



Econometría II

Regresores estocásticos

Miguel Jerez y Sonia Sotoca

Universidad Complutense de Madrid

Marzo 2010

Índice

- Causas
- Consecuencias sobre la estimación MCO
- Estimador de Variables Instrumentales
- Ejemplos
- Anexo: Resultados fundamentales de teoría asintótica

Causas (I): Planteamiento

En Econometría I se planteó el modelo:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} ; \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

En donde se supone que los regresores incluidos en \mathbf{X} son variables deterministas

Se dice que **una variable es determinista si sus valores no cambian en distintas muestras**

Cuando no puedan repetirse las observaciones - como ocurre a menudo en Economía - no es trivial saber si una variable es o no aleatoria. Por ello es importante analizar si el posible incumplimiento de esta hipótesis afecta gravemente a las propiedades de la estimación MCO.

Por otra parte, muchos modelos econométricos implican, por su propia estructura, que algunos regresores son aleatorios. Por ejemplo:

- modelos de ecuaciones simultáneas, o
- modelos dinámicos

En la práctica econométrica, sólo se reconoce que existe este problema cuando la estructura del modelo implica la existencia de regresores estocásticos.

Causas (II): Ecuaciones simultáneas y modelos dinámicos

Los **sistemas de ecuaciones simultáneas** describen el comportamiento de un vector de variables endógenas en función de un vector de variables exógenas. Los regresores estocásticos surgen del hecho de que la variable endógena de una ecuación puede entrar en otra como variable explicativa.

Ejemplo. El modelo de consumo de Haavelmo se define como:

$$\begin{aligned} C_t &= \beta PIB_t + \varepsilon_t \\ PIB_t &= C_t + I_t \end{aligned} \quad \text{o bien,} \quad \begin{bmatrix} 1 & -\beta \\ -1 & 1 \end{bmatrix} \begin{bmatrix} C_t \\ PIB_t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} I_t + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

e implica que PIB_t es un regresor estocástico en la primera ecuación ya que, por la segunda ecuación PIB_t depende de C_t y, por la primera ecuación, C_t depende de ε_t

La variable endógena de un modelo de regresión es siempre estocástica. Por ello en un **modelo de regresión dinámico**, que incluya la variable endógena retardada como variable explicativa, siempre hay regresores estocásticos.

Ejemplo. El modelo AR(1) es un modelo de regresión con variables explicativas estocásticas: $y_t = \phi y_{t-1} + \varepsilon_t$

Hay otras situaciones en las que es necesario reconocer que el MLG tiene regresores estocásticos. Cabe destacar las de variables medidas con error o modelos que incluyen expectativas como regresores.

Consecuencias sobre la estimación MCO (I)

La consecuencia fundamental de la presencia de regresores estocásticos en el MLG, es que resulta imposible obtener los momentos del estimador MCO.

Como ya se vio en Econometría I:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} = \beta + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \varepsilon \Rightarrow E(\hat{\beta}) = \beta + E[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \varepsilon]$$

y para que el estimador fuera insesgado, la esperanza del segundo sumando tendría que ser nula. Sin embargo, como es una función no lineal de variables aleatorias, su esperanza no puede determinarse.

Por tanto, **cuando hay regresores estocásticos no es posible determinar la esperanza ni la varianza del estimador MCO**. Esto impide:

- justificar la insesgadez y eficiencia del estimador MCO, y
- determinar la distribución de los estadísticos de contraste bajo la hipótesis nula

Ello nos obliga a estudiar “propiedades asintóticas”, esto es, las que tendría el estimador en una muestra con infinitas observaciones.

La propiedad asintótica fundamental es la de “consistencia”.

Consecuencias sobre la estimación MCO (II)

Supongamos que todas las variables explicativas están centradas. Trivialmente:

$$\hat{\beta} = \beta + (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^T \varepsilon \Rightarrow \text{plim } \hat{\beta} = \beta + \text{plim} \left(\frac{1}{n} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} \right)^{-1} \left(\frac{1}{n} \tilde{\mathbf{X}}^T \varepsilon \right)$$

y, aplicando sucesivamente los teoremas de Slutsky y Kinchine (ver Apéndice):

$$\text{plim } \hat{\beta} = \beta + \text{plim} \left(\frac{1}{n} \tilde{\mathbf{X}}^T \tilde{\mathbf{X}} \right)^{-1} \text{plim} \left(\frac{1}{n} \tilde{\mathbf{X}}^T \varepsilon \right) = \beta + \Sigma_{\mathbf{X}\mathbf{X}}^{-1} \Sigma_{\mathbf{X}\varepsilon}$$

por lo que **el estimador MCO es consistente si y sólo si los regresores no están correlados con el término de error.**

Es importante tener en cuenta que las propiedades de consistencia e insesgadez son distintas e independientes. Concretamente, es posible que un estimador sea:

- **sesgado y consistente** (p. ej. la varianza muestral),
- **insesgado e inconsistente** (p. ej. $y_i = \tilde{\beta}x_i + \varepsilon_i$; $\tilde{\beta} = y_i / x_i$),
- **sesgado e inconsistente**, (p. ej. el estimador anterior más un sesgo deliberado), o
- **insesgado y consistente** (p. ej. el estimador MCO bajo las hipótesis habituales)

Estimador de Variables Instrumentales (I)

Supongamos que se quiere estimar el modelo $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, en donde \mathbf{X} es una matriz de variables aleatorias tales que $plim \frac{1}{n} \mathbf{X}^T \varepsilon \neq 0$

Asimismo, sea \mathbf{Z} una matriz ($n \times k$) de variables instrumentales (VI) tal que:

- 1) Los instrumentos son independientes del término de error: $plim \frac{1}{n} \mathbf{Z}^T \varepsilon = 0$
- 2) ... están correlados con los regresores: $plim \frac{1}{n} \mathbf{Z}^T \mathbf{X} = \Sigma_{zx} / |\Sigma_{zx}| \neq 0$
- 3) ... y su covarianza con \mathbf{y} es finita: $plim \frac{1}{n} \mathbf{Z}^T \mathbf{y} = \Sigma_{zy}$

en donde 3) es una condición de regularidad que evita resultados absurdos. En estas condiciones, la expresión del estimador de β mediante variables instrumentales, actuando \mathbf{Z} como matriz de instrumentos de \mathbf{X} , es:

$$\hat{\beta}_{VI} = (\mathbf{Z}^T \mathbf{X})^{-1} \mathbf{Z}^T \mathbf{y}$$

y este estimador es consistente, ya que:

$$\hat{\beta}_{VI} = \beta + (\mathbf{Z}^T \mathbf{X})^{-1} \mathbf{Z}^T \varepsilon \Rightarrow plim \hat{\beta} = \beta + plim \left(\frac{1}{n} \mathbf{Z}^T \mathbf{X} \right)^{-1} \left(\frac{1}{n} \mathbf{Z}^T \varepsilon \right) = \beta$$

Estimador de Variables Instrumentales (II)

Por otra parte, la matriz de covarianzas del estimador VI es:

$$\mathbf{cov}_A(\hat{\beta}_{VI}) = \sigma_\varepsilon^2 \frac{1}{n} (\Sigma_{ZX})^{-1} \Sigma_{ZZ} (\Sigma_{XZ})^{-1}$$

... y puede estimarse mediante la expresión:

$$\mathbf{cov}(\hat{\beta}_{VI}) = \hat{\sigma}_\varepsilon^2 (\tilde{\mathbf{Z}}^T \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{Z}}^T \tilde{\mathbf{Z}} (\tilde{\mathbf{X}}^T \tilde{\mathbf{Z}})^{-1}$$

La forma de la matriz de covarianzas pone de manifiesto que **cuanto mayor sea la correlación entre los instrumentos y las variables explicativas, mayor será la eficiencia del estimador VI.**

Por ello, si tenemos k regresores y sólo algunos (digamos que los s primeros) tiene correlación con las perturbaciones, la matriz de instrumentos \mathbf{Z} se construirá:

- sustituyendo las s primeras columnas de \mathbf{X} por los correspondientes instrumentos y
- manteniendo los restantes regresores, que actuarán como instrumentos de sí mismos.

Ejemplos (I): Ecuaciones simultáneas

Consideremos de nuevo el modelo de consumo de Haavelmo, dado por:

$$C_t = \beta PIB_t + \varepsilon_t \quad (1)$$

$$PIB_t = C_t + I_t \quad (2)$$

Para ver que PIB_t es un regresor estocástico de (1), correlado con el término de error, basta sustituir (1) en (2):

$$PIB_t = \beta PIB_t + \varepsilon_t + I_t \Leftrightarrow PIB_t = \frac{I_t}{1-\beta} + \frac{\varepsilon_t}{1-\beta} \quad (3)$$

por tanto $E\left[\left(PIB_t - \frac{I_t}{1-\beta}\right)\varepsilon_t\right] = \frac{\sigma_\varepsilon^2}{1-\beta} \neq 0$ y el estimador MCO de (1) es inconsistente

Para estimar consistentemente (1) puede usarse el estimador VI, usando I_t como instrumento de PIB_t . La expresión del estimador quedaría:

$$\hat{\beta}_{VI} = \frac{\sum I_t \cdot C_t}{\sum I_t \cdot PIB_t}$$

Ejemplos (II): Modelos dinámicos

Sea el modelo dinámico:

$$y_t = \varphi y_{t-1} + \beta x_t + \varepsilon_t \quad (1)$$

con los siguientes términos de error:

$$\varepsilon_t = a_t \quad (2.a)$$

$$\varepsilon_t = a_t - \theta a_{t-1} \quad (2.b)$$

$$\varepsilon_t = \phi \varepsilon_{t-1} + a_t \quad (2.c)$$

En este caso:

- La estimación del modelo (1) por MCO es consistente si el error es (2.a)
- La estimación del modelo (1) por MCO es inconsistente si el error es (2.b) o (2.c), ya que y_{t-1} tiene covarianza no nula con a_t
- El modelo (1)-(2.b) puede estimarse consistentemente usando x_{t-1} como instrumento de y_{t-1} y la variable x_t como instrumento de sí misma. Alternativamente, y_{t-2} es un instrumento válido de y_{t-1} .
- El modelo (1)-(2.c) puede estimarse consistentemente usando x_{t-1} como instrumento de y_{t-1} y la variable x_t como instrumento de sí misma. No hay ningún instrumento válido para estimar por VI el modelo si $\beta = 0$

Anexo: Resultados fundamentales de teoría asintótica (I)

Definición (convergencia en probabilidad): Sea $\{\hat{\beta}_n\}$ una sucesión de variables aleatorias definida en el conjunto de números reales. Se dice que $\{\hat{\beta}_n\}$ converge en probabilidad a β si y sólo si $\exists \delta > 0 / \lim_{n \rightarrow \infty} P(\|\hat{\beta}_n - \beta\| < \delta) = 1$

Otra forma de expresar matemáticamente este concepto es $plim \hat{\beta}_n = \beta$, expresión que se lee “el límite en probabilidad de $\hat{\beta}_n$ es β ”.

Definición (Estimador consistente): Si $\{\hat{\beta}_n\}$ denota una secuencia de estimaciones correspondientes a muestras de tamaño n , se dice que $\hat{\beta}$ es un estimador consistente de β si $plim \hat{\beta}_n = \beta$

Teorema: Una condición suficiente (no necesaria) para que $\hat{\beta}$ sea un estimador consistente de β es que $\lim_{n \rightarrow \infty} E[\hat{\beta}(n)] = \beta$ y $\lim_{n \rightarrow \infty} var[\hat{\beta}(n)] = 0$

Por tanto, el límite en probabilidad es un operador que devuelve el límite asintótico de una secuencia de variables aleatorias.

Su ventaja fundamental en comparación con el operador esperanza matemática es que, al estar basado en límites (en vez de en integrales) tiene propiedades mucho más flexibles.

Anexo: Resultados fundamentales de teoría asintótica (II)

Si $plim \hat{\alpha}_n = \alpha$ y $plim \hat{\beta}_n = \beta$ entonces:

1) $plim \gamma = \gamma$, siendo γ una constante (no depende de n)

2) $plim [\hat{\alpha}_n \pm \hat{\beta}_n] = \alpha \pm \beta$

3) $plim [\hat{\alpha}_n \cdot \hat{\beta}_n] = \alpha \cdot \beta$

4) $plim \left[\frac{\hat{\alpha}_n}{\hat{\beta}_n} \right] = \frac{\alpha}{\beta}$ si $\hat{\beta}_n \neq 0$ y $\beta \neq 0$

5) $plim \sqrt{\hat{\alpha}_n} = \sqrt{\alpha}$ si $\hat{\alpha}_n > 0$ y $\alpha > 0$

...

6) $plim \phi[\hat{\alpha}_n] = \phi(\alpha)$; $\phi(\)$: cualquier función continua (Teorema de Slutsky)

7) Si m_n^r es el momento muestral de orden r calculado con una muestra de n observaciones y m es el correspondiente momento poblacional, entonces $plim m^r = m$ (Teorema de Kinchine)

Estas propiedades se cumplen tanto para escalares como para vectores