

# Mathematics II

Marcos Bujosa

Universidad Complutense de Madrid

22/01/2024

1 / 15

L-19

## 1 Highlights of Lesson 19

### Highlights of Lesson 19

- Mean
- Standard deviation and variance
- Ordinary Least Squares (OLS)

2 / 15

L-19

You can find the last version of these course materials at

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

– Marcos Bujosa. Copyright © 2008–2024

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 / 15

L-19

## 2 Restriction in statistics and probability

Norm of constant vector “one” is 1

This fails using the dot product in  $\mathbb{R}^m$  ( $m > 1$ )

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

New scalar product in  $\mathbb{R}^m$  for statistics

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

$$(\text{so: } \|\mathbf{1}\|^2 = \frac{1}{m} (\mathbf{1} \cdot \mathbf{1}) = 1)$$

3 / 15

### 3 Mean

The mean  $\mu_y$  is the scalar product of  $\mathbf{y}$  and  $\mathbf{1}$

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

The mean  $\mu_y$  is the *value* by which to multiply  $\mathbf{1}$  to get the orthogonal projection of  $\mathbf{y}$  onto  $\mathcal{L}([\mathbf{1}])$

$\bar{\mathbf{y}}$ : projection of  $\mathbf{y} \in \mathbb{R}^m$  onto the line  $\mathcal{L}([\mathbf{1}]) \subset \mathbb{R}^m$

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

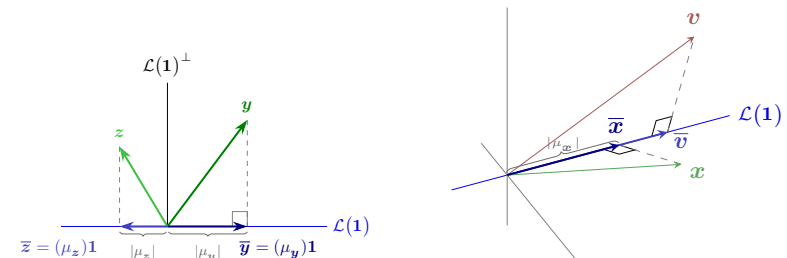
$$\frac{1}{m}(\mathbf{y} - \mathbf{1}\hat{a}) \cdot \mathbf{1} = 0 \Leftrightarrow \frac{1}{m}(\mathbf{y} \cdot \mathbf{1}) - \frac{1}{m}(\mathbf{1} \cdot \mathbf{1})\hat{a} = 0;$$

Therefore

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

4 / 15

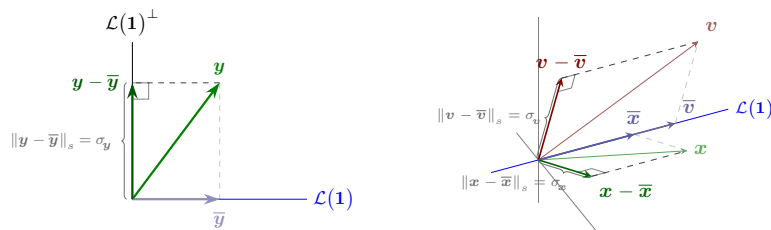
### 4 Mean



5 / 15

### 5 Standard deviation

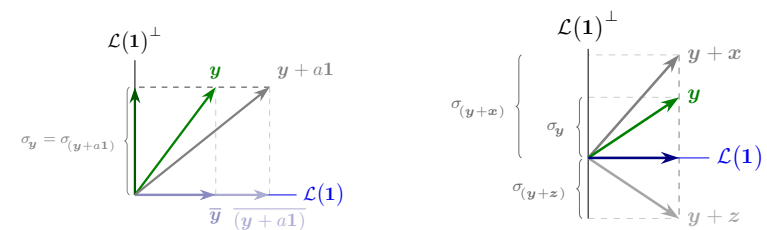
$$\sigma_y = \|\mathbf{y} - \bar{\mathbf{y}}\|.$$



6 / 15

### 6 Constant Vectors and Zero Mean Vectors

Adding a constant vector  $a\mathbf{1}$  to  $\mathbf{y}$  does not change the standard deviation.

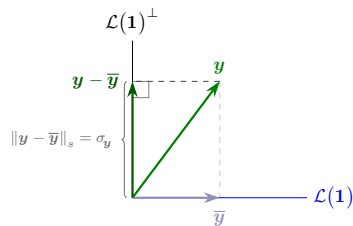


$$\sigma_z = 0 \Leftrightarrow \mathbf{z} = a\mathbf{1}; \quad \mu_z = 0 \Leftrightarrow \mathbf{z} \perp \mathbf{1}$$

7 / 15

## 7 Variance and the Pythagorean theorem

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

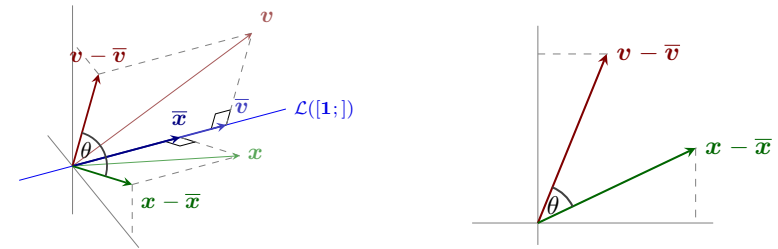


$$\sigma_y^2 = \|\mathbf{y} - \bar{\mathbf{y}}\|^2 = \|\mathbf{y}\|^2 - \|\bar{\mathbf{y}}\|^2 = \frac{1}{m}(\mathbf{y} \cdot \mathbf{y}) - \mu_y^2 = \frac{\sum_i y_i^2}{m} - \mu_y^2.$$

8 / 15

## 8 Covariance and correlation

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



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

9 / 15

## 9 Ordinary Least Squares (OLS)

Let  $\mathbf{X}$  such that  $\mathcal{L}([1;]) \subset \mathcal{C}(\mathbf{X})$ .

$\hat{\mathbf{y}}$  is the orthogonal projection of  $\mathbf{y} \in \mathbb{R}^m$  onto  $\mathcal{C}(\mathbf{X})$

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

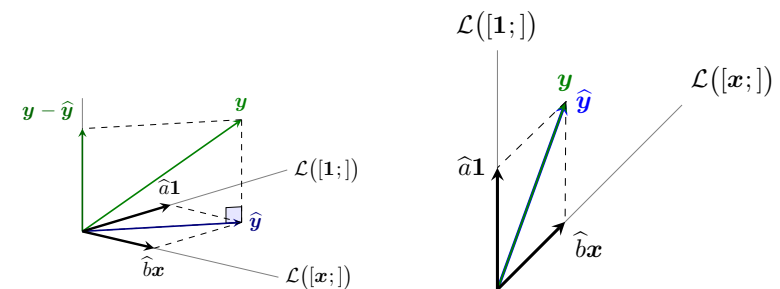
$$\frac{1}{m}\mathbf{X}^\top(\mathbf{y} - \mathbf{X}\hat{\beta}) = \mathbf{0} \iff \frac{1}{m}\mathbf{X}^\top\mathbf{y} - \frac{1}{m}\mathbf{X}^\top\mathbf{X}\hat{\beta} = \mathbf{0}.$$

Therefore

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

## 10 Ordinary Least Squares (OLS)

If  $\mathbf{X} = [1; x;]$  has rank 2.



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

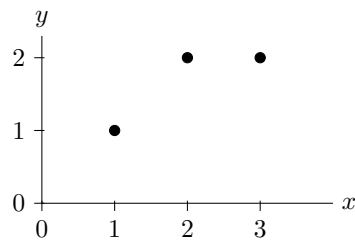
10 / 15

11 / 15

### 11 Application: Least Squares (Fitting by a line)

"looking for the best fitting line  $\hat{y} = \hat{a} + \hat{b}x$ "

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



$$\begin{cases} a + 1b = 1 \\ a + 2b = 2 \\ a + 3b = 2 \end{cases} \rightarrow \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \quad (\mathbf{X}\beta = \mathbf{y} \text{ No solution})$$

12 / 15

### 12 Application: Least Squares (Fitting by a line)

$$\mathbf{X}\beta = \mathbf{y} \quad (\text{No solution}) \rightarrow \left(\frac{1}{m}\mathbf{X}^T\mathbf{X}\right)\hat{\beta} = \frac{1}{m}\mathbf{X}^T\mathbf{y} \rightarrow \hat{\mathbf{y}} = \mathbf{X}\hat{\beta}.$$

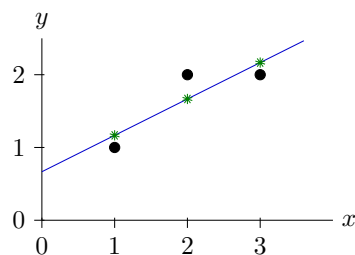
$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$$

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

Best solution:  $\frac{2}{3} + \frac{1}{2}x$

13 / 15

### 13 Application: Least Squares (Fitting by a line)



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

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

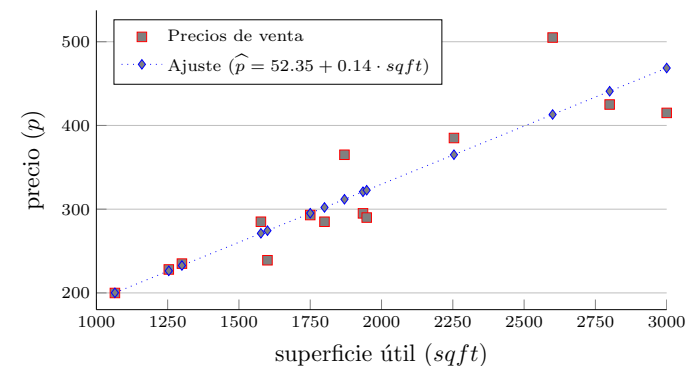
14 / 15

### 14 Application: Least Squares (Fitting by a line)

Selling price and living area of single family homes in University City community of San Diego, in 1990.

price = Sale price is in thousands of dollars

sqft = Square feet of living area (, pp. 78)



15 / 15

## Questions of the Lecture 19

(L-19) QUESTION 1. With the measurements  $\mathbf{y} = (0, 8, 8, 20, )$  at  $\mathbf{x} = (0, 1, 3, 4, )$ ,

- (a) Set up and solve the normal equations  $\mathbf{A}^T \mathbf{A} \hat{\beta} = \mathbf{A}^T \mathbf{y}$ .
  - (b) For the best straight line, find its four fits  $p_i$  and four errors  $e_i$ .
  - (c) What is the value of the square of the norm of the error vector  $\|\mathbf{e}\|^2 = e_1^2 + e_2^2 + e_3^2 + e_4^2$ ?
  - (d) Draw the regression line
  - (e) Change the measurements to  $\mathbf{p} = (1, 5, 13, 17, )$  write down the four equations  $\mathbf{A}\beta = \mathbf{p}$ . Find an exact solution to  $\mathbf{A}\beta = \mathbf{p}$
  - (f) Check that  $\mathbf{e} = \mathbf{y} - \mathbf{p} = (-1, 3, -5, 3, )$  is perpendicular to both columns of the same matrix  $\mathbf{A}$ .
  - (g) What is the shortest distance  $\|\mathbf{e}\|$  from  $\mathbf{y}$  to the column space of  $\mathbf{A}$ ?
- (?, exercise 1–3 from section 4.3.)

(L-19) QUESTION 2.

- (a) Write down three equations  $y = \alpha + \beta x$  given the data:  $y = 7$  at  $x = -1$ ,  $y = 7$  at  $x = 1$ , and  $y = 21$  at  $x = 2$ . Find the least squares solution  $\hat{\beta} = (\hat{\alpha}, \hat{\beta})$  and draw the closest line.
- (b) Find the projection  $\mathbf{p} = \mathbf{A}\hat{\beta}$ . This gives the three heights of the closest line. Show that the error vector is  $\mathbf{e} = (2, -6, 4, )$ . Why is  $\mathbf{P}\mathbf{e} = \mathbf{0}$ ?

(L-19) QUESTION 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.