Modelos Lineales

Regresión Lineal Múltiple

Fernando Massa; Bruno Bellagamba

martes 2 de abril 2024

- Veremos qué quiere decir que los coeficientes del estimador *MCO* sean *parciales*.
- Estudiaremos el caso de omitir variables relevantes en el modelo.
- Analizaremos las repercusiones de que las variables explicativas sean ortogonales.
- Veremos ejemplos prácticos.

Anteriormente mencionamos que la interpretación de los coeficientes β_j en el modelo de *RLM* nos indican:

> *El aumento promedio en Y al aumentar una unidad la variable X_j dejando las demás variables constantes*

y observamos cómo esta interpretación se deducía al comparar dos esperanzas condicionales donde la única variable que cambiaba su valor era X_j.

Anteriormente mencionamos que la interpretación de los coeficientes β_j en el modelo de *RLM* nos indican:

El aumento promedio en Y al aumentar una unidad la variable X_j dejando las demás variables constantes

y observamos cómo esta interpretación se deducía al comparar dos esperanzas condicionales donde la única variable que cambiaba su valor era X_j.

En un ámbito más práctico podemos decir que, a través del modelo del *RLM* se suele decir que el coeficiente de una cierta variable X_j provee el impacto o efecto sobre Y habiendo **ajustado** por las demás variables.

Anteriormente mencionamos que la interpretación de los coeficientes β_j en el modelo de *RLM* nos indican:

El aumento promedio en Y al aumentar una unidad la variable X_j dejando las demás variables constantes

y observamos cómo esta interpretación se deducía al comparar dos esperanzas condicionales donde la única variable que cambiaba su valor era X_j.

En un ámbito más práctico podemos decir que, a través del modelo del *RLM* se suele decir que el coeficiente de una cierta variable X_j provee el impacto o efecto sobre Y habiendo **ajustado** por las demás variables.

Esto se debe a que en la gran mayoría de los casos, las variables explicativas están correlacionadas entre sí y *compiten* por explicar a la Y.

Trabajando con la matriz de proyección H observamos que:

Trabajando con la matriz de proyección H observamos que:

- $\hat{\varepsilon} = (I_n - H_X)Y$ permite obtener los residuos de una regresión de Y sobre las variables en la matriz X. Esto se puede interpretar como que $\hat{\varepsilon}$ contiene la información de la variable Y que NO pudo ser captada por las variables en X.
- $Y^{'}(I_n - H_X)Y$ nos permite trabajar con la variabilidad de Y luego de haber *separado* la porción de variabilidad adjudicable a las columnas de X.
- De manera similar, $Y^{'}(H_X - H_{\mathbb{1}})Y$ permite obtener la porción de variabilidad de Y adjudicable a X luego de remover el efecto del vector de unos.

Descomponiendo X

Para afrontar los próximos puntos, particionaremos a la matriz X de la siguiente manera:

$$X = \left(x_j \vdots X_{-j} \right)$$

Donde x_j representa a la j-ésima columna y X_{-j} es una matriz de n filas y k columnas, o sea, todas menos la j-ésima. A partir de esta descomposición queremos plantear el estimador de MCO:

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

y sustituir a la matriz X (y por ende al vector $\hat{\beta}$ acorde a la representación anterior. Para esto tengamos en cuenta que:

$$X^{'} = \begin{pmatrix} x_j^{'} \\ \dots \\ X_{-j}^{'} \end{pmatrix}$$

Continuando con la idea anterior, $X^{'}X$ es:

$$\begin{pmatrix} x_j^{'}x_j & x_j^{'}X_{-j} \\ X_{-j}^{'}x_j & X_{-j}^{'}X_{-j} \end{pmatrix}$$

Continuando con la idea anterior, $X^{'}X$ es:

$$\begin{pmatrix} x_j^{'}x_j & x_j^{'}X_{-j} \\ X_{-j}^{'}x_j & X_{-j}^{'}X_{-j} \end{pmatrix}$$

Al plantear las ecuaciones normales $X^{'}X\beta = X^{'}Y$ se obtiene:

$$\begin{pmatrix} x_j^{'}x_j & x_j^{'}X_{-j} \\ X_{-j}^{'}x_j & X_{-j}^{'}X_{-j} \end{pmatrix} \begin{pmatrix} \hat{\beta}_j \\ \hat{\beta}_{-j} \end{pmatrix} = \begin{pmatrix} x_j^{'}Y \\ X_{-j}^{'}Y \end{pmatrix}$$

Descomponiendo X

Las dos ecuaciones (matriciales) que surgen son:

$$\begin{aligned} x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \hat{\beta}_{-j} &= x_j^{'} Y \\ X_{-j}^{'} x_j \hat{\beta}_j + X_{-j}^{'} X_{-j} \hat{\beta}_{-j} &= X_{-j}^{'} Y \end{aligned}$$

DESCOMPONIENDO X

Las dos ecuaciones (matriciales) que surgen son:

$$\begin{aligned} x_j^{'}x_j\hat{\beta}_j + x_j^{'}X_{-j}\hat{\beta}_{-j} &= x_j^{'}Y \\ X_{-j}^{'}x_j\hat{\beta}_j + X_{-j}^{'}X_{-j}\hat{\beta}_{-j} &= X_{-j}^{'}Y \end{aligned}$$

Luego, "despejamos" $\hat{\beta}_{-j}$ de la segunda, esto es:

$$\begin{aligned} X_{-j}^{'}X_{-j}\hat{\beta}_{-j} &= X_{-j}^{'}Y - X_{-j}^{'}x_j\hat{\beta}_j \\ X_{-j}^{'}X_{-j}\hat{\beta}_{-j} &= X_{-j}^{'}(Y - x_j\hat{\beta}_j) \\ \hat{\beta}_{-j} &= \left(X_{-j}^{'}X_{-j}\right)^{-1}X_{-j}^{'}(Y - x_j\hat{\beta}_j) \end{aligned}$$

DESCOMPONIENDO X

Las dos ecuaciones (matriciales) que surgen son:

$$\begin{aligned} x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \hat{\beta}_{-j} &= x_j^{'} Y \\ X_{-j}^{'} x_j \hat{\beta}_j + X_{-j}^{'} X_{-j} \hat{\beta}_{-j} &= X_{-j}^{'} Y \end{aligned}$$

Luego, "despejamos" $\hat{\beta}_{-j}$ de la segunda, esto es:

$$\begin{aligned} X_{-j}^{'} X_{-j} \hat{\beta}_{-j} &= X_{-j}^{'} Y - X_{-j}^{'} x_j \hat{\beta}_j \\ X_{-j}^{'} X_{-j} \hat{\beta}_{-j} &= X_{-j}^{'} (Y - x_j \hat{\beta}_j) \\ \hat{\beta}_{-j} &= \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}^{'} (Y - x_j \hat{\beta}_j) \end{aligned}$$

y lo introducimos en la primera

$$\begin{aligned} x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \hat{\beta}_{-j} &= x_j^{'} Y \\ x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}^{'} (Y - x_j \hat{\beta}_j) &= x_j^{'} Y \end{aligned}$$

Descomponiendo X

Retomando la última ecuación de la diapositiva anterior:

$$x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}^{'} (Y - x_j \hat{\beta}_j) = x_j^{'} Y$$

Podemos notar la aparición del término $X_{-j} \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}$ el que podemos llamar H_{-j}, así:

Descomponiendo X

Retomando la última ecuación de la diapositiva anterior:

$$x_j^{'}x_j\hat{\beta}_j + x_j^{'}X_{-j}\left(X_{-j}^{'}X_{-j}\right)^{-1}X_{-j}^{'}(Y - x_j\hat{\beta}_j) = x_j^{'}Y$$

Podemos notar la aparición del término $X_{-j}\left(X_{-j}^{'}X_{-j}\right)^{-1}X_{-j}$ el que podemos llamar H_{-j}, así:

$$x_j^{'}x_j\hat{\beta}_j + x_j^{'}H_{-j}(Y - x_j\hat{\beta}_j) = x_j^{'}Y$$

Si operamos con los términos que cuentan con $x_j^{'}$ y x_j

Retomando la última ecuación de la diapositiva anterior:

$$x_j^{'} x_j \hat{\beta}_j + x_j^{'} X_{-j} \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}^{'} (Y - x_j \hat{\beta}_j) = x_j^{'} Y$$

Podemos notar la aparición del término $X_{-j} \left(X_{-j}^{'} X_{-j} \right)^{-1} X_{-j}$ el que podemos llamar H_{-j}, así:

$$x_j^{'} x_j \hat{\beta}_j + x_j^{'} H_{-j} (Y - x_j \hat{\beta}_j) = x_j^{'} Y$$

Si operamos con los términos que cuentan con $x_j^{'}$ y x_j

$$x_j^{'} x_j \hat{\beta}_j - x_j^{'} H_{-j} x_j \hat{\beta}_j = x_j^{'} Y - x_j^{'} H_{-j} Y$$

Retomando la última ecuación de la diapositiva anterior:

$$x_j^{'}x_j\hat{\beta}_j + x_j^{'}X_{-j}\left(X_{-j}^{'}X_{-j}\right)^{-1}X_{-j}^{'}(Y - x_j\hat{\beta}_j) = x_j^{'}Y$$

Podemos notar la aparición del término $X_{-j}\left(X_{-j}^{'}X_{-j}\right)^{-1}X_{-j}$ el que podemos llamar H_{-j}, así:

$$x_j^{'}x_j\hat{\beta}_j + x_j^{'}H_{-j}(Y - x_j\hat{\beta}_j) = x_j^{'}Y$$

Si operamos con los términos que cuentan con $x_j^{'}$ y x_j

$$x_j^{'}x_j\hat{\beta}_j - x_j^{'}H_{-j}x_j\hat{\beta}_j = x_j^{'}Y - x_j^{'}H_{-j}Y$$

Obteniendo:

$$x_j^{'}(I_n - H_{-j})x_j\hat{\beta}_j = x_j^{'}(I_n - H_{-j})Y$$

En última instancia: $$\hat{\beta}_j = \frac{x_j^{'}(I_n - H_{-j})Y}{x_j^{'}(I_n - H_{-j})x_j}$$

En última instancia:
$$\hat{\beta}_j = \frac{x_j^{'}(I_n - H_{-j})Y}{x_j^{'}(I_n - H_{-j})x_j}$$

Si adicionalmente recordamos que $I_n - H$ es idempotente, podemos plantear:

$$\hat{\beta}_j = \frac{x_j^{'}(I_n - H_{-j})Y}{x_j^{'}(I_n - H_{-j})(I_n - H_{-j})x_j}$$

En última instancia:

$$\hat{\beta}_j = \frac{x_j^{'}(I_n - H_{-j})Y}{x_j^{'}(I_n - H_{-j})x_j}$$

Si adicionalmente recordamos que $I_n - H$ es idempotente, podemos plantear:

$$\hat{\beta}_j = \frac{x_j^{'}(I_n - H_{-j})Y}{x_j^{'}(I_n - H_{-j})(I_n - H_{-j})x_j}$$

De esta manera podemos definir $z_j = x_j(I_n - H_{-j})$ como la *versión* de la variable x_j libre del efecto de las demás variables explcativas. Así:

$$\hat{\beta}_j = \frac{z_{j'}Y}{z_j^{'}z_j}$$

No es más que la estimación correspondiente a un modelo de *RLS* de la variable de respuesta Y sobre la variable explicativa z_j.

Ejemplo

Veamos un ejemplo en R donde se busca relacionar el aumento de peso de un grupo de personas con respecto a su nivel de actividad física.

Mediante un razonamiento similar al anterior podemos indagar sobre el resultado de omitir variables relevantes en el modelo de *RLM*. Entendemos por *variables relevantes* a aquellas variables explicativas que **SI** tienen la capacidad de explicar a la Y.

Mediante un razonamiento similar al anterior podemos indagar sobre el resultado de omitir variables relevantes en el modelo de *RLM*. Entendemos por *variables relevantes* a aquellas variables explicativas que **SI** tienen la capacidad de explicar a la Y.

En esta ocasión partiremos de una situación un poco más general. Particionaremos a la matriz X separándola en dos conjuntos de columnas $X = \left(X_1 \vdots X_2 \right)$. De esta manera

$$
\begin{aligned}
Y &= X\beta + \varepsilon \\
Y &= X_1\beta_1 + X_2\beta_2 + \varepsilon
\end{aligned}
$$

Donde β_1 y β_2 son subvectores del vector de parámetros β de dimensiones acordes a las de X_1 y X_2.

Omisión de variables

Supongamos ahora, que solo contamos con los datos de las variables de X_1 porque:

Omisión de variables

Supongamos ahora, que solo contamos con los datos de las variables de X_1 porque:

- La teoría sobre la que trabajamos para explicar a Y no incluye a las variables en X_2.
- Se sabía que las variables de X_2 eran relevantes pero por distintos motivos no pudieron ser recolectadas.
- Se recolectaron pero contenían grandes cantidades de datos faltantes, por lo que fueron excluídas del análisis.

Supongamos ahora, que solo contamos con los datos de las variables de X_1 porque:

- La teoría sobre la que trabajamos para explicar a Y no incluye a las variables en X_2.
- Se sabía que las variables de X_2 eran relevantes pero por distintos motivos no pudieron ser recolectadas.
- Se recolectaron pero contenían grandes cantidades de datos faltantes, por lo que fueron excluídas del análisis.

Así que en la práctica se trabaja sobre el modelo:

$$Y = X_1 \beta_1^* + \varepsilon^*$$

Supongamos ahora, que solo contamos con los datos de las variables de X_1 porque:

- La teoría sobre la que trabajamos para explicar a Y no incluye a las variables en X_2.
- Se sabía que las variables de X_2 eran relevantes pero por distintos motivos no pudieron ser recolectadas.
- Se recolectaron pero contenían grandes cantidades de datos faltantes, por lo que fueron excluídas del análisis.

Así que en la práctica se trabaja sobre el modelo:

$$Y = X_1 \beta_1^* + \varepsilon^*$$

Nos interesa saber qué repercusiones tiene el hecho de omitir las variables en X_2 sobre la estimación de β_1.

El estimador de *MCO* trabajando con este modelo *reducido* (pero *incorrecto*) es:

$$\hat{\beta}_1^* = \left(X_1^{'} X_1\right)^{-1} X_1^{'} Y$$

El estimador de *MCO* trabajando con este modelo *reducido* (pero *incorrecto*) es:

$$\hat{\beta}_1^* = \left(X_1^{'} X_1\right)^{-1} X_1^{'} Y$$

Veamos su valor esperado

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1^*) &= \mathbb{E}\left[\left(X_1^{'} X_1\right)^{-1} X_1^{'} Y\right] \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} \mathbb{E}(Y) \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} \left(X_1 \beta_1 + X_2 \beta_2\right) \\
&= \beta_1 + \left(X_1^{'} X_1\right)^{-1} X_1^{'} X_2 \beta_2
\end{aligned}$$

El estimador de *MCO* trabajando con este modelo *reducido* (pero *incorrecto*) es:

$$\hat{\beta}_1^* = \left(X_1^{'} X_1\right)^{-1} X_1^{'} Y$$

Veamos su valor esperado

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1^*) &= \mathbb{E}\left[\left(X_1^{'} X_1\right)^{-1} X_1^{'} Y\right] \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} \mathbb{E}(Y) \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} (X_1 \beta_1 + X_2 \beta_2) \\
&= \beta_1 + \left(X_1^{'} X_1\right)^{-1} X_1^{'} X_2 \beta_2
\end{aligned}$$

Se observa que en este caso, el principal problema de omitir información relevante es que se obtienen **estimaciones sesgadas** de los parámetros en β_1...

El estimador de *MCO* trabajando con este modelo *reducido* (pero *incorrecto*) es:

$$\hat{\beta}_1^* = \left(X_1^{'} X_1\right)^{-1} X_1^{'} Y$$

Veamos su valor esperado

$$\begin{aligned}
\mathbb{E}(\hat{\beta}_1^*) &= \mathbb{E}\left[\left(X_1^{'} X_1\right)^{-1} X_1^{'} Y\right] \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} \mathbb{E}(Y) \\
&= \left(X_1^{'} X_1\right)^{-1} X_1^{'} (X_1 \beta_1 + X_2 \beta_2) \\
&= \beta_1 + \left(X_1^{'} X_1\right)^{-1} X_1^{'} X_2 \beta_2
\end{aligned}$$

Se observa que en este caso, el principal problema de omitir información relevante es que se obtienen **estimaciones sesgadas** de los parámetros en β_1... salvo que $X_1^{'} X_2 = 0$

El problema no termina ahí, el estimador de σ^2 también se ve afectado. Bajo el modelo reducido, el estimador es:

$$\hat{\sigma}^2 = \frac{\left(Y - X_1\hat{\beta}_1^*\right)^{'}\left(Y - X_1\hat{\beta}_1^*\right)}{n-k_1}$$

Cuyo valor esperado es:

$$\begin{aligned}
\mathbb{E}(\hat{\sigma}^2) &= \mathbb{E}\left[\frac{\left(Y-X_1\hat{\beta}_1^*\right)\left(Y-X_1\hat{\beta}_1^*\right)^{'}}{n-k_1}\right] \\
&= \frac{\mathbb{E}(Y^{'}(I_n-H_{X_1})Y)}{n-k_1} \\
&= \frac{tr\left[(I_{n\times n}-H_{X_1})\sigma^2 I_n\right]+\beta^{'}X^{'}(I_n-H_{X_1})X\beta}{n-k_1} \\
&= \frac{(n-k_1)\sigma^2}{n-k_1} + \frac{\beta^{'}X^{'}(I_n-H_{X_1})X\beta}{n-k_1} \\
&= \sigma^2 + \frac{\beta_2^{'}X_2^{'}(I_n-H_{X_1})X_2\beta_2}{n-k_1}
\end{aligned}$$

En la medida que esta última matriz es semidefinida positiva (porque $I_n - H_{X_1}$ lo es), este estimador tiene un sesgo positivo.

Veamos otro ejemplo en R donde a partir de simulaciones analizamos el sesgo de las estimaciones de un parámetro al omitir una variable explicativa relevante.

En los apartados anteriores vimos cómo la estimación de un elemento o un subconjunto de elementos de β se ve afectada tanto por las variables que los acompañan como por las demás.

En los apartados anteriores vimos cómo la estimación de un elemento o un subconjunto de elementos de β se ve afectada tanto por las variables que los acompañan como por las demás.

Incluso vimos que los estimadores podrían resultar sesgados si no se incluyen variables que deberían ser incluidas en el modeo de RLM y que el sesgo de la estimación depende de la relación entre las variables presentes y las omitidas.

En los apartados anteriores vimos cómo la estimación de un elemento o un subconjunto de elementos de β se ve afectada tanto por las variables que los acompañan como por las demás.

Incluso vimos que los estimadores podrían resultar sesgados si no se incluyen variables que deberían ser incluidas en el modeo de *RLM* y que el sesgo de la estimación depende de la relación entre las variables presentes y las omitidas.

En esta última sección, veremos que, en el hipotético caso de que $X_1^{'}X_2 = 0$, no sólo se da que el estimador de los coeficientes del modelo reducido son insesgados, sino que adicionalmente son **idénticos** a los del modelo completo.

Ortogonalidad

Esta vez veremos lo que le sucede al estimador de *MCO* cuando particionamos a la matriz $X = \left(X_1 \vdots X_2 \right)$. Donde X_1 y X_2 tienen k_1 y k_2 columnas respectivamente.

Esta vez veremos lo que le sucede al estimador de *MCO* cuando particionamos a la matriz $X = \left(X_1 \vdots X_2\right)$. Donde X_1 y X_2 tienen k_1 y k_2 columnas respectivamente.

$$\begin{aligned}\hat{\beta}_{MCO} &= \left(X^{'}X\right)^{-1} X^{'}Y \\ &= \begin{pmatrix} X_1^{'}X_1 & X_1^{'}X_2 \\ X_2^{'}X_1 & X_2^{'}X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1^{'} \\ X_2^{'} \end{pmatrix} Y\end{aligned}$$

Ortogonalidad

Esta vez veremos lo que le sucede al estimador de *MCO* cuando particionamos a la matriz $X = \left(X_1 \vdots X_2 \right)$. Donde X_1 y X_2 tienen k_1 y k_2 columnas respectivamente.

$$\begin{aligned}\hat{\beta}_{MCO} &= \left(X'X\right)^{-1} X'Y \\ &= \begin{pmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} Y\end{aligned}$$

Al aplicar la condición de ortogonalidad $X_1'X_2 = 0$

$$\hat{\beta}_{MCO} = \begin{pmatrix} \left(X_1'X_1\right)^{-1} & 0 \\ 0 & \left(X_2'X_2\right)^{-1} \end{pmatrix} \begin{pmatrix} X_1' \\ X_2' \end{pmatrix} Y$$

Esta vez veremos lo que le sucede al estimador de *MCO* cuando particionamos a la matriz $X = \left(X_1 \vdots X_2\right)$. Donde X_1 y X_2 tienen k_1 y k_2 columnas respectivamente.

$$\begin{aligned}
\hat{\beta}_{MCO} &= \left(X^{'}X\right)^{-1} X^{'}Y \\
&= \begin{pmatrix} X_1^{'}X_1 & X_1^{'}X_2 \\ X_2^{'}X_1 & X_2^{'}X_2 \end{pmatrix}^{-1} \begin{pmatrix} X_1^{'} \\ X_2^{'} \end{pmatrix} Y
\end{aligned}$$

Al aplicar la condición de ortogonalidad $X_1^{'}X_2 = 0$

$$\hat{\beta}_{MCO} = \begin{pmatrix} \left(X_1^{'}X_1\right)^{-1} & 0 \\ 0 & \left(X_2^{'}X_2\right)^{-1} \end{pmatrix} \begin{pmatrix} X_1^{'} \\ X_2^{'} \end{pmatrix} Y$$

De donde surgen las ecuaciones:

$$\begin{aligned}
\hat{\beta}_{MCO,1} &= \left(X_1^{'}X_1\right)^{-1} X_1^{'}Y \\
\hat{\beta}_{MCO,2} &= \left(X_2^{'}X_2\right)^{-1} X_2^{'}Y
\end{aligned}$$

Llevando el razonamiento anterior al extremo, si todas las columnas de X son ortogonales, es decir:

$$x_r^{'} x_s = 0 \quad \forall r \neq s$$

Llevando el razonamiento anterior al extremo, si todas las columnas de X son ortogonales, es decir:

$$x_r^{'} x_s = 0 \quad \forall r \neq s$$

Es fácil generalizar el razonamiento anterior para obtener que el estimador del vector β del modelo de RLM coincide con $k+1$ estimadores independientes, cada uno de ellos de la forma:

$$\hat{\beta}_j = \frac{X_j^{'} Y}{X_j^{'} X_j} = \frac{\sum_i x_{ij} y_i}{\sum_i x_{ij}^2}$$

Llevando el razonamiento anterior al extremo, si todas las columnas de X son ortogonales, es decir:

$$x_r^{'} x_s = 0 \quad \forall r \neq s$$

Es fácil generalizar el razonamiento anterior para obtener que el estimador del vector β del modelo de RLM coincide con $k+1$ estimadores independientes, cada uno de ellos de la forma:

$$\hat{\beta}_j = \frac{X_j^{'} Y}{X_j^{'} X_j} = \frac{\sum_i x_{ij} y_i}{\sum_i x_{ij}^2}$$

Estos resultados nos permiten determinar que, el único caso donde los coeficientes de regresión **NO** se ven afectados por todas las variables es cuando estas conforman una base ortogonal.

En este último ejemplo en R, veremos la situación donde se quieren comparar los promedios de 3 grupos.

En la próxima clase

La próxima clase tendremos taller:

- Veremos algún ejercicio del práctico 2.
- Veremos un ejemplo completo donde aplicaremos la teoría de MCO.
- Veremos el efecto de omitir variables sobre las predicciones del modelo de RLM.

Bibliografía

- Carmona, Francesc (2003). *Modelos Lineales (notas de curso).* Departament d'Estadística.
- Farraway, Julian (2014). *Linear Models with R, second edition.* Chapman Hall/CRC.
- Rencher, Alvin y Bruce Schaalje (2008). *Linear Models in Statistics, second edition.* John Wiley Sons, Inc.

¿Preguntas?

Muchas Gracias