INTRODUCCIÓN A LA ECONOMETRÍA

3er curso LE y LADE
Tema 2b

Dpto. de Econometría y Estadística (EA3)

UPV—EHU

2.3b MCO en el MRLG.

MRLG: la FRP

- Recuérdese: modelo con K variables explicativas :

$$
\begin{aligned}
Y_t &= \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} + u_t, \\
Y &= X\beta + u
\end{aligned}
\tag{2}
$$

se llama MRLG.

- Función de Regresión de Población (FRP):
$E(u) = 0 \rightsquigarrow$ *parte sistemática* o FRP:

$$
\begin{aligned}
E(Y_t) &= \beta_0 + \beta_1 X_{1t} + \cdots + \beta_K X_{Kt} \\
E(Y) &= X\beta
\end{aligned}
$$

- Interpretación de los coeficientes:
  - $\beta_0 = E(Y_t | X_{1t} = X_{2t} = \cdots = X_{Kt} = 0)$: Valor esperado de Y_t cuando todas las variables explicativas son iguales a cero.
  - $\beta_k = \dfrac{\partial E(Y_t)}{\partial X_{kt}} \simeq \dfrac{\Delta E(Y_t)}{\Delta X_{kt}}, \quad k = 1 \ldots K$: Incremento en el valor (esperado) Y_t cuando $X_k \uparrow$ una unidad (c.p.).

Función de Regresión Muestral (FRM)

- Objetivo del MRLG: obtener estimador $\widehat{\beta} = (\widehat{\beta}_0, \widehat{\beta}_1 \ldots, \widehat{\beta}_K)'$ del vector de parámetros desconocidos en (2).
$\widehat{\beta} \rightsquigarrow$ modelo estimado, ajuste o FRM:

$$
\begin{aligned}
\widehat{Y}_t &= \widehat{\beta}_0 + \widehat{\beta}_1 X_{1t} + \cdots + \widehat{\beta}_K X_{Kt} \\
\widehat{Y} &= X\widehat{\beta}
\end{aligned}
$$

- Notas:
  - Perturbaciones en FRP:

$$
\begin{aligned}
u_t &= Y_t - E(Y_t) = Y_t - \beta_0 - \beta_1 X_{1t} - \cdots - \beta_K X_{Kt} \\
u &= Y - E(Y) = Y - X\beta
\end{aligned}
$$

  - Residuos en FRM:

$$
\begin{aligned}
\widehat{u}_t &= Y_t - \widehat{Y}_t = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 X_{1t} - \cdots - \widehat{\beta}_K X_{Kt} \\
\widehat{u} &= Y - \widehat{Y} = Y - X\widehat{\beta}
\end{aligned}
$$

- Residuos son a la FRM lo que las perturbaciones a la FRP.

Estimación: MCO

- aplicar ajuste Mínimo-Cuadrático al MRLG: $Y = X\beta + u$,
- o bien en forma de observación:

$$
\min_{\beta_0 \ldots \beta_K} \sum_{t=1}^{T} u_t^2 \text{ donde } u_t = Y_t - \beta_0 - \beta_1 X_{1t} - \cdots - \beta_K X_{Kt}
$$

- o en forma matricial:

[recuérdese:

$$
u' = (u_1, u_2, \ldots, u_T) \qquad u = \begin{pmatrix} u_1 \\ u_2 \\ \ldots \\ u_T \end{pmatrix}
$$

entonces $u'u = u_1^2 + u_2^2 + \cdots + u_T^2 = \sum_{t=1}^{T} u_t^2$]

- es decir

$$
\boxed{\min_{\beta} \; u'u \quad \text{donde} \quad u = Y - X\beta}
$$

Nota: derivadas vectoriales

- Sea $u = u(\beta)$: derivs de $c\,u$ y $c\,u^2$ con respecto a β:

$$\frac{\partial}{\partial\beta}(c\,u) = c\,\frac{\partial u}{\partial\beta} \qquad \text{y} \qquad \frac{\partial}{\partial\beta}\,u^2 = 2\,u\frac{\partial u}{\partial\beta}$$

- Con vectores y matrices esto es bastante similar:

- La derivada de la combinación lineal $u'c$

$$\underset{(1\times n)}{u'}\;\underset{(n\times 1)}{c} \qquad (= \textstyle\sum_{i=1}^n c_i u_i, \text{ es decir ¡escalar!!})$$

con respecto a $\underset{(k\times 1)}{\beta}$ es: $\frac{\partial(u'c)}{\partial\beta} = \frac{\partial u'}{\partial\beta}c$

- La derivada de la suma de cuadrados $u'u$

$$\underset{(1\times n)}{u'}\;\underset{(n\times 1)}{u} \qquad (= \textstyle\sum_{i=1}^n u_i^2, \text{ es decir ¡escalar!!})$$

con respecto a $\underset{(k\times 1)}{\beta}$ es: $\frac{\partial(u'u)}{\partial\beta} = 2\,\frac{\partial u'}{\partial\beta}u$

c.1^{er}o. en forma matricial

$$\min_{\beta}(u'u) \quad \text{donde} \quad u = Y - X\beta$$

Primera derivadas de SC $u'u$ con respecto a β:

$$\begin{aligned}\frac{\partial u'u}{\partial\beta} &= 2\,\frac{\partial u'}{\partial\beta}u \\ &= 2\,\frac{\partial(Y' - \beta'X')}{\partial\beta}u \\ &= -2\,X'u\end{aligned}$$

en el mínimo:

$$\boxed{\text{c.}1^{\text{er}}\text{o.}\quad \underset{(K{+}1\times T)}{X'}\;\underset{(T\times 1)}{\widehat{u}} = 0_{K{+}1}}$$

Estimación: Ecuaciones normales y EMC de β

Resolviendo las c.1^{er}o. obtenemos las ecuaciones normales:

$$X'(Y - X\widehat{\beta}) = 0 \;\Rightarrow$$

$$\boxed{\underset{(K{+}1\times 1)}{X'Y} = \underset{(K{+}1\times K{+}1)}{X'X}\;\underset{(K{+}1\times 1)}{\widehat{\beta}}} \tag{3}$$

De donde premultiplicando por $(X'X)^{-1}$ obtenemos el estimador MCO:

$$\boxed{\widehat{\beta}_{\text{MCO}} = (X'X)^{-1}X'Y}$$

Estimación: EMC de β (cont)

- donde $X'X$ es una matriz $[K{+}1 \times K{+}1]$: [¿recuérdese X e Y? $\longrightarrow$]
-

$$\underset{(K{+}1\times K{+}1)}{X'X} = \begin{bmatrix} T & \sum X_{1t} & \sum X_{2t} & \dots & \sum X_{Kt} \\ \sum X_{1t} & \sum X_{1t}^2 & \sum X_{1t}X_{2t} & \dots & \sum X_{1t}X_{Kt} \\ \dots & \dots & \dots & \dots & \dots \\ \sum X_{Kt} & \sum X_{Kt}X_{1t} & \sum X_{Kt}X_{2t} & \dots & \sum X_{Kt}^2 \end{bmatrix}$$

- y $X'Y$ y $\widehat{\beta}$ son vectores $[K{+}1 \times 1]$:

$$\underset{(K{+}1\times 1)}{X'Y} = \begin{bmatrix} \sum Y_t \\ \sum X_{1t}Y_t \\ \dots \\ \sum X_{Kt}Y_t \end{bmatrix} \qquad \underset{(K{+}1\times 1)}{\widehat{\beta}} = \begin{bmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \\ \dots \\ \widehat{\beta}_K \end{bmatrix}$$

Estimador MCO con datos centrados

Una forma alternativa de obtener el estimador MCO es

$$\boxed{\widehat{\beta}^{\star}_{\text{MCO}} = (x'x)^{-1}x'y}$$

para los coeficientes del modelo.

...junto con el intercepto estimado obtenido de la primera ecuación normal

$$\boxed{\widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1\overline{X}_1 - \cdots - \widehat{\beta}_K\overline{X}_K}$$

Nota: caso especial de $K = 1 \rightsquigarrow$ ¡formulas idénticas a las del MRLS!! (Demuéstralo).

2.4b Propiedades de la FRM.

Propiedades de los residuos y la FRM (1)

$$\left.\begin{array}{c} \widehat{\beta} \\ \widehat{\beta}^{\star} \rightsquigarrow \widehat{\beta}_0 \end{array}\right\} \rightsquigarrow \widehat{Y} = X\widehat{\beta} \rightsquigarrow \widehat{u} = Y - \widehat{Y}$$

1. residuos suman cero: $\sum \widehat{u}_t = 0$
   Demo: directamente de las c.1^{er}o.:

$$X'\widehat{u} = 0 \Rightarrow \begin{bmatrix} \sum_1^T \widehat{u}_t \\ \sum_1^T X_{1t}\widehat{u}_t \\ \sum_1^T X_{2t}\widehat{u}_t \\ \cdots \\ \sum_1^T X_{Kt}\widehat{u}_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

$\square$

2. $\overline{\widehat{Y}} = \overline{Y}$
3. la FRM pasa a través del vector $(\overline{X}_1, \ldots \overline{X}_K, \overline{Y})$:
   $\overline{Y} = \widehat{\beta}_0 + \widehat{\beta}_1\overline{X}_1 + \cdots + \widehat{\beta}_K\overline{X}_K$

   Nota: Estas propiedades 1 a 3 son satisfechas si la regresión tiene intercepto; es decir, si X tiene una columna de “unos”.

Propiedades de los residuos y la FRM (2)

4. residuos ortogonales a v. explicativas: $X'\widehat{u} = 0$
   Demo: directamente de las c.1^{er}o. (ver 1) o, alternativamente:

$$\begin{aligned} X'\widehat{u} &= X'(Y - X\widehat{\beta}) = X'Y - X'X\widehat{\beta} \\ &= X'Y - \underbrace{X'X(X'X)^{-1}}_{=I_{K+1}}X'Y = 0 \end{aligned}$$

$\square$

5. residuos ortogonales a la parte explicada de Y: $\widehat{Y}'\widehat{u} = 0$
   Demo: $\widehat{Y}'\widehat{u} = (X\widehat{\beta})'\widehat{u} = \widehat{\beta}'\underbrace{X'\widehat{u}}_{=0} = 0$ $\square$

2.5b Bondad de Ajuste: Coeficiente de Determinación (R^2) y Estimación de la Varianza del Error.

Bondad de Ajuste: R^2 (repaso)

Recuérdese (lo mismo que antes pero ahora lo hacemos con vectores):

$$\begin{aligned} Y'Y &= (\widehat{Y}' + \widehat{u}')(\widehat{Y} + \widehat{u}) \\ &= \widehat{Y}'\widehat{Y} + \widehat{u}'\widehat{u} + 2\widehat{Y}'\widehat{u} \\ &= \widehat{Y}'\widehat{Y} + \widehat{u}'\widehat{u} \quad \text{(desde prop 5)} \end{aligned}$$

$$Y'Y - T\overline{Y}^2 = \widehat{Y}'\widehat{Y} - T\overline{\widehat{Y}}^2 + \widehat{u}'\widehat{u} \qquad \text{(de prop 2)}$$

$$\boxed{\underset{(SCT)}{y'y} = \underset{(SCE)}{\widehat{y}'\widehat{y}} + \underset{(SCR)}{u'u}}$$

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT}$$

$$0 \leq R^2 \leq 1$$

Bondad de Ajuste: R^2 (repaso cont)

Nota 1: R^2 mide la proporción de la variación de la variable dependiente explicada por la variación de (una combinación lineal de) las variables explicativas.

Nota 2:

$$\text{no intercepto} \Rightarrow \begin{cases} \nexists 1^a \text{ fila de c.}1^{\text{er}}\text{o.} \rightsquigarrow \begin{cases} \sum \widehat{u}_t \neq 0, \\ \overline{\widehat{Y}} \neq \overline{Y}, \end{cases} \\ \text{no válido} R^2 \text{ (¡Recuérdese!)} \end{cases}$$

Estimación de $\mathrm{Var}(u_t)$

$$\sigma^2 = \mathrm{Var}(u_t) = \mathrm{E}(u_t^2) \simeq \frac{1}{T}\sum_{t=1}^{T} u_t^2$$

pero con residuos se tienen que satisfacer $K{+}1$ relaciones lineales en $X'\widehat{u} = 0$ así que perdemos $K{+}1$ grados de libertad:

$$\widehat{\sigma}^2 = \frac{1}{T{-}K{-}1}\sum_{t=1}^{T} \widehat{u}_t^2$$

Por lo tanto proponemos el siguiente estimador:

$$\boxed{\widehat{\sigma}^2 = \frac{\text{RSS}}{T{-}K{-}1}}$$

el cual claramente es un estimador insesgado:

Demo:

$$\mathrm{E}(\widehat{\sigma}^2) = \frac{\mathrm{E}(\text{SCR})(^*)}{T{-}K{-}1} = \frac{T{-}K{-}1}{T{-}K{-}1} = \sigma^2$$

□ (* ver libro de texto)

2.6 Propiedades del Estimador de Mínimos Cuadrados en muestras finitas. El Teorema de Gauss-Markov.

Propiedades del Estimador MCO (1)

El estimador $\widehat{\beta}_{\text{MCO}} = (X'X)^{-1}X'Y$ tiene las siguientes propiedades:

- **Lineal:** $\widehat{\beta}_{\text{MCO}}$ es una combinación lineal de perturbaciones:

$$\begin{aligned}\widehat{\beta} &= (X'X)^{-1}X'(X\beta + u)\\ &= (X'X)^{-1}X'X\beta + (X'X)^{-1}X'u\\ &= \beta + (X'X)^{-1}X'u\\ &= \beta + \Gamma' u\end{aligned}$$

- **insesgado:** Dado queo $\text{E}(u) = 0$, $\widehat{\beta}_{\text{MCO}}$ es insesgado:

$$\begin{aligned}\text{E}(\widehat{\beta}) &= \text{E}(\beta + \Gamma' u)\\ &= \beta + \Gamma'\text{E}(u)\\ &= \beta\end{aligned}$$

Propiedades del Estimador MCO (2)

- **Varianza:** Recuérdese:

$$\begin{aligned}\text{Var}(u) &= \sigma^2 I_T,\\ \widehat{\beta} &= \beta + (X'X)^{-1}X'u,\end{aligned}$$

$$\begin{aligned}\text{Var}(\widehat{\beta}) &= \text{E}\left((\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'\right)\\ &= \text{E}\left((X'X)^{-1}X'u\,u'X(X'X)^{-1}\right)\\ &= (X'X)^{-1}X'\,\text{E}(uu')\,X(X'X)^{-1}\\ &= (X'X)^{-1}X'\,\sigma^2 I_T\,X(X'X)^{-1}\\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1}\end{aligned}$$

$$\boxed{\text{Var}(\widehat{\beta}) = \sigma^2(X'X)^{-1}}$$

Propiedades del Estimador MCO (2cont)

$$\text{Var}(\widehat{\beta}) = \begin{bmatrix} \text{Var}(\widehat{\beta}_0) & \text{Cov}(\widehat{\beta}_0,\widehat{\beta}_1) & \dots & \text{Cov}(\widehat{\beta}_0,\widehat{\beta}_K)\\ \text{Cov}(\widehat{\beta}_1,\widehat{\beta}_0) & \text{Var}(\widehat{\beta}_1) & \dots & \text{Cov}(\widehat{\beta}_1,\widehat{\beta}_K)\\ \dots & \dots & \dots & \dots\\ \text{Cov}(\widehat{\beta}_K,\widehat{\beta}_0) & \text{Cov}(\widehat{\beta}_K,\widehat{\beta}_1) & \dots & \text{Var}(\widehat{\beta}_K)\end{bmatrix}$$

$$\sigma^2(X'X)^{-1} = \sigma^2\begin{bmatrix} a_{00} & a_{00} & a_{01} & \dots & a_{0K}\\ a_{10} & a_{11} & a_{12} & \dots & a_{1K}\\ a_{20} & a_{21} & a_{22} & \dots & a_{2K}\\ \dots & \dots & \dots & \dots & \dots\\ a_{K0} & a_{K1} & a_{K2} & \dots & a_{KK}\end{bmatrix}$$

es decir a_{kk} es el $(k+1, k+1)$-elemento del matriz $(X'X)^{-1}$:

$$\begin{aligned}\text{Var}(\widehat{\beta}_k) &= \sigma^2 a_{kk}\\ \text{Cov}(\widehat{\beta}_k,\widehat{\beta}_i) &= \sigma^2 a_{ki}\end{aligned}$$

Teorema Gauss-Markov

“Dados los supuestos básicos del MRLG, el estimador MCO es el de varianza mínima (el mejor) entre todos los estimadores lineales e insesgados”

$$\widehat{\beta}_{\text{MCO}} = \textbf{ELIO} = \textbf{E}_{\text{stimador}}\textbf{L}_{\text{ineal}}\textbf{I}_{\text{nsesgado}}\textbf{O}_{\text{ptimo}}$$

Demo:

Sea $\widetilde{\beta}$ algún otro estimador lineal insesgado:

$$\widetilde{\beta} = D'Y = D'(X\beta + u) = D'X\beta + D'u$$

$$\mathsf{E}(\widetilde{\beta}) = D'X\beta + D'\mathsf{E}(u) = D'X\beta = \beta \Rightarrow \boxed{D'X = I_K}$$

entonces $\widetilde{\beta} = \beta + D'u \quad \rightsquigarrow \quad \widetilde{\beta} - \beta = D'u$
y su varianza:

$$\text{Var}(\widetilde{\beta}) = E\left[(\widetilde{\beta}-\beta)(\widetilde{\beta}-\beta)'\right] = \mathsf{E}(D'u\,u'D)$$

$$= D'\,\mathsf{E}(uu')\,D = D'\,\sigma^2\,I_T\,D = \sigma^2\,D'D$$

Teorema Gauss-Markov (cont)

$$\begin{aligned}
\text{Var}(\widetilde{\beta}) - \text{Var}(\widehat{\beta}) &= \sigma^2\,D'D - \sigma^2\,(X'X)^{-1} \\
&= \sigma^2\,\left[D'D - (X'X)^{-1}\right] \\
&= \sigma^2\,\left[D'D - D'X\,(X'X)^{-1}\,X'D\right] \\
&= \sigma^2\,D'\,\underbrace{\left[I_T - X\,(X'X)^{-1}\,X'\right]}_{M}\,D \\
&= \sigma^2\,D'\,(MM)\,D \\
&= \sigma^2\,(D'M)(M'D) = D^{\star\prime}D^{\star} \\
&> 0
\end{aligned}$$

Esto es, en particular **todas** las varianzas individuales serán mayores que las respectivas por MCO.

2.3c MCO: Expresiones útiles y Cronología.

Expresiones útiles para SC

- $$SCT = \sum(Y_t - \overline{Y})^2 = \sum Y_t^2 - T\overline{Y}^2 = Y'Y - T\overline{Y}^2$$

- $$SCE = \sum(\widehat{Y}_t - \overline{\widehat{Y}})^2 = \sum \widehat{Y}_t^2 - T\overline{\widehat{Y}}^2 = \sum \widehat{Y}_t^2 - T\overline{Y}^2 = \widehat{Y}'\widehat{Y} - T\overline{Y}^2$$
  $$= (X\widehat{\beta})'(X\widehat{\beta}) - T\overline{Y}^2 = \widehat{\beta}'\,\underbrace{X'X\widehat{\beta}}_{X'Y} - T\overline{Y}^2 = \widehat{\beta}'X'Y - T\overline{Y}^2$$

- $$SCR = \sum \widehat{u}_t^2 = \widehat{u}'\widehat{u} \qquad = \sum Y_t^2 - \sum \widehat{Y}_t^2 = Y'Y - \widehat{\beta}'X'Y$$

Principales expresiones y Cronología

- $Y = X\beta + u$
- $(X'X)^{-1} \qquad X'Y$
- $\widehat{\beta} = (X'X)^{-1}X'Y$
- $ESS = \widehat{\beta}'X'Y - T\overline{Y}^2$ (¡necesita $\overline{Y}$!)
- $TSS = Y'Y - T\overline{Y}^2$
- $RSS = Y'Y - \widehat{\beta}'X'Y$ (¡no $\overline{Y}$!)
- $R^2 = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$
- $\widehat{\sigma}^2 = \dfrac{RSS}{T–K–1}$
- $\widehat{\text{Var}}(\widehat{\beta}) = \widehat{\sigma}^2(X'X)^{-1}$