Unidad 3. Modelo de RLS: Propiedades de los estimadores MCO y Teorema de Gauss Markov

Erika R. Badillo
erika.badilloen@unaula.edu.co

Facultad de Economía
Universidad Autónoma Latinoamericana

En este tema

- Un repaso del modelo de RLS
- El Teorema de Gauss-Markov
- Estimador insesgado de la varianza del error
- El Error Cuadrático Medio (ECM)

Referencias

- Wooldridge, Jeffrey (2013). *Introducción a la econometría*. 5a edición, Cengage Learning. Cap. 3
- Gujarati, D. y Porter, D. (2010). *Econometría*. 5a edición, Mc Graw Hill. Cap. 3

Repaso del modelo de RLS

El modelo de RLS tiene la siguiente especificación:

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

Supuestos:

- β_0 y β_1 son coeficientes fijos (parámetros) a estimar
- $E(u_i) = 0 \Longrightarrow$ Modelo completo
- $Var(u_i) = E(u_i^2) = \sigma_u^2 \Longrightarrow$ Homocedasticidad
- $Cov(u_i, u_j) = E(u_i u_j) = 0,\ i \neq j \Longrightarrow$ No autocorrelación
- $Cov(u_i, X_i) = E[(X_i - E(X_i))]E(u_i) = E(X_i u_i) = 0 \Longrightarrow$ Exogeneidad
- $u_i \sim NID(0; \sigma_u^2) \Longrightarrow$ Normalidad

Estimadores del modelo de RLS:

- Minimizando la SCR: encontrar $\widehat{\beta}_0$ y $\widehat{\beta}_1$ que minimicen la SCR

$$\sum \widehat{u}_i^2 = \sum (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)^2$$
$$\frac{\partial \sum \widehat{u}_i^2}{\partial \widehat{\beta}_0} = -2\sum (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i) = 0$$
$$\sum Y_i = n\widehat{\beta}_0 + \widehat{\beta}_1 \sum X_i \text{ (1)}$$
$$\frac{\partial \sum \widehat{u}_i^2}{\partial \widehat{\beta}_1} = -2\sum (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 X_i)X_i = 0$$
$$\sum X_i Y_i = \widehat{\beta}_0 \sum X_i + \widehat{\beta}_1 \sum X_i^2 \text{ (2)}$$

Resolviendo las ecuaciones normales ((1) y (2)) se obtienen las estimaciones de β_0 y β_1:

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$$
$$\widehat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

Los estimadores MCO son los mejores estimadores lineales insesgados (MELI). Miremos entonces qué significa cada componente del acrónimo MELI:

- Ya se sabe que son estimadores: son reglas que pueden aplicarse a cualquier muestra de datos para obtener una estimación
- Ya se sabe que son estimadores insesgados: $E(\widehat{\beta}_0) = \beta_0$ y $E(\widehat{\beta}_1) = \beta_1$
- ¿Qué significa el término lineal? Un estimador es lineal si, y sólo si, se puede expresar como una función lineal de los datos de la variable dependiente:

  $$\widehat{\beta}_j = \sum_{i=1}^{n} w_{ij} Y_i$$

  donde cada w_{ij} puede ser una función de los valores muestrales de todas las variables independientes. Los estimadores MCO son lineales
- ¿Cómo se define mejor? Mejor se define como menor varianza. Dado dos estimadores insesgados, es lógico preferir el que tenga menor varianza

El Teorema de Gauss-Markov

- El teorema de Gauss-Markov justifica el uso del método de MCO en lugar de otros diversos estimadores

- Bajo los supuestos estándar del modelo de RLS, el método de MCO es insesgado. Sin embargo, bajo estos supuestos hay muchos estimadores insesgados de β_0 y β_1. Entonces ¿Cuál estimador elegir? $\Longrightarrow$ el de menor varianza, pero ¿hay otros estimadores cuyas varianzas sean menores que las de los estimadores de MCO?

- Si se limita la clase de los posibles estimadores apropiados, entonces se puede demostrar que el método de MCO es el mejor dentro de su clase

Teorema de Gauss-Markov

Dado cualquier estimador $\widetilde{\beta}$ que sea lineal e insesgado, $Var(\widehat{\beta}_j) \leq Var(\widetilde{\beta}_j)$ y esta desigualdad es por lo general estricta. Es decir, en la clase de los estimadores lineales insesgados, los estimadores MCO tienen la mínima varianza (bajo los supuestos del modelo de RLS)

El Teorema de Gauss-Markov

- La importancia del teorema de Gauss-Markov es que, si el conjunto estándar de supuestos se satisface, no es necesario buscar otros estimadores insesgados: ninguno será mejor que los estimadores MCO
- En otras palabras, que para cualquier otro estimador que sea lineal e insesgado, su varianza será por lo menos tan grande como la varianza de los estimadores MCO
- El teorema justifica el uso del método MCO para estimar los parámetros del modelo de RLS. Si no se satisface alguno de los supuestos del modelo de RLS entonces el teorema no es válido

El Teorema de Gauss-Markov

Demostración del teorema de Gauss-Markov para $\widehat{\beta}_1$ en el modelo de RLS

La intuición de la demostración es la siguiente:

1. Se define un estimador cualquiera de β_1 que sea lineal en Y_i, por ejemplo sea $\widetilde{\beta}_1$
2. Se encuentran las condiciones para que sea insesgado $\Longrightarrow E(\widetilde{\beta}_1) = \beta_1$
3. Se obtiene la varianza del estimador general $\Longrightarrow Var(\widetilde{\beta}_1)$
4. Se minimiza la varianza de $\widetilde{\beta}_1$ sujeta a las restricciones de ser insesgado
5. Se utilizan multiplicadores de Lagrange

El Teorema de Gauss-Markov

De manera similar a lo realizado con $\widehat{\beta}_1$ puede demostrarse iguales propiedades para $\widehat{\beta}_0$

$$\widehat{\beta}_0 = \sum g_i Y_i \text{ donde } g_i = \tfrac{1}{n} - \overline{X} C_i$$
$$\widehat{\beta}_0 = \beta_0 + \sum g_i u_i$$
$$E(\widehat{\beta}_0) = \beta_0$$
$$Var(\widehat{\beta}_0) = \frac{\sigma_u^2 \sum X_i^2}{n \sum x_i^2}$$

Esta varianza es la mínima posible para la clase de estimadores, lineales e insesgados de β_0

Estimador insesgado de la varianza del error

Para efectos prácticos sólo hace falta encontrar un estimador insesgado para la varianza del error:

$$Var(\widehat{\beta}_1) = \frac{\sigma_u^2}{\sum x_i^2} = \frac{\sigma_u^2}{\sum (X_i - \overline{X})^2}$$
$$Var(\widehat{\beta}_0) = \frac{\sigma_u^2 \sum X_i^2}{n \sum x_i^2} = \frac{\sigma_u^2 \sum X_i^2}{n \sum (X_i - \overline{X})^2}$$

Recuerde que el estimador de σ_u^2 por el método de los momentos y por máxima verosimilitud es:

$$\widehat{\sigma}_u^2 = \frac{\sum \widehat{u}_i^2}{n}$$

Pero es este estimador insesgado? Es posible comprobar que

$$E(\widehat{\sigma}_u^2) = E\left(\frac{\sum \widehat{u}_i^2}{n}\right) = \frac{(n-2)\sigma_u^2}{n}$$

luego, es sesgado

De lo anterior, se puede deducir que un estimador insesgado de la varianza del error es:

$$\widehat{\sigma}_u^2 = \frac{\sum \widehat{u}_i^2}{n-2}$$

El Error Cuadrático Medio (ECM)

Qué hacer cuando se tienen dos estimadores:

- uno insesgado con determinada varianza y
- uno sesgado con una varianza menor

$$\left.\begin{array}{c} \widehat{\beta}_1 \text{ es insesgado } E(\widehat{\beta}_1) = \beta_1 \\ \widetilde{\beta}_1 \text{ es sesgado } E(\widetilde{\beta}_1) \neq \beta_1 \end{array}\right\} \text{Var}(\widehat{\beta}_1) > Var(\widetilde{\beta}_1)$$

Se requiere un criterio estadístico que convine varianza y sesgo. El criterio se llama Error Cuadrático Medio (ECM). Si $\widetilde{\theta}$ es estimador de θ entonces el ECM será:

$$\begin{aligned} ECM(\widetilde{\theta}) &= Var(\widetilde{\theta}) + (Sesgo(\widetilde{\theta}))^2 \\ &= Var(\widetilde{\theta}) + [E(\widetilde{\theta}) - \theta]^2 \end{aligned}$$

Lo que hay que hacer es elegir aquel estimador que tiene el menor ECM. Esto puede interpretarse como un análisis costo-beneficio para un economista:

$$ECM(\widehat{\beta}_1) = Var(\widehat{\beta}_1) \text{ por ser } \widehat{\beta}_1 \text{ insesgado}$$
$$ECM(\widetilde{\beta}_1) = Var(\widetilde{\beta}_1) + (Sesgo(\widetilde{\beta}_1))^2 \text{ por ser } \widetilde{\beta}_1 \text{ sesgado}$$

El Error Cuadrático Medio (ECM)

Al compararse los dos estimadores:

$$ECM(\widehat{\beta}_1) - ECM(\widetilde{\beta}_1) = Var(\widehat{\beta}_1) - \underbrace{Var(\widetilde{\beta}_1)}_{\text{ganancia por menor Var}} - \underbrace{(Sesgo(\widetilde{\beta}_1))^2}_{\text{pérdida por sesgo}}$$

Si la ganancia supera la pérdida:

$$ECM(\widehat{\beta}_1) > ECM(\widetilde{\beta}_1) \Longrightarrow \text{se elige } \widetilde{\beta}_1 \text{ el sesgado}$$

Si la ganancia es menor que la pérdida:

$$ECM(\widehat{\beta}_1) < ECM(\widetilde{\beta}_1) \Longrightarrow \text{se elige } \widehat{\beta}_1 \text{ el insesgado}$$