UNIVERSIDAD TORCUATO DI TELLA
MAESTRÍA EN ECONOMÍA — MAESTRÍA EN ECONOMETRÍA
2021

Econometría
Problem Set 2
Tópicos en el Modelo Lineal

---

1 Un breve repaso...

La traza de una matriz cuadrada es la suma de los elementos de su diagonal principal. Algunas propiedades:

1. $tr(A+B) = tr(A) + tr(B)$
2. $tr(A') = tr(A)$
3. $tr(AB) = tr(BA)$
4. El rango de una matriz idempotente es igual a su traza.

Sea A una matriz cuadrada con autovalores $\lambda_1, \lambda_2, ..., \lambda_n$. Entonces:

1. $\det A = \lambda_1 \lambda_2, ... \lambda_n$. Si A es una matriz no singular entonces todos sus autovalores son no nulos.
2. $trA = \lambda_1 + \lambda_2 + ... + \lambda_n$.
3. Si A es una matriz idempotente entonces sus autovalores son 0 o 1.

Algunos Teoremas (ver Apéndice del Greene) que nos serán útiles:

Theorem 1 (B.8) *Si $x \sim N(0, I)$ y A es idempotente, entonces $x'Ax \sim \chi^2_k$ donde $k = rank(A) = \#$ autovalores iguales a 1.*

Theorem 2 (B.12) *Sean $x \sim N(0, I)$, Lx una transformación lineal de x, y $x'Ax$ con A idempotente y simétrica. Entonces Lx y $x'Ax$ son independientes si $LA = 0$.*

Recordar del Problem Set 0 el siguiente resultado:
Sea x un vector aleatorio de $n \times 1$ tal que $x \sim N(\mu, \Sigma)$, donde Σ es una matriz simétrica y definida positiva. Entonces:

$$(x-\mu)' \, \Sigma^{-1} \, (x-\mu) \sim \chi^2_n$$

Recordar también:
Sea x un vector aleatorio con $E(x) = \mu$ y $Var(x) = E[(x-\mu)(x-\mu)'] = \Sigma$ ("matriz de varianzas y covarianzas"), entonces, dada una matriz $A: E[Ax] = A\mu$ y $Var[Ax] = A\Sigma A'$.

---

2 Problemas

---

1. Por el supuesto de normalidad: $u|X \sim N(0, \sigma^2_u I_n)$.

   Sea y :

   $$y = X\beta + u$$

   Puede verse que y es una transformación afín de u, condicional a los regresores. Luego, y tiene distribución normal. A partir del supuesto de exogeneidad estricta,

   $$E\left(y|X\right) = X\beta + E\left(u|X\right) = X\beta$$

Además, por el supuesto de errores esféricos,

$$\begin{aligned}
\Sigma_y &= E\left\{[y-E(y|X)][y-E(y|X)]'/X\right\} \\
&= E\left\{[y-X\beta][y-X\beta]'/X\right\} \\
&= E\left\{uu'/X\right\} \\
&= \sigma_u^2 I_n
\end{aligned}$$

Luego,

$$\boxed{y|X \sim \mathcal{N}\left(X\beta, \sigma_u^2 I_n\right)}$$

Sea $\widehat{\beta}$:

$$\widehat{\beta} = \left(X'X\right)^{-1} X'y = \beta + \left(X'X\right)^{-1} X'u$$

El estimador de MCC se escribe como una transformación afín de u. Entonces es normal. Sus momentos ya han sido calculados en el práctico pasado. Luego,

$$\boxed{\widehat{\beta}/X \sim \mathcal{N}\left(\beta, \sigma_u^2\left(X'X\right)^{-1}\right)}$$

Sea ahora $\widehat{u}$:

$$\begin{aligned}
\widehat{u} &= y - X\widehat{\beta} \\
&= y - X\left(X'X\right)^{-1} X'y \\
&= My \\
&= MX\beta + Mu \\
&= Mu
\end{aligned}$$

En el último paso se usó que $MX = (I - X\left(X'X\right)^{-1} X')X = X - X = 0$. El vector de residuos de MCC es entonces una transformación lineal del vector u. Entonces, es normal. Sus momentos relevantes son, bajo los supuestos de exogeneidad estricta y errores esféricos:

$$\begin{aligned}
E\left(\widehat{u}/X\right) &= ME\left(u|X\right) = 0_n \\
\Sigma_{\widehat{u}} &= E\left[\widehat{u}\widehat{u}'/X\right] \\
&= E\left[Muu'M'/X\right] \\
&= E\left[Muu'M/X\right] \\
&= ME\left[uu'/X\right]M \\
&= \sigma_u^2 MM \\
&= \sigma_u^2 M
\end{aligned}$$

en virtud de que M es simétrica e idempotente. Luego,

$$\boxed{\widehat{u}/X \sim \mathcal{N}\left(0_n, \sigma_u^2 M\right)}$$

Sea $\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2}$:

$$\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2} = \frac{\frac{RSS}{n-k}(n-k)}{\sigma_u^2} = \frac{RSS}{\sigma_u^2}$$

$$\begin{aligned}
\text{Además :} &\quad RSS = \widehat{u}'\widehat{u} = u'M'Mu = u'Mu \\
\text{Entonces} \quad : &\quad \frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2} = \frac{u'Mu}{\sigma_u^2}
\end{aligned}$$

Notar que M es idempotente y que $\frac{u}{\sigma_u} \sim N(0, I)$, entonces podemos aprovechar el Teorema B.8 para afirmar que:

$$\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2} = \frac{u'Mu}{\sigma_u^2}\tilde{}$$

$\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2}$ y $\widehat{\beta}$ son transformaciones de u:

$$\begin{aligned}
\widehat{\beta} &= \beta + \left(X'X\right)^{-1} X'u \\
\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2} &= \frac{\widehat{u}'\widehat{u}}{\sigma_u^2} = \frac{u'Mu}{\sigma_u^2}\tilde{}\chi_\nu^2
\end{aligned}$$

Donde $\upsilon = rank(M) = tr(M)$ dado que M es idempotente. La traza de M es $n-k$:

$$\begin{aligned} tr(M) &= tr(I_n - X\left(X'X\right)^{-1}X') \\ &= tr(I_n) - tr\left[(X)((X'X)^{-1}X')\right] \\ &= tr(I_n) - tr\left[((X'X)^{-1}X')(X))\right] \\ &= tr(I_n) - tr(I_k) \\ &= n-k \end{aligned}$$

Entonces:

$$\boxed{\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2} \sim \chi^2_{n-k}}$$

$\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2}$ y $\widehat{\beta}$ son transformaciones de u:

$$\begin{aligned} \widehat{\beta} &= \beta + \left(X'X\right)^{-1}X'u \\ \frac{\widehat{\sigma}_u^2\left(n-k\right)}{\sigma_u^2} &= \frac{\widehat{u}'\widehat{u}}{\sigma_u^2} = \frac{u'Mu}{\sigma_u^2} \end{aligned}$$

Aprovechando el Teorema 8.12, sabemos que estos dos vectores aleatorios son independientes si y sólo si:

$$\begin{aligned} \left(X'X\right)^{-1}X'M &= O_{k\times n} \\ \left(X'X\right)^{-1}X'M' &= O_{k\times n} \\ \left(X'X\right)^{-1}\left(MX\right)' &= O_{k\times n} \\ \left(X'X\right)^{-1}O_{k\times n} &= O_{k\times n} \checkmark \end{aligned}$$

2. Para este ejercicio se utilizarán las distribuciones derivadas en el punto 1. Sea $\widehat{\beta}$:

$$\begin{aligned} \widehat{\beta}/X &\sim \mathcal{N}\left(\beta, \sigma_u^2\left(X'X\right)^{-1}\right) \\ \left(\widehat{\beta}-\beta\right)'\left[\sigma_u^2\left(X'X\right)^{-1}\right]^{-1}\left(\widehat{\beta}-\beta\right) &\sim \chi^2_k \\ \frac{\left(\widehat{\beta}-\beta\right)'\left(X'X\right)\left(\widehat{\beta}-\beta\right)}{\sigma_u^2} &\sim \chi^2_k \\ \frac{\widehat{\sigma}_u^2\left(n-k\right)}{\sigma_u^2} &\sim \chi^2_{n-k} \\ &\Rightarrow \frac{\left\{\frac{\left[\frac{(\widehat{\beta}-\beta)'(X'X)(\widehat{\beta}-\beta)}{\sigma_u^2}\right]}{k}\right\}}{\left\{\frac{\left[\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2}\right]}{n-k}\right\}} \sim \mathcal{F}\left(k, n-k\right) \\ \frac{\left[\frac{(\widehat{\beta}-\beta)'(X'X)(\widehat{\beta}-\beta)}{k}\right]}{\left[\widehat{\sigma}_u^2\right]} &\sim \mathcal{F}\left(k, n-k\right) \\ \frac{\left(\widehat{\beta}-\beta\right)'\left(X'X\right)\left(\widehat{\beta}-\beta\right)}{k\widehat{\sigma}_u^2} &\sim \mathcal{F}\left(k, n-k\right) \end{aligned}$$

La región de confianza es conceptualmente análoga al intervalo de confianza, para vectores de parámetros. En este caso,

$$P\left(\frac{\left(\widehat{\beta}-\beta\right)'\left(X'X\right)\left(\widehat{\beta}-\beta\right)}{k\widehat{\sigma}_u^2} \leq \mathcal{F}_{1-\alpha}\left(k, n-k\right)\right) = 1-\alpha$$

$$\boxed{C\left(\frac{(\widehat{\beta}-\beta)'(X'X)(\widehat{\beta}-\beta)}{k\widehat{\sigma}_u^2} \leq \mathcal{F}_{1-\alpha}\left(k, n-k\right)\right) = 1-\alpha}$$

La siguiente expresión:

$$\frac{\left(\widehat{\beta}-\beta\right)'(X'X)\left(\widehat{\beta}-\beta\right)}{k\widehat{\sigma}_u^2}=\mathcal{F}_{1-\alpha}(k,n-k)$$

es simplemente una curva de nivel de una forma cuadrática. Para el caso $k=2$, tiene representación gráfica como una elipse (o círculo, según los valores) en el espacio $\mathbb{R}^2$.

Sea ahora $c'\beta$, c no nulo:

$$\begin{aligned}
c'\widehat{\beta}/X &\sim \mathcal{N}\left(c'\beta,\sigma_u^2.c'(X'X)^{-1}c\right)\\
\frac{c'\left(\widehat{\beta}-\beta\right)}{\sqrt{\sigma_u^2.c'(X'X)^{-1}c}} &\sim \mathcal{N}(0,1)\\
\frac{c'\left(\widehat{\beta}-\beta\right)}{\sigma_u\sqrt{.c'(X'X)^{-1}c}} &\sim \mathcal{N}(0,1)\\
\frac{\left[\frac{c'\left(\widehat{\beta}-\beta\right)}{\sigma_u\sqrt{.c'(X'X)^{-1}c}}\right]}{\sqrt{\frac{\left(\frac{\widehat{\sigma}_u^2(n-k)}{\sigma_u^2}\right)}{n-k}}} &\sim \tau_{n-k}\\
\frac{\left[\frac{c'\left(\widehat{\beta}-\beta\right)}{\sqrt{.c'(X'X)^{-1}c}}\right]}{[\widehat{\sigma}_u]} &\sim \tau_{n-k}\\
\frac{c'\left(\widehat{\beta}-\beta\right)}{\widehat{\sigma}_u\sqrt{.c'(X'X)^{-1}c}} &\sim \tau_{n-k}\\
\widehat{\sigma}_u\sqrt{.c'(X'X)^{-1}c} &= \widehat{Se(c'\widehat{\beta})}
\end{aligned}$$

El intervalo de confianza es:

$$\begin{aligned}
P\left(\left|\frac{c'\left(\widehat{\beta}-\beta\right)}{\widehat{Se(c'\widehat{\beta})}}\right|\leq\tau_{1-\alpha,n-k}\right) &= 1-\alpha\\
P\left(-\tau_{1-\alpha,n-k}\leq\frac{c'\left(\widehat{\beta}-\beta\right)}{\widehat{Se(c'\widehat{\beta})}}\leq\tau_{1-\alpha,n-k}\right) &= 1-\alpha\\
P\left(-\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\leq c'\left(\widehat{\beta}-\beta\right)\leq\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\right) &= 1-\alpha\\
C\left(-c'\widehat{\beta}-\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\leq -c'\beta\leq -c'\widehat{\beta}+\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\right) &= 1-\alpha
\end{aligned}$$

$$\boxed{C\left(c'\widehat{\beta}-\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\leq c'\beta\leq c'\widehat{\beta}+\tau_{1-\alpha,n-k}\widehat{Se(c'\widehat{\beta})}\right)=1-\alpha}$$

3. $y_t=\beta_1x_{1t}+\beta_2x_{2t}+\beta_3x_{3t}+u_t, \qquad (t=1,\ldots,n)$

   El primero es un Test de Significatividad Individual:

   $t=\frac{\widehat{\beta}_2}{se(\widehat{\beta}_2)}\sim t_{n-k}$

   Según los resultados obtenidos en el Problem Set 1:

   $t=\frac{\widehat{\beta}_2}{se(\widehat{\beta}_2)}=\frac{1}{\sqrt{\frac{1}{30}}}=\sqrt{30}=5.4772$

   El segundo es un Test de Wald:

   $R=[1\quad 1\quad 0]$

   $r=2$

   $H_0: R\beta=2$

   $W=\frac{(R\widehat{\beta}-r)'(Rs^2(X'X)^{-1}R')^{-1}(R\widehat{\beta}-r)}{\#r}\sim F_{\#r,n-k}$

   Según los resultados obtenidos en el Problem Set 1:

   $W=\frac{10}{3}=3.33.$

4. Algunas consideraciones previas. Considere el modelo:

$$y = X\beta + u = X_1\beta_1 + X_2\beta_2 + u$$

Donde X_1 y X_2 son las observaciones de dos sets de variables.

De la condición de primer orden de minimización de RSS obteníamos:

$$\begin{aligned}
(X'X)\widehat{\beta} &= X'y \\
\left[(X_1 \vdots X_2)'(X_1 \vdots X_2)\right]\widehat{\beta} &= (X_1 \vdots X_2)'y \\
\left[\begin{bmatrix} X_1' \\ \cdots \\ X_2' \end{bmatrix}(X_1 \vdots X_2)\right]\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} &= \begin{bmatrix} X_1' \\ \cdots \\ X_2' \end{bmatrix} y \\
\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}\begin{bmatrix} \widehat{\beta}_1 \\ \widehat{\beta}_2 \end{bmatrix} &= \begin{bmatrix} X_1'y \\ X_2'y \end{bmatrix}
\end{aligned}$$

Tenemos entonces el siguiente sistema de ecuaciones:

$X_1'X_1\widehat{\beta}_1 + X_1'X_2\widehat{\beta}_2 = X_1'y$

$X_2'X_1\widehat{\beta}_1 + X_2'X_2\widehat{\beta}_2 = X_2'y$

Consideremos la primera parte de este sistema:

$X_1'X_1\widehat{\beta}_1 = X_1'y - X_1'X_2\widehat{\beta}_2 \Rightarrow \widehat{\beta}_1 = (X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}X_1'X_2\widehat{\beta}_2 \Rightarrow \widehat{\beta}_1 = (X_1'X_1)^{-1}X_1'(y - X_2\widehat{\beta}_2)$

Notar que si $X_1'X_2 = 0$, entonces $\widehat{\beta}_1 = (X_1'X_1)^{-1}X_1'y$, que es el resultado de regresar y sobre X_1. Esto es un teorema: **en un modelo de regresión lineal múltiple de y sobre dos sets de variables X_1 y X_2, si los dos sets de variables son ortogonales, entonces el vector de coeficientes estimados asociados a X_1 se puede obtener haciendo una regresión de y únicamente sobre X_1; idem para el vector de coeficientes estimados asociados a X_2.**

Ahora tomemos la segunda parte del sistema e insertemos la expresión obtenida para $\widehat{\beta}_1$:

$X_2'X_1\widehat{\beta}_1 + X_2'X_2\widehat{\beta}_2 = X_2'y$

$X_2'X_1(X_1'X_1)^{-1}X_1'(y - X_2\widehat{\beta}_2) + X_2'X_2\widehat{\beta}_2 = X_2'y$

$X_2'X_1(X_1'X_1)^{-1}X_1'y - X_2'X_1(X_1'X_1)^{-1}X_1'X_2\widehat{\beta}_2 + X_2'X_2\widehat{\beta}_2 = X_2'y$

$X_2'(X_1(X_1'X_1)^{-1}X_1'y - y) = X_2'(X_1(X_1'X_1)^{-1}X_1'X_2 - X_2)\widehat{\beta}_2$

$X_2'(X_1(X_1'X_1)^{-1}X_1' - I)y = X_2'(X_1(X_1'X_1)^{-1}X_1' - I)X_2\widehat{\beta}_2$

$X_2'(I - X_1(X_1'X_1)^{-1}X_1')y = X_2'(I - X_1(X_1'X_1)^{-1}X_1')X_2\widehat{\beta}_2$

$\left[X_2'(I - X_1(X_1'X_1)^{-1}X_1')X_2\right]^{-1} X_2'(I - X_1(X_1'X_1)^{-1}X_1')y = \widehat{\beta}_2$

$$\widehat{\beta}_2 = \left[X_2'M_1X_2\right]^{-1} X_2'M_1y$$

Finalmente:

$$\widehat{\beta}_1 = \left[X_1'M_2X_1\right]^{-1} X_1'M_2y$$

Ahora bien, en el ejercicio 1 se vio que, dado $y = X\beta + u$, los residuos de la regresión pueden escribirse: $\widehat{u} = My$, donde $M = I - X\left(X'X\right)^{-1}X'$. Análogamente, podemos ver a M_1X_2 como una matriz de residuos, donde cada columna de M_1X_2 es un vector de residuos provenientes de una regresión de dicha columna de X_2 contra X_1. Esta observación nos lleva al Teorema de Frisch-Waugh:

Theorem 3 *Teorema de Frisch-Waugh. En una regresión lineal de y contra dos sets de variables, X_1 y X_2, el subvector $\widehat{\beta}_2$ se puede obtener regresando los residuos provenientes de regresar y contra X_1, contra los residuos provenientes de regresar cada columna de X_2 contra X_1.*

Demostración:

Se realizan las regresiones de y y de todas las columnas de X_2 sobre X_1:

$$\begin{aligned}
y &= X_1\gamma + v \\
X_2 &= X_1\Delta + W
\end{aligned}$$

El segundo modelo es simplemente una notación compactada de todas las regresiones de cada columna de X_2. Nótese que Δ y W son matrices de coeficientes y residuos, respectivamente. Los estimadores de MCC y residuos asociados son:

$$\begin{aligned}
\widehat{\gamma} &= \left(X_1'X_1\right)^{-1} X_1'y \\
\widehat{v} &= y - X_1\widehat{\gamma} = M_1 y \\
\widehat{\Delta} &= \left(X_1'X_1\right)^{-1} X_1'X_2 \\
\widehat{W} &= X_2 - X_1\widehat{\Delta} = M_1 X_2 \\
\text{Donde } M_1 &= I_n - X_1\left(X_1'X_1\right)^{-1} X_1'
\end{aligned}$$

Sea ahora la regresión de los residuos del primer modelo sobre la matriz de residuos del segundo modelo:

$$\begin{aligned}
\widehat{v} &= \widehat{W}\phi + \varepsilon \\
\widehat{\phi} &= \left(\widehat{W}'\widehat{W}\right)^{-1} \widehat{W}'\widehat{v} \\
\widehat{\phi} &= \left(X_2'M_1'M_1X_2\right)^{-1} X_2'M_1'M_1y \\
\widehat{\phi} &= \left(X_2'M_1X_2\right)^{-1} X_2'M_1y
\end{aligned}$$

Por los resultados de regresión particionada,

$$\widehat{\beta}_2 = \left(X_2'M_1X_2\right)^{-1} X_2'M_1y$$

Luego,

$$\boxed{\widehat{\phi} = \widehat{\beta}_2}$$

que es el resultado de Frisch-Waugh. Intercambiando los subíndices 1 y 2 es claro que también vale para $\widehat{\beta}_1$.

Notar:

$$\begin{aligned}
\widehat{\phi} &= \left(X_2'X_2\right)^{-1} X_2'y \Leftrightarrow M_1X_2 = X_2 \\
&\Leftrightarrow X_2 - X_1\left(X_1'X_1\right)^{-1} X_1'X_2 = X_2 \\
&\Leftrightarrow X_1\left(X_1'X_1\right)^{-1} X_1'X_2 = O \\
&\Leftrightarrow \boxed{X_1'X_2 = O}
\end{aligned}$$

En este caso se dice que los regresores son ortogonales.

Nótese entonces que en un modelo de regresión lineal múltiple con variable dependiente salario y variables explicativas educación y promedio en la carrera de grado, para obtener la estimación de mínimos cuadrados clásicos del coeficiente de educación se puede:

-Empezar regresando salario sobre promedio; y eduación sobre promedio.

-Obtener los residuos de la primera regresión y regresarlos sobre los residuos de la segunda regresión.

Así acabamos haciendo sólamente regresiones simples.

5. Sea el modelo

$$y = X\beta + c.z + u = \left[\begin{array}{ccc} X & \vdots & z \end{array}\right] \left[\begin{array}{c} \beta \\ \cdots \\ c \end{array}\right] + u$$

donde z es un vector columna con las observaciones de uno de los regresores del modelo. Se asume que la primer columna de X es una columna de unos.

Considere las regresiones parciales respecto de X y los residuos asociados:

$$\begin{aligned}
y^* &= y - X\widehat{\theta} \\
z^* &= z - X\widehat{\eta}
\end{aligned}$$

Como ambas regresiones tienen intercepto,

$$\sum_{i=1}^{n} y_i^* = \sum_{i=1}^{n} z_i^* = 0 \Rightarrow \overline{y^*} = \overline{z^*} = 0$$

Del ejercicio anterior sabemos que podemos obtener $\widehat{c}$ de regresar y^* contra z^*: $y^* = cz^* + \varepsilon$.

El coeficiente de correlación (muestral) parcial entre y, z respecto de X es:

$$r_{y,z;X} = r_{y^*,z^*} = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(y_i^* - \overline{y^*}\right)\left(z_i^* - \overline{z^*}\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i^* - \overline{y^*}\right)^2}\sqrt{\frac{1}{n}\left(z_i^* - \overline{z^*}\right)^2}} = \frac{y^{*T}z^*}{\sqrt{y^{*T}y^*}\sqrt{z^{*T}z^*}}$$

De acuerdo a los resultados de regresión particionada y de de MCC,

$$\begin{aligned} y^* &= \widehat{c}z^* + \widehat{\varepsilon} : z^{*T}\widehat{\varepsilon} = 0 \\ y^{*T} &= \widehat{c}z^{*T} + \widehat{\varepsilon}' \\ y^{*T}z^* &= \widehat{c}z^{*T}z^* + \widehat{\varepsilon}'z^* = c\left(z^{*T}z^*\right) \end{aligned}$$

Reemplazando esta expresión en el coeficiente de correlación parcial,

$$\boxed{r_{y,z;X} = \widehat{c}\frac{\sqrt{z^{*T}z^*}}{\sqrt{y^{*T}y^*}}}$$

6. Sea Q_n la matriz sample demeaning:

$$\begin{aligned} Q_n x &= \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} \overline{x} \\ \vdots \\ \overline{x} \end{bmatrix} = \widetilde{x} \\ Q_n &= I_n - \frac{1}{n}1_n 1_n' \end{aligned}$$

Sea el modelo con intercepto:

$$y = X\beta + u$$

y la partición:

$$y = \begin{bmatrix} 1_n & \vdots & X_{-1} \end{bmatrix}\begin{bmatrix} \beta_1 \\ \cdots \\ \beta_{-1} \end{bmatrix} + u$$

donde β_1 es el intercepto del modelo.

Por los resultados de regresión particionada,

$$\begin{aligned} \widehat{\beta}_{-1} &= \left(X_{-1}' M_1 X_{-1}\right)^{-1} X_{-1}' M_1 y \\ \widehat{\beta}_{-1} &= \left(X_{-1}' M_1' M_1 X_{-1}\right)^{-1} X_{-1}' M_1' M_1 y \\ \widehat{\beta}_{-1} &= \left[\left(M_1 X_{-1}\right)'\left(M_1 X_{-1}\right)\right]^{-1}\left(M_1 X_{-1}\right)'\left(M_1 y\right) \\ M_1 &= I_n - 1_n\left(1_n' 1_n\right)^{-1} 1_n' \\ M_1 &= I_n - 1_n\left(n\right)^{-1} 1_n' \\ M_1 &= I_n - \frac{1}{n}1_n 1_n' \\ &\Rightarrow M_1 = Q_n \\ \widehat{\beta}_{-1} &= \left[\left(Q_n X_{-1}\right)'\left(Q_n X_{-1}\right)\right]^{-1}\left(Q_n X_{-1}\right)'\left(Q_n y\right) \\ & \boxed{\widehat{\beta}_{-1} = \left[\widetilde{X}_{-1}'\widetilde{X}_{-1}\right]^{-1}\widetilde{X}_{-1}'\widetilde{y}} \end{aligned}$$

7. El estimador de mínimos cuadrados restringidos es el estimador de β que resuelve el siguiente problema:

$$\begin{aligned} &\min_{\widetilde{\beta}}\left(y - X\widetilde{\beta}\right)'\left(y - X\widetilde{\beta}\right) \\ s.a &: R\widetilde{\beta} = r \end{aligned}$$

El lagrangiano del problema es:

$$\begin{aligned} \mathcal{L} &= \left(y - X\widetilde{\beta}\right)'\left(y - X\widetilde{\beta}\right) - 2\lambda'\left(R\widetilde{\beta} - r\right) \\ &= y'y - 2y'X\widetilde{\beta} + \widetilde{\beta}'X'X\widetilde{\beta} - 2\lambda'\left(R\widetilde{\beta} - r\right) \end{aligned}$$

Las condiciones de primer orden son:

$$
\begin{aligned}
-2X'y + 2X'X\widetilde{\beta} - 2R'\lambda &= 0 \\
R\widetilde{\beta} &= r
\end{aligned}
$$

Trabajamos un poco las expresiones:

$$
\begin{aligned}
-X'y + X'X\widetilde{\beta} - R'\lambda &= 0 \\
-R(X'X)^{-1}X'y + R(X'X)^{-1}X'X\widetilde{\beta} &= R(X'X)^{-1}R'\lambda \\
-R\widehat{\beta} + R\widetilde{\beta} &= R(X'X)^{-1}R'\lambda \\
-R\widehat{\beta} + r &= R(X'X)^{-1}R'\lambda \\
\left(R(X'X)^{-1}R'\right)^{-1}\left(r - R\widehat{\beta}\right) &= \lambda \\
-X'y + X'X\widetilde{\beta} - R'\left(R(X'X)^{-1}R'\right)^{-1}\left(r - R\widehat{\beta}\right) &= 0
\end{aligned}
$$

$$
\begin{aligned}
X'X\widetilde{\beta} &= X'y + R'\left(R(X'X)^{-1}R'\right)^{-1}\left(r - R\widehat{\beta}\right) \\
\widetilde{\beta} &= (X'X)^{-1}X'y + (X'X)^{-1}R'\left(R(X'X)^{-1}R'\right)^{-1}\left(r - R\widehat{\beta}\right) \\
\widetilde{\beta} &= \widehat{\beta} + (X'X)^{-1}R'\left(R(X'X)^{-1}R'\right)^{-1}\left(r - R\widehat{\beta}\right) \\
\widetilde{\beta} &= \widehat{\beta} - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left(R\widehat{\beta} - r\right) \\
\widetilde{\beta} &= \left\{I_k - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\right\}\widehat{\beta} \\
&\quad + (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}r \\
\text{Reescribimos usando } \widehat{\beta} &= \beta + (X'X)^{-1}X'u \\
\widetilde{\beta} &= \left\{I_k - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\right\}\beta \\
&\quad + (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}r \\
&\quad + \left\{I_k - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\right\}(X'X)^{-1}X'u \\
&= \beta - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\beta \\
&\quad + (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}r \\
&\quad + \left\{I_k - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\right\}(X'X)^{-1}X'u \\
&= \beta + \left\{I_k - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R\right\}(X'X)^{-1}X'u \\
&= \beta + Cu
\end{aligned}
$$

que es una transformación afín de u. Luego, es normal. Busquemos la esperanza y varianza:

$$
\begin{aligned}
E(\widetilde{\beta}/X) &= \beta \\
Var(\widetilde{\beta}/X) &= E\left[(\widetilde{\beta} - \beta)(\widetilde{\beta} - \beta)'/X\right] \\
&= E\left[Cu(Cu)'/X\right] \\
&= CE\left[u\,u'/X\right]C \\
&= \sigma_u^2 CC' \\
&= \sigma_u^2\left\{(X'X)^{-1} - (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}R(X'X)^{-1}\right\}
\end{aligned}
$$

La distribución es:

$$\boxed{\widehat{\beta}_R/X \sim \mathcal{N}\left(\beta, \sigma_u^2 \left\{ (X'X)^{-1} - (X'X)^{-1} R' \left[R (X'X)^{-1} R' \right]^{-1} R (X'X)^{-1} \right\}\right)}$$

8. El vector de residuos restringidos y su suma de cuadrados son:

$$\begin{aligned}
\widehat{u}_R &= y - X\widehat{\beta}_R \\
&= y - X\widehat{\beta}_R + X\widehat{\beta} - X\widehat{\beta} \\
&= \widehat{u} - X\left(\widehat{\beta}_R - \widehat{\beta}\right) \\
\widehat{u}_R'\widehat{u}_R &= \left[\widehat{u} - X\left(\widehat{\beta}_R - \widehat{\beta}\right)\right]' \left[\widehat{u} - X\left(\widehat{\beta}_R - \widehat{\beta}\right)\right] \\
&= \widehat{u}'\widehat{u} + \left(\widehat{\beta}_R - \widehat{\beta}\right)' X'X \left(\widehat{\beta}_R - \widehat{\beta}\right) \\
&\quad - \left(\widehat{\beta}_R - \widehat{\beta}\right)' X'\widehat{u} - \widehat{u}'X\left(\widehat{\beta}_R - \widehat{\beta}\right) \\
X'\widehat{u} &= 0_k \Rightarrow \widehat{u}_R'\widehat{u}_R = \widehat{u}'\widehat{u} + \left(\widehat{\beta}_R - \widehat{\beta}\right)' X'X \left(\widehat{\beta}_R - \widehat{\beta}\right) \\
\widehat{\beta}_R - \widehat{\beta} &= -(X'X)^{-1} R' \left[R (X'X)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right) \\
\left(\widehat{\beta}_R - \widehat{\beta}\right)' X'X \left(\widehat{\beta}_R - \widehat{\beta}\right) &= \left(R\widehat{\beta} - r\right)' \left[R (X'X)^{-1} R'\right]^{-1} R (X'X)^{-1} (X'X) \\
&\quad (X'X)^{-1} R' \left[R (X'X)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right) \\
&= \left(R\widehat{\beta} - r\right)' \left[R (X'X)^{-1} R'\right]^{-1} R (X'X)^{-1} \\
&\quad R' \left[R (X'X)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right) \\
&= \left(R\widehat{\beta} - r\right)' \left[R (X'X)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right) \\
&\Rightarrow \boxed{\widehat{u}_R'\widehat{u}_R - \widehat{u}'\widehat{u} = \left(R\widehat{\beta} - r\right)' \left[R (X'X)^{-1} R'\right]^{-1} \left(R\widehat{\beta} - r\right) \geq 0}
\end{aligned}$$

Este resultado implica que el estadístico F equivale al siguiente estadístico de residuos restringidos:

$$\begin{aligned}
\frac{\left[\frac{\widehat{u}_R'\widehat{u}_R - \widehat{u}'\widehat{u}}{q}\right]}{\widehat{\sigma}_u^2} &\sim {}^{H_0}\mathcal{F}(q, n-k) \\
\frac{\left[\frac{\widehat{u}_R'\widehat{u}_R - \widehat{u}'\widehat{u}}{q}\right]}{\left[\frac{\widehat{u}'\widehat{u}}{n-k}\right]} &\sim {}^{H_0}\mathcal{F}(q, n-k) \\
&\boxed{\frac{\left[\frac{SSR_R - SSR}{q}\right]}{\left[\frac{SSR}{n-k}\right]} \sim^{H_0} \mathcal{F}(q, n-k)}
\end{aligned}$$

9. Para probar este resultado, nótese que la inclusión de nuevos regresores a un modelo puede pensarse como un caso particular de una restricción de exclusión de q variables:

$$y = X\beta + \varepsilon \rightarrow y = X\beta + Z\gamma + u \Leftrightarrow \gamma = 0_q$$

donde Z es la matriz de observaciones de q regresores adicionales. Sean RSS_0 la suma de residuos cuadrados del modelo restringido y RSS_1, del modelo ampliado. Estas sumas pueden pensarse como:

$$\begin{aligned}
RSS_1 &= RSS = \widehat{u}'\widehat{u} \\
RSS_0 &= RSS_R = \widehat{\varepsilon}'\widehat{\varepsilon} = \widehat{u}_R'\widehat{u}_R
\end{aligned}$$

Por el resultado del ejercicio 8:

$$\begin{aligned}
\widehat{u}_R'\widehat{u}_R &\geq \widehat{u}'\widehat{u} \\
\frac{\widehat{u}_R'\widehat{u}_R}{y'y-n\overline{y}^2} &\geq \frac{\widehat{u}'\widehat{u}}{y'y-n\overline{y}^2} \\
-\frac{\widehat{u}_R'\widehat{u}_R}{y'y-n\overline{y}^2} &\leq -\frac{\widehat{u}'\widehat{u}}{y'y-n\overline{y}^2} \\
1-\frac{\widehat{u}_R'\widehat{u}_R}{y'y-n\overline{y}^2} &\leq 1-\frac{\widehat{u}'\widehat{u}}{y'y-n\overline{y}^2} \\
& \boxed{R_0^2 \leq R_1^2}\checkmark
\end{aligned}$$

10. Probamos la convergencia en probabilidad. Sea

$$\begin{aligned}
\widehat{\sigma}_u^2 &= \frac{\widehat{u}'\widehat{u}}{n-k} \\
&= \frac{u'Mu}{n-k} \\
&= \frac{u'u-u'X(X'X)^{-1}X'u}{n-k} \\
&= \frac{n}{n-k}\left(\frac{u'u}{n}-\frac{u'X}{n}(\frac{X'X}{n})^{-1}\frac{X'u}{n}\right) \\
p\lim\widehat{\sigma}_u^2 &= 1.\left(E(u_i^2)-0_k'Q^{-1}0_k\right) \\
p\lim\widehat{\sigma}_u^2 &= E(u_i^2) \\
p\lim\widehat{\sigma}_u^2 &= \sigma_u^2
\end{aligned}$$

11. Sea el modelo simple

$$y_i=\alpha+\beta x_i^*+u_i$$

El investigador estima en su lugar el siguiente modelo

$$\begin{aligned}
y_i &= \alpha+\beta\left(x_i-\varepsilon_i\right)+u_i \\
y_i &= \alpha+\beta x_i+u_i-\beta\varepsilon_i \\
y_i &= \alpha+\beta x_i+v_i
\end{aligned}$$

Por tratarse de un modelo simple:

$$\begin{aligned}
\widehat{\beta} &= \frac{\sum_{i=1}^n(x_i-\overline{x})y_i}{\sum_{i=1}^n(x_i-\overline{x})^2} \\
\widehat{\beta} &= \frac{\sum_{i=1}^n(x_i-\overline{x})(\alpha+\beta x_i+v_i)}{\sum_{i=1}^n(x_i-\overline{x})^2} \\
\widehat{\beta} &= \frac{\sum_{i=1}^n(x_i-\overline{x})\alpha}{\sum_{i=1}^n(x_i-\overline{x})^2}+\beta\frac{\sum_{i=1}^n(x_i-\overline{x})x_i}{\sum_{i=1}^n(x_i-\overline{x})^2}+\frac{\sum_{i=1}^n(x_i-\overline{x})v_i}{\sum_{i=1}^n(x_i-\overline{x})^2} \\
\widehat{\beta} &= \beta+\frac{\sum_{i=1}^n(x_i-\overline{x})v_i}{\sum_{i=1}^n(x_i-\overline{x})^2} \\
&= \beta+\frac{\left(\frac{\sum_{i=1}^n(x_i-\overline{x})v_i}{n}\right)}{\left(\frac{\sum_{i=1}^n(x_i-\overline{x})^2}{n}\right)} \\
p\lim\widehat{\beta} &= \beta+\frac{cov(x_i,v_i/X)}{var(x_i/X)}
\end{aligned}$$

En este caso,

$$\begin{aligned}
cov(x_i^*, \varepsilon_i/X) &= 0 \\
var(x_i/X) &= \sigma_{x^*}^2 + \sigma_\varepsilon^2 \\
cov(x_{i,} v_i/X) &= E(x_i.v_i/X) \\
&= E(x_i.u_i/X) - \beta E(x_i.\varepsilon_i/X) \\
&= E(x_i.u_i/X) - \beta E((x_i^* + \varepsilon_i).\varepsilon_i/X) \\
&= 0 - \beta 0 - \beta \sigma_\varepsilon^2 \\
&= -\beta \sigma_\varepsilon^2 \\
&\Rightarrow p\lim \widehat{\beta} = \beta - \frac{\beta \sigma_\varepsilon^2}{\sigma_{x^*}^2 + \sigma_\varepsilon^2} \\
p\lim \widehat{\beta} &= \beta \left[1 - \frac{\sigma_\varepsilon^2}{\sigma_{x^*}^2 + \sigma_\varepsilon^2} \right] \\
&= \beta \left[\frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_\varepsilon^2} \right] \\
&= \beta \left[\frac{1}{1 + \left(\frac{\sigma_\varepsilon^2}{\sigma_{x^*}^2} \right)} \right]
\end{aligned}$$

El estimador de MCC del coeficiente pendiente es inconsistente. El parámetro aparece multiplicado por un factor menos a uno, que se conoce como "sesgo de atenuación". El tamaño de este sesgo depende de la varianza relativa de la variable y el error de medición. Si este error fuera sistemático, $\sigma_\varepsilon^2 = 0$, el estimador preservaría la consistencia.