第 24 章　联立方程模型

24.1　联立方程模型的结构式与简化式

经济理论常常推导出一组相互联系的方程，其中一个方程的解释变量是另一方程的被解释变量，这就是联立方程组。

例　农产品市场均衡模型，由需求函数、供给函数及市场均衡条件组成，参见第 10 章。

例　简单的宏观经济模型，参见第 10 章。

即使我们只关心单个方程，但如果该方程包含内生解释变量，则完整的模型仍然是联立方程组。

由 M 个方程构成的联立方程模型的“结构式”(structural form)：

$$\begin{cases} \gamma_{11} y_{t1} + \gamma_{21} y_{t2} + \cdots + \gamma_{M1} y_{tM} + \beta_{11} x_{t1} + \cdots + \beta_{K1} x_{tK} = \varepsilon_{t1} \\ \gamma_{12} y_{t1} + \gamma_{22} y_{t2} + \cdots + \gamma_{M2} y_{tM} + \beta_{12} x_{t1} + \cdots + \beta_{K2} x_{tK} = \varepsilon_{t2} \\ \cdots\cdots\cdots\cdots \\ \gamma_{1M} y_{t1} + \gamma_{2M} y_{t2} + \cdots + \gamma_{MM} y_{tM} + \beta_{1M} x_{t1} + \cdots + \beta_{KM} x_{tK} = \varepsilon_{tM} \end{cases}$$

$\{y_{ti}\}$为内生变量，$\{x_{tj}\}$为外生变量，第一个下标表示第 t 个观测值 $(t = 1, \cdots, T)$，第二个下标表示第 i 个内生变量 $(i = 1, \cdots, M)$，或第 j 个外生变量 $(j = 1, \cdots, K)$。

内生变量的系数为$\{\gamma_{ik}\}$，其第一个下标表示它是第i个内生变量的系数，而第二个下标表示它在第k个方程中$(k=1,\cdots,M)$。

外生变量的系数为$\{\beta_{jk}\}$，其第一个下标表示它是第j个外生变量的系数，而第二个下标表示它在第k个方程中。

结构方程的扰动项为$\{\varepsilon_{tk}\}$，其第一个下标表示第t个观测值$(t=1,\cdots,T)$，而第二个下标表示它在第k个方程中。

“完整的方程系统”(complete system of equations)要求，内生变量个数等于方程个数M。

将上述方程组写成更简洁的“横排”矩阵形式

$$
\begin{pmatrix} y_{t1} & y_{t2} & \cdots & y_{tM} \end{pmatrix}
\begin{pmatrix}
\gamma_{11} & \gamma_{12} & \cdots & \gamma_{1M} \\
\gamma_{21} & \gamma_{22} & \cdots & \gamma_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
\gamma_{M1} & \gamma_{M2} & \cdots & \gamma_{MM}
\end{pmatrix}
$$

$$
+\begin{pmatrix} x_{t1} & x_{t2} & \cdots & x_{tK} \end{pmatrix}
\begin{pmatrix}
\beta_{11} & \beta_{12} & \cdots & \beta_{1M} \\
\beta_{21} & \beta_{22} & \cdots & \beta_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
\beta_{K1} & \beta_{K2} & \cdots & \beta_{KM}
\end{pmatrix}
=\begin{pmatrix} \varepsilon_{t1} & \varepsilon_{t2} & \cdots & \varepsilon_{tM} \end{pmatrix}
$$

用矩阵来表示即

$$\boldsymbol{y}_t'\boldsymbol{\Gamma} + \boldsymbol{x}_t'\boldsymbol{B} = \boldsymbol{\varepsilon}_t'$$

其中，系数矩阵$\boldsymbol{\Gamma}_{M\times M}$与$\boldsymbol{B}_{K\times M}$的每一列对应于一个方程。

比如，第一个方程为

$$\begin{pmatrix} y_{t1} & y_{t2} & \cdots & y_{tM} \end{pmatrix} \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \\ \vdots \\ \gamma_{M1} \end{pmatrix} + \begin{pmatrix} x_{t1} & x_{t2} & \cdots & x_{tK} \end{pmatrix} \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ \vdots \\ \beta_{K1} \end{pmatrix} = \varepsilon_{t1}$$

扰动项$\boldsymbol{\varepsilon}_t$由第 t 期各方程的扰动项所构成。

假设扰动项$\boldsymbol{\varepsilon}_t$满足$\mathrm{E}(\boldsymbol{\varepsilon}_t \mid \boldsymbol{x}_t) = \mathbf{0}$($\boldsymbol{x}_t$外生)，记其协方差矩阵为，

$$\boldsymbol{\Sigma} \equiv \mathrm{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \mid \boldsymbol{x}_t)$$

由于存在内生变量，如果直接用 OLS 估计每一方程，将导致内生性偏差或联立方程偏差，得不到一致估计。

求解联立方程组：

$$\boldsymbol{y}_t'\boldsymbol{\Gamma} = -\boldsymbol{x}_t'\boldsymbol{B} + \boldsymbol{\varepsilon}_t'$$

假设$\boldsymbol{\Gamma}$非退化，两边同时右乘$\boldsymbol{\Gamma}^{-1}$，

$$\boldsymbol{y}_t' = -\boldsymbol{x}_t'\boldsymbol{B}\boldsymbol{\Gamma}^{-1} + \boldsymbol{\varepsilon}_t'\boldsymbol{\Gamma}^{-1}$$

$$\boldsymbol{y}_t' = \boldsymbol{x}_t'\boldsymbol{\Pi} + \boldsymbol{v}_t'$$

此方程称为“简化式”(reduced form)。

其系数矩阵为$\underbrace{\boldsymbol{\Pi}}_{K\times M} \equiv -\underbrace{\boldsymbol{B}}_{K\times M}\underbrace{\boldsymbol{\Gamma}^{-1}}_{M\times M}$，扰动项为$\boldsymbol{v}_t' \equiv \boldsymbol{\varepsilon}_t'\boldsymbol{\Gamma}^{-1}$，故$\boldsymbol{v}_t \equiv \boldsymbol{\Gamma}^{-1\prime}\boldsymbol{\varepsilon}_t$。

简化式扰动项$\boldsymbol{v}_t$仍与外生变量$\boldsymbol{x}_t$不相关，因为

$$\mathrm{E}(\boldsymbol{v}_t \mid \boldsymbol{x}_t) = \mathrm{E}(\boldsymbol{\Gamma}^{-1\prime}\boldsymbol{\varepsilon}_t \mid \boldsymbol{x}_t) = \boldsymbol{\Gamma}^{-1\prime}\,\mathrm{E}(\boldsymbol{\varepsilon}_t \mid \boldsymbol{x}_t) = \mathbf{0}$$

$\boldsymbol{v}_t$的协方差矩阵为

$$\boldsymbol{\Omega} \equiv \mathrm{E}(\boldsymbol{v}_t\boldsymbol{v}_t' \mid \boldsymbol{x}_t) = \mathrm{E}(\boldsymbol{\Gamma}^{-1\prime}\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'\boldsymbol{\Gamma}^{-1} \mid \boldsymbol{x}_t) = \boldsymbol{\Gamma}^{-1\prime}\,\mathrm{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t' \mid \boldsymbol{x}_t)\boldsymbol{\Gamma}^{-1} = \boldsymbol{\Gamma}^{-1\prime}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1}$$

简化式方程的解释变量全部为外生变量$\boldsymbol{x}_t$，故可用 OLS 得到简化式参数$\boldsymbol{\Pi}$与$\boldsymbol{\Omega}$的一致估计。

但通常我们最终关心的是结构式参数。

在什么情况下，才能从简化式参数$(\boldsymbol{\Pi}, \boldsymbol{\Omega})$反推出结构式参数$(\boldsymbol{\Gamma}, \boldsymbol{B}, \boldsymbol{\Sigma})$呢？

这涉及联立方程模型的“识别问题”(problem of identification)。

24.2 联立方程模型的识别

在对模型的总体参数进行估计之前，其参数必须“可识别”(identified)。

如果一个总体参数可识别，则该参数的任意两个不同取值，都会在随机样本中显示出系统差异，即如果样本容量足够大，则应该能够在统计意义上区分这两个不同的参数值。

反之，如果无论多大的样本都区分不开，即由不同参数值的总体产生的观测数据在统计意义上是一样的，则该参数“不可识别”(unidentified)。

例 考虑以下回归模型：

$$y_i = \alpha_1 + \alpha_2 + \beta x_i + \varepsilon_i$$

仅通过样本数据$\{y_i, x_i\}_{i=1}^n$是无法对α_1与α_2分别进行识别的，但可以识别二者之和$(\alpha_1 + \alpha_2)$。

回到联立方程模型的情形，“可识别”意味着，可以从简化式参数$(\boldsymbol{\Pi}, \boldsymbol{\Omega})$求出结构式参数$(\boldsymbol{\Gamma}, \boldsymbol{B}, \boldsymbol{\Sigma})$的唯一解(unique solution)。

这两组参数之间的关系如下：

$$\boldsymbol{\Pi} \equiv -\boldsymbol{B}\boldsymbol{\Gamma}^{-1}$$

$$\boldsymbol{\Omega} \equiv \boldsymbol{\Gamma}^{-1\prime}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1}$$

如果$\boldsymbol{\Gamma}$已知，则可通过$\boldsymbol{\Pi}$与$\boldsymbol{\Omega}$求得$\boldsymbol{B}$与$\boldsymbol{\Sigma}$。但$\boldsymbol{\Gamma}$一般是由未知参数组成的矩阵。

事实上，结构式的参数个数比简化式的参数个数多出M^2个。

简化式参数$(\boldsymbol{\Pi}, \boldsymbol{\Omega})$的总个数为$\left[K \times M + M(M+1)/2\right]$(其中，$\boldsymbol{\Pi}_{K\times M}$含$K \times M$个参数，而对称矩阵$\boldsymbol{\Omega}_{M\times M}$含$M(M+1)/2$个参数)；

结构式参数$(\boldsymbol{\Gamma}, \boldsymbol{B}, \boldsymbol{\Sigma})$的总个数为$\left[M^2 + K \times M + M(M+1)/2\right]$(其中，$\boldsymbol{\Gamma}_{M \times M}$含$M^2$个参数，$\boldsymbol{B}_{K \times M}$含$K \times M$个参数，对称矩阵$\boldsymbol{\Sigma}$含$M(M+1)/2$个参数)。

一般地，不可能从$(\boldsymbol{\Pi}, \boldsymbol{\Omega})$求出$(\boldsymbol{\Gamma}, \boldsymbol{B}, \boldsymbol{\Sigma})$的唯一解。

如不对结构式参数进行约束，将不可能从简化式参数得到结构式参数的唯一解。

为识别结构方程，常对结构参数施加如下约束。

(1) 标准化(normalization)：在每个结构方程中，可以将一个内生变量视为被解释变量，并将其系数标准化为 1。

(2) 恒等式(identity)：比如，供需相等的均衡条件、会计恒等式、定义式。恒等式中每个变量的系数均为已知，不需要识别或估计。

(3) 排斥约束(exclusion restrictions)：在结构方程中排斥某些内生或外生变量，这相当于对结构矩阵$(\boldsymbol{\Gamma}, \boldsymbol{B})$施以"零约束"(zero restrictions)，即让$(\boldsymbol{\Gamma}, \boldsymbol{B})$中的某些元素为 0。

(4) 线性约束(linear restriction)：比如，在理论上可以假设生产函数为规模报酬不变(constant returns to scale)，则资本的产出弹性与劳动力的产出弹性之和为 1。

(5) 对扰动项协方差矩阵的约束(restrictions on the disturbance covariance matrix)：比如，在某些情况下，可以假设不同方程的扰动项之间不相关。

实践中最重要的约束方法是“排斥变量”(即零约束)。

对于线性约束，可通过重新定义变量转化为“排斥变量”约束。

究竟需要多少零约束才可以保证结构方程可识别呢？

不失一般性，考虑第一个结构方程。

假设在第一个方程中，内生变量y_1的系数已被标准化为 1，另有M_1个内生变量也包括在此方程中，而其余M_1^*个内生变量则被排斥在此方程之外，故$1+M_1+M_1^*=M$。

假设第一个方程包含K_1个外生变量，而其余K_1^*个外生变量则被排斥在此方程之外，故$K_1+K_1^*=K$。

可识别的必要条件为

$$K_1^* \geq M_1$$

称为“阶条件”(order condition)，即结构方程所排斥的外生变量的个数(K_1^*)应大于或等于该方程所包含的内生解释变量的个数(M_1)。

从工具变量法的角度，被第一个结构方程排斥的所有外生变量都是有效工具变量，因为根据外生变量的定义，它们与扰动项不相关(外生性)；而根据简化式，内生变量可以表示为外生变量的函数，故它们与内生解释变量相关(相关性)。

在可识别(即秩条件满足)的情况下，如果恰好$K_1^* = M_1$，则称该结构方程“恰好识别”(just identified)，即工具变量个数正好相等内生解释变量的个数。

如果$K_1^* > M_1$，则称该结构方程“过度识别”(overidentified)，即工具变量个数大于内生解释变量的个数。