§4.3联立方程计量经济学模型的识别

The Identification Problem

一、识别的概念

二、从定义出发识别模型

三、结构式识别条件

四、简化式识别条件

五、实际应用中的经验方法

一、识别的概念

1. 为什么要对模型进行识别？

- 从一个例子看

$$\begin{cases} C_t = \alpha_0 + \alpha_1 Y_t + \mu_{1t} \\ I_t = \beta_0 + \beta_1 Y_t + + \mu_{2t} \\ Y_t = C_t + I_t \end{cases}$$

- 消费方程是包含C、Y和常数项的直接线性方程。
- 投资方程和国内生产总值方程的某种线性组合（消去I）所构成的新方程也是包含C、Y和常数项的直接线性方程。

- 如果利用C、Y的样本观测值并进行参数估计后，很难判断得到的是消费方程的参数估计量还是新组合方程的参数估计量。
- 只能认为原模型中的消费方程是不可估计的。
- 这种情况被称为不可识别。
- 只有可以识别的方程才是可以估计的。

2. 识别的定义

- **3种定义：**

“如果联立方程模型中某个结构方程不具有确定的统计形式，则称该方程为不可识别。”

“如果联立方程模型中某些方程的线性组合可以构成与某一个方程相同的统计形式，则称该方程为不可识别。”

“根据参数关系体系，在已知简化式参数估计值时，如果不能得到联立方程模型中某个结构方程的确定的结构参数估计值，则称该方程为不可识别。”

- 以是否具有确定的统计形式作为识别的基本定义。
- 什么是“统计形式”?
- 什么是“具有确定的统计形式”?

3. 模型的识别

- 上述识别的定义是针对结构方程而言的。
- 模型中每个需要估计其参数的随机方程都存在识别问题。
- 如果一个模型中的所有随机方程都是可以识别的，则认为该联立方程模型系统是可以识别的。反过来，如果一个模型系统中存在一个不可识别的随机方程，则认为该联立方程模型系统是不可以识别的。
- 恒等方程由于不存在参数估计问题，所以也不存在识别问题。但是，在判断随机方程的识别性问题时，应该将恒等方程考虑在内。

4. 恰好识别(Just Identification)与过度识别(Overidentification)

- 如果某一个随机方程具有一组参数估计量，称其为恰好识别；
- 如果某一个随机方程具有多组参数估计量，称其为过度识别。

二、从定义出发识别模型

1. 例题1

$$\begin{cases} C_t = \alpha_0 + \alpha_1 Y_t + \mu_{1t} \\ I_t = \beta_0 + \beta_1 Y_t + + \mu_{2t} \\ Y_t = C_t + I_t \end{cases}$$

- 第2与第3个方程的线性组合得到的新方程具有与消费方程相同的统计形式，所以消费方程也是不可识别的。

- 第**1**与第**3**个方程的线性组合得到的新方程具有与投资方程相同的统计形式，所以投资方程也是不可识别的。
- 于是，该模型系统不可识别。
- 参数关系体系由**3**个方程组成，剔除一个矛盾方程，**2**个方程不能求得4个结构参数的确定值。也证明消费方程与投资方程都是不可识别的。

2. 例题2

$$C_t = \alpha_0 + \alpha_1 Y_t + \mu_{1t}$$

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_{2t}$$

$$Y_t = C_t + I_t$$

- 消费方程是可以识别的，因为任何方程的线性组合都不能构成与它相同的统计形式。
- 投资方程仍然是不可识别的，因为第**1**、第**2**与第**3**个方程的线性组合（消去C）构成与它相同的统计形式。
- 于是，该模型系统仍然不可识别。

- 参数关系体系由**6**个方程组成，剔除**2**个矛盾方程，由**4**个方程是不能求得所有**5**个结构参数的确定估计值。
- 可以得到消费方程参数的确定值，证明消费方程可以识别；因为只能得到它的一组确定值，所以消费方程是恰好识别的方程。
- 投资方程都是不可识别的。
- 注意：与例题1相比，在投资方程中增加了1个变量，消费方程变成可以识别。

3. 例题3

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \mu_{1t}$$

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_{2t}$$

$$Y_t = C_t + I_t$$

- 消费方程仍然是可以识别的，因为任何方程的线性组合都不能构成与它相同的统计形式。
- 投资方程也是可以识别的，因为任何方程的线性组合都不能构成与它相同的统计形式。
- 于是，该模型系统是可以识别的。

- 参数关系体系由9个方程组成，剔除3个矛盾方程，在已知简化式参数估计值时，由6个方程能够求得所有6个结构参数的确定估计值。
- 所以也证明消费方程和投资方程都是可以识别的。
- 而且，只能得到所有6个结构参数的一组确定值，所以消费方程和投资方程都是恰好识别的方程。
- 注意：与例题2相比，在消费方程中增加了1个变量，投资方程变成可以识别。

4. 例题4

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \alpha_3 P_{t-1} + \mu_{1t}$$

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_{2t}$$

$$Y_t = C_t + I_t$$

- 消费方程和投资方程仍然是可以识别的，因为任何方程的线性组合都不能构成与它们相同的统计形式。
- 于是，该模型系统是可以识别的。

- 参数关系体系由**12**个方程组成，剔除**4**个矛盾方程，在已知简化式参数估计值时，由**8**个方程能够求得所有**7**个结构参数的确定估计值。
- 所以也证明消费方程和投资方程都是可以识别的。
- 但是，求解结果表明，对于消费方程的参数，只能得到一组确定值，所以消费方程是恰好识别的方程;
- 而对于投资方程的参数，能够得到多组确定值，所以投资方程是过度识别的方程。

- 注意：
- 在求解线性代数方程组时，如果方程数目大于未知数数目，被认为无解；如果方程数目小于未知数数目，被认为有无穷多解。
- 但是在这里，无穷多解意味着没有确定值，所以，如果参数关系体系中有效方程数目小于未知结构参数估计量数目，被认为不可识别。
- 如果参数关系体系中有效方程数目大于未知结构参数估计量数目，那么每次从中选择与未知结构参数估计量数目相等的方程数，可以解得一组结构参数估计值，换一组方程，又可以解得一组结构参数估计值，这样就可以得到多组结构参数估计值，被认为可以识别，但不是恰好识别，而是过度识别。

5. 如何修改模型使不可识别的方程变成可以识别

- 或者在其它方程中增加变量;
- 或者在该不可识别方程中减少变量。
- 必须保持经济意义的合理性。

三、结构式识别条件

1. 结构式识别条件

- 直接从结构模型出发
- 一种规范的判断方法
- 每次用于1个随机方程
- 具体描述为：

联立方程计量经济学模型的结构式

$$\mathbf{B}\,\mathbf{Y} + \Gamma\,\mathbf{X} = \mathrm{N}$$

中的第 i 个方程中包含 g_i 个内生变量（含被解释变量）和 k_i 个先决变量（含常数项），模型系统中内生变量和先决变量的数目仍用 g 和 k 表示，矩阵 $(\mathrm{B}_0\Gamma_0)$ 表示第 i 个方程中未包含的变量（包括内生变量和先决变量）在其它 $g-1$ 个方程中对应系数所组成的矩阵。于是，判断第 i 个结构方程识别状态的结构式条件为：

如果 $R(\mathrm{B}_0\Gamma_0) < g-1$，则第 i 个结构方程不可识别；

如果 $R(\mathrm{B}_0\Gamma_0) = g-1$，则第 i 个结构方程可以识别，并且

如果 $k-k_i = g_i - 1$，则第 i 个结构方程恰好识别，

如果 $k-k_i > g_i - 1$，则第 i 个结构方程过度识别。

- 一般将该条件的前一部分称为**秩条件（Rank Condition）**，用以判断结构方程是否识别；
- 将后一部分称为**阶条件（Order Conditon）**，用以判断结构方程恰好识别或者过度识别。

2. 例题

$$C_t = \alpha_0 + \alpha_1 Y_t + \alpha_2 C_{t-1} + \alpha_3 P_{t-1} + \mu_{1t}$$

$$I_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + \mu_{2t}$$

$$Y_t = C_t + I_t$$

$$[\mathrm{B}\Gamma] = \begin{bmatrix} 1 & 0 & -\alpha_1 & -\alpha_0 & 0 & -\alpha_2 & -\alpha_3 \\ 0 & 1 & -\beta_1 & -\beta_0 & -\beta_2 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- 判断第1个结构方程的识别状态

$$[\mathrm{B}_0\Gamma_0] = \begin{bmatrix} 1 & -\beta_2 \\ -1 & 0 \end{bmatrix} \longrightarrow R(\mathrm{B}_0\Gamma_0) = 2 = g - 1$$

所以，该方程可以识别。

因为

$$k - k_1 = 1 = g_1 - 1$$

所以，第1个结构方程为恰好识别的结构方程。

- 判断第2个结构方程的识别状态

$$\begin{bmatrix} B_0 \Gamma_0 \end{bmatrix} = \begin{bmatrix} 1 & -\alpha_2 & -\alpha_3 \\ -1 & 0 & 0 \end{bmatrix} \longrightarrow R(B_0 \Gamma_0) = 2 = g - 1$$

所以，该方程可以识别。

因为

$$k - k_2 = 2 > g_2 - 1$$

所以，第2个结构方程为过度识别的结构方程。

- 第3个方程是平衡方程，不存在识别问题。
- 综合以上结果，该联立方程模型是可以识别的。
- 与从定义出发识别的结论一致。

四、简化式识别条件

1. 简化式识别条件

- 如果已经知道联立方程模型的简化式模型参数，那么可以通过对简化式模型的研究达到判断结构式模型是否识别的目的。
- 由于需要首先估计简化式模型参数，所以很少实际应用。

对于简化式模型

$$\boldsymbol{Y}=\Pi\boldsymbol{X}+\mathrm{E}$$

简化式识别条件为：

如果 $R(\Pi_2)<g_i-1$，则第 i 个结构方程不可识别；

如果 $R(\Pi_2)=g_i-1$，则第 i 个结构方程可以识别，并且

如果 $k-k_i=g_i-1$，则第 i 个结构方程恰好识别，

如果 $k-k_i>g_i-1$，则第 i 个结构方程过度识别。

其中 Π_2 是简化式参数矩阵 Π 中划去第 i 个结构方程所不包含的内生变量所对应的行和第 i 个结构方程中包含的先决变量所对应的列之后，剩下的参数按原次序组成的矩阵。

2. 例题

- 需要识别的结构式模型

$$\begin{cases} y_{1i} = \alpha_1 + \alpha_2 x_{1i} + \alpha_3 x_{2i} + \mu_{1i} \\ y_{2i} = \beta_1 y_{3i} + \beta_2 x_{3i} + \mu_{2i} \\ y_{3i} = \gamma_1 y_{1i} + \gamma_2 y_{2i} + \gamma_3 x_{3i} + \mu_{3i} \end{cases}$$

- 已知其简化式模型参数矩阵为

$$\Pi = \begin{bmatrix} 4 & -2 & 3 \\ 2 & -1 & 1 \\ 2 & -1 & 0 \end{bmatrix}$$

- 判断第**1**个结构方程的识别状态

$$\Pi_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

$$R(\Pi_2) = 1 = g_1 - 1$$

所以该方程是可以识别的。又因为

$$k - k_1 = 1 = g_1 - 1$$

所以该方程是恰好识别的。

- 判断第**2**个结构方程的识别状态

$$\Pi_2 = \begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix}$$

$$R(\Pi_2) = 1 = g_2 - 1$$

所以该方程是可以识别的。又因为

$$k - k_2 = 2 > g_2 - 1$$

所以该方程是过度识别的。

- 判断第3个结构方程的识别状态

$$\Pi_2 = \begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 2 & -1 \end{bmatrix} \qquad R(\Pi_2) = 1 < g_3 - 1$$

所以该方程是不可识别的。

- 所以该模型是不可识别的。

- 可以从数学上严格证明，简化式识别条件和结构式识别条件是等价的。

  《计量经济学—方法与应用》（李子奈编著，清华大学出版社，1992年3月）第104—107页。

- 讨论：阶条件是确定过度识别的充分必要条件吗？（李子奈，《数量经济技术经济研究》，1988年第10期）

五、实际应用中的经验方法

- 当一个联立方程计量经济学模型系统中的方程数目比较多时，无论是从识别的概念出发，还是利用规范的结构式或简化式识别条件，对模型进行识别，困难都是很大的，或者说是不可能的。
- 理论上很严格的方法在实际中往往是无法应用的，在实际中应用的往往是一些经验方法。
- 关于联立方程计量经济学模型的识别问题，实际上不是等到理论模型已经建立了之后再进行识别，而是在建立模型的过程中设法保证模型的可识别性。

- “在建立某个结构方程时，要使该方程包含前面每一个方程中都不包含的至少1个变量（内生或先决变量）；同时使前面每一个方程中都包含至少1个该方程所未包含的变量，并且互不相同。”
- 该原则的前一句话是保证该方程的引入不破坏前面已有方程的可识别性。只要新引入方程包含前面每一个方程中都不包含的至少1个变量，那么它与前面方程的任意线性组合都不能构成与前面方程相同的统计形式，原来可以识别的方程仍然是可以识别的。
- 该原则的后一句话是保证该新引入方程本身是可以识别的。只要前面每个方程都包含至少1个该方程所未包含的变量，并且互不相同。那么所有方程的任意线性组合都不能构成与该方程相同的统计形式。

- 在实际建模时，将每个方程所包含的变量记录在如下表所示的表式中，将是有帮助的。

	变量 1	变量 2	变量 3	变量 4	变量 5	变量 6	…
方程 1	×	×		×			
方程 2		×	×	×	×		
方程 3	×		×	×		×	
方程 4		×	×				×
…							