

Pauta Auxiliar 5 - Semestre Otoño 2016

26 de Abril, 2016

Problema 1: Variables en forma funcional

1. Escriba la fórmula de la varianza del estimador β_{MCO} para un modelo con variable irrelevante. Primero de forma matricial y luego de acuerdo a la fórmula. Concluya respecto a qué sucede con la varianza al agregar variables irrelevantes.

R.

1. Para la forma matricial, usaremos las matrices aniquiladoras. Recordemos que la matriz aniquiladora de Z se define como: $M_Z = I - Z'(Z'Z)^{-1}Z'$ y que cumple la propiedad de ser idempotente: $M_Z' M_Z = M_Z$. Para un modelo poblacional $Y = X\beta + U$, pero que se estima un modelo con una variable irrelevante tal que $Y = X\beta_1 + Z\beta_2 + V$, se puede ver que:

$$M_Z Y = M_Z X \beta_1 + M_Z Z \beta_2 + M_Z V = M_Z X \beta_1 + M_Z V \quad (1)$$

Y al calcular este estimador:

$$\hat{\beta}_1 = (X' M_Z Z)^{-1} X' M_Z Y \quad (2)$$

$$\hat{\beta}_1 = (X' M_Z X)^{-1} X' M_Z (X\beta + U) \quad (3)$$

$$\hat{\beta}_1 = \beta + (X' M_Z X)^{-1} X' M_Z U \quad (4)$$

Ahora bien, al calcular la varianza, se debe tener en cuenta que bajo supuesto de no endogeneidad y el de homocedasticidad: $Var(U|X, Z) = E[U'U|X, Z] - E[U|X, Z]^2 = E[U'U|X, Z] = \sigma^2$

$$Var(\hat{\beta}_1|X, Z) = E[(X' M_Z X)^{-1} X' M_Z U' U M_Z X (X' M_Z X)^{-1} | X, Z] \quad (5)$$

$$= (X' M_Z X)^{-1} X' M_Z E[U'U|X, Z] M_Z X (X' M_Z X)^{-1} \quad (6)$$

$$= \sigma^2 (X' M_Z X)^{-1} > \sigma^2 (X' X)^{-1} \quad (7)$$

Por lo que la varianza crece.

2. Para visualizarlo desde la fórmula de varianza multivariable:

$$Var(\hat{\beta}_j) = \frac{Var(Y)(1 - R^2)}{(N - K)Var(X_j)(1 - R_{j,-j}^2)} \quad (8)$$

Aquí uno puede notar que, si agrega variables irrelevantes al modelo, por construcción el $R_{j,-j}^2$ crece, pues representa la explicación de todas las variables explicativas que no son j

en una regresión sobre la variable X_j . Así, al aumentar ese valor, la varianza del estimador también aumentara.

En conclusión, al agregar variables irrelevantes, la varianza del estimador aumenta, lo que dificulta la inferencia con precisión.

2. Muestre y explique por qué una variable omitida en un modelo sesga el estimador MCO y lo hace inconsistente.

R.

La omisión de variables se refiere a estimar a través de una regresión que le falte una de las variables explicitadas en el modelo real. Es decir, para el modelo poblacional (desvíos respecto a la media):

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + u \quad (9)$$

se realiza una regresión a través de MCO pero omitiendo la variable x_2 , es decir se estiman los coeficientes β para el modelo:

$$y_i = \alpha_1 x_{i1} + v \quad (10)$$

Por lo tanto, el estimador vendrá dado por la misma fórmula, pero reemplazando con el modelo real:

$$\begin{aligned} \alpha &= \frac{\sum_i x_{i1} y_i}{\sum_i x_{i1}^2} \\ \alpha &= \frac{\sum_i x_{i1} (\beta_1 x_{i1} + \beta_2 x_{i2} + u)}{\sum_i x_{i1}^2} \\ \alpha &= \frac{\beta_1 \sum_i x_{i1}^2 + \beta_2 \sum_i x_{i1} x_{i2} + \sum_i x_{i1} u}{\sum_i x_{i1}^2} \\ \alpha &= \beta_1 + \beta_2 \frac{\sum_i x_{i1} x_{i2}}{\sum_i x_{i1}^2} + \frac{\sum_i x_{i1} u}{\sum_i x_{i1}^2} \end{aligned}$$

Lo que al tomar valor esperado, (aun se puede asumir que $E(u) = 0$) queda que:

$$\mathbb{E}(\alpha) = \beta_1 + \beta_2 \frac{\sum_i \mathbb{E}(x_{i1}, x_{i2})}{\sum_i x_{i1}^2}$$

Donde se puede ver el sesgo que no desaparece, producto de la relación entre x_{i1} y x_{i2} . A nivel asintótico, esto se convierte en:

$$\mathbb{E}(\alpha) = \beta_1 + \beta_2 \frac{\sum_i Cov(x_1, x_2)}{\sum_i x_{i1}^2}$$

Por lo tanto, se puede ver que en (2), la variable omitida pasa a estar en el error v . Dado esto, se tiene entonces por lo visto, que si existe covarianza no nula entre x_1 y x_2 , entonces existirá también covarianza no nula entre x_1 y v . Por lo tanto, **la variable omitida genera $Cov(x_1, v) \neq 0$, lo que se ve claramente como endogeneidad en el modelo (2).**

3. Concluya respecto a cómo es preferible abordar un modelo.

R. Se debe notar que omitir variables es más preocupante que agregar variables irrelevantes, por lo que por precaución es preferible comenzar con un modelo de muchas variables, y luego ir quitándolas a medida que el modelo lo requiera. Si no, al ir sesgando el estimador, se puede tomar una dirección equivocada omitiendo variables desde el principio. Métodos para realizar una buena forma funcional consisten en LASSO, o ir testeando mediante test de Ramsey.

Problema 2: Series de Tiempo y modelo AR(1)

Considere un modelo AR(1):

$$Y_t = \Phi Y_{t-1} + \varepsilon_t \quad (11)$$

Asuma que ε_t corresponde a un ruido blanco (con media 0 y varianza σ_ε^2),

1. Calcule una expresión para el estimador $\hat{\Phi}_{MCO}$.
2. Analice sesgo y consistencia del estimador $\hat{\Phi}_{MCO}$, para esto tenga en cuenta escribir la recursión de la variable Y_t y además :

$$\text{I. } \mathbb{E}(u'_t u_s) = \begin{cases} \sigma_\varepsilon^2 & \text{si } t = s \\ 0 & \text{si } t \neq s \end{cases}$$

3. Comente sobre alguna aplicación real de modelos AR(1) y la conveniencia de usar rezagos en regresores.

R.

R: 1.

Conociendo la expresión del estimador MCO, y utilizando $X_t = Y_{t-1}$ se puede expresar:

$$\begin{aligned}
\hat{\Phi}_{MCO} &= (X_i' X_i)^{-1} X_i' Y_i \\
&= (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' Y_t \\
&= (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' (\Phi Y_{t-1} + \varepsilon_t) \\
&= \Phi (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' Y_{t-1} + (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' \varepsilon_t \\
&= \Phi + (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' \varepsilon_t
\end{aligned}$$

R: 2.

Para visualizar sesgo, se debe tomar el valor esperado del estimador:

$$\begin{aligned}
\mathbb{E}(\hat{\Phi}_{MCO} | Y_{t-1}) &= \mathbb{E}(\Phi + (Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' \varepsilon_t) \\
&= \Phi + \mathbb{E}((Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' \varepsilon_t)
\end{aligned}$$

Ahora, hay que evaluar qué sucede con ese término para determinar si es un sesgo. Para determinar esa relación, se realizará la recursión de la variable Y_t :

$$\begin{aligned}
Y_1 &= \Phi_1 Y_0 + \varepsilon_1 \\
Y_2 &= \Phi_1 Y_1 + \varepsilon_2 = \Phi_1 (\Phi_1 Y_0 + \varepsilon_1) + \varepsilon_2 = \Phi_1^2 Y_0 + \varepsilon_2 + \Phi_1 \varepsilon_1 \\
Y_3 &= \Phi_1 Y_2 + \varepsilon_3 = \Phi_1 (\Phi_1^2 Y_0 + \varepsilon_2 + \Phi_1 \varepsilon_1) + \varepsilon_3 = \Phi_1^3 Y_0 + \Phi_1^2 \varepsilon_1 + \Phi_1 \varepsilon_2 + \varepsilon_3 \\
&\vdots \\
Y_t &= \Phi_1^t Y_0 + (\varepsilon_t + \Phi_1 \varepsilon_{t-1} + \Phi_1^2 \varepsilon_{t-2} + \dots + \Phi_1^{t-1} \varepsilon_1)
\end{aligned}$$

Notar que al multiplicar la variable Y_t por diferentes errores temporales $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ y tomando los valores esperados desaparecen todos los términos en que $\mathbb{E}(\varepsilon_t, \varepsilon_s) = 0$ si $t \neq s$.

$$\begin{aligned}
\mathbb{E}(Y_t \varepsilon_t) &= 0 + (\mathbb{E}(\varepsilon_t \varepsilon_t) + 0 + 0 + \dots + 0) = \sigma_\varepsilon^2 \\
\mathbb{E}(Y_t \varepsilon_{t-1}) &= 0 + (0 + \Phi_1 \mathbb{E}(\varepsilon_{t-1} \varepsilon_{t-1}) + 0 + \dots + 0) = \Phi_1 \sigma_\varepsilon^2 \\
\mathbb{E}(Y_t \varepsilon_{t-2}) &= 0 + (0 + 0 + \Phi_1^2 \mathbb{E}(\varepsilon_{t-2} \varepsilon_{t-2}) + \dots + 0) = \Phi_1^2 \sigma_\varepsilon^2 \\
&\vdots =
\end{aligned}$$

Este resultado muestra que la variable dependiente se relaciona con todos los errores de tiempos anteriores y el actual. Realizando lo mismo para Y_{t-1} se encuentra que ésta se correlaciona con los errores de tiempos anteriores pero no con el actual (ε_t). Es decir,

$$\begin{aligned}
\mathbb{E}(Y_{t-1}\varepsilon_t) &= 0 \\
\mathbb{E}(Y_{t-1}\varepsilon_{t-1}) &= \sigma_\varepsilon^2 \\
\mathbb{E}(Y_{t-1}\varepsilon_{t-2}) &= \Phi_1\sigma_\varepsilon^2 \\
&\vdots =
\end{aligned}$$

Luego, se puede ver que aunque no exista una endogeneidad propiamente tal (estricta), el regresor se correlaciona con los errores en tiempos anteriores. Esto claramente **sesga el estimador de MCO, pero no necesariamente genera inconsistencia**. Cuando se cumpla esto, se dirá que existe una **exogeneidad contemporánea**, a diferencia de una **exogeneidad estricta** cuando se satisface para todos los tiempos.

Ahora, para analizar su consistencia, conviene expresar el término de sesgo de otra forma y analizar su comportamiento asintótico:

$$\begin{aligned}
(Y'_{t-1}Y_{t-1})^{-1}Y'_{t-1}\varepsilon_t &\Rightarrow \left(\frac{\sum_t^T y_t^2}{T}\right)^{-1}\left(\frac{\sum_t^T y_{t-1}\varepsilon_t}{T}\right) \\
&\Rightarrow plim\left(\left(\frac{\sum_t^T y_t^2}{T}\right)^{-1}\left(\frac{\sum_t^T y_{t-1}\varepsilon_t}{T}\right)\right) \\
&\Rightarrow plim\left(\left(\frac{\sum_t^T y_t^2}{T}\right)^{-1}\right)plim\left(\frac{\sum_t^T y_{t-1}\varepsilon_t}{T}\right) \\
&\Rightarrow plim\left(\left(\frac{\sum_t^T y_t^2}{T}\right)^{-1}\right) \cdot 0
\end{aligned}$$

Donde el término $plim\left(\frac{\sum_t^T y_{t-1}\varepsilon_t}{T}\right) \rightarrow 0$ por la exogeneidad contemporánea vista anteriormente.

Hay que tener en cuenta que el término $plim\left(\left(\frac{\sum_t^T y_t^2}{T}\right)^{-1}\right)$ es acotado siempre que sea un proceso **estacionario**.

En conclusión, bajo exogeneidad estricta, el estimador es insesgado y consistente. Y sólo bajo exogeneidad contemporánea, el estimador es sólo consistente.

R: 3. En general, las series de tiempo se utilizan en estudios macroeconómicos. Un caso importante en Chile es el cálculo del IPC, la inflación y el IMACEC. Se debe tener en cuenta que al agregar rezagos de la variable dependiente se puede caer en procesos no estacionarios. Y al agregar más de un rezago de regresores independientes se puede caer en efectos de multicolinealidad relevantes.