从参数唯一确定的观点看识别和多重共线性的关系

田　国　强

提要：本文从待估参数应是唯一确定的基本观点出发，去掉了外生变量不由任何线性恒等式连接的通常假定[2,4]，在非常一般的假定下探讨了同期模型的识别问题并且得到了一些新结果。其中包括把多重共线性的问题可以归结为识别问题的新观点，通常关于同期模型识别的主要定理可视为我们相应结果的特例[2,4]。

一

经济计量模型的识别问题在计量经济学中是非常重要的。模型识别与否关系到是否能得到模型的参数的一致估计量。人们已经对识别问题进行了大量的研究[2,4]，我们也在文献〔1〕中给出了一些新结果。本文继续利用文献〔1〕中的观点和方法作出进一步的探讨，关于识别的定义，符号和术语均见文献〔1〕。

二

假定由经济理论知道，模型是由下列形式决定的：

$$By_t + \Gamma x_t = u_t \qquad (1)$$

这里y_t，x_t和u_t分别是在时刻t的$G\times1$，$K\times1$和$G\times1$阶可观测内生变量，可观测外生变量和不可观测的扰动项；B和Γ分别是$G\times G$和$G\times K$阶待估系数矩阵。

假定：

1. B是非奇异的；
2. $Eu_t=0$，$Eu_tX_s'=0$*。

令
$$C_x = Ex_tx_t' \qquad (2)$$
$$C_{yx} = Ey_tx_t' \qquad (3)$$
$$C_y = Ey_ty_t' \qquad (4)$$

---

* 当u_t是序列无关时，X_t可以是前定变量即X_t中可以包括滞后的内生变量，因为这时有

$$\operatorname*{Plim}_{T\to\infty} \frac{\sum_{t=0}^{T} x_tu_t'}{T} = 0$$

$$\Sigma = Eu_t u_t' \tag{5}$$

于是关于（1）式有二阶信息：

$$BC_{yx} + \Gamma C_x = 0 \tag{6}$$

$$\Sigma = Eu_t u_t' = (B_{yt} + \Gamma_{xt})(By_t + \Gamma x_t)'$$
$$= BC_y B' + BC_{yx}\Gamma' + \Gamma C_{xy} B' + \Gamma C_x \Gamma'$$

由（6）式，将上式化简，有

$$BC_y B' = \Gamma C_x \Gamma' + \Sigma \tag{7}$$

从（7）式能看出，当B和Γ能识别时，Σ也能识别。

先讨论用二步识别法识别（1）式。作为第一步，定义

$$A = (B, \Gamma) \tag{8}$$

$$W = \begin{bmatrix} C_{yx} \\ C_x \end{bmatrix} \tag{9}$$

这里W包含所有的样本观测信息*（观测信息由样本二阶矩反映）。于是（6）式能够重新写为

$$AW = 0 \tag{10}$$

当然也有

$$A_g W = 0 \tag{11}$$

这里A_g是A的第g行。

假定秩（W）＝K°。由假定1，A是行满秩的。由线性代数知识（或参考〔1〕中引理2、1），秩（W）＝$K^\circ \leqslant K$。当C_x是非奇异时，即$K^\circ = K$，从〔1〕中的引理2、1，C_x构成了W的行核基，当C_x是奇异时，这时$K^\circ < K$，A不是W的行核基。因而（可参考〔1〕中引理2、2）存在$K-K^\circ$个与A的行线性无关的方程$xW=0$的解向量$C_1, C_2, \cdots, C_{K-K_0}$使得

$$\bar{A} = \begin{pmatrix} A \\ C_1 \\ C_1 \\ \vdots \\ C_{K-K_0} \end{pmatrix} \tag{12}$$

构成了W的行核基。有

$$\bar{A}W = 0 \tag{13}$$

由〔1〕中的定理2、1，相对于（B）而言，观测信息只能将A_g与不是$\bar{A}$的行向量的线性组合区分开来，而不能将A_g与$\bar{A}$的行向量的线性组合区分开来。这样，如果对（1）式的参数没有任何约束的话，任何一个$(G+K-K^\circ)\times(G+K-K^\circ)$可逆矩阵按照〔2〕的定义都是允许变换矩阵，于是有

$$F\bar{A}W = 0 \tag{14}$$

---

* 这里W和文献〔2，4〕中$W = \begin{bmatrix} \pi \\ I \end{bmatrix}$不一样，π是（1）式的诱导型中的参数。注意$C_x$也许是奇异的。

因此A_g（$g=1、2，\cdots，G$）是完全不能识别的。

作为第二步，如果还知道有A_g约束

$$A_g\Phi_g=d_g \tag{15}$$

式中Φ_g是元素已知的（$G+K$）$\times R_g$阶矩阵；d_g是元素已知的$1\times R_g$阶列向量，于是有

定理1、（秩条件）在假定1—2，及约束（11）和（15）下，A_g能识别的充分必要条件是

$$\text{秩}(\bar{A}\Phi_g)=\begin{cases}G+K-K^{\circ}, & d_g\neq 0;\\ G+K-K^{\circ}-1, & d_g=0。\end{cases}$$

证：见〔1〕中定理2、4。

推论1、（阶条件）在上定理的假定和约束下，A_g能识别的必要条件是

$$R_g\geqslant\begin{cases}G+K-K^{\circ}, & d_g\neq 0;\\ G+K-K^{\circ}-1, & d_g=0。\end{cases}$$

推论2，当C_x非奇异时，在上定理的假定和约束下，A_g能识别的充分必要条件是

$$\text{秩}(A\Phi_g)=\begin{cases}G, & d_g\neq 0;\\ G-1, & d_g=0,\end{cases}$$

A_g能识别的必要条件是

$$R_g\geqslant\begin{cases}G, & d_g\neq 0;\\ G-1, & d_g=0\end{cases}$$

这样当C_x非奇异时，就回到通常的秩条件和阶条件[2,4]。

定理2假如秩（$\bar{A}\Phi_g$）$=\begin{cases}G+K-K^{\circ}, & d_g\neq 0;\\ G+K-K^{\circ}-1, & d\bar{A}=0。\end{cases}$

在定理1中的假定和约束下，A_g能恰好识别的充分必要条件是

$$R_g=\begin{cases}G+K-K^{\circ}, & d_g\neq 0;\\ G+K-K^{\circ}-1, & d_g=0,\end{cases}$$

A_g能过渡识别的充分必要条件是

$$R_g>\begin{cases}G+K-K^{\circ}, & d_g\neq 0;\\ G+K-K^{\circ}-1, & d_g=0。\end{cases}$$

当C_x是非奇异时，这个定理就化为通常的结果[2,4]。

现在利用一步识别法来讨论参数A的识别，它特别对于约束条件涉及不同方程中的参数时适用。将A按行拉直，并用$\bar{A}$表示，即$\bar{A}=〔A_1，A_2，\cdots，A_G〕'$，于是$AW=0$可以重新写为

$$(I_G\otimes W')\bar{A}=0 \tag{16}$$

式中（$I_G\otimes W'$）是$GK\times G(G+K)$阶矩阵；$\otimes$是Kronecker积，即

$$(I_G\otimes W')=\begin{pmatrix}W' & & & \\ & W' & & \\ & & \ddots & \\ & & & W'\end{pmatrix} \tag{17}$$

假定还知道$\bar{A}$具有约束：

$$\Phi \bar{A} = d \qquad (18)$$

式中Φ是元素为已知的$R \times G(G+K)$阶矩阵；d是元素为已知的$R \times 1$列向量且$d \neq 0$。令

$$Q = \begin{pmatrix} I \otimes W' \\ \Phi \end{pmatrix}, \qquad \bar{d} = \begin{pmatrix} 0 \\ d \end{pmatrix}, \qquad (19)$$

有 $Q\bar{A} = \bar{d}$ (20)

定理3、在假定1—2下，$\bar{A}$能识别的充分必要条件是秩$(Q) = G(G+K)$

推论3、在假定1—2和约束(20)下，$\bar{A}$能识别的必要条件是

$$R \geqslant G(G+K-K^{\circ})。$$

推论4、当C_x非奇异时，在假定1—2和约束(20)下，$\bar{A}$能识别的必要条件是$R \geqslant G^2$，这样，得到通常的结果[4]。

当$\bar{A}$不能识别，但$\bar{A}$的部分元素能识别时，可以利用文献〔4〕的证明方法，得到类似〔4〕的结果。

三

所谓多重共线性问题就是外生变量之间，存在着线性关系或近似线性关系。人们已经对如何消除多重共线性进行了大量研究并且有了一些具体的方法和措施[2,5]。

这些方法各有不同的含义。本文从识别的观点来看多重共线性问题，得出了有意义的分析结论。

当C_x非奇异时，将(6)式左乘B^{-1}，右乘C_x^{-1}，有

$$C_{yx}C_x^{-1} = -B^{-1}\Gamma = \pi \qquad (21)$$

这时后验信息(样本观测信息)完全决定了诱导型的参数π，π能够一致地估计。然而当C_x是奇异时，它不满足Promity定理的条件[2]，仅根据样本观测信息，不能得到π的一致估计量。π将对应着无数个它的估计量，从识别的观点看，这时π不能识别。于是得出结论：当C_x奇异时，对π(或A)也要有事前信息，π才能识别，否则不能。

如果$A = (B, \Gamma)$能识别，由$\pi = -B^{-1}\Gamma$，当然π也能识别。当C_x不可逆时，不能利用一般最小二乘法和二段最小二乘法估计π。但当A能识别时，能够结合工具变量和A的事前信息来估计A。π的估计量$\hat{\pi}$也能由$-\hat{B}^{-1}\hat{\Gamma}$来得到，这时$\hat{A}$和$\hat{\pi}$都是一致估计量，于是导致一个重要观点：多重共线性问题可以看作识别问题。只要有足够的事前信息，不管多重共线性多么严重，总能够得到A的一致估计量。也可以从另外一个角度考虑多重共线性问题：如果存在着多重共线性，必定有某些外生变量能够由其余的外生变量来线性表示或近似线性表示。不妨把外生变量看为内生变量。这种新的关系式连同结构式(1)构成了一个新的结构式，因此A的识别就需要更多的先验信息。为了说明上述观点，设有如下模型：

$$P_t = a_0 + a_1 W_t + a_2 L_t + a_3 I_t + e_t \qquad (22)$$

式中P_t，W_t，L_t，I_t和e_t分别表示在时刻t时的产量，工资、劳力、投资和误差项。

假定工资和劳力有完全线性关系：$W_t=KL_t$，则对任意的常数 a_1^*，都给出了相同的值 P_t：

$$P_t=a_0+a_1W_t+a_2L_t+a_3I_t; \tag{23}$$

$$P_t=a_0+(a_1-a_1^*)W_t+(a_2+Ka_1^*)La_{t3}I_t \tag{24}$$

即有无穷多组参数估计量给出相同的 P_t，从识别的观点看，（22）式不能识别。将 $W_t=KL_t$ 代入（22）式，有

$$P_t=a_0+(a_1K+a_2)L_t+a_3I_t+e_t \tag{25}$$

假定剩下的变量没有多重共线性，对（25）式回归，能得到 a_0，$b\triangleq(a_1k+a_2)$ 和 a_3 的一致估计量。于是对（22）式的参数，有约束：

$$a_1k+a_2=b \tag{26}$$

还知道（22）式的参数已规范化（P_t 前面的系数 $a_4\equiv1$）。这时，$G=1$，$K=3$，$K^\circ=2$，且 Φ_1 的秩等于2，而 $G+K-K^\circ=1+3-2=2$，所以由定理1，（22）式在约束（26）下能识别*。

还可以再从另一角度看这个问题，把

$$P_t=a_0+a_1W_t+a_2L_t+a_3I_t+e_{1t} \tag{27}$$

$$W_t=KL_t+e_{2t} \qquad (e_{2t}\equiv0)$$

看成一个联立方程组，于是

$$B=\begin{pmatrix}1 & -a_1\\ 0 & 1\end{pmatrix},\quad \Gamma=\begin{pmatrix}-a_0 & -a_2 & -a_3\\ 0 & -K & 0\end{pmatrix}.$$

显然 $B=1\doteq0$，如果没有任何事前约束，（22）式不能与（27）式的线性组合区分开来，因此(27)式的第一个方程是完全不能识别的，除非加上某些先验信息（例如 $a_1k+a_2=b$）。

参 考 文 献

〔1〕田国强，“计量经济学中联立方程组模型的识别”，全国数量经济学会第一次讨论会论文，1982．2．

〔2〕Fisher, F, M. “The Identification Problem in Econometries”, New york: Mcgraw-Hill, 1966.

〔3〕Hockings, R, R. “The Analysis and Selection of Variables in Linear Regression”, Biometrics, 32, 1—49, 1976.

〔4〕Hsiao, C. “Identification”, Technical Report, No. 311, Stangord, California, 1980

〔5〕Kendall, M, Q. “Multivariate Analgsis, Charles Griffin and Company LTD, London and High Wycombe, 1975.

---

* 既然已经知道 W_t 和 L_t 有完全线性关系，估计时，可以只取 W_t 或 L_t 作为工具变量。