# Chapter 14
# 彷彿無相關迴歸模型

彷彿無相關迴歸模型(seemingly unrelated regression model, 簡稱 SUR 模型, 或稱爲 SURE 模型, 代表 seemingly unrelated rgression equations 的縮寫) 是 Arnold Zellner於1962年所提出, 與追蹤資料模型 (panel data model) 都是屬於同時考慮資料的橫斷面與時間序列特性的「系統方程式模型 (system of equations)」。假若 $N$ 爲方程式 (或個別模型) 的數目, $T$ 爲觀察值數目 (且爲時間序列資料)。SUR模型基本上是處理時間序列長度"大於"橫斷面 (即模型數目) 的情形 (即 $T > N$ ), 而追蹤資料模型則是處理時間序列長度"小於"橫斷面 (即模型數目) 的情形 (即 $T < N$ )。在本章中, 我們考慮的是前者。我們也將探討 SUR 模型在 CAPM 檢定上的應用。

## 14.1 模型設定

考慮 $N$ 個方程式如下:

$$y_{it} = x_{it}'\beta_i + \varepsilon_{it}, \qquad i = 1, \cdots, N; t = 1, \cdots, T. \tag{1}$$

上式亦可表達爲:

$$y_i = x_i\beta_i + \varepsilon_i, \qquad i = 1, \cdots, N$$

其中 $y_i$ 爲 $(T \times 1)$ 的向量, $x_i$ 爲 $(T \times k_i)$ , $\beta_i$ 爲 $(k_i \times 1)$ 的模型參數。令 $k \equiv \sum_{i=1}^{N} k_i$ 爲模型的總參數數目。假若各個模型間的誤差項沒有相關性, 亦即

$$E(\varepsilon_{is}\varepsilon_{jt}) = 0, \qquad \forall\ i \neq j,$$

則以 OLS 估計各個迴歸模型即可。但是若各個模型的誤差項間有相關呢, 例如, $E(\varepsilon_{it}\varepsilon_{jt}) \neq 0$ ? 概念上, 若各方程式間的誤差項存在相關, 表示有額外的資訊可

提供我們在估計上能更有效率, 我們可將模型推疊如下:

$$\begin{aligned} y_1 &= x_1\beta_1 + \varepsilon_1 \\ y_2 &= x_2\beta_2 + \varepsilon_2 \\ &\vdots \\ y_N &= x_N\beta_N + \varepsilon_N \end{aligned}$$

我們可進一步將其合併成一"大"模型如下:

$$Y = X\beta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \Omega),$$

其中

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}_{TN\times 1}, X = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & x_N \end{pmatrix}_{TN\times k},$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix}_{k\times 1}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}_{TN\times 1}$$

且 $k \equiv \sum_i k_i$。上式中, $X$ 是一個"區塊對角"矩陣 (block-diagonal matrix)。Zellner(1962) 假設各模型誤差項間只有存在同期相關 (contemporaneous correlation), 亦即

$$E(\varepsilon_{is}\varepsilon_{jt}) = \begin{cases} \sigma_{ij} & \text{若 } s = t\,, \\ 0 & \text{其他。} \end{cases}$$

在此假設下, 我們可證明聯合的誤差項 $\varepsilon$ 的共變異數矩陣如下:

$$
\begin{aligned}
\Omega &= E(\varepsilon\varepsilon') \\
&= E\left(\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{pmatrix}\begin{pmatrix} \varepsilon_1' & \cdots & \varepsilon_N' \end{pmatrix}\right) \\
&= E\begin{pmatrix} \varepsilon_1\varepsilon_1' & \varepsilon_1\varepsilon_2' & \cdots & \varepsilon_1\varepsilon_N' \\ \varepsilon_2\varepsilon_1' & \varepsilon_2\varepsilon_2' & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_N\varepsilon_1' & \cdots & \cdots & \varepsilon_N\varepsilon_N' \end{pmatrix} \\
&= \begin{pmatrix} \sigma_1^2 I_T & \sigma_{12} I_T & \cdots & \sigma_{1N} I_T \\ \sigma_{21} I_T & \cdots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \sigma_{N1} I_T & \cdots & \cdots & \sigma_N^2 I_T \end{pmatrix} \\
&= \Sigma \otimes I_T,
\end{aligned}
$$

其中 $\Sigma$ 如下:

$$
\Sigma \equiv \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \cdots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & \cdots & \sigma_N^2 \end{pmatrix}
$$

前式第四個等式的結果是因為 $E(\varepsilon_i \varepsilon_j) = \sigma_{ij} I_T$ , 證明如下:

$$
\begin{aligned}
E(\varepsilon_i \varepsilon_j') &= E\left( \begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iT} \end{pmatrix} \begin{pmatrix} \varepsilon_{j1} & \cdots & \varepsilon_{jT} \end{pmatrix} \right) \\
&= E \begin{pmatrix} \varepsilon_{i1}\varepsilon_{j1} & \varepsilon_{i1}\varepsilon_{j2} & \cdots & \varepsilon_{i1}\varepsilon_{jT} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{iT}\varepsilon_{j1} & \cdots & \cdots & \varepsilon_{iT}\varepsilon_{jT} \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{ij} & 0 & \cdots & 0 \\ 0 & \sigma_{ij} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_{ij} \end{pmatrix} \\
&= \sigma_{ij} I_T.
\end{aligned}
$$

非對角線的值為零是因為我們假設非同期間之相關性為零。由以上的分析可知, $\Omega$ 非為對角矩陣, 而對角的變異數也不盡相同 (亦即 $\Omega \neq \sigma^2 I_T$ ), 因此以OLS 估計此一模型是沒有效率的。不過, 由於在此, $\Omega$ 的結構為已知, 因此我們可以以 FGLS 來估計之。其估計子為:

$$
\begin{aligned}
\hat{\beta}_{GLS} &= (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}Y \\
&= (X'(\hat{\Sigma}^{-1} \otimes I_T)X)^{-1}X'(\hat{\Sigma}^{-1} \otimes I_T)Y \qquad (2)
\end{aligned}
$$

上式中, $\hat{\Sigma}$ 為 $\Sigma$ 的樣本估計子, 其各元素可估計如下:

$$\hat{\sigma}_{ij} = \frac{1}{T}\hat{\varepsilon}_i'\hat{\varepsilon}_j = \frac{1}{T}\sum_{t=1}^{T}\hat{\varepsilon}_{it}\hat{\varepsilon}_{jt}$$

其中 $\hat{\varepsilon}_i, \hat{\varepsilon}_j$ 為以OLS 估計各子模型所得到的殘差項。上式估計子未經自由度調整, 一些學者建議可調整如下:

$$\hat{\sigma}_{ij} = \frac{\hat{\varepsilon}_i'\hat{\varepsilon}_j}{\sqrt{(T-k_i)(T-k_j)}}$$

或是

$$\hat{\sigma}_{ij} = \frac{\hat{\varepsilon}_i\hat{\varepsilon}_j}{T - max(k_i, k_j)}$$

在文獻上,(2) 式稱爲 Zellner 的 SUR 估計子 (Zellner's seemingly unrelated regression estimator), 而此一模型即稱爲 SUR 模型。

Zellner提出的是我們先前所提的二階段估計子。事實上, 我們也可以利用重覆的估計子, 得到所謂的重覆 SUR 估計子 (iterated SUR estimator), 而此一估計子會與假設常態分配下的最大概似法 (MLE) 估計子相同。如同我們先前提過的, 在此不論是 Zellner 的估計子或是重覆估計子, 其大樣本性質皆相同, 因此沒有孰優孰劣的問題。

另外, 由以上的分析可知, 當不同模型間誤差項間的相關性愈高時, 以 SUR 來估計的效率會愈高。另外, 由於在估計的過程中需要估計 $\Sigma$, 而 $\Sigma$ 中總共有 $N(N+1)/2$ 的參數, 也就是當模型的數目過多時, SUR估計的效率會因爲估計太多參數所造成的自由度喪失而下降。因此, 一般會要求 $T$ 要遠大於 $N$。

## 14.2 假說檢定

由於 SUR 估計子是一 FGLS 估計子, 因此其假說檢定僅能以 $Z$ 或是 $\chi^2$ 檢定。不過, 文獻中似乎發現以 $F$ 檢定較保守 (亦即比較不過有過度拒絕虛無假說的情況發生), 因此部份套裝軟體也會顯示 $F$ 檢定的結果。其作法大致如下:

假若我們有興趣的線性假說如下:

$$H_0 : R\beta = q,$$

其中 $q$ 爲 $(J \times 1)$ 的向量。根據 FGLS 的結果, 我們知道

$$\hat{\beta}_{SUR} \overset{a}{\sim} \mathcal{N}(\beta, (X'\Omega^{-1}X)^{-1}),$$

$$R\hat{\beta}_{SUR} \overset{a}{\sim} \mathcal{N}(R\beta, R(X'\Omega^{-1}X)^{-1}R'),$$

因此以下的二次式有 $\chi^2$ 的漸近分配

$$(R\hat{\beta} - q)'(R(X'\hat{\Omega}^{-1}X)^{-1}R')^{-1}(R\hat{\beta} - q) \overset{a}{\sim} \chi^2_J, \tag{3}$$

另外, 因爲 $\varepsilon \sim \mathcal{N}(0, \Omega)$, 因此

$$\hat{\varepsilon}'\hat{\Omega}^{-1}\hat{\varepsilon} \overset{a}{\sim} \chi^2_{NT-k}$$

亦有卡方的近似分配,將以上二式相除,並除以各自之自由度,得一"近似"(approximate)的 $F$ 分配:

$$\hat{F} = \frac{(R\hat{\beta}-q)'(R(X'\hat{\Omega}^{-1}X)^{-1}R')^{-1}(R\hat{\beta}-q)/J}{\hat{\varepsilon}'\hat{\Omega}^{-1}\hat{\varepsilon}/(NT-k)} \sim F_{J,NT-k}$$

在樣本趨近於無窮大時, 我們知道 $\hat{F}$ 的分母會趨近於1。這是因爲

1. $\lim_{T\to\infty} E(\frac{\hat{\varepsilon}'\hat{\Omega}^{-1}\hat{\varepsilon}}{NT-k}) = 1$ .
2. $\lim_{T\to\infty} var(\frac{\hat{\varepsilon}'\hat{\Omega}^{-1}\hat{\varepsilon}}{NT-k}) = 0$ .

因此, $\lim(\frac{\varepsilon'\hat{\Omega}^{-1}\hat{\varepsilon}}{NT-k}) = 1$ , 利用 Slutsky 定理, 我們可知

$$\text{plim } J\hat{F} = \text{plim } (R\hat{\beta}-q)'(R(X'\hat{\Omega}^{-1}X)^{-1}R')^{-1}(R\hat{\beta}-q).$$

根據 (3) 式的結果, 可知 $J\hat{F}$ 會有 $\chi^2_J$ 的漸近分配。$\hat{F}$ 與 $J\hat{F}$ 即一般軟體會顯示的二種主要的檢定。

**例 14.1.** 我們檢定匯率曝險。[1]

本表報告下列模型的估計結果:

$$R_{it} = a_i + b_i R_{mt} + \gamma_i X_{it} + e_{it} \quad \forall\ i,\ t,$$

其中 $X_{it}$ 爲國家 $i$ 以貿易加權的匯率。$\gamma_i = 0$ 的聯合檢定是以 $F$ 檢定。

[1]這個例子取自 Chou, Chou and Shyy (2002), Capital Market Integration and Exchange Risk Exposure of the Asian Emerging Markets, 台灣管理學刊 1:2, 165-182。

表 14.1: 已開發國家與亞洲新興國家的匯率風險暴露: SUR 估計

本表報告下列模型的估計結果:

$$R_{it} = a_i + b_i R_{mt} + \gamma_i ex_{it} + e_{it} \quad \forall\ i,\ t,$$

其中 $R_{it}$ 爲國家 $i$ 在時點 $t$ 的月報酬; $ex_{it}$ 爲國家 $i$ 在時點 $t$ 以貿易加權的匯率。$H_0 : \gamma_i = 0\ \forall\ i \in$ Group 的聯合檢定是以 $F$ 檢定。Group 包括 AEMs, OECD, 以及 AEMs+OECD。樣本期間爲 1985 年 3 月至 1996 年 5 月,

| | 全期間 | | 1985.3-1990.12 | | 1991.1-1996.5 | |
|---|---|---|---|---|---|---|
| 國家 | $\gamma$ | (P-value) | $\gamma$ | (P-value) | $\gamma$ | (P-value) |
| Korea | 0.0303 | (0.9197) | -0.1495 | (0.7281) | 0.4284 | (0.2409) |
| Malaysia | 0.2647 | (0.2230) | 0.0196 | (0.9476) | 0.3804 | (0.0934) |
| Philippines | -0.0761 | (0.7794) | 0.1598 | (0.6949) | -0.2218 | (0.3461) |
| Taiwan | -1.2287 | (0.0149) | -2.9873 | (0.0001) | -0.2085 | (0.6812) |
| Thailand | 0.1322 | (0.5433) | 0.0560 | (0.8125) | 0.3499 | (0.1976) |
| Australia | -0.1859 | (0.0945) | -0.5869 | (0.0048) | -0.1088 | (0.1901) |
| Belgium | -0.0363 | (0.4016) | -0.0518 | (0.2943) | 0.3838 | (0.0053) |
| Denmark | -0.8123 | (0.0001) | -1.5032 | (0.0001) | -0.3276 | (0.0519) |
| Germany | 0.0888 | (0.3927) | 0.0335 | (0.8855) | 0.2474 | (0.0015) |
| France | 0.1879 | (0.1227) | 0.0975 | (0.6466) | 0.2400 | (0.0142) |
| U. K. | -0.5167 | (0.0001) | -0.5860 | (0.0009) | -0.4793 | (0.0001) |
| H. K. | 0.2798 | (0.2890) | 0.7340 | (0.0995) | 0.3223 | (0.2508) |
| Italy | -0.2952 | (0.0995) | 0.0299 | (0.8961) | -0.6531 | (0.0077) |
| Japan | -0.6963 | (0.0001) | -0.9206 | (0.0001) | -0.5606 | (0.0149) |
| Canada | -0.7139 | (0.0001) | -1.2436 | (0.0001) | -0.6265 | (0.0001) |
| Netherlands | 0.1136 | (0.0930) | 0.0790 | (0.4537) | 0.3119 | (0.0001) |
| Norway | -0.1949 | (0.1154) | 0.0115 | (0.9576) | -0.0764 | (0.5172) |
| Austria | -0.2446 | (0.5609) | 0.3631 | (0.7063) | -0.3185 | (0.2297) |
| Sweden | -0.2673 | (0.0742) | -0.1250 | (0.5810) | -0.2636 | (0.1238) |
| Swiss | 0.1664 | (0.1160) | 0.2951 | (0.0698) | 0.2573 | (0.0420) |
| Spain | -0.1474 | (0.2906) | -0.1071 | (0.5911) | -0.3739 | (0.0070) |
| USA | -0.3224 | (0.0001) | -0.4548 | (0.0001) | -0.1920 | (0.0023) |

| Hypothesis | F stat. | P-value | F stat. | P-value | F stat. | P-value |
|---|---|---|---|---|---|---|
| $H_0 : \gamma_i = 0$ $\forall\ i \in$ AEMs | 1.7388 | (0.1223) | 3.8121 | (0.0020) | 1.3083 | (0.2578) |
| $H_0 : \gamma_i = 0$ $\forall\ i \in$ OECD17 | 8.4222 | (0.0001) | 8.4470 | (0.0001) | 8.6883 | (0.0001) |
| $H_0 : \gamma_i = 0$ $\forall\ i \in$ AEMs and OECD17 | 6.9253 | (0.0001) | 7.4539 | (0.0001) | 7.0271 | (0.0001) |

## 14.3 SUR 估計子與 OLS 估計子相等的二種情況

### 14.3.1 各模型間誤差項不存在同期相關下

當 $\sigma_{ij} = 0$ , $\forall\ i \neq j$ 時, $\Sigma$ 變成一對角矩陣, 亦即

$$\Sigma = diag(\sigma_1^2, \cdots, \sigma_N^2)$$

這時,SUR 估計子會變成與 OLS 估計子相同, 證明如下:

$$\begin{aligned}
\hat{\beta}_{SUR} &= (X'(\Sigma^{-1} \otimes I_T)X)^{-1} X'(\Sigma^{-1} \otimes I_T)^{-1} Y \\
&= \left[ \begin{pmatrix} x_1' & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_N' \end{pmatrix} \begin{pmatrix} \sigma_1^{-2} I_T & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_N^{-2} I_T \end{pmatrix} \begin{pmatrix} x_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_N \end{pmatrix} \right]^{-1} \\
&\times \begin{pmatrix} x_1' & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_N' \end{pmatrix} \begin{pmatrix} \sigma_1^{-2} I_T & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_N^{-2} I_T \end{pmatrix} \\
&= \begin{pmatrix} \sigma_1^{-2}(x_1'x_1)^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_N^{-2}(x_N'x_N)^{-1} \end{pmatrix} \begin{pmatrix} \sigma_1^{-2} x_1' y_1 \\ \vdots \\ \sigma_N^{-2} x_N' y_N \end{pmatrix} \\
&= \begin{pmatrix} (x_1'x_1)^{-1} x_1' y_1 \\ \vdots \\ (x_N'x_N)^{-1} x_N' y_N \end{pmatrix} \\
&= \begin{pmatrix} \hat{\beta}_{1,OLS} \\ \vdots \\ \hat{\beta}_{N,OLS} \end{pmatrix}
\end{aligned}$$

### 14.3.2 當各模型的解釋變數相同時

當 $x_1 = x_2 = \cdots = x_N = Z$ 時, $X = I_N \otimes Z$ , SUR 估計子也會變成與 OLS 估計子相

同, 證明如下:

$$\begin{aligned}
\hat{\beta}_{SUR} &= (X'(\Sigma \otimes I_T)^{-1}X)^{-1}X'(\Sigma \otimes I_T)^{-1}Y \\
&= ((I_N \otimes Z)'(\Sigma^{-1} \otimes I_T)(I_N \otimes Z))^{-1}(I_N \otimes Z)'(\Sigma^{-1} \otimes I_T)Y \\
&= (\Sigma^{-1} \otimes (Z'Z))^{-1}(\Sigma^{-1} \otimes Z')Y \\
&= (\Sigma \otimes (Z'Z)^{-1})(\Sigma^{-1} \otimes Z')Y \\
&= (I_N \otimes (Z'Z)^{-1}Z')Y \\
&= \begin{pmatrix} (Z'Z)^{-1}Z' & 0 & 0 & 0 \\ 0 & (Z'Z)^{-1}Z' & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & (Z'Z)^{-1}Z' \end{pmatrix} \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix} \\
&= \begin{pmatrix} (Z'Z)^{-1}Z'Y_1 \\ \vdots \\ (Z'Z)^{-1}Z'Y_N \end{pmatrix} \\
&= \begin{pmatrix} \hat{\beta}_{1,OLS} \\ \vdots \\ \hat{\beta}_{N,OLS} \end{pmatrix}
\end{aligned}$$

儘管在這樣的情況下,SUR 估計子與 OLS 估計子相等, 但這並不表示 SUR 的分析無效。基本上, 的確在此我們知道將各模型合併並不能進一步改善參數估計上的效率。但是, 在假說檢定上, 尤其是在"跨模型係數 (cross-equation restrictions)"檢定上, 我們仍有必要考慮模型誤差項間的相關性。

由模型的設定, 我們知道在此 $\hat{\beta}_{SUR} = \hat{\beta}_{OLS}$ 的共變異數矩陣爲:

$$\begin{aligned}
cov(\hat{\beta}_{OLS}) &= (X'\Omega^{-1}X)^{-1} \\
&= [(I_N \otimes Z)'(\Sigma^{-1} \otimes I_T)(I_N \otimes Z)]^{-1} \\
&= (\Sigma^{-1} \otimes Z'Z)^{-1} \\
&= \Sigma \otimes (Z'Z)^{-1}
\end{aligned}$$

這表示對各模型的參數估計子而言, 其共變異矩陣與分別估計時仍是相同的:

$$cov(\hat{\beta}_{i,OLS}) = \sigma_i^2(Z'Z)^{-1},$$

其中 $\sigma_i^2$ 爲 $\Sigma$ 的第 $(i,i)$ 個元素。因此, 如果我們只對個別模型的參數之檢定, 並無需將模型合併來估計, 也不必考慮其間之相關性。但在跨模型係數檢定上, 仍須考慮此一相關性, 作法則與前節相同。比較特殊的是在一些情況下, 檢定統計量在小樣本下服從 F 分配, 可以"exact F test"檢定。

本節討論的解釋變數相同 (或稱爲相同迴歸元 (identical regressors)) 的情形, 在統計文獻上亦稱爲多變量迴歸模型 (multivariate regression model, 簡稱 MVRM)。[2]

## 14.4 MVRM 之應用: CAPM 檢定

令 $r_{it} = R_{it} - R_{ft}, r_{mt} = R_{mt} - R_{ft}$ 分別代表資產 $i$ 與市場投資組合的超額報酬 (excess return), 其中 $R_{it}$ , $R_{mt}$ , $R_{ft}$ 分別代表資產 $i$ 、市場投資組合、無風險資在時點 $t$ 的報酬。考慮市場模型如下:

$$r_{it} = \alpha_i + \beta_i r_{mt} + \varepsilon_{it}, \qquad i = 1,\cdots,N; t = 1,\cdots,T. \tag{4}$$

上式可寫爲:

$$r_i = \bar{x}\beta_i + \varepsilon_i \tag{5}$$

其中 $r_i$ 爲 $(T\times 1)$ 的向量, $\bar{x}$ 爲 $(T\times 2)$ 的矩陣, $\beta_i$ 爲 $(2\times 1)$ 的向量, $\varepsilon_i$ 爲 $(T\times 1)$ 的誤差向量。

在 CAPM 成立下, 我們知道如下關係:

$$E(r_i) = \beta_i E(r_m) \qquad \forall_i,$$

將(4) 式兩邊取期望值, 得

$$E(r_{it}) = \alpha_i + \beta_i E(r_{mt}),$$

[2] 文獻上常有研究將 multiple regression model 與 multivariate regression model 二個名詞混爲一用。在實證上, 因爲單變量分析常以 univariate analysis 表示, 所以在多變數的分析自然就以 multivariate analysis 來表示。不過, 在迴歸分析中,「單變量」或「多變量」係針對被解釋變數的個數 (也就是方程式的數目) 而言。模型中「解釋變數」的多寡則分爲「簡單迴歸」與「複迴歸」。這二種差異, 讀者應有所分別。

比較上式與 (4) 式, 我們知道 CAPM 成立下,(4) 式的截距項應爲 0, 亦即我們可檢定以下的虛無假說:

$$H_0 : \alpha_i = 0 \quad \forall\ i$$

我們可進一步將 (5) 式推疊成 SUR 模型如下:

$$\begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix} = \begin{bmatrix} \bar{x} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \bar{x} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}, \tag{6}$$

亦即

$$R = XB + \varepsilon \tag{7}$$

其中 $R$ 爲 $(NT \times 1)$ , $X$ 爲 $(NT \times 2N)$ , $B$ 爲 $(2N \times 1)$ , $\varepsilon$ 爲 $(NT \times 1)$ 。在 iid 常態假設下, 可知

$$\varepsilon \sim \mathcal{N}(0, \Sigma \otimes I_T)$$

(7) 式可以 OLS 估計之, 其分配如下:

$$\hat{B}_{OLS} = (X'X)^{-1}X'R \sim \mathcal{N}(B, \Sigma \otimes (\bar{X}'\bar{X})^{-1})$$

令 $\hat{\varepsilon}$ 爲 OLS 殘差項, 即

$$\hat{\varepsilon} = R - X\hat{B}_{OLS},$$

則 $\Sigma$ 可以估計如下:

$$\hat{\Sigma} = \frac{1}{T-2}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$$

進一步, 可知 $(T-2)\hat{\Sigma}$ 會服從一 Wishart 分配, 表達爲:

$$A \equiv (T-2)\hat{\Sigma} \sim W_N(T-2, \Sigma)$$

Wishart分配基本上類似卡方分配的多變量一般式 (multivariate generalization)。例如, 當 $N = 1$ 時, $A$ 即爲殘差平方合:

$$A = (T-2)\hat{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon}$$

而我們知道 $A/\sigma^2$ 服從卡方分配, 自由度爲 $T-2$ 。Wishart 分配與卡方分配不同的是 Wishart 分配不只是考慮變異數估計子 $\hat{\sigma}_i^2$ 個別的分配, 還同時考慮共變異數估計子 $\hat{\sigma}_{ij}$ 的分配。在此,Wishart 分配所描述的是所有的變異數與共變異數估計子的聯合分配。

在更進一步推導 $H_0 : \alpha_i = 0$ 的檢定之前, 我們先引用一個定理如下:

**定理 14.1.** 若 $X \sim \mathcal{N}(\mu, \Omega)$ , 且 $A \sim W_m(n, \Omega)$, $n \geq m$ , 且 $X$ 與 $A$ 獨立, 則下式有非中央 $F$ 分配 *(noncentral F distribution)*:[3]

$$\frac{(n-m+1)}{m} X' A^{-1} X \sim F_{m,n-m+1}(\lambda)$$

其中 $\lambda = \mu' \Sigma^{-1} \mu$ 爲非中央參數。

證明. 參見 Muirhead (1982), 第211-212頁或第98-99頁。 ♠

以上之統計量, 在多變量統計裡稱爲 Hotelling's $T^2$ 。令 $H = I_N \otimes (1 \quad 0)'$ , 則 $H_0 : \alpha_i = 0,\ \forall_i$ 可表達爲

$$H_0 : HB = 0,$$

因爲

$$\begin{aligned} HB &= \left( I_N \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \vdots \\ \alpha_N \\ \beta_N \end{pmatrix} \\ &= \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{pmatrix} \\ &\equiv \alpha \end{aligned}$$

[3] 若 $x \sim \mathcal{N}(\mu, I_n)$ , 則 $x^2$ 的分配即爲非中央卡方分配 (noncentral chi-square), 表達爲 $x^2 \sim \chi_n^2(\lambda)$ , 其中 $\lambda = \mu' \mu$ 爲此分配的非中央參數 (noncentrality parameter)。進一步, 如果 $x_1 \sim \chi_m^2(\lambda)$ , $x_2 \sim \chi_n^2$ , 且 $x_1$ 與 $x_2$ 互相獨立, 則 $f = \frac{x_1/m}{x_2/n}$ 的分配即爲非中央 $F$ 分配 (noncentral $F$ distribution), 表達爲 $f \sim F_{m,n}(\lambda)$ 。

因此

$$\hat{\alpha} = H\hat{B} \sim \mathcal{N}(\alpha, \Sigma_{\hat{\alpha}}),$$

其中, $\Sigma_{\hat{\alpha}}$ 爲

$$\begin{aligned}
\Sigma_{\hat{\alpha}} &= Hcov(\hat{B})H' \\
&= \left(I_N \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) (\Sigma \otimes (\bar{x}'\bar{x})^{-1}) \left(I_N \otimes \begin{pmatrix} 1 & 0 \end{pmatrix}\right) \\
&= \Sigma \otimes a_{11},
\end{aligned}$$

$a_{11}$ 爲 $(\bar{x}'\bar{x})^{-1}$ 的第(1,1) 個元素, 進一步推導, 可證明

$$a_{11} = \frac{1}{T}(1 + \hat{\theta}_m^2),$$

其中, $\hat{\theta}_M = \frac{\bar{r}_m}{s_m}$, $s_m^2 = \frac{1}{T}\Sigma(r_{mt} - \bar{r}_m)^2$ 。因此, 可知

$$\begin{aligned}
\hat{\alpha} &\sim \mathcal{N}(\alpha, a_{11}\Sigma) \\
(T-2)\hat{\Sigma} &\sim W_N(T-2, \Sigma),
\end{aligned}$$

令 $X = \frac{\hat{\alpha}}{\sqrt{\frac{1}{T}(1+\hat{\theta}_m^2)}}$ , $A = (T-2)\hat{\Sigma}$ , 則

$$\begin{aligned}
X &\sim \mathcal{N}(\alpha, \Sigma) \\
A &\sim W_N(T-2, \Sigma),
\end{aligned}$$

利用上述定理, 可知 $m = N$ , $n = T - 2$ , 且下式有 $F$ 分配

$$\begin{aligned}
J_1 &\equiv \frac{(T-2) - N + 1}{N}\left(\frac{1}{T}(1 + \hat{\theta}_m^2)\right)^{-1}\hat{\alpha}'((T-2)\hat{\Sigma})^{-1}\hat{\alpha} \\
&= \frac{T(T-N-1)}{N(T-2)}(1 + \hat{\theta}_m^2)^{-1}\hat{\alpha}'\Sigma^{-1}\hat{\alpha} \\
&\sim F_{N,T-N-1}(\lambda)
\end{aligned}$$

在虛無假說下, $H_0 : \alpha = 0$ , 所以非中央參數 $\lambda$ 會等於零, 因此 $J_1$ 檢定的分配即爲一般熟知之(中央) $F$ 分配, 這就是著名的 Gibbons, Ross, and Shanken (GRS, 1989) 檢定。[4]

**問題 14.1.** 證明 $a_{11} = \frac{1}{T}(1 + \hat{\theta}_m^2)$ 。

[4]GRS檢定於 1989 年發表於*Econometrica*, 不過這篇文章可回溯至 1980 年時, Steven

## 14.5 實例分析

在這一節裡, 我們以1995年1月1日至1997年2月28日止252家台灣證券交易所 (TSE) 上市公司的普遍股日報酬率和週報酬率, 參考 Fama and MacBeth (1973) 的方法將其重組爲10組和20組的規模投資組合; 因此10組投資組合中除了第一組和第10組包含了26家股票外, 其餘8組各包含了25家股票; 20組投資組合中除了第一組和第20組包含了18家股票外, 其餘18組各包含了12家股票。此10或 20組規模投資組合代表 GRS 檢定中的 $N$ 項資產。[5]此外, 我們亦選取 TSE 所編制的七種產業指數作爲此 $N$ 項資產, 檢定四指數是否爲均異有效率。本研究的資料除了台灣100指數取自台灣經濟新報資料庫外, 其餘皆取自教育部之 AREMOS 資料庫。[6]

表14.2與表14.3報告分別以日資料及週資料爲樣本, 以 GRS 檢定檢定四種台股指數均異效率的結果。由表14.2的結果顯示, 就日資料而言, 不論是相對於規模投資組合或產業投資組合, 在5% 的顯著水準下, GRS檢定皆無法拒絕四種台股指數的均異效率性; 除了相對於10組規模投資組合, 加權股價指數的 P 値爲 0.0739 外, 其餘的檢定之 P 値皆遠大於 0.10。就週資料而言, 由表14.3的實證結果可知, 即使在 10% 的顯著水準下, GRS檢定亦無法拒絕四種台股指數的均異效率性, 顯示四種指數皆可爲市場投資組合適當的替代指標。即便如此, 我們仍可計算各指數的相對效率性, 從中挑選出最具效率的市場投資組合替代指標。

進一步探究, 我們發現, 就日報酬率而言, 不論是相對於10組規模投資組合或20組規模投資組合, 摩根台股指數的相對效率指標値 値和 値皆爲四指數中最小者, 表示其效率性最佳, 道瓊指數次之,T-100 指數再次之, 而加權指數則最差。有趣的

---

Ross 首度發表於 World Econometric Society 的年會 (當時作者只有 Ross 一人)。我在就讀博士班時, 曾研讀過 Ross 在1983年版本的 working paper。在這篇文章中,Steven Ross 顯然並不知道多變量統計中的 Hotelling $T^2$ 統計量, 而是自行以「概度比檢定」推導出此 $F$ 檢定。我記得 Ross 在文中指出, 因爲有些地方他仍不清楚, 因此選擇不發表這篇文章。不過, 稍後在1986年流通的版本就加入了 Michael Gibbons 與 Jay Shanken 二位, 原來文中許多的推導也都刪除, 而最後才在1989年刊登。整個過程, 如果加上 Ross 原先的撰寫過程, 歷經十年以上。而實際上, 早在1989年正式發表之前, 就已有相當多的論文引用這篇文章而發表在 top-tier 期刊了。

[5]這些被選來作爲檢定的資產或投資組合, 文獻上也稱爲「測試資產 (test assets)」。

[6]這一節的結果, 主要摘錄自我與劉貽芳、林惠雪於 1998 年發表於證券市場發展季刊 10:4, 頁 1-26, 的一篇論文。

是, 相對於產業投資組合而言, 其順序則有所改變; 加權指數最佳,T-100 指數次之, 最差者爲道瓊指數。

就週報酬率而言, 不論是相對於 10 組規模投資組合或 20 組規模投資組合, 摩根台股指數、道瓊指數、與 T-100 指數三指數的值皆分常接近, 加權指數則仍然爲四者中效率性最差的。如同日資料的結果, 相對於產業投資組合而言, 仍以加權指數最佳, 而摩根台股指數反而成爲四者中最差的。

綜合日資料和週資料的結果, 雖然相對於規模投資組合而言, 摩根台股指數的相對均異效率性較佳, 但相對於產業投資組合而言, 卻未必較佳。相對於產業投資組合而言, 加權指數的相對效性皆優於其它三者。探究其原因, 摩根台股指數、道瓊台股指數和 T-100 指數三者雖然根據市值自各產業中挑選樣本股票, 但可能因爲受限於指數可選入的資產數目, 而無法充分反映出產業的特性。相反的, 加權指數因爲涵蓋了所有上市公司的股票, 故較能反映出產業的狀況。綜合言之, 四指數皆具均異效率性, 但未有一指數顯著優於其他指數。

表 14.2: 台股指數均異效率檢定 (1995.1.1 ~ 1997.2.28 日資料)

本表報告以 Gibbons, Ross, and Shanken (1989) 的 GRS 統計量檢定四種台股指數的相對於臺灣股市的均異效率性。我們分別考慮以 10 組規模投資組合,20 組規模投資組合, 及產業投資組合代表市場的產產。$\theta^*$ 代表基準投資組合或事後效率前緣的最大Sharpe 指標。

| | 加權指數 | | | T-100 指數 | | | 摩根台股指數 | | | 道瓊台股指數 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 投資組合 | $\theta^*$ | GRS 統計量 | P 值 | $\theta^*$ | GRS | P 值 | $\theta^*$ | GRS | P 值 | $\theta^*$ | GRS | P 值 |
| 基準投資組合 | 0.0046 | | | 0.0129 | | | 0.0178 | | | 0.0252 | | |
| 10 組規模投資組合 | 0.2019 | 2.4404 | 0.0739 | 0.1243 | 0.9152 | 0.5186 | 0.0980 | 0.5560 | 0.8499 | 0.1105 | 0.6929 | 0.7315 |
| 20 組規模投資組合 | 0.2171 | 1.3870 | 0.1213 | 0.1568 | 0.7188 | 0.8086 | 0.1340 | 0.5190 | 0.9593 | 0.1433 | 0.5852 | 0.9239 |
| 產業投資組合 | 0.0753 | 0.4630 | 0.8614 | 0.0804 | 0.5422 | 0.8026 | 0.0969 | 0.7810 | 0.6034 | 0.1274 | 1.3405 | 0.2285 |

表 14.3: 台股指數均異效率檢定 (1995.1.1 ~ 1997.2.28 週資料)

本表報告以 Gibbons, Ross, and Shanken (1989) 的 GRS 統計量檢定四種台股指數的相對於臺灣股市的均異效率性。我們分別考慮以 10 組規模投資組合,20 組規模投資組合, 及產業投資組合代表市場的產產。$\theta^*$ 代表基準投資組合或事後效率前緣的最大Sharpe 指標。

| | 加權指數 | | | T-100 指數 | | | 摩根台股指數 | | | 道瓊台股指數 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 投資組合 | $\theta^*$ | GRS 統計量 | P 值 | $\theta^*$ | GRS | P 值 | $\theta^*$ | GRS | P 值 | $\theta^*$ | GRS | P 值 |
| 基準投資組合 | 0.0127 | | | | 0.0496 | | 0.0602 | | | 0.0533 | | |
| 10 組規模投資組合 | 0.3267 | 1.0743 | 0.3892 | 0.2168 | 0.4479 | 0.9188 | 0.2178 | 0.4402 | 0.9232 | 0.2168 | 0.4440 | 0.9211 |
| 20 組規模投資組合 | 0.3812 | 0.6577 | 0.8563 | 0.3499 | 0.5425 | 0.9396 | 0.3561 | 0.5564 | 0.9318 | 0.3525 | 0.5487 | 0.9362 |
| 產業投資組合 | 0.1727 | 0.4401 | 0.8747 | 0.1957 | 0.5308 | 0.8094 | 0.2094 | 0.5950 | 0.7587 | 0.1979 | 0.5375 | 0.8043 |

## 14.6 本章習題

1. (10%) 假若你認為兩家公司的月報酬與某一變數有關,

$$R_{it} = a_i + b_i Z_{it} + e_{it},$$

$i = 1, 2$。你想估計此一模型並檢定 $b_1 = b_2$。回答以下問題:

   (a) 請以 SUR 模型表達此一模型 (表達為矩陣形式)。

   (b) 若 $Z_{it}$ 為公司的市場佔有率, 則此檢定有何分配? 又、如果 $Z_{it}$ 為通貨膨脹率呢?(指出是哪一分配並簡單說明即可)。

2. 說明哪二種情況下,SUR 估計子會等於 OLS 估計子。另外, 當二者相等時,SUR 無用嗎?

3. 假若你認為兩家公司的月報酬與其前月的銷貨成長率有關,

$$R_{it} = a_i + b_i \text{Sales-Growth}_{i,t-1} + e_{it}$$

$i = 1, 2$。你應如何估計此一模型並檢定 $b_1 = b_2$? 請盡可能說明, 例如如何以矩陣表達此一模型。無須證明或推導。

4. Let $r_{it}$ be the return on the $i$ -th asset in excess of a riskfree return. Consider the market model:

$$r_{it} = \alpha_i + \beta_i r_{pt} + \varepsilon_{it}, \quad i = 1, \ldots, 3, \quad t = 1, \ldots, T,$$

where $r_{pt}$ is the excess return on the market portfolio. $\varepsilon_{it}$ is the disturbance term for asset $i$ in period $t$ , distributed **normally** as follows:

$$E(\varepsilon_{is}, \varepsilon_{jt}) = \begin{cases} \sigma_{ij} & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

The CAPM implies the following hypothesis:

$$H_0 : \alpha_i = 0 \;\; \forall \;\; i.$$

(a) (20 points) Write down the model in matrix form, and the EGLS estimator for the parameters $\alpha_i$ 's and $\beta$ 's based on Zellner's seemingly unrelated regression model. What is the distribution of the estimator?

(b) (5 points) What statistic and testing distribution can you use to test the null hypothesis?

(c) (5 points) Under what situations will SUR estimators be the same as OLS estimators? Does it mean that SUR is useless for estimation and inference in such cases?

5. Let $r_{it}$ be the return on the $i$ -th asset in excess of a riskfree return. Consider the market model:

$$r_{it} = \alpha_i + \beta_i r_{pt} + \varepsilon_{it}, \quad i = 1, 2, \quad t = 1, \ldots, T,$$

where $r_{pt}$ is the excess return on the market portfolio. $\varepsilon_{it}$ is the disturbance term for asset $i$ in period $t$ , distributed **normally** as follows:

$$E(\varepsilon_{is}, \varepsilon_{jt}) = \begin{cases} \sigma_{ij} & \text{if } s = t \\ 0 & \text{otherwise.} \end{cases}$$

Answer the following questions:

(a) (10 points) If $\sigma_{11} = \sigma_{22}$ and $\sigma_{12} = 0$ , how will use dummy variables to test the following hypothesis?

$$H_0 : \alpha_1 = \alpha_2, \quad and \quad \beta_1 = \beta_2.$$

Write down the model, and report the related statistic and its distribution.

(b) (10 points) If $\sigma_{11} \neq \sigma_{22}$ and $\sigma_{12} \neq 0$ , how will you test the above hypothesis? Write down the model in matrix form, and report the related statistic and its distribution.

(c) (10 points) If $r_{pt} \sim \mathcal{N}(\mu_p, \sigma_p^2)$ , what are the mean and variance of $r_{1t}$ ?