Estimación MC3E, MVIC en Modelos de Ecuaciones Simultáneas

OCW 2013 Marta Regúlez Castillo

Economía Aplicada III (UPV/EHU)

OCW 2013

Contenidos

Información Completa

- Un método de información completa considera estimar conjuntamente el sistema de ecuaciones de la forma estructural.
- Requiere que todas ellas estén identificadas.
- Hay que especificar de forma explícita todas y cada una de las ecuaciones de la forma estructural del sistema.
- Utiliza también la matriz de varianzas y covarianzas de los términos de error estructurales.

Información Completa

- Ventaja de utilizar un método de información completa: en general los estimadores así obtenidos son más eficientes asintóticamente ya que incorporan toda la información del sistema
- Hay situaciones en los que estimar una ecuación con información completa no mejora en eficiencia frente a estimarla de forma aislada con información limitada.
- Desventaja: Si alguna de las ecuaciones del sistema está mal especificada esto produce estimadores no consistentes de los parámetros de otras ecuaciones.

Mínimos Cuadrados en 3 etapas

Supongamos que normalizamos las G ecuaciones cada una con respecto a una de las variables endógenas y consideramos el siguiente sistema de las G ecuaciones:

$$\underbrace{\mathbf{y}_j}_{Tx1} = \underbrace{Y_j}_{TxG_j} \underbrace{\beta_j}_{G_jx1} + \underbrace{X_j}_{TxK_j} \underbrace{\gamma_j}_{K_jx1} + \underbrace{\mathbf{u}_j}_{Tx1} \qquad j = 1, \ldots, G$$

o de forma más compacta

$$\underbrace{\mathbf{y}_j}_{Tx1} = \underbrace{Z_j}_{Tx(G_j+K_j)} \underbrace{\delta_j}_{(G_j+K_j)x1} + \underbrace{\mathbf{u}_j}_{Tx1} \qquad j = 1, \ldots, G$$

siendo

$$Z_j = [Y_j \;\; X_j]\,,\; \delta_j = \begin{bmatrix} \beta_j \\ \gamma_j \end{bmatrix},\; \mathbf{u}_j \sim (\mathbf{0}, \sigma_j^2 I_T),\; E(\mathbf{u}_i\mathbf{u}_j') = \sigma_{ij} I_T\,.$$

Mínimos Cuadrados en 3 etapas

Matricialmente,

$$\underbrace{\mathbf{y}}_{GTx1} = \underbrace{Z}_{GTxJ}\underbrace{\delta}_{Jx1} + \underbrace{\mathbf{u}}_{GTx1}$$

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_G \end{bmatrix} = \begin{bmatrix} Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_G \end{bmatrix}$$

donde $J = \sum_{j=1}^{G}(G_j + K_j)$ y $\mathbf{u} \sim (\mathbf{0}, (\Sigma \bigotimes I_T))$.

$$E(\mathbf{u}\mathbf{u}') = (\Sigma \bigotimes I_T) = \begin{bmatrix} \sigma_1^2 I_T & \sigma_{12} I_T & \cdots & \sigma_{1G} I_T \\ \sigma_{12} I_T & \sigma_2^2 I_T & \cdots & \sigma_{2G} I_T \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{G1} I_T & \sigma_{G2} I_T & \cdots & \sigma_G^2 I_T \end{bmatrix}$$

Mínimos Cuadrados en 3 etapas

Este es un sistema de ecuaciones aparentemente no relacionadas donde en la matriz Z hay variables endógenas que están correlacionadas con los términos de error.

Premultiplicamos cada una de las ecuaciones por la matriz de variables exógenas X que son ortogonales a los errores estructurales:

$$X'\mathbf{y}_j = X'Z_j\delta_j + X'\mathbf{u}_j \quad j = 1, \ldots, G$$

$$\begin{bmatrix} X'\mathbf{y}_1 \\ X'\mathbf{y}_2 \\ \vdots \\ X'\mathbf{y}_G \end{bmatrix} = \begin{bmatrix} X'Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & X'Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & X'Z_G \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_G \end{bmatrix} + \begin{bmatrix} X'\mathbf{u}_1 \\ X'\mathbf{u}_2 \\ \vdots \\ X'\mathbf{u}_G \end{bmatrix}$$

Mínimos Cuadrados en 3 etapas

Posteriormente, aplicamos MCG(F) al sistema de ecuaciones resultante:

$$\underbrace{\mathbf{y}^*}_{KGx1} = \underbrace{Z^*}_{KGxJ}\underbrace{\delta}_{Jx1} + \underbrace{\mathbf{u}^*}_{KGx1}$$

teniendo en cuenta la estructura de varianzas y covarianzas que tiene $\mathbf{u}^*$. Si consideramos a X fija entonces $\mathbf{u}^* \sim (\mathbf{0}, \Sigma \bigotimes (X'X))$. El estimador MCG, si se conociera Σ, vendría definido como

$$\hat{\delta} = \left(Z^{*\prime}(\Sigma^{-1} \bigotimes (X'X)^{-1})Z^*\right)^{-1} \left(Z^{*\prime}(\Sigma^{-1} \bigotimes (X'X)^{-1})\mathbf{y}^*\right)$$

Esta expresión se puede desarrollar de la forma siguiente:

Mínimos Cuadrados en 3 etapas

$$Z^* = \begin{bmatrix} X'Z_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & X'Z_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & X'Z_G \end{bmatrix} = (I \bigotimes X')Z$$

$$\mathbf{y}^* = \begin{bmatrix} X'\mathbf{y}_1 \\ X'\mathbf{y}_2 \\ \vdots \\ X'\mathbf{y}_G \end{bmatrix} = (I \bigotimes X')\mathbf{y}$$

Por lo tanto, podemos desarrollar los siguientes términos:

$$Z^{*\prime}(\Sigma^{-1} \bigotimes (X'X)^{-1}) = Z'(I \bigotimes X')' \left(\Sigma^{-1} \bigotimes (X'X)^{-1}\right)$$

$$Z'(I \bigotimes X) \left(\Sigma^{-1} \bigotimes (X'X)^{-1}\right) = Z' \left(\Sigma^{-1} \bigotimes X(X'X)^{-1}\right)$$

Mínimos Cuadrados en 3 etapas

$$Z^{*\prime}(\Sigma^{-1}\bigotimes(X'X)^{-1})Z^{*} = Z'\left(\Sigma^{-1}\bigotimes X(X'X)^{-1}X'\right)Z = Z'\left(\Sigma^{-1}\bigotimes P_X\right)Z$$

$$Z^{*\prime}(\Sigma^{-1}\bigotimes(X'X)^{-1})\mathbf{y}^{*} = Z'\left(\Sigma^{-1}\bigotimes X(X'X)^{-1}X'\right)\mathbf{y} = Z'\left(\Sigma^{-1}\bigotimes P_X\right)\mathbf{y}$$

$$Z'\left(\Sigma^{-1}\bigotimes P_X\right)Z = W'Z \qquad Z'\left(\Sigma^{-1}\bigotimes P_X\right)\mathbf{y} = W'\mathbf{y}$$

Si no se conoce Σ un posible estimador consistente de sus elementos es

$$\hat{\sigma}_{ij} = \frac{1}{T}\hat{\mathbf{u}}_i'\hat{\mathbf{u}}_j \qquad i = 1, \ldots, G \quad j = 1, \ldots, G$$

donde $\hat{\mathbf{u}}_i = \mathbf{y}_i - Z_i\hat{\delta}_i^{MC2E}$ y $\hat{\delta}^{MC2E}$ es el estimador de los coeficientes de la F.E por MC2E ecuación por ecuación.

Mínimos Cuadrados en 3 etapas

Sustituyendo obtenemos la expresión del estimador de MC3E:

$$\hat{\delta}_{MC3E} = \left(Z'\left(\widehat{\Sigma}^{-1} \bigotimes P_X\right) Z\right)^{-1} \left(Z'\left(\widehat{\Sigma}^{-1} \bigotimes P_X\right) \mathbf{y}\right)$$

El estimador MC3E se puede interpretar como un estimador de VI

$$\hat{\delta}_{MC3E} = (\hat{W}'Z)^{-1}\,\hat{W}'\mathbf{y}$$

donde la matriz de instrumentos es $\hat{W} = Z'\left(\widehat{\Sigma}^{-1} \bigotimes P_X\right)$.

Equivalencia MC2E y MC3E

Los estimadores de MC2E ecuación por ecuación y MC3E coinciden en los siguientes casos:

1) Si $\sigma_{ij} = 0 \quad \forall i \neq j$. Como ocurre en un SURE, si no existe correlación contemporánea entre las perturbaciones de las distintas ecuaciones de la forma estructural (Σ diagonal), entonces no se gana en eficiencia de estimar el sistema conjunto por MCG(F) que estimar separadamente ecuación por ecuación.
2) Si **todas** las ecuaciones del sistema están exactamente identificadas.

La demostración de estos resultados se encuentra en la lectura recomendada "Documento Sarriko-On 05-08".

Máxima Verosimilitud Información Completa

- El método de estimación por Máxima Verosimilitud Información Completa (MVIC) considera la estimación de los parámetros estructurales de todo el sistema conjuntamente.
- El estimador MVIC es la solución al problema de maximizar la función de verosimilitud Y dado X, sujeto a las restricciones que la sobreidentificación de algunas ecuaciones estructurales puedan impone sobre los parámetros de la Forma Reducida. La función de verosimilitud bajo Normalidad de los errores es:

$$L = L(\Pi, \Omega; Y, X) = \prod_{t=1}^{T} L(\Pi, \Omega; \mathbf{y}_t, \mathbf{x}_t) =$$

$$= \prod_{t=1}^{T} \left((2\pi)^{-G/2} |\Omega|^{-1/2} \exp\left[(-1/2) \mathbf{v}_t' \Omega^{-1} \mathbf{v}_t \right] \right) =$$

$$= \prod_{t=1}^{T} \left((2\pi)^{-G/2} |\Omega|^{-1/2} \exp\left[(-1/2)(\mathbf{y}_t - \Pi \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \Pi \mathbf{x}_t) \right] \right)$$

Máxima Verosimilitud Información Completa

Tomando logaritmos, obtenemos

$$lnL = -((TG/2)\, ln\, 2\pi) - (T/2)\, ln\, |\Omega| \; - 1/2 \left[\sum_{t=1}^{T} (\mathbf{y}_t - \Pi \mathbf{x}_t)' \Omega^{-1} (\mathbf{y}_t - \Pi \mathbf{x}_t) \right] =$$

$$= -((TG/2)\, ln\, 2\pi) - (T/2)\, ln\, |\Omega| \; - 1/2 \left[\sum_{t=1}^{T} \mathbf{v}_t' \Omega^{-1} \mathbf{v}_t \right] =$$

$$= -(T/2) \left[G\, ln\, 2\pi + \; ln\, |\Omega| \; + \; tr \left(\Omega^{-1}\, W \right) \right]$$

donde $W = \frac{1}{T} \sum_{t=1}^{T} \mathbf{v}_t\, \mathbf{v}_t' = \frac{1}{T} \sum_{t=1}^{T} (\mathbf{y}_t - \Pi \mathbf{x}_t)(\mathbf{y}_t - \Pi \mathbf{x}_t)'$. Sustituyendo

$\Pi = -B^{-1}\Gamma$ y $\Omega = B^{-1}\Sigma(B^{-1})'$ por lo que $\Omega^{-1} = B'\Sigma^{-1}B$ se obtiene

$$lnL = -(1/2)\; TG\;\; ln\, (2\pi) - 1/2\, Tln |B^{-1}\Sigma(B^{-1})'| - (1/2)\, Ttr \left(B'\Sigma^{-1} B\, W \right)$$

donde $W = \frac{1}{T} \sum_{t=1}^{T} (\mathbf{y}_t + B^{-1}\Gamma \mathbf{x}_t)(\mathbf{y}_t + B^{-1}\Gamma \mathbf{x}_t)'$.

Máxima Verosimilitud Información Completa

- El estimador MVIC maximiza esta función sujeto a todas las restricciones sobre B, Γ y sobre la matriz de varianzas y covarianzas Σ si las hubiera.
- Esto se puede llevar a cabo calculando las derivadas parciales de lnL con respecto a los parámetros que quedan sin restringir e igualar estos gradientes a cero.
- En general, salvo algunas excepciones, la resolución analítica de este sistema para todos los parámetros no es factible ya que es un problema altamente no lineal, y por lo tanto se requiere de la utilización de métodos de optimización numérica.

Máxima Verosimilitud Información Completa

Dos excepciones son las siguientes:

1) El modelo es recursivo. En ese caso B es triangular con unos en la diagonal principal y Σ diagonal. Se puede demostrar que el estimador MVIC en este caso coincide con el estimador MCO de cada ecuación estructural.
2) Si todas las ecuaciones del sistema están exactamente identificadas. En este caso, hay una correspondencia uno a uno entre (B, Γ) y Π. Al no haber restricciones sobre Π debidas a la sobreidentificación, MVIC se obtiene estimando Π por MCO y resolviendo para B, Γ de la relación $B\Pi + \Gamma = 0$. Apelando al principio de invarianza de los estimadores máximo-verosímiles, en un modelo exactamente identificado bajo Normalidad,

$$MCI = MC2E = MVIL = MC3E = MVIC$$