Transparencias para Métodos Cuantitativos I

Marcos Bujosa

29 de mayo de 2023

Transparencias para Métodos Cuantitativos I

Marcos Bujosa

29 de mayo de 2023

¿Por qué modelizar?

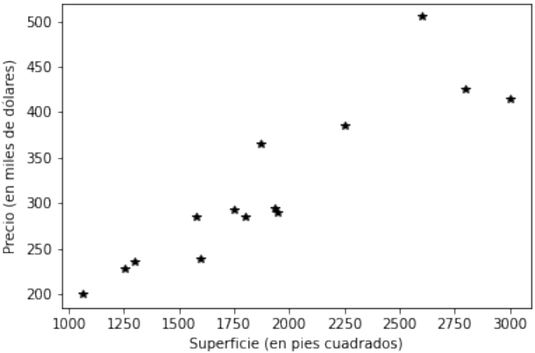
Modelado consiste en intentar ajustar un modelo matemático (estadístico) a un conjunto de datos (“la muestra”).

Un modelo es útil cuando (pese a ser *simple*) *capta las características* de los datos que consideramos más interesantes.

Ejemplos de objetivos por los que construir modelos

- Estimación: sensibilidad de un valor financiero a movimientos de un índice de referencia (evaluación de exposición al riesgo y cobertura con derivados sobre el índice).
- Previsión: probabilidad de impago de préstamos (función de las características de la operación y del solicitante).
- Simulación: rendimiento de una cartera de valores en diferentes escenarios.
- Control: (*bancos centrales*) intervención de tipos para controlar la inflación.

¿Hay relación entre tamaño y precio de una vivienda?



$$Precio_n = a + b(Superficie_n) + OtrasCosas_n$$

Función de consumo

Supongamos que consumo (con) y renta disponible (rd) de las familias siguen la relación:

$$con \;=\; \beta_1 \;+\; \beta_2\, rd \;+\; otrascosas$$

donde $otrascosas$ son otros aspectos distintos de la renta (activos financieros, estado de ánimo, edad, lugar de residencia, etc.).

Si disponemos datos de *consumo* y *renta disponible* de N familias como vectores de $\mathbb{R}^N$, *podemos construir una aproximación* $(\widetilde{\boldsymbol{con}})$ *del consumo* mediante una combinación lineal de la renta disponible $(\boldsymbol{rd})$ y de un término cte. $(\boldsymbol{1})$ ignorando las $otrascosas$:

$$\widetilde{\boldsymbol{con}} \;=\; \widetilde{\beta_1}\boldsymbol{1} \;+\; \widetilde{\beta_2}\boldsymbol{rd} \;=\; \begin{bmatrix}\boldsymbol{1}; & \boldsymbol{rd};\end{bmatrix}\begin{pmatrix}\widetilde{\beta_1}\\ \widetilde{\beta_2}\end{pmatrix}.$$

Nomenclatura y notación

- *regresando*: vector de datos de *consumo* ($\boldsymbol{con}$)
- *regresores*: vector de unos ($\boldsymbol{1}$) y de rentas disp. ($\boldsymbol{rd}$):

$$\mathbf{X} = \begin{bmatrix}\boldsymbol{1}; & \boldsymbol{rd};\end{bmatrix}, \quad \text{donde} \quad \mathbf{X}_{|1} = \boldsymbol{1} \quad \text{y} \quad \mathbf{X}_{|2} = \boldsymbol{rd}.$$

- *vector de parámetros*: $\quad \widetilde{\boldsymbol{\beta}} = \begin{pmatrix}\widetilde{\beta_1}\\ \widetilde{\beta_2}\end{pmatrix}$

Otro ejemplo: Un modelo para los salarios

$$salario = \beta_1 + \beta_2\, educ + \beta_3\, exper + \beta_4\, IQ + otrascosas$$

(disponiendo de datos de N trabajadores) el **ajuste** es

$$\widetilde{\boldsymbol{salario}} = \widetilde{\beta_1}\boldsymbol{1} + \widetilde{\beta_2}\boldsymbol{educ} + \widetilde{\beta_3}\boldsymbol{exper} + \widetilde{\beta_4}\boldsymbol{iq}$$

Ajuste MCO. Función lineal en los parámetros

- La aproximación o ajuste $\widetilde{\boldsymbol{y}}$ es una combinación lineal de los *regresores* $\mathbf{X}_{|j}$:

$$\begin{pmatrix} \widetilde{y_1} \\ \vdots \\ \widetilde{y_N} \end{pmatrix} = \widetilde{\beta_1} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} + \widetilde{\beta_2} \begin{pmatrix} x_{12} \\ \vdots \\ x_{N2} \end{pmatrix} + \cdots + \widetilde{\beta_k} \begin{pmatrix} x_{1k} \\ \vdots \\ x_{Nk} \end{pmatrix}$$

ó

$$\begin{aligned} \widetilde{\boldsymbol{y}} &= \widetilde{\beta_1}\mathbf{1} + \widetilde{\beta_2}\mathbf{X}_{|2} + \widetilde{\beta_3}\mathbf{X}_{|3} + \cdots + \widetilde{\beta_k}\mathbf{X}_{|k} \\ &= \left[\mathbf{1};\ \mathbf{X}_{|2};\ldots\ \mathbf{X}_{|k};\right] \begin{pmatrix} \widetilde{\beta_1} \\ \vdots \\ \widetilde{\beta_k} \end{pmatrix} = \mathbf{X}\widetilde{\boldsymbol{\beta}}. \end{aligned}$$

Así los valores ajustados son: $\widetilde{\boldsymbol{y}} = \mathbf{X}\widetilde{\boldsymbol{\beta}} \quad \in \mathbb{R}^N$

Datos del ejemplo del precio de las viviendas

Precio viviendas (miles de $) y superficie útil (pies al cuadrado)

14 casas unifamiliares en *University City*. San Diego, California. Año 1990.

	price (y)	sqft (x)	$\widetilde{\text{price}}$ ($\widetilde{y}$)
0	199.9	1065	
1	228	1254	
2	235	1300	
3	285	1577	
4	239	1600	
5	293	1750	
6	285	1800	
7	365	1870	
8	295	1935	
9	290	1948	
10	385	2254	
11	505	2600	
12	425	2800	
13	415	3000	

Si asumimos que el precio y se relaciona con la superficie x del siguiente modo:

$$y_n = a + b\, x_n + otrascosas_n,$$

podemos "*aproximar*" el vector de precios, $\boldsymbol{y}$, con una combinación lineal de los regresores:

$$\widetilde{\boldsymbol{y}} = \widetilde{\beta_1}\,\mathbf{1} + \widetilde{\beta_2}\,\boldsymbol{x} \quad = \quad \begin{bmatrix}\mathbf{1}; & \boldsymbol{x};\end{bmatrix} \begin{pmatrix}\widetilde{\beta_1}\\ \widetilde{\beta_2}\end{pmatrix} = \quad \mathbf{X}\widetilde{\boldsymbol{\beta}}.$$

Los precios ajustados como combinación lineal de los regresores

$$\widetilde{\boldsymbol{y}} = (\mathbf{X}_{|1})\widetilde{\beta_1} + (\mathbf{X}_{|2})\widetilde{\beta_2} = \begin{pmatrix} \widetilde{y_1} \\ \widetilde{y_2} \\ \widetilde{y_3} \\ \widetilde{y_4} \\ \widetilde{y_5} \\ \widetilde{y_6} \\ \widetilde{y_7} \\ \widetilde{y_8} \\ \widetilde{y_9} \\ \widetilde{y_{10}} \\ \widetilde{y_{11}} \\ \widetilde{y_{12}} \\ \widetilde{y_{13}} \\ \widetilde{y_{14}} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \widetilde{\beta_1} + \begin{pmatrix} 1065 \\ 1254 \\ 1300 \\ 1577 \\ 1600 \\ 1750 \\ 1800 \\ 1870 \\ 1935 \\ 1948 \\ 2254 \\ 2600 \\ 2800 \\ 3000 \end{pmatrix} \widetilde{\beta_2} = \begin{bmatrix} 1 & 1065 \\ 1 & 1254 \\ 1 & 1300 \\ 1 & 1577 \\ 1 & 1600 \\ 1 & 1750 \\ 1 & 1800 \\ 1 & 1870 \\ 1 & 1935 \\ 1 & 1948 \\ 1 & 2254 \\ 1 & 2600 \\ 1 & 2800 \\ 1 & 3000 \end{bmatrix} \begin{pmatrix} \widetilde{\beta_1} \\ \widetilde{\beta_2} \end{pmatrix} = \mathbf{X}\widetilde{\boldsymbol{\beta}};$$

El precio ajustado para el séptimo piso de la muestra será

$$\widetilde{y_7} = {}_{7|}\widetilde{\boldsymbol{y}} = {}_{7|}\mathbf{X}\widetilde{\boldsymbol{\beta}} = (1,\ 1800,) \begin{pmatrix} \widetilde{\beta_1} \\ \widetilde{\beta_2} \end{pmatrix} = (1)\widetilde{\beta_1} + (1800)\widetilde{\beta_2} \neq y_7.$$

Pero *¿qué criterio empleamos para elegir $\widetilde{\beta_1}$ y $\widetilde{\beta_2}$ en el ajuste $\widetilde{\boldsymbol{y}} = \mathbf{X}\widetilde{\boldsymbol{\beta}}$?*

Término de error

Dados $\mathbf{X}$ e $\boldsymbol{y}$, el *"error de ajuste"* empleando $\widetilde{\boldsymbol{\beta}}$ es

$$\widetilde{\boldsymbol{e}} \quad = \quad \boldsymbol{y} - \mathbf{X}\widetilde{\boldsymbol{\beta}} \quad = \quad \boldsymbol{y} - \widetilde{\boldsymbol{y}};$$

Así, descomponemos los datos observados $\boldsymbol{y}$ en: $\quad \boldsymbol{y} = \widetilde{\boldsymbol{y}} + \widetilde{\boldsymbol{e}}$.

Llamamos "Suma de los Residuos al Cuadrado" (SRC) del ajuste $\widetilde{\boldsymbol{y}}$ a

$$SRC(\widetilde{\boldsymbol{\beta}}) \quad \equiv \quad \sum_{n=1}^{N}(\widetilde{e}_n)^2 \quad = \quad \widetilde{\boldsymbol{e}} \cdot \widetilde{\boldsymbol{e}} \quad = \quad \|\widetilde{\boldsymbol{e}}\|^2$$

es decir, al cuadrado de la longitud del vector $\widetilde{\boldsymbol{e}} \;=\; (\boldsymbol{y} - \widetilde{\boldsymbol{y}})$.

Criterio del ajuste MCO

- Objetivo. Encontrar una combinación lineal de los regresores

  Sean $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$ y $\mathbf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_N \end{bmatrix}$. Buscamos una $\mathbf{X}\widetilde{\boldsymbol{\beta}}$.

- Criterio de búsqueda (ajuste MCO)

  Elegir el vector $\widetilde{\boldsymbol{\beta}}$ tal que $\mathbf{X}\widetilde{\boldsymbol{\beta}}$ esté *lo más cerca* de $\boldsymbol{y}$; es decir, tal que

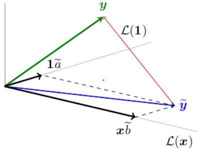
  la componente $\widetilde{\boldsymbol{e}}$ sea lo más pequeña posible en la descomposición:

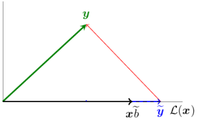
$$\begin{aligned} \boldsymbol{y} &= \mathbf{X}\widetilde{\boldsymbol{\beta}} + \widetilde{\boldsymbol{e}} \\ &= \widetilde{\boldsymbol{y}} + \widetilde{\boldsymbol{e}}. \end{aligned}$$

Geometría de un mal ajuste lineal

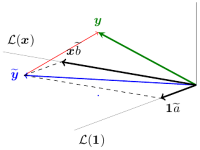
Un $\widetilde{a}$ demasiado pequeño y un $\widetilde{b}$ demasiado grande.

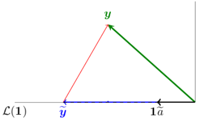
Desde un lado



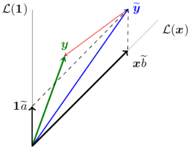


Desde el otro





Desde arriba



$$\mathbf{X} = [\mathbf{1};\ \boldsymbol{x};]; \quad \widetilde{\boldsymbol{\beta}} = \begin{pmatrix}\widetilde{a}\\ \widetilde{b}\end{pmatrix}; \quad \boxed{\widetilde{\boldsymbol{y}} = \mathbf{X}\widetilde{\boldsymbol{\beta}}}; \quad \boldsymbol{y} = \widetilde{\boldsymbol{y}} + \widetilde{\boldsymbol{e}}; \quad \widetilde{\boldsymbol{e}} = \boldsymbol{y} - \widetilde{\boldsymbol{y}}$$

Ecuaciones normales

Como el vector $\widehat{\boldsymbol{e}} = (\boldsymbol{y} - \widehat{\boldsymbol{y}})$ es mínimo si es *perpendicular* a cada regresor. Es decir, si:

$$\widehat{\boldsymbol{e}} \perp \mathbf{X}_{|j} \quad \Leftrightarrow \quad \mathbf{0} = \mathbf{X}^\intercal \widehat{\boldsymbol{e}} = \mathbf{X}^\intercal (\boldsymbol{y} - \widehat{\boldsymbol{y}}).$$

Tenemos que,

$$\widehat{\boldsymbol{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} \quad \Leftrightarrow \quad \mathbf{X}^\intercal \left(\boldsymbol{y} - \mathbf{X}\widehat{\boldsymbol{\beta}}\right) = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{X}^\intercal \boldsymbol{y} - \mathbf{X}^\intercal \mathbf{X}\widehat{\boldsymbol{\beta}} = \mathbf{0}$$

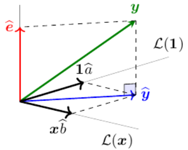
$$\widehat{\boldsymbol{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}} \qquad \text{si} \qquad \boxed{(\mathbf{X}^\intercal \mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}^\intercal \boldsymbol{y}} \tag{1}$$

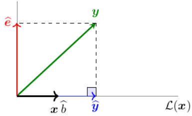
Las soluciones $\widehat{\boldsymbol{\beta}}$ son los parámetros del ajuste MCO: $\widehat{\boldsymbol{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}$

Ajuste MCO (Geometría de la proyección ortogonal)

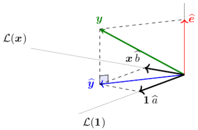
$$\widehat{\boldsymbol{e}} \perp \mathbf{X} \quad \Longleftrightarrow \quad \widehat{\boldsymbol{\beta}} \text{ es tal que } \quad (\mathbf{X}^\intercal\mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}^\intercal\boldsymbol{y}$$

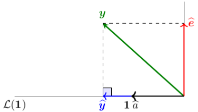
Desde un lado



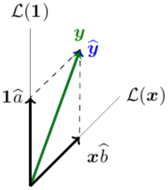


Desde el otro





Desde arriba



$$\mathbf{X} = [\mathbf{1};\ \boldsymbol{x};];\quad \widehat{\boldsymbol{\beta}} = \begin{pmatrix}\widehat{a}\\ \widehat{b}\end{pmatrix};\quad \boxed{\widehat{\boldsymbol{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}};\quad \boldsymbol{y} = \widehat{\boldsymbol{y}} + \widehat{\boldsymbol{e}};\quad \widehat{\boldsymbol{e}} = \boldsymbol{y} - \widehat{\boldsymbol{y}}$$

¿Es el sistema de ecuaciones normales determinado?

Puesto que

$$\mathbf{X}\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{y}} \qquad \Longleftrightarrow \qquad (\mathbf{X}^{\intercal}\mathbf{X})\widehat{\boldsymbol{\beta}} = \mathbf{X}^{\intercal}\boldsymbol{y}, \qquad \text{donde } \underset{N\times k}{\mathbf{X}};$$

ambos sistemas tendrán *solución única si y sólo* si sus matrices de coeficientes son de *rango* k.

En tal caso, multiplicando ambos lados de las ecuaciones normales por $\left(\mathbf{X}^{\intercal}\mathbf{X}\right)^{-1}$ tenemos que

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{X}^{\intercal}\mathbf{X}\right)^{-1}\mathbf{X}^{\intercal}\boldsymbol{y} \tag{2}$$

es *la única solución*.

Ecuación de salarios:

Supongamos el siguiente modelo

$$Salar_n = e^{\beta_1 + \beta_2 \left(educ_n\right) + \beta_3 \left(antig_n\right) + \beta_4 \left(exper_n\right) + otrascosas_n};$$

Tomando logaritmos tenemos una nueva variable $\ln(Salar_n)$ que podremos ajustar con un modelo lineal en los parámetros pues,

$$\ln(Salar_n) = \beta_1 + \beta_2 \left(educ_n\right) + \beta_3 \left(antig_n\right) + \beta_4 \left(exper_n\right) + otrascosas_n.$$

Pero ¿qué pasará si jamás ningún trabajador cambió de empresa?

La matriz $\mathbf{X}^\intercal\mathbf{X}$ será singular. De hecho, como *experiencia* y *antigüedad* coinciden, como mucho sólo podemos calcular su efecto conjunto:

$$\ln(Salar_n) = \beta_1 + \beta_2 \left(educ_n\right) + (\beta_3 + \beta_4) exper_n + otrascosas_n,$$