ECONOMETRÍA

JOSÉ ALBERTO MAURICIO

2

REGRESIÓN LINEAL MÚLTIPLE I

2.2 MCO - PROPIEDADES ALGEBRAICAS

Departamento de Análisis Económico y Economía Cuantitativa
Universidad Complutense de Madrid

EctrGr-JAM-2-2.pdf

ucm.randomshock.com/ectrgr

Versión 2.0 - Enero 2023

MATERIAL AUXILIAR Y COMPLEMENTARIO

El instrumental matemático que se utiliza en esta Sección 2.2 tiene que ver con sumatorios, estadísticos muestrales, operaciones con vectores y matrices, y optimización matemática (mínimos de funciones de dos o más variables). Se puede repasar todo ello en Wooldridge (2020): Secciones A-1 y D-1 - D-6. En la PARTE 1 y en el APÉNDICE de esta Sección 2.2 se presentan, respectivamente, un breve resumen sobre sumatorios y estadísticos muestrales, y algunas fórmulas para la inversión de matrices.

Adicionalmente, en esta Sección 2.2 se hace referencia a los Ejercicios 1 - 3 disponibles en la página de la asignatura en Internet. A su vez, en esos ejercicios se mencionan varias secciones de la guía *Introducción al Uso de EViews 4.1* de cara a la resolución de las preguntas que se plantean en ellos.

BIBLIOGRAFÍA PARA TODO EL TEMA 2

Hayashi (2000): Capítulo 1.
Hill, Griffiths, Lim (2018): Capítulos 2 - 6.
Wooldridge (2020): Capítulos 2 - 6 y Apéndice E.

PARTE 1 - ESTADÍSTICOS MUESTRALES

Datos (Números):

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{bmatrix}.$$

Sumatorios:

$$\textstyle\sum_{i=1}^{N} w_i = w_1 + w_2 + \ldots + w_N .$$

$$\textstyle\sum_{i=1}^{N} a = Na.$$

$$\textstyle\sum_{i=1}^{N} aw_i = a \sum_{i=1}^{N} w_i .$$

$$\textstyle\sum_{i=1}^{N} (w_i \pm v_i) = \sum_{i=1}^{N} w_i \pm \sum_{i=1}^{N} v_i .$$

$$\textstyle\sum_{i=1}^{N} (aw_i \pm bv_i) = a \sum_{i=1}^{N} w_i \pm b \sum_{i=1}^{N} v_i .$$

$$\textstyle\sum_{i=1}^{N} (a \pm bw_i) = Na \pm b \sum_{i=1}^{N} w_i .$$

$$\textstyle\sum_{i=1}^{N} (w_i v_i) \neq \left(\sum_{i=1}^{N} w_i \right) \times \left(\sum_{i=1}^{N} v_i \right).$$

$$\textstyle\sum_{i=1}^{N} w_i^2 \neq \left(\sum_{i=1}^{N} w_i \right)^2 .$$

$$\textstyle\sum_{i=1}^{N} (w_i / v_i) \neq \left(\sum_{i=1}^{N} w_i \right) \big/ \left(\sum_{i=1}^{N} v_i \right).$$

Medias Muestrales: $\bar{w} = \frac{1}{N} \sum_{i=1}^{N} w_i , \ \bar{v} = \frac{1}{N} \sum_{i=1}^{N} v_i ,$

$$\textstyle\sum_{i=1}^{N} (w_i - \bar{w}) = \sum_{i=1}^{N} (v_i - \bar{v}) = 0.$$

Varianza Muestral [1]: $\widehat{\text{var}}[\mathbf{w}] = \frac{1}{N} \sum_{i=1}^{N} (w_i - \bar{w})^2 ,$

$$\begin{aligned} \textstyle\sum_{i=1}^{N} (w_i - \bar{w})^2 &= \textstyle\sum_{i=1}^{N} w_i (w_i - \bar{w}) \\ &= \textstyle\sum_{i=1}^{N} w_i^2 - N\bar{w}^2 = \mathbf{w}'\mathbf{w} - N\bar{w}^2 . \end{aligned} \qquad [1.1]$$

En adelante, el símbolo $'$ representa la operación (vectorial o matricial) de trasposición.

Covarianza Muestral [1]: $\widehat{\text{cov}}[\mathbf{w}, \mathbf{v}] = \frac{1}{N} \sum_{i=1}^{N} (w_i - \bar{w})(v_i - \bar{v}),$

$$\begin{aligned} \textstyle\sum_{i=1}^{N} (w_i - \bar{w})(v_i - \bar{v}) &= \textstyle\sum_{i=1}^{N} w_i (v_i - \bar{v}) = \sum_{i=1}^{N} (w_i - \bar{w}) v_i \\ &= \textstyle\sum_{i=1}^{N} w_i v_i - N\bar{w}\bar{v} = \mathbf{w}'\mathbf{v} - N\bar{w}\bar{v}. \end{aligned} \qquad [2.1]$$

Correlación Simple Muestral: $\widehat{\text{corr}}[\mathbf{w}, \mathbf{v}] = \dfrac{\widehat{\text{cov}}[\mathbf{w}, \mathbf{v}]}{\widehat{\text{dvt}}[\mathbf{w}] \times \widehat{\text{dvt}}[\mathbf{v}]} ,$

con $\widehat{\text{dvt}}[\cdot] = \sqrt{\widehat{\text{var}}[\cdot]}$ (desviación típica muestral), $-1 \leq \widehat{\text{corr}}[\mathbf{w}, \mathbf{v}] \leq +1$.

Datos en Desviaciones con respecto a la Media:

$$\tilde{\mathbf{w}} = \begin{bmatrix} \tilde{w}_1 \\ \tilde{w}_2 \\ \vdots \\ \tilde{w}_N \end{bmatrix} = \begin{bmatrix} w_1 - \overline{w} \\ w_2 - \overline{w} \\ \vdots \\ w_N - \overline{w} \end{bmatrix}, \; \tilde{\mathbf{v}} = \begin{bmatrix} \tilde{v}_1 \\ \tilde{v}_2 \\ \vdots \\ \tilde{v}_N \end{bmatrix} = \begin{bmatrix} v_1 - \overline{v} \\ v_2 - \overline{v} \\ \vdots \\ v_N - \overline{v} \end{bmatrix},$$

$$\tilde{w}_i = w_i - \overline{w},\; \tilde{v}_i = v_i - \overline{v} \;\; (i = 1, 2, ..., N),$$

$$\textstyle\sum_{i=1}^N \tilde{w}_i = \sum_{i=1}^N \tilde{v}_i = 0.$$

Varianza Muestral [2]: $$\widehat{\mathrm{var}}[\mathbf{w}] = \tfrac{1}{N} \textstyle\sum_{i=1}^N \tilde{w}_i^2, \qquad [1.2]$$

$$\textstyle\sum_{i=1}^N \tilde{w}_i^2 = \tilde{\mathbf{w}}'\tilde{\mathbf{w}} = \mathbf{w}'\tilde{\mathbf{w}}.$$

Covarianza Muestral [2]: $$\widehat{\mathrm{cov}}[\mathbf{w}, \mathbf{v}] = \tfrac{1}{N} \textstyle\sum_{i=1}^N \tilde{w}_i \tilde{v}_i, \qquad [2.2]$$

$$\textstyle\sum_{i=1}^N \tilde{w}_i \tilde{v}_i = \tilde{\mathbf{w}}'\tilde{\mathbf{v}} = \mathbf{w}'\tilde{\mathbf{v}} = \tilde{\mathbf{w}}'\mathbf{v}.$$

Observación Importante: $\mathbf{w}'\mathbf{w}$ $[\tilde{\mathbf{w}}'\tilde{\mathbf{w}}]$ es el cuadrado de la **norma euclídea** (el **módulo**, o el **tamaño**) del vector $\mathbf{w}$ $[\tilde{\mathbf{w}}]$, que se puede interpretar como la cantidad de información que contiene $\mathbf{w}$ medida con respecto a cero [con respecto a su media muestral].

Demostración de [1.1]:

$$\textstyle\sum_{i=1}^N (w_i - \overline{w})^2 = \sum_{i=1}^N (w_i^2 + \overline{w}^2 - 2w_i\overline{w}) =$$
$$= \textstyle\sum_{i=1}^N w_i^2 + \sum_{i=1}^N \overline{w}^2 + \sum_{i=1}^N (-2w_i\overline{w}) = \sum_{i=1}^N w_i^2 + \sum_{i=1}^N \overline{w}^2 - 2\overline{w}\sum_{i=1}^N w_i =$$
$$= \textstyle\sum_{i=1}^N w_i^2 + N\overline{w}^2 - 2\overline{w}N\overline{w} = \sum_{i=1}^N w_i^2 + N\overline{w}^2 - 2N\overline{w}^2 = \sum_{i=1}^N w_i^2 - N\overline{w}^2,$$

$$\textstyle\sum_{i=1}^N w_i(w_i - \overline{w}) = \sum_{i=1}^N (w_i^2 - w_i\overline{w}) = \sum_{i=1}^N w_i^2 + \sum_{i=1}^N (-w_i\overline{w}) =$$
$$= \textstyle\sum_{i=1}^N w_i^2 - \overline{w}\sum_{i=1}^N w_i = \sum_{i=1}^N w_i^2 - \overline{w}N\overline{w} = \sum_{i=1}^N w_i^2 - N\overline{w}^2.$$

Demostración de [2.1]:

$$\textstyle\sum_{i=1}^N (w_i - \overline{w})(v_i - \overline{v}) = \sum_{i=1}^N (w_i v_i - \overline{w}v_i - w_i\overline{v} + \overline{w}\overline{v}) =$$
$$= \textstyle\sum_{i=1}^N w_i v_i - \sum_{i=1}^N \overline{w}v_i - \sum_{i=1}^N w_i\overline{v} + \sum_{i=1}^N \overline{w}\overline{v} =$$
$$= \textstyle\sum_{i=1}^N w_i v_i - \overline{w}\sum_{i=1}^N v_i - \overline{v}\sum_{i=1}^N w_i + \sum_{i=1}^N \overline{w}\overline{v} =$$
$$= \textstyle\sum_{i=1}^N w_i v_i - \overline{w}N\overline{v} - \overline{v}N\overline{w} + N\overline{w}\overline{v} = \sum_{i=1}^N w_i v_i - N\overline{w}\overline{v},$$

$$\textstyle\sum_{i=1}^N w_i(v_i - \overline{v}) = \sum_{i=1}^N (w_i v_i - w_i\overline{v}) = \sum_{i=1}^N w_i v_i - \sum_{i=1}^N w_i\overline{v} =$$
$$= \textstyle\sum_{i=1}^N w_i v_i - \overline{v}\sum_{i=1}^N w_i = \sum_{i=1}^N w_i v_i - \overline{v}N\overline{w} = \sum_{i=1}^N w_i v_i - N\overline{w}\overline{v},$$

$$\sum_{i=1}^{N}(w_i - \bar{w})v_i = \sum_{i=1}^{N}(w_i v_i - \bar{w}v_i) = \sum_{i=1}^{N} w_i v_i - \sum_{i=1}^{N} \bar{w}v_i =$$
$$= \sum_{i=1}^{N} w_i v_i - \bar{w}\sum_{i=1}^{N} v_i = \sum_{i=1}^{N} w_i v_i - \bar{w}N\bar{v} = \sum_{i=1}^{N} w_i v_i - N\bar{w}\bar{v}.$$

PARTE 2 - ESTIMACIÓN MCO DE MODELOS RLS

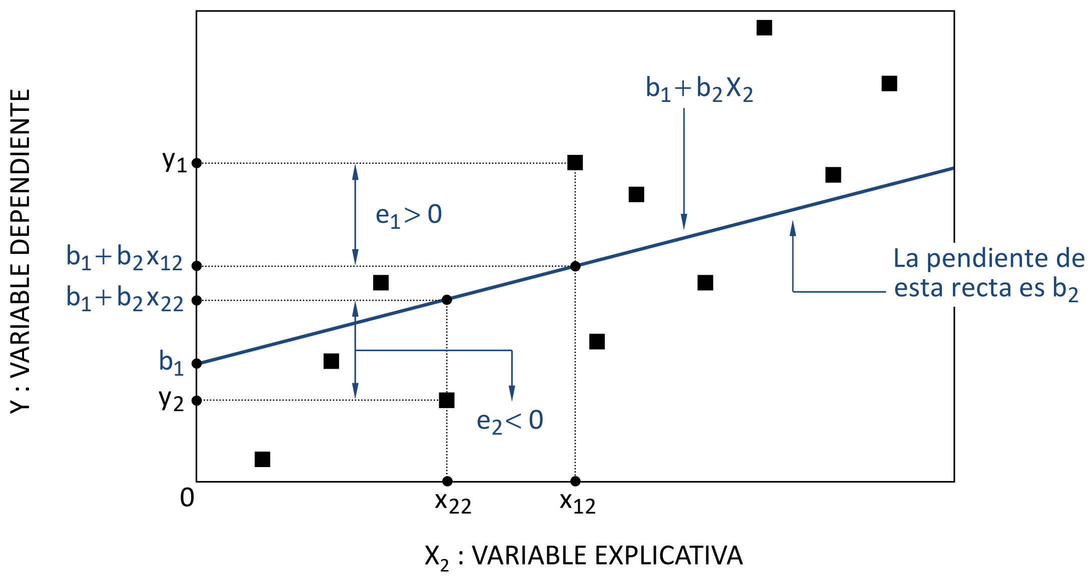
Para estimar los parámetros β_1 y β_2 en un modelo RLS,

Modelo: $$Y = \beta_1 + \beta_2 X_2 + U,$$

disponemos de una colección de datos compuesta por N observaciones sobre cada una de las variables (Y, X_2), que puede representarse como

Datos: $$\begin{bmatrix} y_1 & x_{12} \\ y_2 & x_{22} \\ \vdots & \vdots \\ y_N & x_{N2} \end{bmatrix} = [\mathbf{y}, \mathbf{x}_2].$$

Las N observaciones (puntos muestrales, filas) sobre Y, X_2 pueden referirse a N entidades observables en un momento dado (datos de sección cruzada), o bien a N momentos consecutivos de la historia de una única entidad observable (datos de series temporales).



Residuos: $$e_i(b_1, b_2) = y_i - (b_1 + b_2 x_{i2}) \quad (i = 1, ..., N), \tag{3}$$

donde b_1, b_2 son dos estimaciones cualesquiera de β_1, β_2, respectivamente.

Criterio MCO: $$\text{Minimizar } S(b_1, b_2) = \sum_{i=1}^{N} e_i(b_1, b_2)^2. \tag{4}$$

Dos números $\hat{\beta}_1, \hat{\beta}_2$ son la solución (única) del problema [4] sí y sólo sí las derivadas parciales de primer orden de $S(\cdot)$ evaluadas en $\hat{\beta}_1, \hat{\beta}_2$ son iguales a cero, y la matriz de derivadas parciales de segundo orden de $S(\cdot)$ evaluadas en $\hat{\beta}_1, \hat{\beta}_2$ es definida positiva.

SCR

$$\sum_{i=1}^{N} e_i(b_1, b_2)^2 = \sum_{i=1}^{N} [y_i - (b_1 + b_2 x_{i2})]^2 =$$
$$\sum_{i=1}^{N} y_i^2 + \sum_{i=1}^{N} (b_1 + b_2 x_{i2})^2 - 2\sum_{i=1}^{N} (b_1 + b_2 x_{i2}) y_i .$$

DPO

$$\begin{aligned}
\frac{\partial S}{\partial b_1} &= 2\sum_{i=1}^{N} (b_1 + b_2 x_{i2}) - 2\sum_{i=1}^{N} y_i \\
&= 2Nb_1 + 2b_2 \sum_{i=1}^{N} x_{i2} - 2\sum_{i=1}^{N} y_i \\
&= 2N(b_1 + b_2 \overline{x}_2 - \overline{y}), \\
\frac{\partial S}{\partial b_2} &= 2\sum_{i=1}^{N} (b_1 + b_2 x_{i2}) x_{i2} - 2\sum_{i=1}^{N} x_{i2} y_i \\
&= 2b_1 \sum_{i=1}^{N} x_{i2} + 2b_2 \sum_{i=1}^{N} x_{i2}^2 - 2\sum_{i=1}^{N} x_{i2} y_i \\
&= 2\left(b_1 N\overline{x}_2 + b_2 \sum_{i=1}^{N} x_{i2}^2 - \sum_{i=1}^{N} x_{i2} y_i \right).
\end{aligned}$$

$$\hat{\beta}_1 + \hat{\beta}_2 \overline{x}_2 - \overline{y} = 0,$$
$$\hat{\beta}_1 N\overline{x}_2 + \hat{\beta}_2 \sum_{i=1}^{N} x_{i2}^2 - \sum_{i=1}^{N} x_{i2} y_i = 0.$$

$$\hat{\beta}_1 = \overline{y} - \hat{\beta}_2 \overline{x}_2 \ [*],$$
$$\hat{\beta}_1 N\overline{x}_2 + \hat{\beta}_2 \sum_{i=1}^{N} x_{i2}^2 = \sum_{i=1}^{N} x_{i2} y_i .$$

$$\begin{aligned}
(\overline{y} - \hat{\beta}_2 \overline{x}_2) N\overline{x}_2 + \hat{\beta}_2 \sum_{i=1}^{N} x_{i2}^2 &= \sum_{i=1}^{N} x_{i2} y_i , \\
N\overline{x}_2 \overline{y} - \hat{\beta}_2 N\overline{x}_2^2 + \hat{\beta}_2 \sum_{i=1}^{N} x_{i2}^2 &= \sum_{i=1}^{N} x_{i2} y_i , \\
\hat{\beta}_2 \left(\sum_{i=1}^{N} x_{i2}^2 - N\overline{x}_2^2 \right) &= \sum_{i=1}^{N} x_{i2} y_i - N\overline{x}_2 \overline{y}, \\
\hat{\beta}_2 \sum_{i=1}^{N} (x_{i2} - \overline{x}_2)^2 &= \sum_{i=1}^{N} (x_{i2} - \overline{x}_2)(y_i - \overline{y}) \ [**].
\end{aligned}$$

DSO

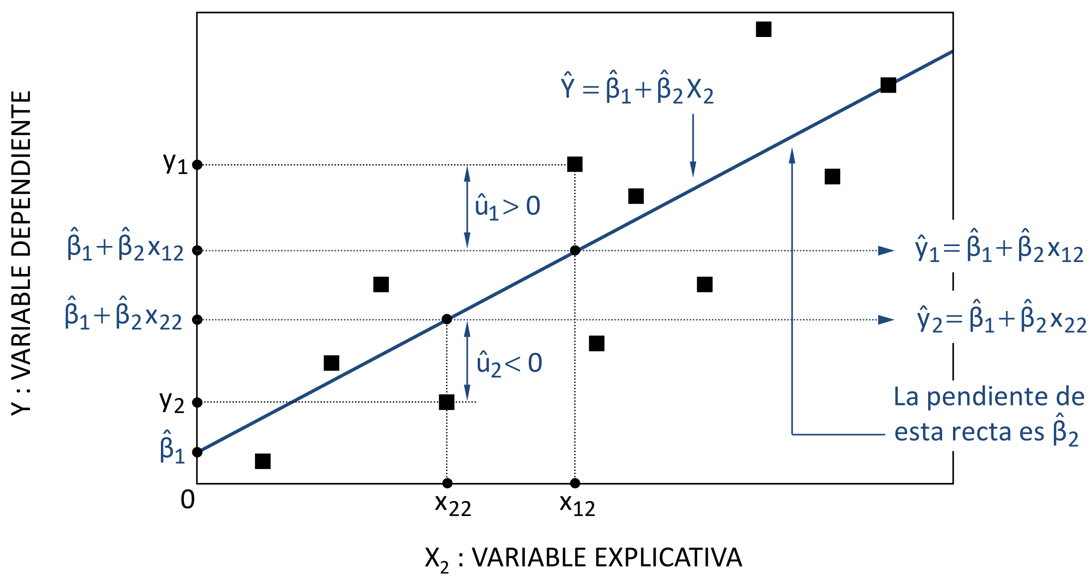
$$\begin{bmatrix} \frac{\partial^2 S}{\partial b_1^2} & \frac{\partial^2 S}{\partial b_1 \partial b_2} \\ \frac{\partial^2 S}{\partial b_2 \partial b_1} & \frac{\partial^2 S}{\partial b_2^2} \end{bmatrix} = 2 \times \begin{bmatrix} N & N\overline{x}_2 \\ N\overline{x}_2 & \sum_{i=1}^{N} x_{i2}^2 \end{bmatrix}, \text{ definida positiva sí y sólo sí}$$
$$N\sum_{i=1}^{N} x_{i2}^2 - N^2 \overline{x}_2^2 = N\left(\sum_{i=1}^{N} x_{i2}^2 - N\overline{x}_2^2 \right) = N\sum_{i=1}^{N} (x_{i2} - \overline{x}_2)^2 > 0.$$

La solución de [4] es

$$\hat{\beta}_1 = \overline{y} - \hat{\beta}_2 \overline{x}_2 ,$$
$$\hat{\beta}_2 = \frac{\sum_{i=1}^{N} (x_{i2} - \overline{x}_2)(y_i - \overline{y})}{\sum_{i=1}^{N} (x_{i2} - \overline{x}_2)^2} = \frac{\widehat{\text{cov}}[\mathbf{x}_2, \mathbf{y}]}{\widehat{\text{var}}[\mathbf{x}_2]} = \frac{\widehat{\text{dvt}}[\mathbf{y}]}{\widehat{\text{dvt}}[\mathbf{x}_2]} \times \widehat{\text{corr}}[\mathbf{x}_2, \mathbf{y}]. \quad [5]$$

Valores Ajustados MCO: $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_{i2} \; (i = 1, ..., N).$ [6]

Residuos MCO: $\hat{u}_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_{i2}) \; (i = 1, ..., N).$ [7]



EJEMPLO

Datos:

$$N = 5, \; \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} 84 \\ 108 \\ 92 \\ 110 \\ 106 \end{bmatrix}, \; \mathbf{x}_2 = \begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \\ x_{42} \\ x_{52} \end{bmatrix} = \begin{bmatrix} 32 \\ 40 \\ 36 \\ 44 \\ 48 \end{bmatrix}.$$

Medias Muestrales:

$$\bar{y} = \tfrac{1}{5} \textstyle\sum_{i=1}^{5} y_i = \tfrac{1}{5} \times 500 = 100,$$
$$\bar{x}_2 = \tfrac{1}{5} \textstyle\sum_{i=1}^{5} x_{i2} = \tfrac{1}{5} \times 200 = 40.$$

Estimaciones MCO:

$$\hat{\beta}_2 = \frac{\sum_{i=1}^{N} (x_{i2} - \bar{x}_2)(y_i - \bar{y})}{\sum_{i=1}^{N} (x_{i2} - \bar{x}_2)^2} = \frac{248}{160}$$
$$= \frac{\widehat{\text{cov}}[\mathbf{x}_2, \mathbf{y}]}{\widehat{\text{var}}[\mathbf{x}_2]} = \frac{49.6}{32}$$
$$= \frac{\widehat{\text{dvt}}[\mathbf{y}]}{\widehat{\text{dvt}}[\mathbf{x}_2]} \times \widehat{\text{corr}}[\mathbf{x}_2, \mathbf{y}] = \frac{\sqrt{104}}{\sqrt{32}} \times \frac{49.6}{\sqrt{32} \times \sqrt{104}} = 1.55,$$

$$\hat{\beta}_1 = \bar{y} - \hat{\beta}_2 \bar{x}_2 = 100 - 1.55 \times 40 = 38.$$

Modelo Estimado: $\hat{Y} = 38 + 1.55 X_2 .$

Valores Ajustados MCO:

$$\hat{y}_i = 38 + 1.55 x_{i2} \; (i = 1, ..., 5)$$
$$= 87.6, 100, 93.8, 106.2, 112.4.$$

Residuos MCO:

$$\hat{u}_i = y_i - \hat{y}_i = y_i - (38 + 1.55 x_{i2}) \ (i = 1, ..., N)$$
$$= -3.6, 8, -1.8, 3.8, -6.4.$$

PARTE 3 - PROPIEDADES ALGEBRAICAS FUNDAMENTALES

Modelo:

$$Y = \beta_1 + \beta_2 X_2 + ... + \beta_K X_K + U. \qquad [8]$$

Parámetros:

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{bmatrix} = [\beta_1, \beta_2, ..., \beta_K]'. \qquad [9]$$

Datos:

$$[\mathbf{y}, \mathbf{x}_2, ..., \mathbf{x}_K] = \begin{bmatrix} y_1 & x_{12} & \cdots & x_{1K} \\ y_2 & x_{22} & \cdots & x_{2K} \\ \vdots & \vdots & & \vdots \\ y_N & x_{N2} & \cdots & x_{NK} \end{bmatrix}. \qquad [10]$$

Criterio MCO.1: Minimizar $S(\mathbf{b}) = \sum_{i=1}^{N} e_i(\mathbf{b})^2$, con

$$\mathbf{b} = [b_1, b_2, ..., b_K]' \quad \text{(cualquier estimación de } \boldsymbol{\beta}\text{)},$$
$$e_i(\mathbf{b}) = y_i - (b_1 + b_2 x_{i2} + ... + b_K x_{iK}) \quad (i = 1, 2, ..., N).$$

El vector de residuos $\mathbf{e}(\mathbf{b})$ asociado con cualquier estimación $\mathbf{b}$ de $\boldsymbol{\beta}$ es

$$\mathbf{e}(\mathbf{b}) = \begin{bmatrix} e_1(\mathbf{b}) \\ e_2(\mathbf{b}) \\ \vdots \\ e_N(\mathbf{b}) \end{bmatrix} = \begin{bmatrix} y_1 - (b_1 + b_2 x_{12} + ... + b_K x_{1K}) \\ y_2 - (b_1 + b_2 x_{22} + ... + b_K x_{2K}) \\ \vdots \\ y_N - (b_1 + b_2 x_{N2} + ... + b_K x_{NK}) \end{bmatrix}$$

$$= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} 1 & x_{12} & \dots & x_{1K} \\ 1 & x_{22} & \dots & x_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & x_{N2} & \dots & x_{NK} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} = \mathbf{y} - \mathbf{X}\mathbf{b}, \text{ con}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \ \mathbf{X} = [\mathbf{i}, \mathbf{x}_2, ..., \mathbf{x}_K] = \begin{bmatrix} 1 & x_{12} & \dots & x_{1K} \\ 1 & x_{22} & \dots & x_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & x_{N2} & \dots & x_{NK} \end{bmatrix}. \qquad [11]$$

En [11], $\mathbf{i} = [1, 1, ..., 1]'$ es un vector columna de N unos asociado con el término constante β_1 del modelo [8], y, para $j = 2, ..., K$, $\mathbf{x}_j$ es un vector columna que contiene los N datos disponibles $(x_{1j}, x_{2j}, ..., x_{Nj})$ sobre la variable explicativa X_j (como en [10]).

Criterio MCO.2: Minimizar $S(\mathbf{b}) = \mathbf{e}(\mathbf{b})'\mathbf{e}(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b})$. [12]

PRA.01: El vector gradiente $(K \times 1)$ y la matriz hessiana $(K \times K)$ de la función $S(\mathbf{b})$ en [12] son, respectivamente, $\nabla S(\mathbf{b}) = 2\mathbf{X}'\mathbf{X}\mathbf{b} - 2\mathbf{X}'\mathbf{y}$ y $\nabla^2 S(\mathbf{b}) = 2\mathbf{X}'\mathbf{X}$, por lo que un vector $\hat{\boldsymbol{\beta}}$ $(K \times 1)$ es la solución (única) de [12] sí y sólo sí $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}'\mathbf{y} = \mathbf{0}$ y $\mathbf{X}'\mathbf{X}$ es una matriz definida positiva.

Demostración: La función $S(\mathbf{b})$ en [12] se puede escribir como

$$S(\mathbf{b}) = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) =$$

$$= \mathbf{y}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} - \mathbf{y}'\mathbf{X}\mathbf{b} - \mathbf{b}'\mathbf{X}'\mathbf{y} = \mathbf{y}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} - 2\mathbf{b}'\mathbf{X}'\mathbf{y},$$

donde (ver [11]):

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{i}' \\ \mathbf{x}_2' \\ \vdots \\ \mathbf{x}_K' \end{bmatrix} [\mathbf{i}, \mathbf{x}_2, ..., \mathbf{x}_K] = \begin{bmatrix} \mathbf{i}'\mathbf{i} & \mathbf{i}'\mathbf{x}_2 & \cdots & \mathbf{i}'\mathbf{x}_K \\ \mathbf{x}_2'\mathbf{i} & \mathbf{x}_2'\mathbf{x}_2 & \cdots & \mathbf{x}_2'\mathbf{x}_K \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_K'\mathbf{i} & \mathbf{x}_K'\mathbf{x}_2 & \cdots & \mathbf{x}_K'\mathbf{x}_K \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1K} \\ q_{21} & q_{22} & \cdots & q_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1} & q_{K2} & \cdots & q_{KK} \end{bmatrix} \Rightarrow$$

$\Rightarrow \mathbf{X}'\mathbf{X}$ simétrica $\Rightarrow q_{ji} = q_{ij}$ (ver también PRA.02),

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} \mathbf{i}'\mathbf{y} \\ \mathbf{x}_2'\mathbf{y} \\ \vdots \\ \mathbf{x}_K'\mathbf{y} \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_K \end{bmatrix},$$

$$\mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} = [b_1, b_2, ..., b_K] \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1K} \\ q_{21} & q_{22} & \cdots & q_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1} & q_{K2} & \cdots & q_{KK} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} =$$

$$= b_1 \sum_{j=1}^{K} q_{1j} b_j + b_2 \sum_{j=1}^{K} q_{2j} b_j + ... + b_K \sum_{j=1}^{K} q_{Kj} b_j,$$

$$\mathbf{b}'\mathbf{X}'\mathbf{y} = [b_1, b_2, ..., b_K] \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_K \end{bmatrix} = \sum_{j=1}^{K} b_j q_j.$$

Por lo tanto,

$$S(\mathbf{b}) = \mathbf{y}'\mathbf{y} + \mathbf{b}'\mathbf{X}'\mathbf{X}\mathbf{b} - 2\mathbf{b}'\mathbf{X}'\mathbf{y} =$$
$$= \left[\textstyle\sum_{j=1}^{N} y_j^2\right] + \left[b_1 \textstyle\sum_{j=1}^{K} q_{1j} b_j + b_2 \sum_{j=1}^{K} q_{2j} b_j + \dots + b_K \sum_{j=1}^{K} q_{Kj} b_j\right] - \left[2 \textstyle\sum_{j=1}^{K} b_j q_j\right].$$

Teniendo en cuenta la simetría de $\mathbf{X}'\mathbf{X}$ $(\Rightarrow q_{ji} = q_{ij})$,

$$\begin{aligned}
\frac{\partial S(\mathbf{b})}{\partial b_1} &= \left[\textstyle\sum_{j=1}^{K} q_{1j} b_j + q_{11} b_1\right] + b_2 q_{21} + \dots + b_K q_{K1} - 2q_1 \\
&= \left[\textstyle\sum_{j=1}^{K} q_{1j} b_j\right] + q_{11} b_1 + q_{12} b_2 + \dots + q_{1K} b_K - 2q_1 \\
&= 2\textstyle\sum_{j=1}^{K} q_{1j} b_j - 2q_1, \\
\frac{\partial S(\mathbf{b})}{\partial b_2} &= b_1 q_{12} + \left[\textstyle\sum_{j=1}^{K} q_{2j} b_j + q_{22} b_2\right] + \dots + b_K q_{K2} - 2q_2 \\
&= \left[\textstyle\sum_{j=1}^{K} q_{2j} b_j\right] + q_{21} b_1 + q_{22} b_2 + \dots + q_{2K} b_K - 2q_2 \\
&= 2\textstyle\sum_{j=1}^{K} q_{2j} b_j - 2q_2, \\
&\vdots \\
\frac{\partial S(\mathbf{b})}{\partial b_K} &= b_1 q_{1K} + b_2 q_{2K} + \dots + \left[\textstyle\sum_{j=1}^{K} q_{Kj} b_j + q_{KK} b_K\right] - 2q_K \\
&= \left[\textstyle\sum_{j=1}^{K} q_{Kj} b_j\right] + q_{K1} b_1 + q_{K2} b_2 + \dots + q_{KK} b_K - 2q_K \\
&= 2\textstyle\sum_{j=1}^{K} q_{Kj} b_j - 2q_K,
\end{aligned}$$

de manera que,

$$\nabla S(\mathbf{b}) = \begin{bmatrix} \frac{\partial S(\mathbf{b})}{\partial b_1} \\ \frac{\partial S(\mathbf{b})}{\partial b_2} \\ \vdots \\ \frac{\partial S(\mathbf{b})}{\partial b_K} \end{bmatrix} = \begin{bmatrix} 2\sum_{j=1}^{K} q_{1j} b_j - 2q_1 \\ 2\sum_{j=1}^{K} q_{2j} b_j - 2q_2 \\ \vdots \\ 2\sum_{j=1}^{K} q_{Kj} b_j - 2q_K \end{bmatrix} = 2\mathbf{X}'\mathbf{X}\mathbf{b} - 2\mathbf{X}'\mathbf{y},$$

y teniendo en cuenta de nuevo la simetría de $\mathbf{X}'\mathbf{X}$ $(\Rightarrow q_{ji} = q_{ij})$,

$$\nabla^2 S(\mathbf{b}) = \begin{bmatrix} \frac{\partial^2 S(\mathbf{b})}{\partial b_1^2} & \frac{\partial^2 S(\mathbf{b})}{\partial b_1 \partial b_2} & \cdots & \frac{\partial^2 S(\mathbf{b})}{\partial b_1 \partial b_K} \\ \frac{\partial^2 S(\mathbf{b})}{\partial b_2 \partial b_1} & \frac{\partial^2 S(\mathbf{b})}{\partial b_2^2} & \cdots & \frac{\partial^2 S(\mathbf{b})}{\partial b_2 \partial b_K} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 S(\mathbf{b})}{\partial b_K \partial b_1} & \frac{\partial^2 S(\mathbf{b})}{\partial b_K \partial b_2} & \cdots & \frac{\partial^2 S(\mathbf{b})}{\partial b_K^2} \end{bmatrix} = 2 \times \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1K} \\ q_{21} & q_{22} & \cdots & q_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ q_{K1} & q_{K2} & \cdots & q_{KK} \end{bmatrix} = 2\mathbf{X}'\mathbf{X},$$

lo que prueba la primera parte del enunciado. Para la segunda parte, $\hat{\boldsymbol{\beta}}$ es la única solución de [12] sí y sólo sí $\nabla S(\hat{\boldsymbol{\beta}}) = \mathbf{0}$ y $\nabla^2 S(\hat{\boldsymbol{\beta}})$ es una matriz definida positiva, lo que equivale a que $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}'\mathbf{y} = \mathbf{0}$ y $\mathbf{X}'\mathbf{X}$ sea definida positiva. ■

Ecuaciones "Normales": $\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$, con [13]

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{12} & x_{22} & \cdots & x_{N2} \\ \vdots & \vdots & & \vdots \\ x_{1K} & x_{2K} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} 1 & x_{12} & \dots & x_{1K} \\ 1 & x_{22} & \dots & x_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & x_{N2} & \dots & x_{NK} \end{bmatrix}$$

$$= \begin{bmatrix} N & \sum_{i=1}^{N} x_{i2} & \cdots & \sum_{i=1}^{N} x_{iK} \\ \sum_{i=1}^{N} x_{i2} & \sum_{i=1}^{N} x_{i2}^2 & \cdots & \sum_{i=1}^{N} x_{i2}x_{iK} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} x_{iK} & \sum_{i=1}^{N} x_{iK}x_{i2} & \cdots & \sum_{i=1}^{N} x_{iK}^2 \end{bmatrix}, \ \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix},$$

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{12} & x_{22} & \cdots & x_{N2} \\ \vdots & \vdots & & \vdots \\ x_{1K} & x_{2K} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{i2}y_i \\ \vdots \\ \sum_{i=1}^{N} x_{iK}y_i \end{bmatrix}.$$

PRA.02: Si $\text{rango}(\mathbf{X}) = K < N$ (ausencia de multicolinealidad exacta), entonces $\mathbf{X}'\mathbf{X}$ es una matriz simétrica definida positiva y, en consecuencia, no singular.

Demostración: En primer lugar, $(\mathbf{X}'\mathbf{X})' = (\mathbf{X})'(\mathbf{X}')' = \mathbf{X}'\mathbf{X}$ (simétrica). Adicionalmente, si $\mathbf{d} \neq \mathbf{0}$ es un vector $K \times 1$, entonces $\mathbf{c} = \mathbf{X}\mathbf{d}$ es una combinación lineal de las K columnas de $\mathbf{X}$, que son linealmente independientes entre sí [$\Leftarrow \text{rango}(\mathbf{X}) = K < N$]. Por lo tanto, $\mathbf{c} \neq \mathbf{0}$ y $\mathbf{d}'(\mathbf{X}'\mathbf{X})\mathbf{d} = (\mathbf{X}\mathbf{d})'(\mathbf{X}\mathbf{d}) = \mathbf{c}'\mathbf{c} = \sum_{i=1}^{N} c_i^2 > 0$, de manera que $\mathbf{X}'\mathbf{X}$ es simétrica y definida positiva, lo que implica, en particular, que los autovalores de $\mathbf{X}'\mathbf{X}$ son todos números reales positivos $\Rightarrow |\mathbf{X}'\mathbf{X}| > 0 \Rightarrow \mathbf{X}'\mathbf{X}$ es no singular. ■

Estimación MCO de β: $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$ [14]

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix} = \begin{bmatrix} N & \sum_{i=1}^{N} x_{i2} & \cdots & \sum_{i=1}^{N} x_{iK} \\ \sum_{i=1}^{N} x_{i2} & \sum_{i=1}^{N} x_{i2}^2 & \cdots & \sum_{i=1}^{N} x_{i2}x_{iK} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} x_{iK} & \sum_{i=1}^{N} x_{iK}x_{i2} & \cdots & \sum_{i=1}^{N} x_{iK}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_{i2}y_i \\ \vdots \\ \sum_{i=1}^{N} x_{iK}y_i \end{bmatrix}.$$

EJ2 HASTA A

Observación: La expresión [14] sugiere que para calcular $\hat{\boldsymbol{\beta}}$ en la práctica es necesario (i) calcular explícitamente $(\mathbf{X}'\mathbf{X})^{-1}$, y (ii) multiplicar $(\mathbf{X}'\mathbf{X})^{-1}$ por $\mathbf{X}'\mathbf{y}$. Aunque la ejecución de las operaciones (i)-(ii) tiene cierto interés en algunos ejemplos sencillos, lo cierto es que $\hat{\boldsymbol{\beta}}$ se calcula realmente (utilizando algún programa de ordenador) como la solución del sistema de ecuaciones lineales [13] sin necesidad de invertir $\mathbf{X}'\mathbf{X}$ de manera explícita (por ejemplo, mediante el "método de eliminación de Gauss-Jordan", o a través de alguna "factorización" de la matriz $\mathbf{X}$). Los elementos de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$ serán necesarios más adelante para otras operaciones, pero no para calcular $\hat{\boldsymbol{\beta}}$. Para más detalles sobre estas cuestiones, ver, por ejemplo, la Sección 1.5 en Davidson, R., MacKinnon, J.M. (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

Valores Ajustados MCO: $$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}\mathbf{y}. \quad [15]$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \begin{bmatrix} \hat{\beta}_1 + \hat{\beta}_2 x_{12} + \ldots + \hat{\beta}_K x_{1K} \\ \hat{\beta}_1 + \hat{\beta}_2 x_{22} + \ldots + \hat{\beta}_K x_{2K} \\ \vdots \\ \hat{\beta}_1 + \hat{\beta}_2 x_{N2} + \ldots + \hat{\beta}_K x_{NK} \end{bmatrix} = \begin{bmatrix} 1 & x_{12} & \ldots & x_{1K} \\ 1 & x_{22} & \ldots & x_{2K} \\ \vdots & \vdots & & \vdots \\ 1 & x_{N2} & \ldots & x_{NK} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix} = \mathbf{X}\hat{\boldsymbol{\beta}}.$$

Residuos MCO: $$\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{H}\mathbf{y} = (\mathbf{I} - \mathbf{H})\mathbf{y} = \mathbf{M}\mathbf{y}. \quad [16]$$

$$\hat{\mathbf{u}} = \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_N \end{bmatrix} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ \vdots \\ y_N - \hat{y}_N \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \mathbf{y} - \hat{\mathbf{y}}.$$

Ls matrices $\mathbf{H}$ y $\mathbf{M}$ en [15]-[16] son simétricas ($\mathbf{H}' = \mathbf{H}$, $\mathbf{M}' = \mathbf{M}$) e idempotentes ($\mathbf{HH} = \mathbf{H}$, $\mathbf{MM} = \mathbf{M}$). Además, $\mathbf{HX} = \mathbf{X}$ ("projection/hat matrix") y $\mathbf{MX} = \mathbf{0}$ ("annihilator").

PRA.03:

$$\mathbf{X}'\hat{\mathbf{u}} = \mathbf{0}, \quad [17]$$

$$\hat{\mathbf{y}}'\hat{\mathbf{u}} = 0, \quad [18]$$

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\mathbf{u}}'\hat{\mathbf{u}}, \quad [19]$$

$$\hat{\mathbf{u}}'\hat{\mathbf{u}} = \mathbf{y}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{y}} = \mathbf{y}'\mathbf{y} - \hat{\boldsymbol{\beta}}'\mathbf{X}'\hat{\mathbf{y}}, \quad [20]$$

$$\hat{\mathbf{y}}'\hat{\mathbf{y}} = \hat{\mathbf{y}}'\mathbf{y}. \quad [21]$$

Demostración: [13] y [16] implican que $\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0} \Leftrightarrow \mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{0} \Leftrightarrow \mathbf{X}'\hat{\mathbf{u}} = \mathbf{0}$, lo que prueba [17]. [15] y [17] implican que $\hat{\mathbf{y}}'\hat{\mathbf{u}} = (\mathbf{X}\hat{\boldsymbol{\beta}})'\hat{\mathbf{u}} = \hat{\boldsymbol{\beta}}'(\mathbf{X}'\hat{\mathbf{u}}) = \hat{\boldsymbol{\beta}}'\mathbf{0} = 0$, lo que prueba [18]. [16] implica que $\mathbf{y}'\mathbf{y} = (\hat{\mathbf{y}} + \hat{\mathbf{u}})'(\hat{\mathbf{y}} + \hat{\mathbf{u}}) = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\mathbf{u}}'\hat{\mathbf{u}} + \hat{\mathbf{y}}'\hat{\mathbf{u}} + \hat{\mathbf{u}}'\hat{\mathbf{y}}$, lo que junto con [18] prueba [19]. La primera parte de [20] se deduce directamente de [19]; la segunda parte se deduce de que $\hat{\mathbf{y}}'\hat{\mathbf{y}} = (\mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{X}\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$. [16] implica que $\hat{\mathbf{y}}'\hat{\mathbf{y}} = \hat{\mathbf{y}}'(\mathbf{y} - \hat{\mathbf{u}}) = \hat{\mathbf{y}}'\mathbf{y} - \hat{\mathbf{y}}'\hat{\mathbf{u}}$, lo que junto con [18] prueba [21]. ■

La propiedad [17] se puede escribir explícitamente (ver [11], donde $\mathbf{i} = [1, 1, \ldots, 1]'$) como

$$\mathbf{X}'\hat{\mathbf{u}} = \begin{bmatrix} \mathbf{i}'\hat{\mathbf{u}} \\ \mathbf{x}_2'\hat{\mathbf{u}} \\ \vdots \\ \mathbf{x}_K'\hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{12} & x_{22} & \cdots & x_{N2} \\ \vdots & \vdots & & \vdots \\ x_{1K} & x_{2K} & \cdots & x_{NK} \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_N \end{bmatrix} = \begin{bmatrix} \sum \hat{u}_i \\ \sum x_{i2}\hat{u}_i \\ \vdots \\ \sum x_{iK}\hat{u}_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad [22]$$

de manera que cada vector de datos sobre cada variable explicativa (cada columna de $\mathbf{X}$) es perpendicular (ortogonal, o normal) al vector de residuos MCO.

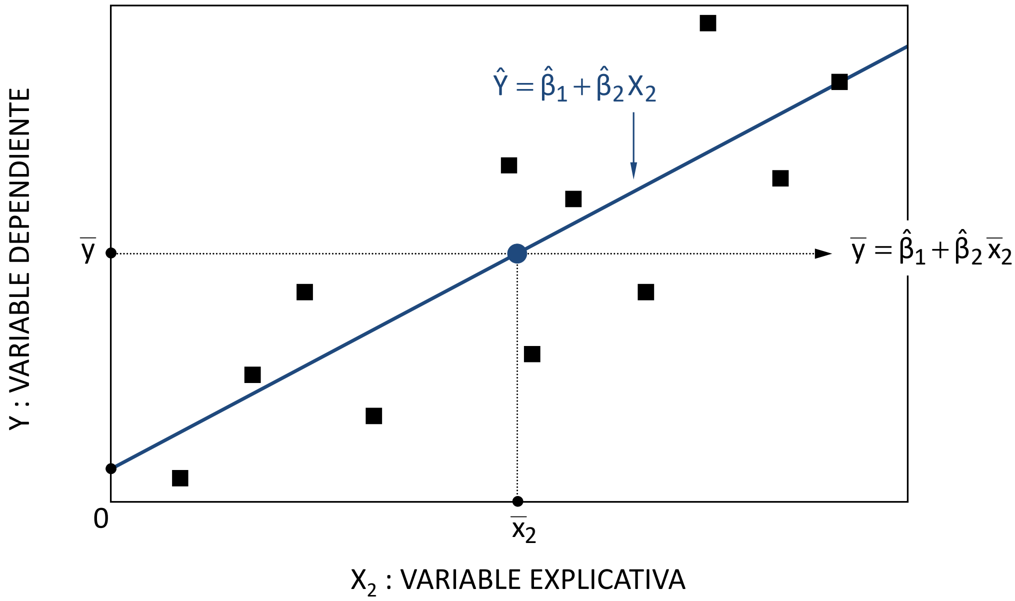
En particular, cuando la primera columna de la matriz $\mathbf{X}$ es una columna de unos, la primera relación de ortogonalidad en [22] ($\mathbf{i}'\hat{\mathbf{u}} = 0$) implica que

$$\textstyle\sum_{i=1}^{N} \hat{u}_i = 0 \Rightarrow \sum_{i=1}^{N} (y_i - \hat{y}_i) = 0 \Rightarrow \sum_{i=1}^{N} y_i = \sum_{i=1}^{N} \hat{y}_i \Rightarrow \bar{y} = \bar{\hat{y}}. \qquad [23]$$

En un modelo RLS, la última relación en [23] establece que

$$\begin{aligned} \bar{y} &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \hat{y}_i = \tfrac{1}{N}\sum_{i=1}^{N} (\hat{\beta}_1 + \hat{\beta}_2 x_{i2}) \\ &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} \hat{\beta}_1 + \tfrac{1}{N}\sum_{i=1}^{N} \hat{\beta}_2 x_{i2} = \hat{\beta}_1 + \hat{\beta}_2 \bar{x}_2, \end{aligned}$$

lo que puede representarse gráficamente como en este diagrama:



EJ2 HASTA B

En la PARTE 5 se consideran otras implicaciones importantes de [23] en modelos con término constante estimados por MCO.

PRA.04: El mero hecho de añadir variables explicativas en un modelo de regresión implica automáticamente una reducción (o, a lo sumo, ningún cambio) en el valor de la suma de cuadrados de los residuos MCO, con independencia de si las variables explicativas añadidas son relevantes en algún sentido o no lo son en ninguno.

Demostración:

$$[\text{M1}]\ \mathbf{y} = \underbrace{\mathbf{X}_1\hat{\boldsymbol{\beta}}}_{\hat{\mathbf{y}}} + \hat{\mathbf{u}},\ [\text{M2}]\ \mathbf{y} = [\,\mathbf{X}_1, \mathbf{X}_2\,]\begin{bmatrix}\hat{\boldsymbol{\beta}}_1 \\ \hat{\boldsymbol{\beta}}_2\end{bmatrix} + \hat{\mathbf{v}} = \underbrace{\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2}_{\hat{\mathbf{w}}} + \hat{\mathbf{v}}. \qquad [24]$$

$$[17] \Rightarrow \qquad \mathbf{X}_1'\hat{\mathbf{u}} = \mathbf{0},\ [\,\mathbf{X}_1, \mathbf{X}_2\,]'\,\hat{\mathbf{v}} = \mathbf{0} \Leftrightarrow \mathbf{X}_1'\hat{\mathbf{v}} = \mathbf{0}, \mathbf{X}_2'\hat{\mathbf{v}} = \mathbf{0}. \qquad [25]$$

$$[24] \Rightarrow \qquad \mathbf{X}_1\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} = \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{v}}. \qquad [26]$$

$$[25]\ [26] \Rightarrow \qquad \begin{aligned} \hat{\mathbf{u}}'\mathbf{X}_1\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}}'\hat{\mathbf{u}} &= \hat{\mathbf{u}}'\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \hat{\mathbf{u}}'\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{u}}'\hat{\mathbf{v}}, \\ \hat{\mathbf{u}}'\hat{\mathbf{u}} &= \hat{\mathbf{u}}'\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{u}}'\hat{\mathbf{v}}. \end{aligned} \qquad [27]$$

$$[25]\ [26] \Rightarrow \qquad \begin{aligned} \hat{\mathbf{v}}'\mathbf{X}_1\hat{\boldsymbol{\beta}} + \hat{\mathbf{v}}'\hat{\mathbf{u}} &= \hat{\mathbf{v}}'\mathbf{X}_1\hat{\boldsymbol{\beta}}_1 + \hat{\mathbf{v}}'\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{v}}'\hat{\mathbf{v}}, \\ \hat{\mathbf{v}}'\hat{\mathbf{u}} &= \hat{\mathbf{v}}'\hat{\mathbf{v}}. \end{aligned} \qquad [28]$$

$$[27]\ [28] \Rightarrow \qquad \hat{\mathbf{u}}'\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 = \hat{\mathbf{u}}'\hat{\mathbf{u}} - \hat{\mathbf{v}}'\hat{\mathbf{v}}. \qquad [29]$$

Haciendo uso de [24], [25] y [29],

$$\begin{aligned} \hat{\mathbf{v}}'\hat{\mathbf{v}} &= (\mathbf{y} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)'(\mathbf{y} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2) \\ &= (\mathbf{X}_1\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2)'(\mathbf{X}_1\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} - \mathbf{X}_1\hat{\boldsymbol{\beta}}_1 - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2) \\ &= [\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{u}}]'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\mathbf{u}}] \\ &= [\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2]'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2] + \hat{\mathbf{u}}'\hat{\mathbf{u}} + 2\hat{\mathbf{u}}'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2] \\ &= [\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2]'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2] + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2\hat{\mathbf{u}}'\mathbf{X}_2\hat{\boldsymbol{\beta}}_2 \\ &= [\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2]'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2] + \hat{\mathbf{u}}'\hat{\mathbf{u}} - 2(\hat{\mathbf{u}}'\hat{\mathbf{u}} - \hat{\mathbf{v}}'\hat{\mathbf{v}}), \end{aligned}$$

lo que implica que

$$\hat{\mathbf{v}}'\hat{\mathbf{v}} = \hat{\mathbf{u}}'\hat{\mathbf{u}} - [\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2]'[\mathbf{X}_1(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_1) - \mathbf{X}_2\hat{\boldsymbol{\beta}}_2] \leq \hat{\mathbf{u}}'\hat{\mathbf{u}}. \qquad [30]$$

La expresión [30] prueba el enunciado. Adicionalmente, [24], [19] y [30] implican que

$$[\mathrm{M1}]\ \mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\mathbf{u}}'\hat{\mathbf{u}},\ [\mathrm{M2}]\ \mathbf{y}'\mathbf{y} = \hat{\mathbf{w}}'\hat{\mathbf{w}} + \hat{\mathbf{v}}'\hat{\mathbf{v}} \Rightarrow \hat{\mathbf{w}}'\hat{\mathbf{w}} \geq \hat{\mathbf{y}}'\hat{\mathbf{y}}. \qquad [31]$$

por lo que el mero hecho de añadir variables explicativas en un modelo de regresión implica automáticamente un aumento (o, a lo sumo, ningún cambio) en el valor de la suma de cuadrados de los valores ajustados MCO, con independencia de si las variables explicativas añadidas son relevantes en algún sentido o no lo son en ninguno. ■

PARTE 4 - GRADO DE AJUSTE

Coeficiente de Determinación No Centrado:

$$\begin{aligned} R^2_{\mathrm{NC}} &= \frac{\sum_{i=1}^N \hat{y}_i^2}{\sum_{i=1}^N y_i^2} = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} \underset{[19]}{\underset{\uparrow}{=}} 1 - \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y}} = 1 - \frac{\sum_{i=1}^N \hat{u}_i^2}{\sum_{i=1}^N y_i^2}, \\ R^2_{\mathrm{NC}} &= \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} = \frac{\hat{\mathbf{y}}'\mathbf{y}}{\hat{\mathbf{y}}'\mathbf{y}} \times \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}}}{\mathbf{y}'\mathbf{y}} \underset{[21]}{\underset{\uparrow}{=}} \frac{(\hat{\mathbf{y}}'\mathbf{y})^2}{(\hat{\mathbf{y}}'\hat{\mathbf{y}})(\mathbf{y}'\mathbf{y})} = \frac{\left(\sum_{i=1}^N \hat{y}_i y_i\right)^2}{\left(\sum_{i=1}^N \hat{y}_i^2\right)\left(\sum_{i=1}^N y_i^2\right)}. \end{aligned} \qquad [32]$$

Coeficiente de Determinación Centrado:

$$R^2 = \frac{\left[\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})\right]^2}{\left[\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2\right]\left[\sum_{i=1}^{N}(y_i - \bar{y})^2\right]} = \frac{\widehat{\text{cov}}[\hat{\mathbf{y}},\mathbf{y}]^2}{\widehat{\text{var}}[\hat{\mathbf{y}}]\widehat{\text{var}}[\mathbf{y}]} = \left(\frac{\widehat{\text{cov}}[\hat{\mathbf{y}},\mathbf{y}]}{\widehat{\text{dvt}}[\hat{\mathbf{y}}]\widehat{\text{dvt}}[\mathbf{y}]}\right)^2 = \widehat{\text{corr}}[\hat{\mathbf{y}},\mathbf{y}]^2,$$

$$R^2 = \frac{\left[\hat{\mathbf{y}}'\mathbf{y} - N\bar{\hat{y}}\bar{y}\right]^2}{\left[\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2\right]\left[\mathbf{y}'\mathbf{y} - N\bar{y}^2\right]} \underset{\substack{\uparrow \\ [21]}}{=} \frac{\left[\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}\bar{y}\right]^2}{\left[\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2\right]\left[\mathbf{y}'\mathbf{y} - N\bar{y}^2\right]}. \qquad [33]$$

Tanto [32] como [33] están siempre comprendidos entre 0 y 1. En los apartados siguientes se consideran otras propiedades importantes de los coeficientes de determinación.

4.1 EL R^2 EN MODELOS CON TÉRMINO CONSTANTE

En cualquier modelo RLM con término constante estimado por MCO ocurre ([22]-[23]) que

$$\sum_{i=1}^{N}\hat{u}_i = 0 \Rightarrow \bar{y} = \bar{\hat{y}}. \qquad [34]$$

En este caso, los símbolos $\bar{y}$ e $\bar{\hat{y}}$ pueden utilizarse indistintamente en [33]. En particular,

$$\hat{\mathbf{y}}'\mathbf{y} - N\bar{\hat{y}}\bar{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2 \Rightarrow R^2 = \frac{\left[\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2\right]^2}{\left[\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2\right]\left[\mathbf{y}'\mathbf{y} - N\bar{y}^2\right]} = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2}{\mathbf{y}'\mathbf{y} - N\bar{y}^2} = \frac{\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{y}^2}{\mathbf{y}'\mathbf{y} - N\bar{y}^2}. \qquad [35]$$

La penúltima expresión para R^2 en [35] implica que

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} = \frac{\widehat{\text{var}}[\hat{\mathbf{y}}]}{\widehat{\text{var}}[\mathbf{y}]}, \qquad [36]$$

y la última expresión para R^2 en [35] que

$$R^2 = 1 - \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{\mathbf{y}'\mathbf{y} - N\bar{y}^2} = 1 - \frac{\sum_{i=1}^{N}\hat{u}_i^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}, \qquad [37]$$

donde se ha tenido en cuenta que [19] implica, en particular, que en cualquier modelo estimado por MCO ocurre que $[\mathbf{y}'\mathbf{y} - N\bar{y}^2] = [\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{y}^2] + \hat{\mathbf{u}}'\hat{\mathbf{u}}$.

4.2 SUMAS DE CUADRADOS

$$\begin{aligned} \text{SCT} &= \textstyle\sum_{i=1}^{N}(y_i - \bar{y})^2 = \tilde{\mathbf{y}}'\tilde{\mathbf{y}} = \mathbf{y}'\mathbf{y} - N\bar{y}^2 = N \times \widehat{\text{var}}[\mathbf{y}], \\ \text{SCE} &= \textstyle\sum_{i=1}^{N}(\hat{y}_i - \bar{\hat{y}})^2 = \hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{\hat{y}}^2 = N \times \widehat{\text{var}}[\hat{\mathbf{y}}], \\ \text{SCR} &= \textstyle\sum_{i=1}^{N}\hat{u}_i^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}}. \end{aligned} \qquad [38]$$

En modelos con término constante, $\bar{\hat{y}} = \bar{y}$ implica que $\text{SCE} = \hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{y}^2$. Por otro lado, teniendo en cuenta [19], $[\mathbf{y}'\mathbf{y} - N\bar{y}^2] = [\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{y}^2] + \hat{\mathbf{u}}'\hat{\mathbf{u}}$ en cualquier modelo estimado por MCO. Por lo tanto, en modelos con término constante, ocurre que

$$\text{SCT} = \text{SCE} + \text{SCR},\ R^2 = \frac{\text{SCE}}{\text{SCT}} = 1 - \frac{\text{SCR}}{\text{SCT}}\,. \qquad [39]$$

Si las definiciones dadas en [38] se aplican en modelos SIN término constante, entonces SCE no coincide con $\hat{\mathbf{y}}'\hat{\mathbf{y}} - N\bar{y}^2$, por lo que las tres igualdades en [39] no son válidas.

PRA.05: El valor del coeficiente de determinación en [32] puede variar al aplicar algunas **transformaciones lineales** en los datos. El valor del coeficiente de determinación en [35]-[37] es invariante frente a cualesquiera transformaciones lineales en los datos.

Demostración: A continuación se expone la demostración para modelos RLS. Para modelos RLM (con término constante), la demostración es completamente análoga (aunque más laboriosa) y proporciona exactamente las mismas conclusiones.

Modelo original: $$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{u}_i\,. \qquad [40]$$

Modelo transformado: $$y_i^* = \hat{\beta}_1^* + \hat{\beta}_2^* x_i^* + \hat{u}_i^*,\ \text{con} \qquad [41]$$

Transformaciones lineales: $$\begin{aligned} y_i^* &= \delta y_i + a \Rightarrow \bar{y}^* = \delta\bar{y} + a, \\ x_i^* &= \mu x_i + b \Rightarrow \bar{x}^* = \mu\bar{x} + b. \end{aligned} \qquad [42]$$

Por ejemplo, si y_i son datos de temperaturas expresados en grados centígrados (Celsius), entonces $y_i^* = 1.8y_i + 32$ $(\delta = 1.8,\, a = 32)$ son esos mismos datos en grados Fahrenheit.

$$[41]\ [42] \Rightarrow \quad \begin{aligned} \delta y_i + a &= \hat{\beta}_1^* + \hat{\beta}_2^*(\mu x_i + b) + \hat{u}_i^*, \\ y_i &= \tfrac{1}{\delta}(\hat{\beta}_1^* - a) + \tfrac{1}{\delta}\hat{\beta}_2^*(\mu x_i + b) + \tfrac{1}{\delta}\hat{u}_i^*, \\ y_i &= \underbrace{\tfrac{1}{\delta}(\hat{\beta}_1^* - a + \hat{\beta}_2^* b)}_{\hat{\beta}_1} + \underbrace{\tfrac{\mu}{\delta}\hat{\beta}_2^*}_{\hat{\beta}_2} x_i + \underbrace{\tfrac{1}{\delta}\hat{u}_i^*}_{\hat{u}_i} \quad (\text{ver } [40]). \end{aligned} \qquad [43]$$

$$[43] \Rightarrow \quad \begin{aligned} \hat{\beta}_2^* &= \tfrac{\delta}{\mu}\hat{\beta}_2,\ \hat{u}_i^* = \delta\hat{u}_i, \\ \hat{\beta}_1^* &= \delta\hat{\beta}_1 + a - \hat{\beta}_2^* b = \delta\hat{\beta}_1 + a - \tfrac{\delta}{\mu}\hat{\beta}_2 b. \end{aligned} \qquad [44]$$

$$[42] \Rightarrow \quad \begin{aligned} \textstyle\sum_{i=1}^N (y_i^*)^2 &= \textstyle\sum_{i=1}^N \delta^2 y_i^2 + \sum_{i=1}^N a^2 + \sum_{i=1}^N 2\delta a y_i \\ &= \delta^2 \textstyle\sum_{i=1}^N y_i^2 + Na^2 + 2\delta a N\bar{y} \\ &= \delta^2 \textstyle\sum_{i=1}^N y_i^2 + Na(a + 2\delta\bar{y}). \end{aligned} \qquad [45]$$

$$[42] \Rightarrow \quad \begin{aligned} &\textstyle\sum_{i=1}^N (y_i^* - \bar{y}^*)^2 = \sum_{i=1}^N [\delta(y_i - \bar{y})]^2 = \delta^2 \sum_{i=1}^N (y_i - \bar{y})^2 \Rightarrow \\ &\Rightarrow \text{SCT}^* = \delta^2 \times \text{SCT}. \end{aligned} \qquad [46]$$

$$[44] \Rightarrow \quad \begin{aligned} &\textstyle\sum_{i=1}^N (\hat{u}_i^*)^2 = \sum_{i=1}^N \delta^2\hat{u}_i^2 = \delta^2 \sum_{i=1}^N \hat{u}_i^2 \Rightarrow \\ &\Rightarrow \text{SCR}^* = \delta^2 \times \text{SCR}. \end{aligned} \qquad [47]$$

$$[45]\ [47] \Rightarrow \qquad R^{2*}_{\text{NC}} = 1 - \frac{\sum_{i=1}^N (\hat{u}_i^*)^2}{\sum_{i=1}^N (y_i^*)^2} = 1 - \frac{\delta^2 \sum_{i=1}^N \hat{u}_i^2}{\delta^2 \sum_{i=1}^N y_i^2 + Na(a+2\delta\overline{y})}, \qquad [48]$$

por lo que $R^{2*}_{\text{NC}} \neq R^2_{\text{NC}}$ a menos que $Na(a + 2\delta\overline{y}) = 0$ ($\Leftrightarrow a = 0$, o bien $a = -2\delta\overline{y}$). La expresión [48] implica en particular que R^{2*}_{NC} puede ser arbitrariamente próximo a 1 si a es lo suficientemente grande. Por el contrario,

$$[46]\ [47] \Rightarrow \qquad R^{2*} = 1 - \frac{\text{SCR}^*}{\text{SCT}^*} = 1 - \frac{\delta^2 \times \text{SCT}}{\delta^2 \times \text{SCR}} = 1 - \frac{\text{SCT}}{\text{SCR}} = R^2, \qquad [49]$$

de manera que el coeficiente de determinación centrado es invariante ante cualesquiera cambios de escala y/o de origen. ■

PRA.06: El mero hecho de añadir variables explicativas en un modelo de regresión implica automáticamente un aumento (o, a lo sumo, ningún cambio) en el valor de los coeficientes de determinación considerados en [32] y en [35]-[37], con independencia de si las variables explicativas añadidas son relevantes en algún sentido o no lo son en ninguno.

Demostración: El enunciado es una implicación directa de la propiedad PRA.04 y de las expresiones [32] y [35]-[37]. ■

4.3 EL COEFICIENTE DE DETERMINACIÓN AJUSTADO EN MODELOS CON TÉRMINO CONSTANTE

$$\overline{R}^2 = 1 - \frac{\text{SCR}/(N-K)}{\text{SCT}/(N-1)} = 1 - \frac{N-1}{N-K}(1 - R^2). \qquad [50]$$

En [50] cabe la posibilidad de que al incluir nuevas variables explicativas en un modelo, la reducción consiguiente en SCR no sea tan grande como la reducción en $N - K$, de manera que $\overline{R}^2$ puede disminuir. Por otro lado (excepto cuando $K = 1$), $\overline{R}^2 < R^2$, por lo que $\overline{R}^2 < 1$; sin embargo, $\overline{R}^2$ puede ser negativo cuando un modelo tiene muy poca capacidad explicativa [cuando SCR/SCT es mayor que $(N-K)/(N-1)$].

EJ2 HASTA C

4.4 EJEMPLO - REGRESIÓN LINEAL SIMPLE: Y = β_1 + β_2X$_2$ + U

$$\begin{aligned} \hat{y}_i &= \hat{y}_i + \hat{u}_i, \text{ con } \hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_{i2}, \\ \hat{\beta}_1 &= \overline{y} - \overline{x}_2\hat{\beta}_2, \\ \hat{\beta}_2 &= \frac{\sum_{i=1}^N (x_{i2} - \overline{x}_2)(y_i - \overline{y})}{\sum_{i=1}^N (x_{i2} - \overline{x}_2)^2} = \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2}. \end{aligned} \qquad [51]$$

$$\begin{aligned} \hat{y}_i &= \overline{y} - \overline{x}_2\hat{\beta}_2 + \hat{\beta}_2 x_{i2} = \overline{y} + \hat{\beta}_2(x_{i2} - \overline{x}_2), \\ \hat{u}_i &= y_i - \hat{y}_i = (y_i - \overline{y}) - \hat{\beta}_2(x_{i2} - \overline{x}_2) = \tilde{y}_i - \tilde{x}_{i2}\hat{\beta}_2, \\ \hat{\mathbf{u}} &= \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\hat{\beta}_2. \end{aligned} \qquad [52]$$

$$\hat{y}_i - \overline{y} = \hat{\beta}_2(x_{i2} - \overline{x}_2) \Rightarrow \text{SCE} = \textstyle\sum_{i=1}^{N}(\hat{y}_i - \overline{y})^2 =$$
$$= \hat{\beta}_2^2 \times \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 = \hat{\beta}_2 \times \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} \times \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 = \hat{\beta}_2\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} = \hat{\beta}_2 \textstyle\sum_{i=1}^{N}(x_{i2} - \overline{x}_2)(y_i - \overline{y}). \quad [53]$$

$$R^2 = \frac{\text{SCE}}{\text{SCT}} = \hat{\beta}_2 \times \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{y}}'\tilde{\mathbf{y}}} = \frac{(\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}})^2}{(\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2)(\tilde{\mathbf{y}}'\tilde{\mathbf{y}})} = \left(\frac{\hat{\text{cov}}[\mathbf{x}_2, \mathbf{y}]}{\hat{\text{dvt}}[\mathbf{x}_2]\hat{\text{dvt}}[\mathbf{y}]}\right)^2 = \hat{\text{corr}}[\mathbf{x}_2, \mathbf{y}]^2. \quad [54]$$

EJ2 HASTA D

4.5 EJEMPLO - REGRESIÓN SOBRE UNA CONSTANTE: Y = β₁ + U

$$\hat{y}_i = \hat{y}_i + \hat{u}_i, \text{ con } \hat{y}_i = \hat{\beta}_1 = \overline{y},$$
$$\hat{y}_i - \overline{y} = 0 \Rightarrow \text{SCE} = \textstyle\sum_{i=1}^{N}(\hat{y}_i - \overline{y})^2 = 0 \Rightarrow R^2 = 0. \quad [55]$$

EJ1 - PARTE PRÁCTICA

PARTE 5 - ALGUNAS PROPIEDADES ALGEBRAICAS ADICIONALES

PRA.07: En cualquier modelo RLM con término constante estimado por MCO,

$$\tilde{\mathbf{y}} = \tilde{\mathbf{x}}_2\hat{\beta}_2 + \tilde{\mathbf{x}}_3\hat{\beta}_3 + \ldots + \tilde{\mathbf{x}}_K\hat{\beta}_K + \hat{\mathbf{u}}. \quad [56]$$

Demostración: De acuerdo con [16] y [11], en cualquier modelo RLM con término constante estimado por MCO,

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} = \mathbf{i}\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{x}_3\beta_3 + \ldots + \mathbf{x}_K\beta_K + \hat{\mathbf{u}}, \quad [57]$$

que se puede escribir observación por observación (fila por fila) como

$$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3} + \ldots + \hat{\beta}_K x_{iK} + \hat{u}_i \quad (i = 1, 2, \ldots, N). \quad [58]$$

Por otro lado, [23] implica que

$$\overline{y} = \hat{\beta}_1 + \hat{\beta}_2\overline{x}_2 + \hat{\beta}_3\overline{x}_3 + \ldots + \hat{\beta}_K\overline{x}_K, \quad [59]$$

y restando esta expresión de [58] se obtiene finalmente que

$$\tilde{y}_i = \hat{\beta}_2\tilde{x}_{i2} + \hat{\beta}_3\tilde{x}_{i3} + \ldots + \hat{\beta}_K\tilde{x}_{iK} + \hat{u}_i \quad (i = 1, 2, \ldots, N).$$

5.1 ESTIMACIÓN MCO CON DATOS EN DESVIACIONES

La expresión [56] se puede escribir como

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}}_{\text{b}}\hat{\boldsymbol{\beta}}_{\text{b}} + \hat{\mathbf{u}}, \text{ con} \quad [60]$$

$$\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_N \end{bmatrix}, \tilde{\mathbf{X}}_{\text{b}} = [\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3, \ldots, \tilde{\mathbf{x}}_K] = \begin{bmatrix} \tilde{x}_{12} & \tilde{x}_{13} & \cdots & \tilde{x}_{1K} \\ \tilde{x}_{22} & \tilde{x}_{23} & \cdots & \tilde{x}_{2K} \\ \vdots & \vdots & & \vdots \\ \tilde{x}_{N2} & \tilde{x}_{N3} & \cdots & \tilde{x}_{NK} \end{bmatrix}, \hat{\boldsymbol{\beta}}_{\text{b}} = \begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_K \end{bmatrix}. \quad [61]$$

[60] ⇒ $$\tilde{\mathbf{X}}_{\text{b}}'\tilde{\mathbf{y}} = \tilde{\mathbf{X}}_{\text{b}}'\tilde{\mathbf{X}}_{\text{b}}\hat{\boldsymbol{\beta}}_{\text{b}} + \tilde{\mathbf{X}}_{\text{b}}'\hat{\mathbf{u}},$$

donde, teniendo en cuenta [22],

$$\begin{aligned}\tilde{\mathbf{X}}_b'\hat{\mathbf{u}} &= [\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3, ..., \tilde{\mathbf{x}}_K]'\hat{\mathbf{u}} \\ &= [(\mathbf{x}_2 - \overline{x}_2\mathbf{i}), (\mathbf{x}_3 - \overline{x}_3\mathbf{i}), ..., (\mathbf{x}_K - \overline{x}_K\mathbf{i})]'\hat{\mathbf{u}} = \mathbf{0},\end{aligned}$$

de manera que

$$\tilde{\mathbf{X}}_b'\tilde{\mathbf{X}}_b\hat{\boldsymbol{\beta}}_b = \tilde{\mathbf{X}}_b'\tilde{\mathbf{y}}. \qquad [62]$$

PRA.08: Si $\text{rango}(\mathbf{X}) = K < N$, entonces $\text{rango}(\tilde{\mathbf{X}}_b) = K - 1 < N$, por lo que $\tilde{\mathbf{X}}_b'\tilde{\mathbf{X}}_b$ es una matriz simétrica definida positiva y, en consecuencia, no singular.

Demostración: En primer lugar, $(\tilde{\mathbf{X}}_b'\tilde{\mathbf{X}}_b)' = (\tilde{\mathbf{X}}_b)'(\tilde{\mathbf{X}}_b')' = \tilde{\mathbf{X}}_b'\tilde{\mathbf{X}}_b$ (simétrica). Por otro lado, $\text{rango}(\mathbf{X}) = K < N$ implica (ver PRA.02) que $\mathbf{Xd} \neq \mathbf{0}$ para todo vector $\mathbf{d} \neq \mathbf{0}$ de tamaño $K \times 1$, de manera que

$$\mathbf{Xd} = d_1\mathbf{i} + \textstyle\sum_{i=2}^K d_i\mathbf{x}_i \neq \mathbf{0} \Rightarrow \textstyle\sum_{i=2}^K d_i\mathbf{x}_i \neq -d_1\mathbf{i},$$

lo que significa que cualquier combinación lineal de las columnas $\mathbf{x}_2, ..., \mathbf{x}_K$ de $\mathbf{X}$ da como resultado un vector cuyos elementos NO son todos iguales entre sí (constantes). Por su parte, cualquier combinación lineal de las columnas de $\tilde{\mathbf{X}}_b$ puede escribirse como

$$\tilde{\mathbf{X}}_b\mathbf{h} = \textstyle\sum_{i=2}^K h_i\tilde{\mathbf{x}}_i = \sum_{i=2}^K h_i(\mathbf{x}_i - \overline{x}_i\mathbf{i}) = \sum_{i=2}^K h_i\mathbf{x}_i - \left(\sum_{i=2}^K h_i\overline{x}_i\right)\mathbf{i},$$

con $\mathbf{h} \neq \mathbf{0}$ de tamaño $(K-1)\times 1$. Como $\left(\sum_{i=2}^K h_i\overline{x}_i\right)\mathbf{i}$ es un vector cuyos elementos son todos iguales entre sí, la conclusión anterior con respecto a $\mathbf{X}$ implica que $\tilde{\mathbf{X}}_b\mathbf{h} \neq \mathbf{0}$, por lo que las $K-1$ columnas de $\tilde{\mathbf{X}}_b$ son linealmente independientes. El resto del enunciado se puede demostrar análogamente a la propiedad PRA.02. ■

PRA.09: Si en un modelo RLM con término constante estimado por MCO consideramos la partición siguiente de $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$:

$$\underset{K\times 1}{\hat{\boldsymbol{\beta}}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix} \begin{matrix} \}\, 1 \\ \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} K-1 \end{matrix} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\boldsymbol{\beta}}_b \end{bmatrix}, \text{ con } \underset{(K-1)\times 1}{\hat{\boldsymbol{\beta}}_b} = \begin{bmatrix} \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_K \end{bmatrix},$$

entonces

$$\begin{aligned}\hat{\beta}_1 &= \overline{y} - \overline{x}_2\hat{\beta}_2 - \overline{x}_3\hat{\beta}_3 - ... - \overline{x}_K\hat{\beta}_K, \\ \hat{\boldsymbol{\beta}}_b &= (\tilde{\mathbf{X}}_b'\tilde{\mathbf{X}}_b)^{-1}\tilde{\mathbf{X}}_b'\tilde{\mathbf{y}}.\end{aligned} \qquad [63]$$

Adicionalmente,

$$\begin{aligned}\hat{\mathbf{u}} &= \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\hat{\beta}_2 - \tilde{\mathbf{x}}_3\hat{\beta}_3 - ... - \tilde{\mathbf{x}}_K\hat{\beta}_K = \tilde{\mathbf{y}} - \tilde{\mathbf{X}}_b\hat{\boldsymbol{\beta}}_b, \\ \underbrace{[\,\hat{\mathbf{u}}'\hat{\mathbf{u}}\,]}_{\text{SCR}} &= \underbrace{[\,\tilde{\mathbf{y}}'\tilde{\mathbf{y}}\,]}_{\text{SCT}} - \underbrace{[\,\hat{\boldsymbol{\beta}}_b'\tilde{\mathbf{X}}_b'\tilde{\mathbf{y}}\,]}_{\text{SCE}} = \underbrace{[\,\text{SCT} \times (1 - R^2)\,]}_{\text{SCR}}.\end{aligned} \qquad [64]$$

Demostración: [Ver también el Apartado 5.5.] Las dos líneas en [63] resultan de [59] y [62], respectivamente. La primera línea en [64] resulta de [56] y [60]. Por otro lado,

$$\begin{aligned}\hat{\mathbf{u}}'\hat{\mathbf{u}} &= (\tilde{\mathbf{y}} - \tilde{\mathbf{X}}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}})'(\tilde{\mathbf{y}} - \tilde{\mathbf{X}}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}}) \\ &= \tilde{\mathbf{y}}'\tilde{\mathbf{y}} + \hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}} - \tilde{\mathbf{y}}'\tilde{\mathbf{X}}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}} - \hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}} \\ &= \tilde{\mathbf{y}}'\tilde{\mathbf{y}} + \hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}}_{\mathrm{b}} - 2\hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}} = \tilde{\mathbf{y}}'\tilde{\mathbf{y}} - \hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}},\end{aligned}$$

lo que, teniendo en cuenta [38]-[39], proporciona la segunda línea en [64]. ■

En [62]-[64],

$$\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}} = \begin{bmatrix}\tilde{\mathbf{x}}_2' \\ \tilde{\mathbf{x}}_3' \\ \vdots \\ \tilde{\mathbf{x}}_K'\end{bmatrix}[\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3, ..., \tilde{\mathbf{x}}_K] = \begin{bmatrix}\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_K \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_K \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_3 & \cdots & \tilde{\mathbf{x}}_K'\tilde{\mathbf{x}}_K\end{bmatrix},$$

$$\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}} = \begin{bmatrix}\tilde{\mathbf{x}}_2' \\ \tilde{\mathbf{x}}_3' \\ \vdots \\ \tilde{\mathbf{x}}_K'\end{bmatrix}\tilde{\mathbf{y}} = \begin{bmatrix}\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{y}} \\ \vdots \\ \tilde{\mathbf{x}}_K'\tilde{\mathbf{y}}\end{bmatrix}. \qquad [65]$$

5.2 EJEMPLO - MODELO RLM Y = $\beta_1 + \beta_2 X_2 + \beta_3 X_3$ + U

EJ3 PR1

En relación con la estimación por MCO del modelo $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U$, el vector de parámetros, el vector de datos sobre la variable dependiente, la matriz de datos sobre las variables explicativas, y la estimación MCO del vector de parámetros, pueden representarse, respectivamente, como:

$$\boldsymbol{\beta} = \begin{bmatrix}\beta_1 \\ \beta_2 \\ \beta_3\end{bmatrix}, \ \mathbf{y} = \begin{bmatrix}y_1 \\ y_2 \\ \vdots \\ y_N\end{bmatrix}, \mathbf{X} = \begin{bmatrix}1 & x_{12} & x_{13} \\ 1 & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ 1 & x_{N2} & x_{N3}\end{bmatrix}, \hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3\end{bmatrix}.$$

En [61],

$$\tilde{\mathbf{y}} \equiv \begin{bmatrix}\tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_N\end{bmatrix}, \tilde{\mathbf{X}}_{\mathrm{b}} \equiv [\tilde{\mathbf{x}}_2, \tilde{\mathbf{x}}_3] \equiv \begin{bmatrix}\tilde{x}_{12} & \tilde{x}_{13} \\ \tilde{x}_{22} & \tilde{x}_{23} \\ \vdots & \vdots \\ \tilde{x}_{N2} & \tilde{x}_{N3}\end{bmatrix}, \hat{\boldsymbol{\beta}}_{\mathrm{b}} \equiv \begin{bmatrix}\hat{\beta}_2 \\ \hat{\beta}_3\end{bmatrix}.$$

En [65],

$$\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}} = \begin{bmatrix} \tilde{\mathbf{x}}_2' \\ \tilde{\mathbf{x}}_3' \end{bmatrix} [\tilde{\mathbf{x}}_2 , \tilde{\mathbf{x}}_3] = \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N \tilde{x}_{i2}^2 & \sum_{i=1}^N \tilde{x}_{i2}\tilde{x}_{i3} \\ \sum_{i=1}^N \tilde{x}_{i3}\tilde{x}_{i2} & \sum_{i=1}^N \tilde{x}_{i3}^2 \end{bmatrix} =$$

$$= N \times \begin{bmatrix} \hat{\mathrm{var}}[\mathbf{x}_2] & \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{x}_3] \\ \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{x}_2] & \hat{\mathrm{var}}[\mathbf{x}_3] \end{bmatrix},$$

$$\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{x}}_2' \\ \tilde{\mathbf{x}}_3' \end{bmatrix} \tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N \tilde{x}_{i2}\tilde{y}_i \\ \sum_{i=1}^N \tilde{x}_{i3}\tilde{y}_i \end{bmatrix} = N \times \begin{bmatrix} \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{y}] \\ \hat{\mathrm{cov}}[\mathbf{x}_3 , \mathbf{y}] \end{bmatrix}.$$

En [63],

$$\hat{\beta}_1 = \bar{y} - \bar{x}_2\hat{\beta}_2 - \bar{x}_3\hat{\beta}_3 ,$$

$$\begin{bmatrix} \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2 & \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} \\ \tilde{\mathbf{x}}_3'\tilde{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \hat{\mathrm{var}}[\mathbf{x}_2] & \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{x}_3] \\ \hat{\mathrm{cov}}[\mathbf{x}_3 , \mathbf{x}_2] & \hat{\mathrm{var}}[\mathbf{x}_3] \end{bmatrix}^{-1} \begin{bmatrix} \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{y}] \\ \hat{\mathrm{cov}}[\mathbf{x}_3 , \mathbf{y}] \end{bmatrix}.$$

Nótese que si $\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 (= \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2) = 0$ $[\Leftrightarrow \hat{\mathrm{cov}}[\mathbf{x}_2 , \mathbf{x}_3] \ (= \hat{\mathrm{cov}}[\mathbf{x}_3 , \mathbf{x}_2]) = 0]$, entonces $\hat{\beta}_2$ coincide con la estimación de la pendiente en la RLS de Y sobre X_2, y $\hat{\beta}_3$ coincide con la estimación de la pendiente en la RLS de Y sobre X_3.

Por último, en [64],

$$\hat{\mathbf{u}} = \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\hat{\beta}_2 - \tilde{\mathbf{x}}_3\hat{\beta}_3 ,$$

$$\underbrace{[\,\hat{\mathbf{u}}'\hat{\mathbf{u}}\,]}_{\text{SCR}} = \underbrace{[\,\tilde{\mathbf{y}}'\tilde{\mathbf{y}}\,]}_{\text{SCT}} - \underbrace{[\,\hat{\beta}_2\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} + \hat{\beta}_3\tilde{\mathbf{x}}_3'\tilde{\mathbf{y}}\,]}_{\text{SCE}} = \underbrace{[\,\textstyle\sum_{i=1}^N \tilde{y}_i^2\,]}_{\text{SCT}} - \underbrace{[\,\hat{\beta}_2 \times \textstyle\sum_{i=1}^N \tilde{x}_{i2}\tilde{y}_i + \hat{\beta}_3 \times \textstyle\sum_{i=1}^N \tilde{x}_{i3}\tilde{y}_i\,]}_{\text{SCE}}.$$

5.3 ESTIMACIONES INDIVIDUALES DE LAS PENDIENTES

En relación con $\hat{\beta}_2$ (para la estimación MCO de cualquier otra pendiente se puede seguir un argumento completamente análogo), premultiplicando [56] por $\tilde{\mathbf{x}}_2'$,

$$\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}} = \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2\hat{\beta}_2 + \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3\hat{\beta}_3 + ... + \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_K\hat{\beta}_K + \tilde{\mathbf{x}}_2'\hat{\mathbf{u}},$$

donde, por [22], $\tilde{\mathbf{x}}_2'\hat{\mathbf{u}} = (\mathbf{x}_2 - \bar{x}_2\mathbf{i})'\hat{\mathbf{u}} = 0$. Por lo tanto, despejando $\hat{\beta}_2$,

$$\hat{\beta}_2 = \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} - \frac{\tilde{\mathbf{x}}_2'}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} [\tilde{\mathbf{x}}_3 , ..., \tilde{\mathbf{x}}_K] \begin{bmatrix} \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_K \end{bmatrix}, \qquad [66]$$

que NO coincide con la estimación de la pendiente en la RLS (ver [51]) de Y sobre X_2, excepto, por ejemplo, cuando $\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 = ... = \tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_K = 0$, o bien cuando $\hat{\beta}_3 = ... = \hat{\beta}_K = 0$.

PRA.10: En un modelo RLM del tipo $Y = \beta_1 + \beta_2 X_2 + ... + \beta_K X_K + U$ con $K \geq 3$, la estimación MCO de cualquier pendiente β_j $(2 \leq j \leq K)$ puede expresarse como

$$\hat{\beta}_j = \frac{\hat{\mathbf{r}}_j'\mathbf{y}}{\hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j}, \qquad [67]$$

donde $\hat{\mathbf{r}}_j$ es el vector de residuos MCO en la regresión con término constante de X_j sobre las demás variables explicativas del modelo RLM considerado.

Demostración: En relación con $\hat{\beta}_2$ (para la estimación MCO de cualquier otra pendiente la demostración es completamente análoga), la regresión con término constante de X_2 sobre las demás variables explicativas estimada por MCO se puede escribir como

$$\mathbf{x}_2 = \hat{\mathbf{x}}_2 + \hat{\mathbf{r}}_2,$$

donde

$$\hat{\mathbf{x}}_2 = \hat{\alpha}_1\mathbf{i} + \hat{\alpha}_3\mathbf{x}_3 + ... + \hat{\alpha}_K\mathbf{x}_K,$$
$$\hat{\mathbf{r}}_2 = \mathbf{x}_2 - \hat{\mathbf{x}}_2 = \mathbf{x}_2 - \hat{\alpha}_1\mathbf{i} - \hat{\alpha}_3\mathbf{x}_3 - ... - \hat{\alpha}_K\mathbf{x}_K.$$

Si $\hat{\mathbf{u}}$ es el vector de residuos MCO en la regresión con término constante de Y sobre todas las variables explicativas, entonces, por la ortogonalidad entre variables explicativas y residuos, ocurre que

$$\hat{\mathbf{r}}_2'\hat{\mathbf{u}} = 0 \Leftrightarrow \hat{\mathbf{r}}_2'(\mathbf{y} - \hat{\beta}_1\mathbf{i} - \hat{\beta}_2\mathbf{x}_2 - \hat{\beta}_3\mathbf{x}_3 - ... - \hat{\beta}_K\mathbf{x}_K) = \hat{\mathbf{r}}_2'\mathbf{y} - \hat{\beta}_2\hat{\mathbf{r}}_2'\mathbf{x}_2 = 0. \qquad [68]$$

Por otro lado, la ortogonalidad entre valores ajustados y residuos implica que

$$\hat{\mathbf{r}}_2'\mathbf{x}_2 = \hat{\mathbf{r}}_2'(\hat{\mathbf{r}}_2 + \hat{\mathbf{x}}_2) = \hat{\mathbf{r}}_2'\hat{\mathbf{r}}_2. \qquad [69]$$

Juntando las expresiones [68] y [69],

$$\hat{\mathbf{r}}_2'\mathbf{y} - \hat{\beta}_2\hat{\mathbf{r}}_2'\hat{\mathbf{r}}_2 = 0 \Leftrightarrow \hat{\beta}_2 = \frac{\hat{\mathbf{r}}_2'\mathbf{y}}{\hat{\mathbf{r}}_2'\hat{\mathbf{r}}_2}. \qquad \blacksquare$$

El numerador en [67] es

$$\hat{\mathbf{r}}_j'\mathbf{y} = \textstyle\sum_{i=1}^N \hat{r}_{ij}y_i,$$

que, como $\sum_{i=1}^N \hat{r}_{ij} = 0$, se puede expresar y calcular de manera equivalente como

$$\hat{\mathbf{r}}_j'\tilde{\mathbf{y}} = \textstyle\sum_{i=1}^N \hat{r}_{ij}(y_i - \bar{y}).$$

Por lo tanto,

$$\hat{\beta}_j = \frac{\hat{\mathbf{r}}_j'\tilde{\mathbf{y}}}{\hat{\mathbf{r}}_j'\hat{\mathbf{r}}_j}, \qquad [70]$$

que coincide con la estimación MCO de la pendiente en la RLS de $\mathbf{y}$ sobre $\hat{\mathbf{r}}_j$.

5.4 EJEMPLO - MODELO RLM $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U$

EJ3 PR1

En el modelo $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U$, [70] y [66] implican (en este orden) que

$$\hat{\beta}_2 = \frac{\hat{\mathbf{r}}_2'\tilde{\mathbf{y}}}{\hat{\mathbf{r}}_2'\hat{\mathbf{r}}_2} = \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} - \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2}\,\hat{\beta}_3\,,$$

donde $\hat{\mathbf{r}}_2$ es el vector de residuos MCO en la RLS de X_2 sobre X_3. Análogamente,

$$\hat{\beta}_3 = \frac{\hat{\mathbf{r}}_3'\tilde{\mathbf{y}}}{\hat{\mathbf{r}}_3'\hat{\mathbf{r}}_3} = \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3} - \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}\,\hat{\beta}_2\,,$$

donde $\hat{\mathbf{r}}_3$ es el vector de residuos MCO en la RLS de X_3 sobre X_2. Nótese que si $\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3 (= \tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2) = 0$, entonces $\hat{\beta}_2$ coincide con la estimación de la pendiente en la RLS de Y sobre X_2, y $\hat{\beta}_3$ coincide con la estimación de la pendiente en la RLS de Y sobre X_3.

Como añadido a lo anterior, el vector de residuos MCO en $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + U$ se puede expresar (ver [64]) como

$$\begin{aligned}
\hat{\mathbf{u}} &= \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\hat{\beta}_2 - \tilde{\mathbf{x}}_3\hat{\beta}_3 \\
&= \tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\left[\frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} - \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2}\,\hat{\beta}_3\right] - \tilde{\mathbf{x}}_3\hat{\beta}_3 \\
&= \left[\tilde{\mathbf{y}} - \tilde{\mathbf{x}}_2\,\frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2}\right] - \left[\tilde{\mathbf{x}}_3 - \tilde{\mathbf{x}}_2\,\frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2}\right]\hat{\beta}_3 = \hat{\mathbf{u}}_{\bar{3}} - \hat{\mathbf{r}}_3\hat{\beta}_3\,,
\end{aligned}$$

donde $\hat{\mathbf{u}}_{\bar{3}}$ es el vector de residuos MCO en la regresión con término constante de Y sobre todas las variables explicativas excepto X_3 (la RLS de Y sobre X_2). Análogamente,

$$\hat{\mathbf{u}} = \left[\tilde{\mathbf{y}} - \tilde{\mathbf{x}}_3\,\frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}\right] - \left[\tilde{\mathbf{x}}_2 - \tilde{\mathbf{x}}_3\,\frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3}\right]\hat{\beta}_2 = \hat{\mathbf{u}}_{\bar{2}} - \hat{\mathbf{r}}_2\hat{\beta}_2\,,$$

donde $\hat{\mathbf{u}}_{\bar{2}}$ es el vector de residuos MCO en la regresión con término constante de Y sobre todas las variables explicativas excepto X_2 (la RLS de Y sobre X_3). En resumen:

$$\begin{array}{ccccccccccccccc}
\text{RLM} & & \text{RLS Y X2} & & \text{RLS X3 X2} & & \text{RLM} & & \text{RLM} & & \text{RLS Y X3} & & \text{RLS X2 X3} & & \text{RLM} \\
\hat{\beta}_2 & = & \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} & - & \frac{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_3}{\tilde{\mathbf{x}}_2'\tilde{\mathbf{x}}_2} & \times & \hat{\beta}_3, & & \hat{\beta}_3 & = & \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{y}}}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3} & - & \frac{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_2}{\tilde{\mathbf{x}}_3'\tilde{\mathbf{x}}_3} & \times & \hat{\beta}_2. \\
 & & \downarrow & & \downarrow & & & & & & \downarrow & & \downarrow & & \\
\hat{\mathbf{u}} & = & \hat{\mathbf{u}}_{\bar{3}} & - & \hat{\mathbf{r}}_3 & \times & \hat{\beta}_3, & & \hat{\mathbf{u}} & = & \hat{\mathbf{u}}_{\bar{2}} & - & \hat{\mathbf{r}}_2 & \times & \hat{\beta}_2.
\end{array}$$

5.5 OTRO VISTAZO A MCO CON DATOS EN DESVIACIONES

Un modelo RLM con término constante y $K \geq 2$ estimado por MCO se puede escribir como

$$\mathbf{y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}} = [\mathbf{i}\ \ \mathbf{X}_{\mathrm{b}}]\begin{bmatrix}\hat{\beta}_1 \\ \hat{\boldsymbol{\beta}}_{\mathrm{b}}\end{bmatrix} + \hat{\mathbf{u}} = \mathbf{i}\hat{\beta}_1 + \mathbf{X}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}} + \hat{\mathbf{u}}. \qquad [71]$$

donde $\mathbf{i}$ es la primera columna de $\mathbf{X}$ (un vector de unos), $\mathbf{X}_{\mathrm{b}} = [\mathbf{x}_2, ..., \mathbf{x}_K]$ (las restantes

$K-1$ columnas de $\mathbf{X}$), y $\hat{\boldsymbol{\beta}}_{\mathrm{b}} = [\hat{\beta}_2, ..., \hat{\beta}_K]'$. La partición de $\mathbf{X}$ como $\mathbf{X} = [\mathbf{i}, \mathbf{X}_{\mathrm{b}}]$ en [71] implica (ver [A1.1]) que

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} \mathbf{i}'\mathbf{i} & \mathbf{i}'\mathbf{X}_{\mathrm{b}} \\ \mathbf{X}_{\mathrm{b}}'\mathbf{i} & \mathbf{X}_{\mathrm{b}}'\mathbf{X}_{\mathrm{b}} \end{bmatrix}, \ \mathbf{X}'\mathbf{y} = \begin{bmatrix} \mathbf{i}'\mathbf{y} \\ \mathbf{X}_{\mathrm{b}}'\mathbf{y} \end{bmatrix},$$

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} (\mathbf{i}'\mathbf{i})^{-1} + (\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\mathbf{X}_{\mathrm{b}}(\mathbf{X}_{\mathrm{b}}'\mathbf{D}\mathbf{X}_{\mathrm{b}})^{-1}\mathbf{X}_{\mathrm{b}}'\mathbf{i}(\mathbf{i}'\mathbf{i})^{-1} & -(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\mathbf{X}_{\mathrm{b}}(\mathbf{X}_{\mathrm{b}}'\mathbf{D}\mathbf{X}_{\mathrm{b}})^{-1} \\ -(\mathbf{X}_{\mathrm{b}}'\mathbf{D}\mathbf{X}_{\mathrm{b}})^{-1}\mathbf{X}_{\mathrm{b}}'\mathbf{i}(\mathbf{i}'\mathbf{i})^{-1} & (\mathbf{X}_{\mathrm{b}}'\mathbf{D}\mathbf{X}_{\mathrm{b}})^{-1} \end{bmatrix},$$

donde $\mathbf{D} \equiv \left[\mathbf{I} - \mathbf{i}(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\right]$ es una matriz simétrica $(\mathbf{D}' = \mathbf{D})$ e idempotente $(\mathbf{D}\mathbf{D} = \mathbf{D})$. De las expresiones anteriores para $(\mathbf{X}'\mathbf{X})^{-1}$ y $\mathbf{X}'\mathbf{y}$, resulta que

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{N} + \overline{\mathbf{x}}'(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\overline{\mathbf{x}} & -\overline{\mathbf{x}}'(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1} \\ -(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\overline{\mathbf{x}} & (\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1} \end{bmatrix}, \qquad [72]$$

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\boldsymbol{\beta}}_{\mathrm{b}} \end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \begin{bmatrix} \overline{y} - \overline{\mathbf{x}}'\hat{\boldsymbol{\beta}}_{\mathrm{b}} \\ (\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}} \end{bmatrix}, \qquad [73]$$

donde $\overline{y} = (\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\mathbf{y}$ es la media muestral de $\mathbf{y}$, $\overline{\mathbf{x}}' = [\overline{x}_2, ..., \overline{x}_K] = (\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\mathbf{X}_{\mathrm{b}}$ es el vector $1 \times (K-1)$ que contiene las medias muestrales de $\mathbf{x}_2, ..., \mathbf{x}_K$, y la matriz $\tilde{\mathbf{X}}_{\mathrm{b}} = \mathbf{D}\mathbf{X}_{\mathrm{b}}$ $[N \times (K-1)]$ y el vector $\tilde{\mathbf{y}} = \mathbf{D}\mathbf{y}$ $[N \times 1]$ contienen datos en desviaciones con respecto a sus medias muestrales correspondientes:

$$\tilde{\mathbf{y}} = \underbrace{\left[\mathbf{I} - \mathbf{i}(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\right]}_{\mathbf{D}}\mathbf{y} = \mathbf{y} - \mathbf{i}\underbrace{\tfrac{1}{N}\mathbf{i}'\mathbf{y}}_{\overline{y}} = \begin{bmatrix} y_1 - \overline{y} \\ \vdots \\ y_N - \overline{y} \end{bmatrix},$$

$$\tilde{\mathbf{X}}_{\mathrm{b}} = [\tilde{\mathbf{x}}_2, ..., \tilde{\mathbf{x}}_K], \text{ con} \qquad [74]$$

$$\tilde{\mathbf{x}}_j = \underbrace{\left[\mathbf{I} - \mathbf{i}(\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'\right]}_{\mathbf{D}}\mathbf{x}_j = \mathbf{x}_j - \mathbf{i}\underbrace{\tfrac{1}{N}\mathbf{i}'\mathbf{x}_j}_{\overline{x}_j} = \begin{bmatrix} x_{1j} - \overline{x}_j \\ \vdots \\ x_{Nj} - \overline{x}_j \end{bmatrix} \ (j = 2, ..., K).$$

Dado que $\mathbf{D}\mathbf{i} = \mathbf{0}$ y $\mathbf{D}\hat{\mathbf{u}} = \hat{\mathbf{u}}$, los residuos MCO en [71] se pueden expresar como

$$\hat{\mathbf{u}} = \tilde{\mathbf{y}} - \tilde{\mathbf{X}}_{\mathrm{b}}\hat{\boldsymbol{\beta}}_{\mathrm{b}} = \left[\mathbf{I} - \tilde{\mathbf{X}}_{\mathrm{b}}(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\tilde{\mathbf{X}}_{\mathrm{b}}'\right]\tilde{\mathbf{y}} = \tilde{\mathbf{M}}_{\mathrm{b}}\tilde{\mathbf{y}}, \qquad [75]$$

con $\tilde{\mathbf{M}}_{\mathrm{b}} = \mathbf{I} - \tilde{\mathbf{X}}_{\mathrm{b}}(\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}})^{-1}\tilde{\mathbf{X}}_{\mathrm{b}}'$ simétrica e idempotente. Por lo tanto (ver [38]-[39]),

$$\underbrace{[\,\hat{\mathbf{u}}'\hat{\mathbf{u}}\,]}_{\text{SCR}} = \underbrace{[\,\tilde{\mathbf{y}}'\tilde{\mathbf{M}}_{\mathrm{b}}\tilde{\mathbf{y}}\,]}_{\text{SCR}} = \underbrace{[\,\tilde{\mathbf{y}}'\tilde{\mathbf{y}}\,]}_{\text{SCT}} - \underbrace{[\,\hat{\boldsymbol{\beta}}_{\mathrm{b}}'\tilde{\mathbf{X}}_{\mathrm{b}}'\tilde{\mathbf{y}}\,]}_{\text{SCE}} = \underbrace{[\,\text{SCT} \times (1 - R^2)\,]}_{\text{SCR}}. \qquad [76]$$

Las fórmulas en [73], [75], [76] son exactamente las mismas que en [63], [64].

APÉNDICE - ALGUNAS FÓRMULAS PARA LA INVERSIÓN DE MATRICES

INV.1: Si

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

es una matriz cuadrada no singular, entonces

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22} \\ -\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{B}_{22} \end{bmatrix}, \text{ con} \qquad \text{[A1.1]}$$

$$\mathbf{B}_{22} = (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1},$$

siempre que $\mathbf{A}_{11}$ sea no singular. Alternativamente,

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{B}_{11} & -\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11} & \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \end{bmatrix}, \text{ con} \qquad \text{[A1.2]}$$

$$\mathbf{B}_{11} = (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1},$$

siempre que $\mathbf{A}_{22}$ sea no singular.

INV.2: Si $\mathbf{A}$ y $\mathbf{D}$ son matrices cuadradas no singulares de orden N y M, respectivamente, y $\mathbf{B}$ y $\mathbf{C}$ son matrices de orden $N \times M$ y $M \times N$, respectivamente, entonces

$$(\mathbf{A} + \mathbf{BDC})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}^{-1} + \mathbf{CA}^{-1}\mathbf{B})^{-1}\mathbf{CA}^{-1}. \qquad \text{[A2]}$$

INV.3: Si $\mathbf{A}$ es una matriz cuadrada no singular de orden N, y $\mathbf{b}$ y $\mathbf{c}$ son vectores de orden $N \times 1$, entonces

$$(\mathbf{A} + \mathbf{bc}')^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{bc}'\mathbf{A}^{-1}}{1 + \mathbf{c}'\mathbf{A}^{-1}\mathbf{b}}. \qquad \text{[A3]}$$

INV.4: Si $\mathbf{A}$, $\mathbf{B}$ y $\mathbf{A} + \mathbf{B}$ son matrices cuadradas no singulares, entonces

$$\begin{aligned} (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1} &= \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B} = \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{A} \\ &= \mathbf{A} - \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{A} = \mathbf{B} - \mathbf{B}(\mathbf{A} + \mathbf{B})^{-1}\mathbf{B}. \end{aligned} \qquad \text{[A4]}$$