Ejercicio 1

Sea el modelo de regresión lineal

$$y = X\beta + u$$

donde y es un vector de $n \times 1$, X es una matriz de $n \times k$ observaciones de variables con $ran(X) = k < n$ con probabilidad 1, β es un vector de $k \times 1$ de parámetros desconocidos, y u es un vector de perturbaciones estocásticas de $n \times 1$ con $\mathbf{E}[u/X] = 0$ y $\mathbf{E}[uu'] = \sigma^2 I_n$. El estimador OLS de β es

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

De aquí sale directamente la propiedad de linealidad. Para que dicho estimador exista es necesario que exista la inversa de $(X'X)$, lo cual se asegura con el supuesto de rango completo. Considerando que $y = \beta X + u$ se puede reescribir al estimador como

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

Por lo tanto

$$\mathbf{E}\left[\hat{\beta}/X\right] = \beta + (X'X)^{-1} X'\mathbf{E}[u/X] = \beta$$

por el supuesto de exogeneidad estricta. Entonces el estimador de OLS es insesgado. El Teorema de Gauss-Markov afirma que el estimador OLS es el de mínima varianza entre todos los estimadores linealmente insesgados. Con este fin debemos comparar su matriz de varianzas y covarianzas con la del resto de los estimadores linealmente insesgados. Sabemos que

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

Entonces, sea $\Sigma_{\hat{\beta}}$ la matriz de varianzas y covarianzas del estimador OLS condicional en X

$$\Sigma_{\hat{\beta}} = \mathbf{E}\left[\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)'/X\right] = \mathbf{E}\left[(X'X)^{-1} X'uu'X (X'X)^{-1}/X\right] = (X'X)^{-1} X' \underbrace{\mathbf{E}[uu'/X]}_{\sigma^2 I_n} X (X'X)^{-1} = \sigma^2 (X'X)^{-1}$$

Sea $\tilde{\beta}$ otro estimador linealmente insesgado. Por linealidad $\tilde{\beta} = \tilde{D}y$, con $\tilde{D}$ siendo una matriz de $k \times n$. Se puede reescribir como $\tilde{\beta} = \tilde{D}X\beta + \tilde{D}u$, para que sea insesgado necesitamos que

$$\mathbf{E}\left[\tilde{\beta}/X\right] = \mathbf{E}\left[\tilde{D}X\beta + \tilde{D}u/X\right] = \tilde{D}X\beta = \beta \Longleftrightarrow \tilde{D}X = I_k$$

La matriz de varianzas y covarianzas de $\tilde{\beta}$ condicional en X, $\Sigma_{\tilde{\beta}}$ es

$$\Sigma_{\tilde{\beta}} = \mathbf{E}\left[\left(\tilde{\beta} - \beta\right)\left(\tilde{\beta} - \beta\right)'/X\right] = \mathbf{E}\left[\tilde{D}uu'\tilde{D}'/X\right] = \sigma^2 \tilde{D}\tilde{D}'$$

Para probar que $\hat{\beta}$ es de varianza mínima basta con probar que $\Sigma_{\tilde{\beta}} - \Sigma_{\hat{\beta}}$ es una matriz semidefinida positiva. Con este fin definamos $D = \tilde{D} - (X'X)^{-1} X'$. Notar que $DX = 0$ (esto implica también que $D'X' = 0$. $\Sigma_{\tilde{\beta}}$ se puede reescribir, entonces

$$\begin{aligned}\Sigma_{\tilde{\beta}} &= \sigma^2 \left(D + (X'X)^{-1} X'\right)\left(D + (X'X)^{-1} X'\right)' \\ &= \sigma^2 \left(DD' + (X'X)^{-1} X'D' + DX (X'X)^{-1} + (X'X)^{-1}\right) = \sigma^2 DD' + \sigma^2 (X'X)^{-1}\end{aligned}$$

Como DD' es una matriz semidefinida positiva entonces $\Sigma_{\tilde{\beta}} - \Sigma_{\hat{\beta}} = DD'\sigma^2$ es semidefinida positiva. Esto prueba que el estimador OLS de β es el de mínimza varianza, y completa la prueba que son BLUE. Sea $\theta = c'\beta$, odnde c es un vector de $k \times 1$ variables, se desea construir un estimador BLUE De θ. Queremos probar que el estimador BLUE es $c'\hat{\beta}$, con $\hat{\beta}$ como el estiamdor OLS de β. Como es lineal considere el estiamdor $\hat{\theta} = a'y$, donde a es un vector de $n \times 1$ variables. Como $\hat{\theta}$ debe ser insesgado

$$\mathbf{E}\left[\hat{\theta}/X\right] = \mathbf{E}[a'(X\beta + u)] = a'X\beta + a'\mathbf{E}[u/X] = a'X\beta = \beta \Longleftrightarrow a'X = c' \implies X'a = c$$

De aquí se desprende que

$$\hat{\theta} = a'X\beta + a'u \implies \hat{\theta} - \theta = a'u$$

La matriz de varianzas y covarianzas condicional en X, $\Sigma_{\hat{\theta}}$ en este caso es un escalar, y tiene la forma

$$\Sigma_{\hat{\theta}} = \mathbf{E}\left[\left(\hat{\theta} - \theta\right)\left(\hat{\theta} - \theta\right)'/X\right] = \mathbf{E}[a'uu'a/X] = \sigma^2 a'a$$

Encontrar el estimador BLUE implica minimizar la varianza condicional $\Sigma_{\hat{\theta}}$ sujeto a que sea insesgado. Como σ^2 es un escalar positivo este problema equivale a

$$\begin{aligned} &\max_a a'a \\ &s.a. X'a = c \end{aligned}$$

El lagrangiano de este problema y las condiciones de primer orden (necesarias y suficientes) son

$$\mathcal{L} = a'a - 2\lambda' \left(X'a - c\right)$$

$$\begin{aligned} a' = \lambda' X' &\implies a = X\lambda \\ X'a &= c \end{aligned}$$

Utilizanndo ambas ecuaciones llegamos a que $\lambda = \left(X'X\right)^{-1} c$. Esto implica que $a = X\left(X'X\right)^{-1} c$. Utilizando esto en $\hat{\theta}$ llegamos a que $\hat{\theta} = c'\left(X'X\right)^{-1} X'y = c'\hat{\beta}$, que es lo que queríamos probar.

Ejercicio 2

Sea $\hat{\beta}$ el estimador OLS de β, la condición de primer orden del problema es la siguiente:

$$\left(\nabla_{\hat{\beta}} \left(y - X\hat{\beta}\right)\left(y - X\hat{\beta}\right)'\right)' = 0 \implies -2X'y + 2\left(X'X\right)\hat{\beta} = 0$$

La condición de segundo orden del problema implica que la matriz Hessiana sea semidefinida positiva. En este caso

$$H\left(\left(y - X\hat{\beta}\right)\left(y - X\hat{\beta}\right)'\right) = D_{\hat{\beta}}\left(\nabla_{\hat{\beta}}\left(y - \hat{\beta}X\right)\left(y - \hat{\beta}X\right)'\right) = 2\left(X'X\right)$$

Como X es una matriz de rango completo entonces $(X'X)$ es una matriz semidefinida positiva, lo cual verifica la condición de segundo orden del problema.

Ejercicio 3

Sea el modelo $y = X\beta + u$. El estimador OLS de β es $\hat{\beta}$. Consideremos primero un cambio de escala. Un cambio de escala puede representarse como una transformación lineal que afecta a las observaciones de la misma variable por igual. Sean $\tilde{y}$ y $\tilde{X}$ las variables transformadas esto implica que, con $\lambda, \lambda_j \neq 0$

$$\begin{cases} \tilde{y} = \lambda y \\ \tilde{x}_j = \lambda_j x_j \end{cases}$$

En notación matricial esto implica que

$$\tilde{y} = \lambda y$$

$$\tilde{X} = X \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \lambda_k \end{bmatrix}}_{\Lambda} = X\Lambda$$

$$\Lambda = diag\left(\lambda_1, \ldots, \lambda_k\right) \; det(\Lambda) \neq 0 \; \Lambda' = \Lambda$$

El nuevo modelo es, entonces

$$\lambda y = X\Lambda\beta^* + w$$

El estimador OLS de β^* en el nuevo modelo será

$$\hat{\beta}^* = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{y} = \left((X\Lambda)'(X\Lambda)\right)^{-1}(X\Lambda)'\lambda y = \lambda\left(\Lambda' X' X\Lambda\right)^{-1}\Lambda' X' y = \lambda\Lambda^{-1}\left(X'X\right)^{-1}\Lambda^{-1}\Lambda X' y = \lambda\Lambda^{-1}\left(X'X\right)^{-1}X'y$$

$$\hat{\beta}^* = \lambda\Lambda^{-1}\hat{\beta}$$

$$\begin{bmatrix} \hat{\beta}_1^* \\ \hat{\beta}_2^* \\ \vdots \\ \vdots \\ \hat{\beta}_k^* \end{bmatrix} = \begin{bmatrix} \frac{\lambda}{\lambda_1}\hat{\beta}_1 \\ \frac{\lambda}{\lambda_2}\hat{\beta}_2 \\ \vdots \\ \frac{\lambda}{\lambda_k}\hat{\beta}_k \end{bmatrix}$$

Se puede probar que el ajuste del modelo es el mismo

$$\hat{\tilde{y}} = \tilde{X}\hat{\beta}^* \iff \hat{\tilde{y}} = X\Lambda\lambda\Lambda^{-1}\hat{\beta} = \lambda\hat{y} \iff \tilde{y} = \lambda\hat{y} + \hat{w} \iff \lambda\left(y - \hat{y}\right) = \hat{w} \implies \lambda\hat{u} = \lambda\hat{w}$$

Entonces para calcular el ajuste del modelo

$$\tilde{R}^2 = 1 - \frac{\hat{w}'\hat{w}}{\tilde{y}'\tilde{y}} = 1 - \frac{\lambda^2\hat{u}'\hat{u}}{\lambda^2 y'y} = 1 - \frac{\hat{u}'\hat{u}}{y'y} = R^2$$

Consideremos ahora una traslación en los regresores. Sea d un vector columna de k constantes, la traslación consiste en sumar a todas las observaciones de un mismo regresor una misma constante. Es decir

$$\tilde{x}_{ji} = x_{ji} + d_j$$

El nuevo modelo es, particionano el vector de coeficientes e intercepto y pendientes

$$y_i = \left[1 \vdots x_i' + d'\right]\beta + v_i = \beta_1 + \left(x_i' + d'\right)\beta_{-1} + v_i$$

Donde β_1 es el intercepto y β_{-1} es el vector de los $k-1$ coeficientes pendientes. En notación matricial, sean 0_n y 1_n vectores columna de $n \times 1$ variables y X_{-1} la matrix de observaciones de los regresores de la variable

$$\tilde{X} = X + \left[0 \vdots 1_n \cdot d'\right]$$

$$y = \left\{X + \left[0 \vdots 1_n \cdot d'\right]\right\}\beta + v = \left\{\left[1_n \vdots X_1\right] + \left[0_n \vdots 1_n \cdot d'\right]\right\}\begin{bmatrix}\beta_1 \\ \beta_{-1}\end{bmatrix} + v = 1_n\beta_1 + \left[X_{-1} + 1_n d'\right]\beta_{-1} + v$$
$$= \left[1_n + 1_n \cdot d'\beta_{-1}\right] + X_{-1}\beta_{-1} + v = 1_n\left[\beta_1 + d'\beta_{-1}\right] + X_{-1}\beta_{-1} = \left[1_n \vdots X_{-1}\right]\begin{bmatrix}\beta_1 + d'\beta_{-1} \\ \beta_1\end{bmatrix} + v$$

Sea $\beta_1^* = \beta_1 + d'\beta_{-1}$, el nuevo modelo resulta

$$y = X\beta^* + v \quad \beta^* = \left[\beta_1 + d'\beta_{-1}\right]$$

Este último cambio afecta sólo al intercepto del modelo. El estimador OLS para β_{-1}^* es el mismo que en el modelo original. Por lo hallado en el ejercicio 1A sabemos que el estimador OLS de una combinación lineal de los parámetros originales es la combinación lineal de los estimadores OLS de los parámetros. De este modo

$$\hat{\beta}^* = \begin{bmatrix}\hat{\beta}_1 + d'\hat{\beta}_{-1} \\ \hat{\beta}_{-1}\end{bmatrix}$$

Ejercicio 4

Sea $\hat{\beta}$ el estimador OLS de β, entonces

$$\hat{\beta} \in arg\min_{\tilde{\beta}}\left\{SSR\left(\tilde{\beta}\right)\right\}$$

$$\hat{\beta} \in arg\min_{\tilde{\beta}}\left\{\left(y - X\tilde{\beta}\right)\left(y - X\tilde{\beta}\right)'\right\}$$

$$y'y - n\bar{y}^2 \geq 0 \implies \hat{\beta} \in arg\min_{\tilde{\beta}}\left\{\frac{\left(y - X\tilde{\beta}\right)\left(y - X\tilde{\beta}\right)'}{y'y - n\bar{y}^2}\right\}$$

$$\hat{\beta} \in arg\max_{\tilde{\beta}}\left\{-\frac{\left(y - X\tilde{\beta}\right)\left(y - X\tilde{\beta}\right)'}{y'y - n\bar{y}^2}\right\}$$

$$\hat{\beta} \in arg\max_{\tilde{\beta}}\left\{1 - \frac{\left(y - X\tilde{\beta}\right)\left(y - X\tilde{\beta}\right)'}{y'y - n\bar{y}^2}\right\} \therefore \hat{\beta} \in arg\max_{\tilde{\beta}}\left\{R^2\right\}$$

Ejercicio 5

Considerando el modelo en forma matricial con $n = 33$ observaciones

$$y = X\beta + u$$

donde

$$Y_{n\times 1} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \; X_{n\times 3} = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \\ \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & x_{3n} \end{bmatrix} \; \beta_{3\times 1} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \; u_{n\times 1} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

Estimando por OLS llegamos a los siguientes resultados

$$\hat{\beta} = (X'X)^{-1} X'y$$
$$\Sigma_{\hat{\beta}} = \hat{\sigma}^2 (X'X)^{-1}$$
$$\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{n-3}$$

Sabemos que

$$(X'X) = \begin{bmatrix} \sum_{i=1}^{33} x_{1i}^2 & \sum_{i=1}^{33} x_{1i}x_{2i} & \sum_{i=1}^{33} x_{1i}x_{3i} \\ \sum_{i=1}^{33} x_{1i}x_{2i} & \sum_{i=1}^{33} x_{2i}^2 & \sum_{i=1}^{33} x_{2i}x_{3i} \\ \sum_{i=1}^{33} x_{1i}x_{3i} & \sum_{i=1}^{n} x_{2i}x_{3i} & \sum_{i=1}^{33} x_{3i}^2 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 10 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Luego

$$(X'X)^{-1} = \frac{1}{9} \begin{bmatrix} 10 & -1 & -10 \\ -1 & 1 & 1 \\ -10 & 1 & 19 \end{bmatrix}$$

También sabemos que

$$X'y = \begin{bmatrix} \sum_{i=1}^{n} x_{1i}y_i \\ \sum_{i=1}^{n} x_{2i}y_i \\ \sum_{i=1}^{n} x_{3i}y_i \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 4 \end{bmatrix}$$

Entonces podemos computar los estimadores OLS

$$\hat{\beta} = \frac{1}{9} \begin{bmatrix} 10 & -1 & -10 \\ -1 & 1 & 1 \\ -10 & 1 & 19 \end{bmatrix} \begin{bmatrix} 5 \\ 10 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}$$

Ahora para computar la varianza de los errores sabemos que

$$\hat{u} = y - X\hat{\beta} = y - X (X'X)^{-1} X'y = \underbrace{\left(I - X (X'X)^{-1} X'\right)}_{M} y$$

M es una matriz idempotente, entonces

$$\hat{u}'\hat{u} = y'MMy = y'My = y' \left(I - X (X'X)^{-1} X'\right) y = y'y - y'X (X'X)^{-1} Xy$$

Sabemos que $y'y = \sum_{i=1}^{33} y_i^2 = 35$. De pasos anteriores sabemos que

$$y'X = (X'y)' = \begin{bmatrix} \sum_{i=1}^{33} x_{1i}y_i & \sum_{i=1}^{33} x_{2i}y_i & \sum_{i=1}^{33} x_{3i}y_i \end{bmatrix} = \begin{bmatrix} 5 & 10 & 4 \end{bmatrix}$$

Entonces

$$\hat{u}'\hat{u} = 35 - \begin{bmatrix} 5 & 10 & 4 \end{bmatrix} \frac{1}{9} \begin{bmatrix} 10 & -1 & -10 \\ -1 & 1 & 1 \\ -10 & 1 & 19 \end{bmatrix} \begin{bmatrix} 5 \\ 10 \\ 4 \end{bmatrix} = 9 \therefore \hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{n-3} = \frac{9}{30}$$

Finalmente

$$\Sigma_{\hat{\beta}} = \hat{\sigma}^2 \left(X'X\right)^{-1} = \frac{9}{30}\frac{1}{9} \begin{bmatrix} 10 & -1 & -10 \\ -1 & 1 & 1 \\ -10 & 1 & 19 \end{bmatrix} = \frac{1}{30} \begin{bmatrix} 10 & -1 & -10 \\ -1 & 1 & 1 \\ -10 & 1 & 19 \end{bmatrix}$$