* MRLM:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i$$

Parte determinista — Parte aleatoria

- $k \equiv$ nº de variables independientes
- $k+1 \equiv$ nº de parámetros del modelo
- $n \equiv$ nº de muestras

$$n=1 \quad y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} + \cdots + \beta_k x_{k1} + \varepsilon_1$$
$$n=2 \quad y_2 = \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} + \cdots + \beta_k x_{k2} + \varepsilon_2$$
$$\vdots$$
$$n \quad y_n = \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} + \cdots + \beta_k x_{kn} + \varepsilon_n$$

$$Y = X \cdot \beta + \varepsilon$$

$$y = X \cdot \beta + \varepsilon$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{k1} \\ 1 & x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & & & & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$n \times 1$ $\quad$ $n \times (k+1)$ $\quad$ $(k+1) \times 1$ $\quad$ $n \times 1$

$n \times 1$

Objetivo MCO: minimizar $\sum \hat{\varepsilon}_i^2$

$$\frac{2 \sum \hat{\varepsilon}_i^2}{2 \beta_j} = 0 \Rightarrow \hat{\beta} = (X'X)^{-1} X' y = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Y el mínimo de esta función:

$$\min_{\hat{\beta}_0,\hat{\beta}_1,\ldots,\hat{\beta}_k} \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 = \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \ldots. - \hat{\beta}_k X_{ki}\right)^2$$

se consigue derivando respecto a cada parámetro e igualando a cero y operando se llega a las k+1 **ecuaciones normales**:

$$\sum_{i=1}^{n} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \ldots. - \hat{\beta}_k X_{ki}\right) = \sum_{i=1}^{n} \hat{\varepsilon}_i = 0$$

$$\sum_{i=1}^{n} X_{ji} \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i} - \ldots. - \hat{\beta}_k X_{ki}\right) = \sum_{i=1}^{n} X_{ij} \hat{\varepsilon}_i = 0$$

A partir de las ecuaciones normales se despejan los estimadores de los coeficientes y tenemos:

tienen β_0

$$\begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} & \cdots & \sum X_{ki} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i}X_{2i} & \cdots & \sum X_{1i}X_{ki} \\ \sum X_{2i} & \sum X_{2i}X_{1i} & \sum X_{2i}^2 & \cdots & \sum X_{2i}X_{ki} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum X_{ki} & \sum X_{ki}X_{1i} & \sum X_{ki}X_{2i} & \cdots & \sum X_{ki}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_n \end{bmatrix}$$

$$(X'X) \quad \hat{\beta} = X' \quad y$$

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

$$\boldsymbol{\hat{\beta}_{MCO}} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y} = \begin{pmatrix} n & \sum_{i=1}^{n} X_{1i} & \cdots & \sum_{i=1}^{n} X_{ki} \\ \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{1i}^2 & \cdots & \sum_{i=1}^{n} X_{1i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} X_{ki} & \sum_{i=1}^{n} x_{ki}x_{1i} & \cdots & \sum_{i=1}^{n} X_{ki}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} Y_i X_{1i} \\ \vdots \\ \sum_{i=1}^{n} Y_i X_{ki} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_k \end{pmatrix}$$

La matriz X'X tiene dimensiones (k+1)x(k+1)

$$X'X = \begin{pmatrix} n & \sum_{i=1}^{n} X_{1i} & \cdots & \sum_{i=1}^{n} X_{ki} \\ \sum_{i=1}^{n} X_{1i} & \sum_{i=1}^{n} X_{1i}^2 & \cdots & \sum_{i=1}^{n} X_{1i}X_{ki} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} X_{ki} & \sum_{i=1}^{n} X_{ki}X_{1i} & \cdots & \sum_{i=1}^{n} X_{ki}^2 \end{pmatrix}$$

- Donde n es el tamaño muestral.
- Las expresiones: $\sum_{i=1}^{n} X_{1i}^2 \quad \cdots \quad \sum_{i=1}^{n} X_{ki}^2$ de la diagonal principal son las sumas de los cuadrados de las variable explicativas.
- Las expresiones: $\sum_{i=1}^{n} X_{1i}X_{ki} \quad \ldots \quad \sum_{i=1}^{n} X_{ki}X_{1i}$ son las sumas de los productos cruzados de las variables explicativas.

En general las características que vimos en el análisis de regresión lineal simple se pueden extender al múltiple:

- $\sum_{i=1}^{n} \hat{y}_i \hat{\varepsilon}_i = 0$ La estimación de la variable regresada y los residuos no están correlados.

Lo que implica que la covarianza es nula $\text{cov}(\hat{Y}_i, \hat{\varepsilon}_i) = 0$

- La variables independientes X y los residuos también están incorrelacionados.

$\text{cov}(X_{1i}, \hat{\varepsilon}_i) = 0$

Con los supuestos del modelo clásico de regresión lineal estos estimadores de MCO son lineales e insesgados, y dentro de éstos son los de mínima varianza (MELI).

ELIO / BLUE } Insesgados Lineal Eficiente

Calculamos el coeficiente de determinación R^2 , que se interpreta igual que en el modelo de regresión lineal simple pero además incluiremos la formulación matricial.

$$R^2 = \frac{var(\hat{Y})}{var(Y)} = \frac{var(Y) - var(\hat{\varepsilon})}{var(Y)} = \frac{SCT - SCR}{SCT} = \frac{\mathbf{y'y} - (\mathbf{y'y} - \hat{\boldsymbol{\beta}}'\mathbf{X'y})}{\mathbf{y'y}} = \frac{\hat{\boldsymbol{\beta}}'\mathbf{X'y}}{\mathbf{y'y}}.$$

$$R^2 = \frac{SCE}{SCT} = 1 - \frac{SCR}{SCT} = \frac{\hat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - n\bar{Y}^2}{\boldsymbol{y}'\boldsymbol{y} - n\bar{Y}^2}$$

↳ $y_i = \beta_0 + \beta_1 x_1 \dots$

Existe una relación entre el R^2 y la varianza de los estimadores $\hat{\beta}_j$:

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{\sum x_j^2}\left(\frac{1}{1 - R_j^2}\right) = \frac{\sigma^2}{SCT_j(1 - R_j^2)}$$

$R_j^2 \rightarrow x_j = \alpha_0 + \alpha_2 x_2 + \alpha x_3 + \dots$

Donde R_j^2 es el coeficiente de determinación de la regresión de X_j sobre las (k - 1) variables independientes restantes y SCT_j es la suma cuadrática total de la variable independiente j.

El coeficiente de determinación indica que proporción de variabilidad total queda explicada por la regresión. Si el modelo tiene término independiente, entonces R^2 toma valores entre 0 y 1. Este coeficiente permite, además, seleccionar entre modelos clásicos que tengan el mismo número de regresores, ya que la capacidad explicativa de un modelo es mayor cuanto más elevado sea el valor que tome este coeficiente.

2.1.2 Coeficiente de determinación R^2 corregido

Una característica del modelo de regresión múltiple es que a medida que aumentamos el número de regresores X_j el coeficiente de determinación R^2 necesariamente aumenta salvo que el coeficiente estimado sea *exactamente* nulo. Entonces el R^2 nunca disminuye al incorporar nuevos regresores. Así que un incremento del R^2 no significa necesariamente que añadir una nueva variable realmente haya mejorado la calidad del ajuste de nuestro modelo. En realidad incluso si la nueva variable incluida en el modelo mejora nuestro ajuste, sabemos que necesariamente el R^2 de la nueva regresión estará artificialmente «inflado» por el mero hecho de incorporar un nuevo regresor. Por este motivo se utiliza el **R^2 corregido,** que ajusta por el número de coeficientes estimados y cuya definición es:

$$\bar{R}^2 = 1 - \frac{\frac{SCR}{n-k-1}}{\frac{SCT}{n-1}} = 1 - \frac{\hat{\sigma}^2}{S_Y^2}$$

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k-1}$$

$< R^2$

Donde:

- $\hat{\sigma}^2$ es el estimador ~~ingresado~~ insegado de la verdadera varianza de los residuos.
- S_Y^2 es la varianza muestral de Y.

A tener en cuenta:

- $\bar{R}^2 < R^2$ El coeficiente de determinación corregido siempre será menor que el coeficiente de determinación.

$\bar{R}^2$ → Criterio de selección de modelos restringidos

- Añadir un nuevo regresor tiene dos efectos opuestos; el resultado final sobre $\bar{R}^2$, el resultado final dependerá de cuál de los dos efectos es mayor:
  - Al disminuir la SCR → se incrementa la $\bar{R}^2$
  - El cociente (n-1)/(n-k-1) aumenta → disminuir $\bar{R}^2$
- $\bar{R}^2$ **puede ser negativo** si los regresores en conjunto reducen la SCR en una cantidad tan pequeña que dicha reducción no logre superar el efecto del factor (n-1)/(n-k-1).

Ejemplo 7: Consumo de las familias catalanas dedicadas a la hostelería y el turismo

Con datos de Cataluña y del sector de la hostelería , nos proponemos analizar el consumo de las familias catalanas.

Modelo poblacional: $\ln consumo = \beta_0 + \beta_1(\ln ingresos) + \varepsilon$

Y su estimación (FRM): $\widehat{\ln consumo} = 3,89 + 0,615(\ln ingresos)$

$n = 95, R^2 = 0,3292.$

$\hat{\beta}_0 = 3'89$ ≠ $x=0$

$\hat{\beta}_1 = 0'615 = \varepsilon_{ingresos}$

Comparar → ~~R^2~~

↳ $\bar{R}^2$

Este modelo es una regresión lineal simple. Es un modelo log-log, por tanto los coeficiente se interpretan directamente como las elasticidades. Un incremento del 1 % en los ingresos provoca que el consumo se incremente un 0,615 %.

Ampliando el modelo, ya que es muy probable que el consumo también tenga relación con el número de miembros en la familia (tamaño), ya que se espera que a medida que el tamaño aumenta, el consumo también aumenta, la nueva estimación:

$$\widehat{\ln consumo_i} = 5,15 + 0,443 \cdot (\ln ingreso_i) + 0,1420 \cdot tamaño_i,$$

$$n = 95, R^2 = 0,4149.$$

$\hat{\beta}_0 = 5'15$

$\hat{\beta}_1 = 0'443$

$\hat{\beta}_2 = 0'142$

ingresos → log-log

tamaño → log-lin

Al introducir la nueva variable, observamos que los coeficientes han cambiado, como hemos explicado anteriormente. Así que el término independiente y la elasticidad la variable ingresos cambia.

La nueva estimación nos aporta información sobre cómo influye el incremento, o decremento, del número de miembros en el consumo de las familias, dado un nivel determinado de ingresos.

Respecto a la variable tamaño el modelo es un modelo log-lineal así que la interpretación de los coeficientes estimados nos indica que, manteniendo constante el nivel de ingresos, es decir controlando el efecto del ingreso en el consumo, entonces el incremento de un miembro en la familia se prevé un incremento medio del 14,20 % del consumo familiar (100x0,1420 = 14,20).

Por otro lado, el incremento de los ingresos en un 1 %, dado un tamaño familiar determinado, solo produce un incremento del 0,443 % del consumo, que contrasta con el 0,615 % de la expresión del modelo anterior de regresión lineal simple. Por tanto la introducción de nuevas variables (*tamaño*) afecta al resto de coeficientes de las variables del modelo (*Iningreso).*

Comparando los dos modelos podemos ver que el coeficiente de determinación del segundo modelo R^2 = 0,4149 es mayor que el del primero R^2 = 0,3239. Esto no podía ser de otro modo ya que como hemos comentado anteriormente el coeficiente de determinación aumenta al añadir regresores al modelo. Así que en este caso, no tiene sentido comparar estos coeficientes para decidir si la nueva variable añadida (tamaño) ha mejorado el ajuste del modelo, para poder determinar si el nuevo ajuste es mejor, necesitaríamos conocer el coeficiente de determinación corregido.