

## § 4.5 联立方程计量经济学模型的系统 估计方法

the Systems Estimation Methods

一、联立方程模型随机误差项方差—协  
方差矩阵

二、三阶段最小二乘法简介

三、完全信息最大似然法简介

# 一、联立方程模型随机误差项方差—协方差矩阵

## 1. 随机误差项的同期相关性

- 随机误差项的相关性不仅存在于每个结构方程不同样本点之间，而且存在于不同结构方程之间。
- 对于不同结构方程的随机误差项之间，不同时期互不相关，只有同期的随机误差项之间才相关，称为具有同期相关性。

## 2. 具有同期相关性的方差—协方差矩阵

$$\mathbf{YB} + \mathbf{X}\Gamma = \mathbf{N} \longrightarrow \mathbf{Y} = \mathbf{Z}\Delta + \tilde{\mathbf{N}}$$

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

$$Y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in} \end{bmatrix}$$

$$Y_i = \mathbf{Z}_i \Delta_i + \tilde{N}_i$$

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{Y}_0^i & \mathbf{X}_0^i \end{pmatrix}$$

$$\Delta_i = \begin{pmatrix} \mathbf{B}_0^i \\ \Gamma_0^i \end{pmatrix}$$

假设:

- 对于一个结构方程的随机误差项，在不同样本点之间，具有同方差性和序列不相关性。即

$$Cov(\tilde{N}_i) = \sigma_{ii}^2 \mathbf{I}$$

- 对于不同结构方程的随机误差项之间，具有且仅具有同期相关性。即

$$Cov(\tilde{N}_i, \tilde{N}_j) = \sigma_{ij} \mathbf{I}$$

于是，联立方程模型系统随机误差项方差—协方差矩阵为：

$$\begin{aligned} Cov(\tilde{\mathbf{N}}) &= \begin{bmatrix} \sigma_{11}^2 \mathbf{I} & \sigma_{12} \mathbf{I} & \cdots & \sigma_{1g} \mathbf{I} \\ \sigma_{21} \mathbf{I} & \sigma_{22}^2 \mathbf{I} & \cdots & \sigma_{2g} \mathbf{I} \\ \vdots & & & \\ \sigma_{g1} \mathbf{I} & \sigma_{g2} \mathbf{I} & & \sigma_{gg}^2 \mathbf{I} \end{bmatrix} \\ &= \Sigma \otimes \mathbf{I} \\ &= \Omega \end{aligned}$$

## 二、三阶段最小二乘法简介 (3SLS, Three Stages Least Squares)

## 1. 概念

- **3SLS是由Zellner和Theil于1962年提出的同时估计联立方程模型全部结构方程的系统估计方法。**
- **其基本思路是  $3SLS=2SLS+GLS$**   
即首先用**2SLS**估计模型系统中每一个结构方程，然后再用**GLS**估计模型系统。



## 2. 三阶段最小二乘法的步骤

(1) 用2SLS估计结构方程

$$Y_i = \mathbf{Z}_i \Delta_i + \tilde{N}_i$$

得到方程随机误差项的估计值。

$$\mathbf{Z}_i = \begin{pmatrix} \mathbf{Y}_0^i & \mathbf{X}_0^i \end{pmatrix} \longrightarrow \mathbf{Y}_0^i = \mathbf{X}\Pi_0^i + \mathbf{E}_0^i$$

*OLS*  
估计

$$\hat{\mathbf{Y}}_0^i = \mathbf{X}\hat{\Pi}_0^i = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}_0^i$$

$$\hat{\mathbf{Z}}_i = \begin{pmatrix} \hat{\mathbf{Y}}_0^i & \mathbf{X}_0^i \end{pmatrix}$$

*OLS* 估计

$$\hat{\Delta}_i = (\hat{\mathbf{Z}}_i' \hat{\mathbf{Z}}_i)^{-1} \hat{\mathbf{Z}}_i' \mathbf{Y}_i$$

$$\hat{\hat{\mathbf{Y}}}_i = \hat{\mathbf{Z}}_i \hat{\Delta}_i \longrightarrow e_{il} = y_{il} - \hat{\hat{y}}_{il}$$

(2) 求随机误差项方差—协方差矩阵的估计量

$$\mathbf{e}_i = \begin{pmatrix} e_{i1} & e_{i2} & \cdots & e_{in} \end{pmatrix}'$$

$$\hat{\sigma}_{ij} = \frac{\mathbf{e}_i' \mathbf{e}_j}{\sqrt{(n - g_i + 1 - k_i)(n - g_j + 1 - k_j)}}$$

$$\hat{\Sigma} = (\hat{\sigma}_{ij})$$

$$\hat{\Sigma} \otimes \mathbf{I} = \hat{\Omega}$$

(3) 用GLS估计原模型系统

$$\mathbf{Y} = \mathbf{Z} \Delta + \tilde{\mathbf{N}}$$

得到结构参数的3SLS估计量为：

$$\begin{aligned}\hat{\hat{\Delta}} &= (\hat{\mathbf{Z}}' \hat{\Omega}^{-1} \hat{\mathbf{Z}})^{-1} \hat{\mathbf{Z}}' \hat{\Omega}^{-1} \mathbf{Y} \\ &= (\hat{\mathbf{Z}}' (\hat{\Sigma} \otimes \mathbf{I})^{-1} \hat{\mathbf{Z}})^{-1} \hat{\mathbf{Z}}' (\hat{\Sigma} \otimes \mathbf{I})^{-1} \mathbf{Y}\end{aligned}$$

### 3. 三阶段最小二乘法估计量的统计性质

- (1) 如果联立方程模型系统中所有结构方程都是可以识别的，并且非奇异，则**3SLS**估计量是一致性估计量。
- (2) **3SLS**估计量比**2SLS**估计量更有效。为什么？
- (3) 如果  $\Sigma$  是对角矩阵，即模型系统中不同结构方程的随机误差项之间无相关性，那么可以证明**3SLS**估计量与**2SLS**估计量是等价的。
- (4) 这反过来说明，**3SLS**方法主要优点是考虑了模型系统中不同结构方程的随机误差项之间的相关性。

### 三、完全信息最大似然法简介

( FIML, Full Information  
Maximum Likelihood )

## 1. 概念

- 另一种已有实际应用的联立方程模型的系统估计方法。
- **Rothenberg**和**Leenders**于**1964**年提出一个线性化的**FIML**估计量。
- **FIML**是**ML**的直接推广，是在已经得到样本观测值的情况下，使整个联立方程模型系统的或然函数达到最大以得到所有结构参数的估计量。

## 2. 复习：多元线性单方程模型的最大似然估计

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \mu_i$$

$$i=1,2,\dots,n$$

$$Y = XB + N$$



- Y的随机抽取的n组样本观测值的联合概率

$$L(\hat{B}, \sigma_{\mu}^2) = P(y_1, y_2, \dots, y_n)$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_{\mu}^n} e^{-\frac{1}{2\sigma_{\mu}^2} \sum (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \dots + \hat{\beta}_k x_{ki}))^2}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_{\mu}^n} e^{-\frac{1}{2\sigma_{\mu}^2} (Y - X\hat{B})'(Y - X\hat{B})}$$

- 对数或然函数为

$$\begin{aligned} L^* &= Ln(L) \\ &= -nLn(\sqrt{2\pi}\sigma_{\mu}) - \frac{1}{2\sigma_{\mu}^2} (Y - X\hat{B})' (Y - X\hat{B}) \end{aligned}$$

- 参数的最大或然估计

$$\hat{B} = (X'X)^{-1} X'Y$$

### 3. 复习：有限信息最大或然法 (LIML, **Limited Information Maximum Likelihood**)

- 以最大或然为准则、通过对简化式模型进行最大或然估计，以得到结构方程参数估计量的**联立方程模型的单方程估计方法**。
- 由Anderson和Rubin于1949年提出，早于两阶段最小二乘法。
- 适用于恰好识别和过度识别结构方程的估计。

- 在该方法中，以下两个概念是重要的：
  - 一是这里的“有限信息”指的是每次估计只考虑一个结构方程的信息，而没有考虑模型系统中其它结构方程的信息；
  - 二是这里的“最大或然法”是针对结构方程中包含的内生变量的简化式模型的，即应用最大或然法求得的是简化式参数估计量，而不是结构式参数估计量。

$$Y_1 = (Y_0, X_0) \begin{pmatrix} B_\theta \\ \Gamma_\theta \end{pmatrix} + N_1$$

$$(Y_0^1, X_0) \begin{pmatrix} B_0^1 \\ \Gamma_\theta \end{pmatrix} + N_1 = 0$$

$$Y_0^1 = X \Pi_0^1 + E_0^1$$

$$LnL(Y_0^1) = c + \frac{n}{2} Ln|\Omega_0^{-1}| - \frac{1}{2} tr\Omega_0^{-1} (Y_0^1 - X\Pi_0^1)'(Y_0^1 - X\Pi_0^1)$$

## 4. 完全信息最大似然函数

$$\mathbf{Y} = \mathbf{Z}\Delta + \tilde{\mathbf{N}}$$

$$\tilde{\mathbf{N}} \sim \text{正态}(\mathbf{0}, \Sigma \otimes \mathbf{I})$$

*ML的直接推广*

$$L(\mathbf{Y}) = \frac{1}{(2\pi)^{gn/2} |\Sigma \otimes \mathbf{I}|^{-\frac{1}{2}}} \left| \frac{\partial \tilde{\mathbf{N}}}{\partial \mathbf{Y}} \right| e^{-\frac{1}{2}(\mathbf{Y} - \mathbf{Z}\Delta)'(\Sigma^{-1} \otimes \mathbf{I})(\mathbf{Y} - \mathbf{Z}\Delta)}$$

$$= \frac{1}{(2\pi)^{gn/2} |\Sigma \otimes \mathbf{I}|^{-\frac{1}{2}}} |\mathbf{B}|^n e^{-\frac{1}{2}(\mathbf{Y} - \mathbf{Z}\Delta)'(\Sigma^{-1} \otimes \mathbf{I})(\mathbf{Y} - \mathbf{Z}\Delta)}$$

$$L(\mathbf{Y}) = \frac{1}{(2\pi)^{gn/2} |\Sigma \otimes \mathbf{I}|^{-\frac{1}{2}}} \left| \frac{\partial \tilde{\mathbf{N}}}{\partial \mathbf{Y}} \right| e^{-\frac{1}{2}(\mathbf{Y} - \mathbf{Z}\Delta)'(\Sigma^{-1} \otimes \mathbf{I})(\mathbf{Y} - \mathbf{Z}\Delta)}$$

$$-\frac{1}{2}(\mathbf{Y} - \mathbf{Z}\Delta)'(\Sigma^{-1} \otimes \mathbf{I})(\mathbf{Y} - \mathbf{Z}\Delta)$$

- 对数或然函数对于协方差逆矩阵的元素取极大值的一阶条件，得到协方差矩阵的元素的 **FIML** 估计量；
- 对数或然函数对于待估计参数取极大值的一阶条件，求解该方程系统，即可得到结构参数的 **FIML** 估计量。
- 研究的重点是如何求解非线性方程系统。