Taller de Reforzamiento

Introducción a la Econometría

Sesión 4

Chamorro Rodriguez, Gianfranco David

E-mail: gianfranco.chamorror@gmail.com

Enero 2023

El presente documento sirve de apoyo para la sesión 4 del taller de reforzamiento de Econometría, durante el cual se trabajarán conceptos que serán necesarios para iniciar un curso introductorio de pregrado o reforzar conceptos olvidados. Se brindará una perspectiva amable junto a material que permitara a los participantes abarcar temas mas complejos adelante.

---

1. Método de Mínimos Cuadrados Ordinarios

Como vimos en nuestra primera sesión uno de los puntos de inicio es el método de Mínimos Cuadrados Ordinarios (MCO).

Este método presenta muchas ventajas en cuanto a lo fácil de su uso y por lo adecuado del planteamiento estadístico matemático que permite adecuarse a los supuestos para los modelos econométricos.El término de MCO esta vinculado con la regresión y la correlación, ambas determinan la existencia de relación entre dos o mas variables. Encontrar nuestro vector de estimadores ($\hat{\beta}$) es el inicio del cálculo y análisis a realizar:

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

Recordemos:

$$Y = \hat{Y} + e$$

Siendo :

$$\hat{Y} = X\hat{\beta}$$

$$e = Y - X\hat{\beta}$$

1.1. Bondad del Ajuste y Análisis de la Varianza

La suma cuadrada de residuos nos brinda una medida del ajuste de la recta de regresión de los datos. Las medidas de bondad en general resumen la discrepancia entre los valores observados y los valores esperados en el modelo de estudio. La variación total de Y es la suma de las desviaciones al cuadrado. La variación de la variable dependiente se define en términos de desviaciones respecto de su media ($Y_i - \bar{Y}$).[1]

$$STC = min\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

[1] Greene, William -Análisis econométrico(2012)

Si la regresión contiene un término constante, los residuos sumarán cero y la media de los valores predichos de y_i se igualará a la media de los valores actuales :

$$\mathbf{y}_i - \bar{y} = \widehat{y_i} - \bar{y} + e_i$$

$$\mathbf{y}_i - \bar{y} = (\mathbf{x}_i - \bar{x})b' + e_i$$

Debemos tener en cuenta que una regresión se ajustaría mejor si las desviaciones de y con respecto a su media estuviera representado en su mayor parte por las desviaciones de x con respecto a su media que por los errores. todos en términos de sumas al cuadrado. Desde un punto de vista matricial, haremos uso del concepto de la Matriz idemponente **Mº**.

$$\mathbf{M^o y = M^o X b + M^o e}^2$$

Considerando los elementos al cuadrado:

$$\mathbf{y' M^o y = b' X M^o X b + e' e}$$

$$\mathbf{STC = SEC + SRC}$$

Ahora calculamos una medida de bondad del ajuste mediante la utilización del coeficiente de determinación R^2.

$$R^2 = \frac{SEC}{STC} = \frac{b'XMXb}{y'My}$$

$$= 1 - \frac{e'e}{y'My}$$

Otro enfoque :

SUMAS AL CUADRADO			
	Matricial	Matricial II	Sumatoria
SEC	b'XMºXb	$\hat{Y}'\hat{Y} - N\bar{Y}^2$	$\sum \hat{Y}^2 - N\bar{Y}^2$
SRC	e'e	$Y'Y - Y'X\hat{\beta}$	$\sum e_i^2$
STC	Y'MºY	$Y'Y - N\bar{Y}^2$	$\sum Y^2 - N\bar{Y}^2$

Análisis de la Varianza:

Análisis de la Variación			
	Índice	Grados de Libertad	M.C.
MODELO	SEC	k - 1	SEC / k -1
RESIDUO	SRC	N - k	SRC / N - k
TOTAL	STC	N - 1	STC / N - 1

[2] Mºe=e

El R^2 presenta un problema para medir la bondad de ajuste, primero con respecto al número de grados de libertad en la estimación.Este indicador no decrecerá cuando se aumente una variable adicional incluso cuando esta no presente una variabilidad explicativa significativa sin embargo a medida que aumentamos el número de variables los parámetros estimados se hacen progresivamente menos precisos.
Ante ello se propone el coeficiente de determinación R ajustado ($R^2_{aj.}$) que se calcula :

$$R^2_{aju} = \mathbf{1} \text{ - } \frac{\frac{e'e}{N-K}}{\frac{y'My}{N-1}}$$

$$R^2_{aju} = \mathbf{1} \text{ - } \frac{N-1}{N-k}(1 - R^2)$$

1.2. Matrices: Varianza y Covarianza del Error

$$\textbf{VarCov(e)= E [(e - E(e)) (e - E(e))}'] \text{ [3]}$$

$$\textbf{VarCov(e)= E [(e) (e)}']$$

$$\textbf{VarCov(e)=} \text{ E} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \begin{bmatrix} e_1 & e_2 & \dots & e_n \end{bmatrix}$$

$$\textbf{VarCov(e)=} \text{ E} \begin{bmatrix} e_1^2 & e_1e_2 & \dots & e_1e_n \\ e_2e_1 & e_2^2 & \dots & e_2e_n \\ \vdots & \vdots & \ddots & \vdots \\ e_ne_1 & e_ne_2 & \dots & e_n^2 \end{bmatrix} = \mathbf{E} \begin{bmatrix} \sigma_e^2 & 0 & \dots & 0 \\ 0 & \sigma_e^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_e^2 \end{bmatrix} \text{ [4]}$$

[3] E(e)=0
[4] Supuestos:
No Heterocedasticidad : E(e_i^2) $= \sigma_e^2 ... \forall i$
No Autocorrelación: $E(e_ie_j) = 0 ... \forall i \neq j$

$$= \sigma_e^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} = \sigma_e^2 I$$

1.3. Matrices: Varianza y Covarianza del Estimador

$$\textbf{VarCov}(\ \hat{\beta}\ )\textbf{=}\ \textbf{E}\ [(\hat{\beta}\ \textbf{-}\ \textbf{E}(\hat{\beta}))\ (\hat{\beta}\ \textbf{-}\ \textbf{E}(\hat{\beta}))']$$

Recordando :

$$\hat{\beta} = (\mathbf{X'X})^{-1} X'Y$$

$$\hat{\beta} = (\mathrm{X'X})^{-1} X'(X\beta + \mathrm{e})$$

$$\hat{\beta} = (\mathrm{X'X})^{-1} (X'X)\beta + (\mathrm{X'X})^{-1} X'e$$

$$\boxed{\hat{\beta} = \beta + (\mathrm{X'X})^{-1} X'e}$$

y :

$$\mathbf{E}(\hat{\beta}) = \mathbf{E}\ (\ \beta + (\mathbf{X'X})^{-1} \underbrace{X'e}_{E(X'e)=0}\ )$$

$$\boxed{\mathbf{E}(\hat{\beta}) = \beta}$$

Entonces :

$$\textbf{VarCov}(\ \hat{\beta}\ )\textbf{=}\ \textbf{E}\ [(\hat{\beta}\ \textbf{-}\ \beta)\ (\hat{\beta}\ \textbf{-}\ \beta)']$$

De $\hat{\beta}$: $\hat{\beta}$ - $\beta = (\mathbf{X}'\mathbf{X})^{-1}X'e$

$$\textbf{VarCov } (\hat{\beta}) = \mathbf{E}\ [(\mathbf{X}'\mathbf{X})^{-1}\ \mathbf{X'e}]\ [(\mathbf{X}'\mathbf{X})^{-1}X'e]'$$

$$\textbf{VarCov } (\hat{\beta}) = \mathbf{E}\ [(\mathbf{X}'\mathbf{X})^{-1}\ \mathbf{X'}ee'\mathbf{X}\ (\mathbf{X}'\mathbf{X})^{-1}]$$

$$\textbf{VarCov } (\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\ \mathbf{X'}VarCov(e)\mathbf{X}\ (\mathbf{X}'\mathbf{X})^{-1}$$

$$\textbf{VarCov } (\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\ \mathbf{X'}\ \sigma_e^2 I\ \mathbf{X}\ (\mathbf{X}'\mathbf{X})^{-1}$$

$$\textbf{VarCov } (\hat{\beta}) = \sigma_e^2(\mathbf{X'X})^{-1}X'X(X'X)^{-1}$$

$$\boxed{\textbf{VarCov } (\hat{\beta}) = \sigma_e^2(\mathbf{X'X})^{-1}}$$

Consideremos que :

$$\boxed{\sigma_e^2 = \frac{SCR}{N-k}}$$

RESUMEN	
Estimación	**Supuestos MCR**
$\hat{\beta} = (X'X)^{-1}X'Y$	$Y = X\beta + e$
$\hat{Y} = X\hat{\beta}$	**X en n x K con rango K**
$\hat{e} = Y - \hat{Y}$	**E** $[\mathbf{e}\vert\mathbf{X}]$**=0**
$SCR = \hat{e}'\hat{e}$	**E** $[\mathbf{ee'}\vert\mathbf{X}] = \sigma_e^2 I$
$\sigma_e^2 = \frac{SRC}{N-k}$	**X es una matriz no estocástica**
$Var(\hat{\beta}) = \sigma_e^2(X'X)^{-1}$	$\mathbf{e}\vert\mathbf{X} \sim N(\mu, \sigma^2 I)$

2. Ejercicios Propuestos

1. Se cuenta con la siguiente información de dos variables X e Y:

Y	15	30	20	12	15	28	26	25	33	17	19	23	26	19	31	30	13	22	22	26
X	8	14	12	8	9	13	16	14	18	11	12	15	11	9	20	19	9	14	15	10

Se solicita calcular :

- **Hallar Los estimadores por MCO de forma matricial.**
- **El estimador por MCO sin considerar una constante en el Modelo.**
- **Encontrar la diferencia de los errores estimados de cada modelo.**
- **Determinar el R^2 y $R^2_{ajustado}$ del modelo insesgado.**
- **Hallar la matriz de Varianza y Covarianzas de β.**
- **Hallar la desviación estándar de los estimadores.**