

Matemáticas II

Marcos Bujosa

Universidad Complutense de Madrid

30/03/2023

1 / 14

L-19

1 Esquema de la Lección 19

Esquema de la Lección 19

- Media
- Desviación típica y varianza
- Mínimos Cuadrados Ordinarios (MCO)

2 / 14

L-19

Puede encontrar la última versión de este material en

<https://github.com/mbujosab/MatematicasII/tree/main/Esp>

– Marcos Bujosa. Copyright © 2008–2023

Algunos derechos reservados. Esta obra está bajo una licencia de Creative Commons Reconocimiento-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-sa/4.0/> o envíe una carta a Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

1 / 14

L-19

2 Una restrcción en estadística y probabilidad

La norma del vector constante “uno” es 1

Esto no se cumple con el producto punto de \mathbb{R}^m ($m > 1$)

$$\|\mathbf{1}\|^2 = \langle \mathbf{1}, \mathbf{1} \rangle = \mathbf{1} \cdot \mathbf{1} = \sum_{i=1}^m 1 = m.$$

Nuevo producto escalar en \mathbb{R}^m para la estadística

$$\langle \mathbf{x}, \mathbf{y} \rangle_s = \frac{1}{m} (\mathbf{x} \cdot \mathbf{y})$$

(de manera que: $\|\mathbf{1}\|^2 = \frac{1}{m} (\mathbf{1} \cdot \mathbf{1}) = 1$)

3 / 14

3 La media aritmética

La media aritmética μ_y es el producto escalar de \mathbf{y} con $\mathbf{1}$

$$\mu_y = \frac{1}{m}(\mathbf{1} \cdot \mathbf{y}), \quad \text{es decir,} \quad \mu_y = \frac{1}{m} \sum_{i=1}^m y_i$$

La media aritmética μ_y es el **valor** por el que multiplicar $\mathbf{1}$ para obtener la proyección ortogonal de \mathbf{y} sobre $\mathcal{L}([\mathbf{1};])$

\bar{y} : proyección de $\mathbf{y} \in \mathbb{R}^m$ sobre la recta $\mathcal{L}([\mathbf{1};]) \subset \mathbb{R}^m$

$$\bar{y} = \mathbf{1}\hat{a} \quad \text{y} \quad (\mathbf{y} - \bar{y}) \perp \mathbf{1} \Rightarrow \frac{1}{m}(\mathbf{y} - \bar{y}) \cdot \mathbf{1} = 0$$

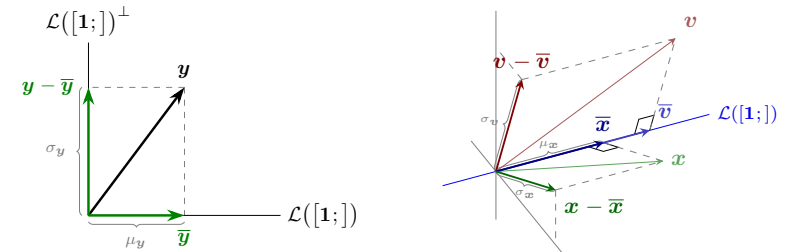
$$\frac{1}{m}(\mathbf{y} - \mathbf{1}\hat{a}) \cdot \mathbf{1} = 0 \Leftrightarrow \frac{1}{m}(\mathbf{y} \cdot \mathbf{1}) - \frac{1}{m}(\mathbf{1} \cdot \mathbf{1})\hat{a} = 0;$$

Por tanto

$$\hat{a} = \frac{1}{m}(\mathbf{y} \cdot \mathbf{1}) = \mu_y$$

4 / 14

4 La media aritmética

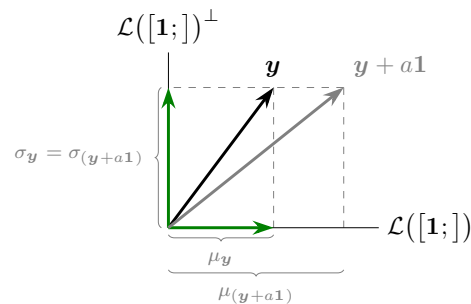


5 / 14

5 Desviación típica

$$\sigma_y = \|\mathbf{y} - \bar{y}\|.$$

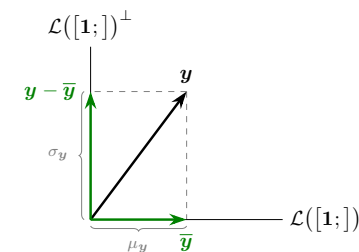
Sumar a \mathbf{y} un vector constante $a\mathbf{1}$ no cambia la desviación típica.



6 / 14

6 Varianza y el Teorema de Pitagoras

$$\sigma_y^2 = \|\mathbf{y} - \bar{y}\|^2 = \frac{1}{m}(\mathbf{y} - \bar{y}) \cdot (\mathbf{y} - \bar{y}) = \frac{1}{m} \sum_i (y_i - \mu_y)^2.$$

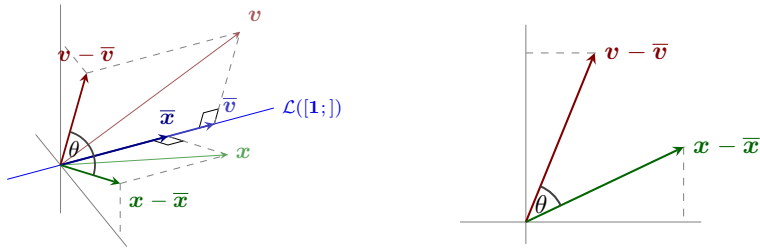


$$\sigma_y^2 = \|\mathbf{y} - \bar{y}\|^2 = \|\mathbf{y}\|^2 - \|\bar{y}\|^2 = \frac{1}{m}(\mathbf{y} \cdot \mathbf{y}) - \mu_y^2 = \frac{\sum_i y_i^2}{m} - \mu_y^2.$$

7 / 14

7 Covarianza y correlación

$$\sigma_{xy} = \frac{1}{m}(\mathbf{x} - \boldsymbol{\mu}_x) \cdot (\mathbf{y} - \bar{\mathbf{y}});$$



$$\rho_{xy} = \frac{\frac{1}{m}(\mathbf{x} - \boldsymbol{\mu}_x) \cdot (\mathbf{y} - \bar{\mathbf{y}})}{\|(\mathbf{x} - \boldsymbol{\mu}_x)\| \cdot \|(\mathbf{y} - \bar{\mathbf{y}})\|} = \frac{\sigma_{xy}}{\sqrt{\sigma_x \sigma_y}} = \cos(\theta).$$

8 / 14

8 Mínimos Cuadrados Ordinarios (MCO)

Sea \mathbf{X} tal que $\mathcal{L}([1;]) \subset \mathcal{C}(\mathbf{X})$.

Denotamos con $\hat{\mathbf{y}}$ la proyección ortogonal de $\mathbf{y} \in \mathbb{R}^m$ sobre $\mathcal{C}(\mathbf{X})$

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} \quad \text{y} \quad (\mathbf{y} - \hat{\mathbf{y}}) \perp \mathcal{C}(\mathbf{X}) \Rightarrow \frac{1}{m}\mathbf{X}^\top(\mathbf{y} - \hat{\mathbf{y}}) = \mathbf{0}$$

$$\frac{1}{m}\mathbf{X}^\top(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = \mathbf{0} \iff \frac{1}{m}\mathbf{X}^\top\mathbf{y} - \frac{1}{m}\mathbf{X}^\top\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{0}.$$

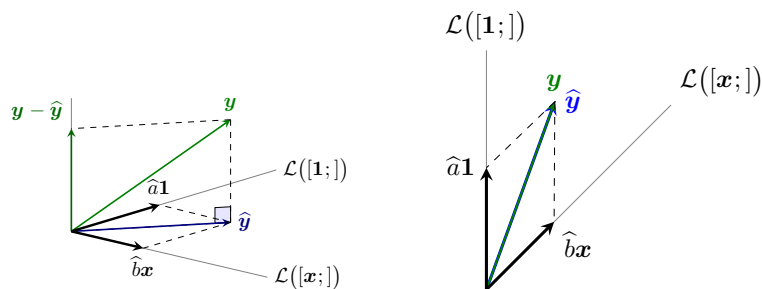
Por tanto

$$\left(\frac{1}{m}\mathbf{X}^\top\mathbf{X}\right)\hat{\boldsymbol{\beta}} = \frac{1}{m}\mathbf{X}^\top\mathbf{y}.$$

9 / 14

9 Mínimos Cuadrados Ordinarios (MCO)

Si $\mathbf{X} = [1; x;]$ es de rango 2.



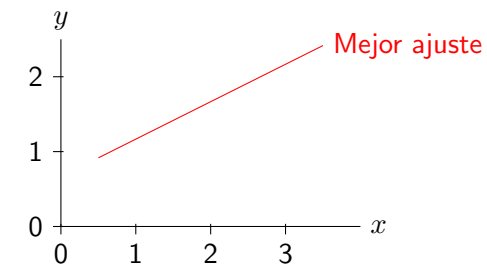
$$\left(\frac{1}{m}\mathbf{X}^\top\mathbf{X}\right)\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{1}{m}\mathbf{X}^\top\mathbf{y}.$$

10 / 14

10 Aplicación: Mínimos cuadrados (Ajuste lineal)

“buscando la mejor recta de ajuste $\hat{y} = \hat{a} + \hat{b}x$ ”

Puntos (x, y) : $(1, 1)$; $(2, 2)$; $(3, 2)$



$$\begin{cases} a + 1b = 1 \\ a + 2b = 2 \\ a + 3b = 2 \end{cases} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \quad (\mathbf{X}\boldsymbol{\beta} = \mathbf{y} \text{ Sin solución})$$

11 / 14

11 Aplicación: Mínimos cuadrados (Ajuste lineal)

$$\mathbf{X}\beta = \mathbf{y} \quad (\text{Sin solución}) \rightarrow \left(\frac{1}{m}\mathbf{X}^T\mathbf{X}\right)\hat{\beta} = \frac{1}{m}\mathbf{X}^T\mathbf{y} \rightarrow \hat{\mathbf{y}} = \mathbf{X}\hat{\beta}.$$

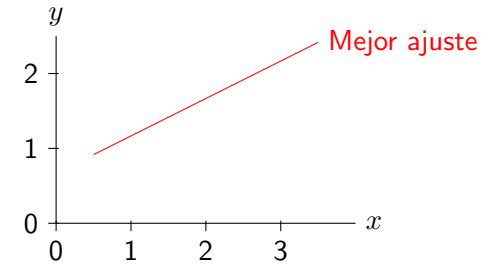
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 5 \\ 11 \end{bmatrix} \Rightarrow \hat{a} = \frac{2}{3}; \quad \hat{b} = \frac{1}{2}.$$

Mejor solución: $\frac{2}{3} + \frac{1}{2}x$

12 / 14

12 Aplicación: Mínimos cuadrados (Ajuste lineal)



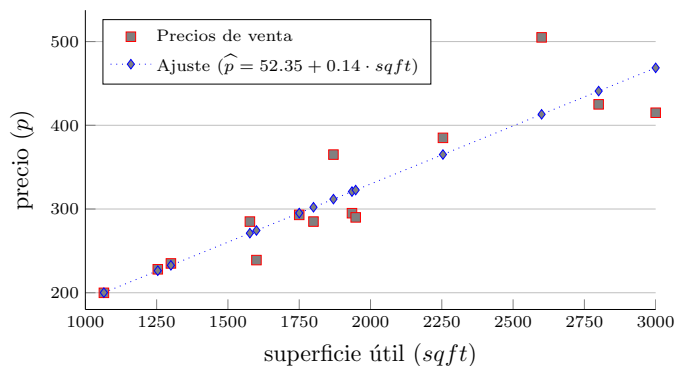
$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \rightarrow \hat{\mathbf{y}} = \begin{bmatrix} 7/6 \\ 10/6 \\ 13/6 \end{bmatrix} \rightarrow \hat{\mathbf{e}} = \begin{bmatrix} -1/6 \\ 2/6 \\ -1/6 \end{bmatrix}$$

$$\mathbf{y} = \hat{\mathbf{y}} + \hat{\mathbf{e}} \quad \text{y} \quad \begin{cases} \hat{\mathbf{e}} \cdot \hat{\mathbf{y}} = 0 \\ \hat{\mathbf{e}}\mathbf{X} = \mathbf{0} \end{cases}.$$

13 / 14

13 Aplicación: ajustando por mínimos cuadrados

Precios de venta (miles de dólares) y superficie útil (pies al cuadrado) de 14 casas unifamiliares en la *University City* de la ciudad de San Diego en California en 1990 (? , pp. 78)



14 / 14

Problemas de la Lección 19

(L-19) PROBLEMA 1. Con las medidas $\mathbf{y} = (0, 8, 8, 20,)$ tomadas en los instantes $\mathbf{x} = (0, 1, 3, 4,)$,

- Plantee y resuelva las ecuaciones normales $\mathbf{A}^T\mathbf{A}\hat{\beta} = \mathbf{A}^T\mathbf{y}$.
 - Para el mejor ajuste lineal, encuentre los ajustes p_i y los cuatro errores e_i .
 - ¿Cuál es el cuadrado de la norma del vector de errores $\|\mathbf{e}\|^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2$?
 - Dibuje la recta de regresión
 - Sustituya las medidas \mathbf{y} por los valores ajustados $\mathbf{p} = (1, 5, 13, 17,)$ escriba las cuatro ecuaciones $\mathbf{A}\beta = \mathbf{p}$. Encuentre la solución exacta a $\mathbf{A}\beta = \mathbf{p}$
 - Verifique que $\mathbf{e} = \mathbf{y} - \mathbf{p} = (-1, 3, -5, 3,)$ es perpendicular a las dos columnas de \mathbf{A} .
 - ¿Cuál es la distancia más corta $\|\mathbf{e}\|$ desde \mathbf{y} al espacio columna de \mathbf{A} ?
- (?, ejercicio 1-3 del conjunto de problemas 4.3.)

(L-19) PROBLEMA 2.

- Escriba las tres ecuaciones $y = \alpha + \beta x$ dado el conjunto de datos: $y = 7$ para $x = -1$, $y = 7$ para $x = 1$, y $y = 21$ para $x = 2$. Encuentre la solución de mínimos cuadrados $\hat{\beta} = (\hat{\alpha}, \hat{\beta})$ y pinte el mejor ajuste lineal.
- Encuentre la proyección $\mathbf{p} = \mathbf{A}\hat{\beta}$. Es decir, tres valores del mejor ajuste lineal. Demuestre que el vector de error es $\mathbf{e} = (2, -6, 4,)$. ¿Por qué es $\mathbf{P}\mathbf{e} = \mathbf{0}$?

14 / 14

(L-19) **PROBLEMA 3.** Our measurements at times $t = 1, 2, 3$ are $b = 1, 4$, and b_3 . We want to fit those points by the nearest line $C + Dt$, using least squares.

- (a) Which value for b_3 will put the three measurements on a straight line? *Which line is it?* Will least squares choose that line if the third measurement is $b_3 = 9$? (Yes or no).
- (b) What is the linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ that would be solved exactly for $\mathbf{x} = (C, D)$ if the three points do lie on a line? Compute the projection matrix \mathbf{P} onto the column space of \mathbf{A} .
- (c) What is the rank of that projection matrix \mathbf{P} ? How is the column space of \mathbf{P} related to the column space of \mathbf{A} ? (You can answer with or without the entries of \mathbf{P} computed in (b).)
- (d) Suppose $b_3 = 1$. Write down the equation for the best least squares solution $\hat{\mathbf{x}}$, and show that the best straight line is horizontal.