，(3.1.5)式中的 a_j 可表示成

$$a_j = 1 - V_{kj}\,(\sum_{j=1}^{s} V_{kj}^2/\lambda_j)^{-1}\,(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)/\lambda_j\hat{\gamma}_j$$

其中 $V_k^{'} = (V_{k1}, V_{k2}, \cdots, V_{ks})$ 是 V_s 的最後一列。

故

$$\begin{aligned}
&\sum_{j=1}^{s} a_j(1-a_j)\lambda_j\hat{\gamma}_j^2 \\
&= \sum_{j=1}^{s}\lambda_j\,[1 - V_{kj}\,(\sum_{j=1}^{s} V_{kj}^2/\lambda_j)^{-1}(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)/\lambda_j\hat{\gamma}_j]\,[V_{kj}\,(\sum_{j=1}^{s} V_{kj}^2/\lambda_j)^{-1}(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)/\lambda_j\hat{\gamma}_j]\,\hat{\gamma}_j^2 \\
&= \sum_{j=1}^{s}[V_{kj}(\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-1}(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)]\hat{\gamma}_j - \sum_{j=1}^{s}[V_{kj}^2(\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-2}(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)^2]/\lambda_j \\
&= (\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-1}\,(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)^2 - (\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-2}\,(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)^2\,(\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j) \\
&= (\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-1}\,(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)^2 - (\sum_{j=1}^{s} V_{kj}^2\big/\lambda_j)^{-1}\,(\sum_{j=1}^{s} V_{kj}\hat{\gamma}_j)^2 \\
&= 0
\end{aligned}$$

3.4　加上限制條件的 M.W.G.變數刪除法

在 3.1 節裡，(3.1.1)式 $\tilde{\gamma}_s = A_s\hat{\gamma}_s$ 或(3.1.2)式 $\tilde{\beta}=V_s\tilde{\gamma}_s=V_sA_s\hat{\gamma}_s$ 的 M.W.G.估計中, 對角線矩陣 A_s 的對角線上元素 a_j 並無任何限制，因此採用這種估計時其變異數有可能會膨脹的很大。在 3.2 節裡，比較了 M.W.G.估計與廣義壓縮最小平方估計後，我們發現兩者的差別僅在於對(3.2.1)式最廣義脊估計 $\hat{\beta}_{MGR}$ 中 A 的元素如何加以設限而已。因此，M.W.G.估計以及廣義壓縮最小平方估計均係最廣義脊估計的特例。

本文主要的構想，即是把廣義壓縮最小平方估計中，將對角線上元素 a_j 設限在[0,1]範圍裡的概念引用在 M.W.G.估計上，並探討如此設限後對整個變數刪除的過程與結果會產生何種影響?

現在將(3.2.1)式 $\hat{\beta}_{MGR}=VAV'\hat{\beta}$ 中對角線矩陣 A 的對角線上元素限制在區間[0,1]裡。多加上這個限制條件後的 M.W.G.變數刪除法其步驟和未加上時相仿，進行方式如下:

步驟一: 每次退出一個自變數後，在限制條件下，求出殘差平方和的增量。令這 k 個增量的最小值為 u_{1m}^* 。隨後根據仿 F -型檢定統計量 $F=u_{1m}^*/MSE$ 值的大小決定對應的變數是否要保留或刪除。當 F 的值足夠大時，表示該變數能解釋

的總變異夠多，理當保留在模型中。此時本刪除法便宣告結束。反之，若F的值不足夠大，表示該變數能解釋的總變異不夠多，可以刪除此多餘變數。接著進行步驟二。

步驟二：將剩餘的$k-1$個變數分別和步驟一中被刪除的變數配對。每次退出一對，並在限制條件下計算殘差平方和的增量。令這些增量的最小值為u_{2m}^*。若此時仿F-型統計量$F=(u_{2m}^*-u_{1m}^*)/MSE$的值足夠大，則保留對應增量為$u_{2m}^*$的變數，並結束本刪除法，反之，若$F$的值不足夠大，則刪除掉該變數，並進行步驟三。

一般而言，在第$j+1$個步驟時，衡量殘差平方和增量是否顯著的仿F-型檢定統計量為:

$$F=(u_{(j+1),m}^*-u_{jm}^*)/MSE \text{ ， } j=0,1,2,\cdots,l \text{，}$$

其中$u_{0m}=0$， l 為被刪除變數的個數，而MSE為 3.1 節中所定義的最小平方估計之均方誤。

在上述的變數刪除過程中，有涉及到在某些限制條件下殘差平方和增量的計算問題，也就是說要解決下類的數學問題：

$$\min \quad u_r = \sum_{j=1}^{s}(1-a_j)^2\lambda_j\hat{\gamma}_j^2 \qquad (3.4.1)$$

$$\text{s.t.} \quad \sum_{j=1}^{s} a_j V_{tj}\hat{\gamma}_j = 0 \quad , \; t=k-r+1,\cdots,k \qquad (3.4.2)$$

$$0 \le a_j \le 1 \text{ , } j=1,2,\cdots,s \qquad (3.4.3)$$

其中V_{tj}是矩陣V_2中的元素，s是被保留對應較大特徵值的主成份個數，r是被刪除變數的個數。對這個求極值的問題，在此提出兩種處理的方法。方法一:先不管限制式(3.4.3)，在僅考慮(3.4.2)式的限制下，利用 Lagrange 乘數法求出使(3.4.1)式最小的解a_j，如(3.1.5)式所示。然後再利用 Rao-Ghangurde 修正法修正a_j的解使他們符合限制式(3.4.3)。Rao-Ghangurde 的修正過程為:若解出的a_j全部在[0,1]的範圍內，則這一組a_j即符合所求，再將他們代入(3.4.1)式中就可得到殘差平方和的增量u_r；如果解出的a_j並非全部在[0,1]的範圍內，則做如下的修正。當$a_j>1$ 時，則令$a_j=1$；當$a_j<0$ 時，則令$a_j=0$。將這些已被修正過的a_j代入(3.4.1)式與(3.4.2)式後，可得更新過的式子。對更新後的(3.4.1)式與(3.4.2)式再次以 Lagrange 乘數法求解。若所求得的解a_j都在[0,1]的範圍內，則這組解就符合所求；否則需要再做修正。這過程一直繼續下去，直到由 Lagrange 乘數法所求得的解a_j全都在[0,1]範圍內才告結束。方法二:先利用 LINGO 軟體求a_j的解，再搭配 MATLAB 軟體求殘差平方和增量。在第四章分析實例時，我們將採用方法二。

在比較 3.1 節與本節刪除變數的過程後，可察覺到矩陣A中對角線上的元素a_j是否有設限確實會影響到殘差平方和的增量，因此在變數刪除上可能產生不同的效果。一般來說，在考慮設

限後，變數被刪除的個數會較不設限時少，其原因為：在步驟一中，設限時的最小殘差平方和增量常常遠大於不設限時的殘差平方和增量，因此造成設限時的仿F-型檢定統計量F之值是顯著的，而不設限時的仿F-型檢定統計量F之值是不顯著的。於是，設限將使變數刪除的動作較為保守。

3.5 影響自變數刪除的三個主要因素

在利用 Mansfield et al.(1977)所提出的方法進行變數刪除時，被刪除變數的順序與最終被保留下來的變數通常會受到下面三個因素的影響。

一、主成分個數的決定。

二、在刪除某個變數後，$X'X$ 與 $X'Y$ 中對應的行與列是否也跟著刪除?即採用前面所提到的改良法。

三、是否將對角線矩陣A中對角線上的元素a_j設限在[0,1]的範圍內?

通常在進行主成分分析時，對應於較小特徵值的主成分所能解釋的變異不大，因此常被剔除掉，故主成分的個數大多會小於變數的個數。在第四章我們將以實例說明剔除或不剔除對應於較小的特徵值之主成份對於變數刪除的過程與結果會產生怎樣的影響。

在 3.1 節中已經說明過 M.W.G.刪除變數的原方法和改良法在過程上有何不同，在第四章中我們要更具體的從實例中觀察採用原方法和改良法在最後所獲致的結果有何差異。

在 3.4 節中已經解釋了考慮對矩陣A之對角線上元素a_j設限的理由，同時也了解到不可能導出a_j的公式解。在分析實例時，我們將先用 LINGO 軟體求出a_j之值，以便繼續進行變數的刪除過程。從理論的觀點來說，對a_j設限將會導致變數刪除過程較為保守，在下一章的實例分析結果中的確呈現出這樣的現象。