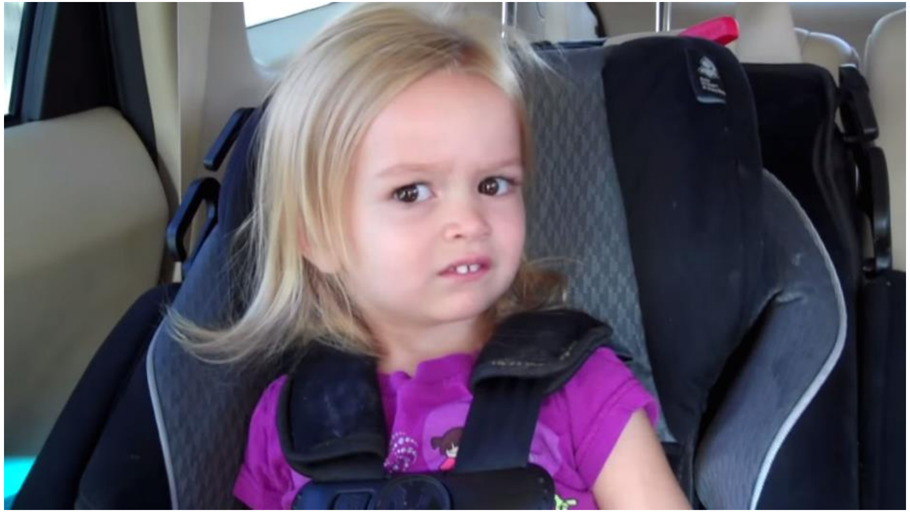
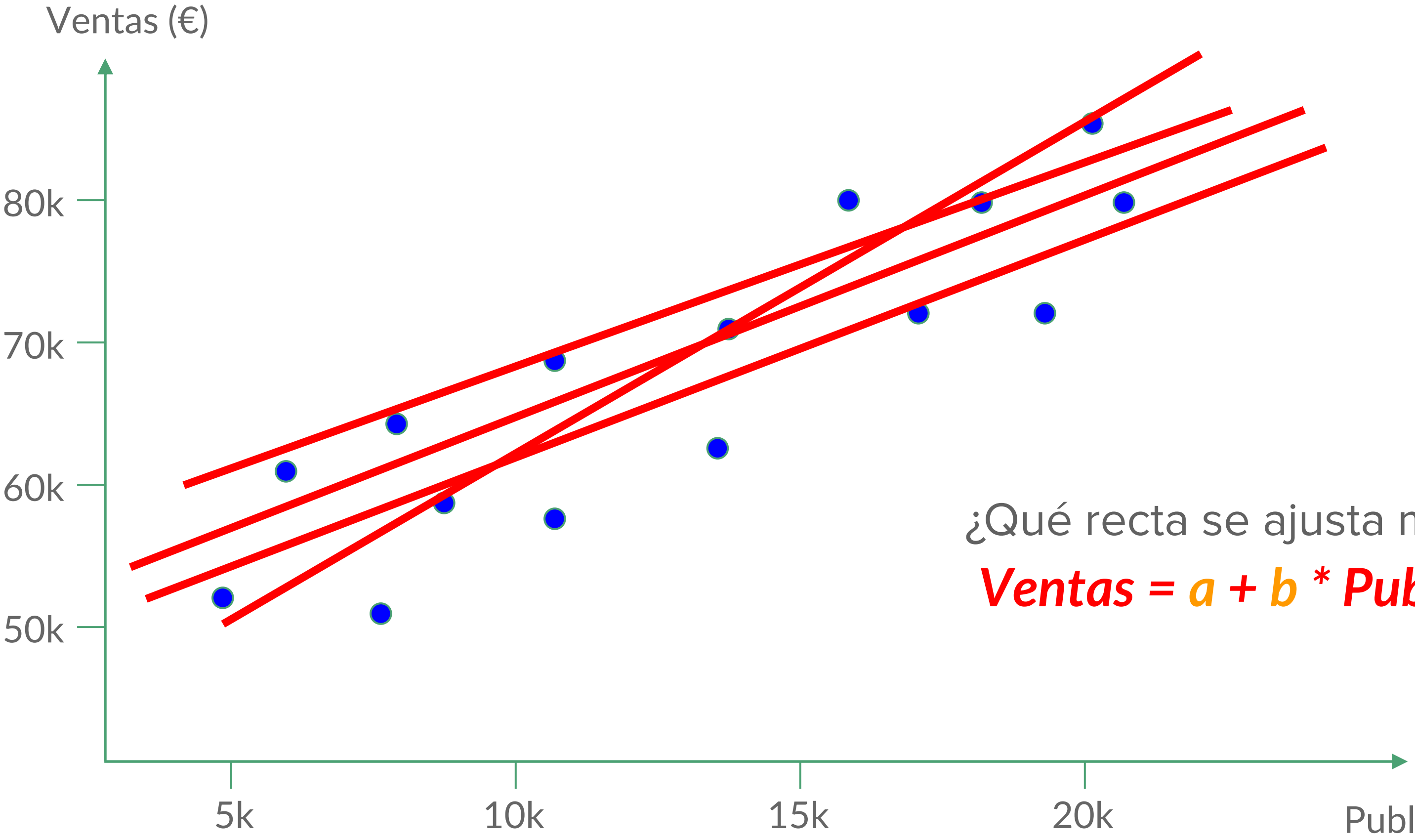
Mínimos cuadrados

Mínimos cuadrados

Mínimos cuadrados es una técnica de análisis numérico enmarcada dentro de la optimización matemática, en la que, dados un conjunto de puntos, se intenta encontrar la función que mejor se aproxime a los datos, de acuerdo con el criterio de mínimo error cuadrático.



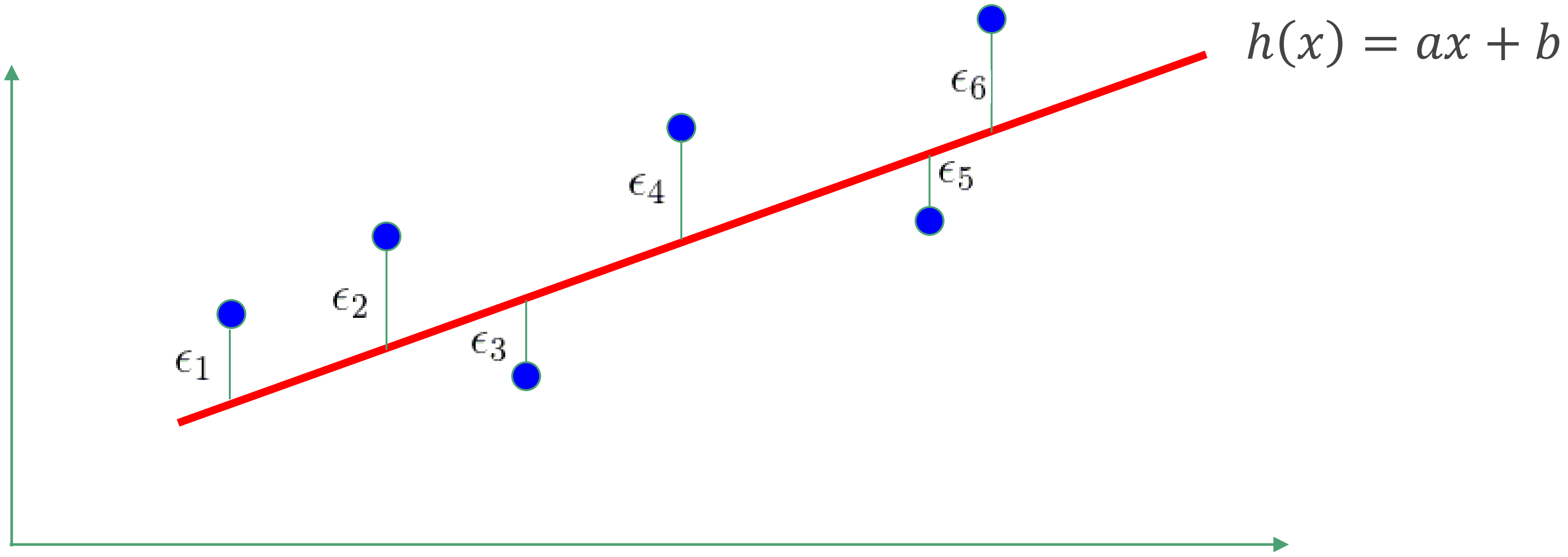
Mínimos cuadrados



¿Qué recta se ajusta mejor a los puntos?

Ventas = a + b * Publicidad

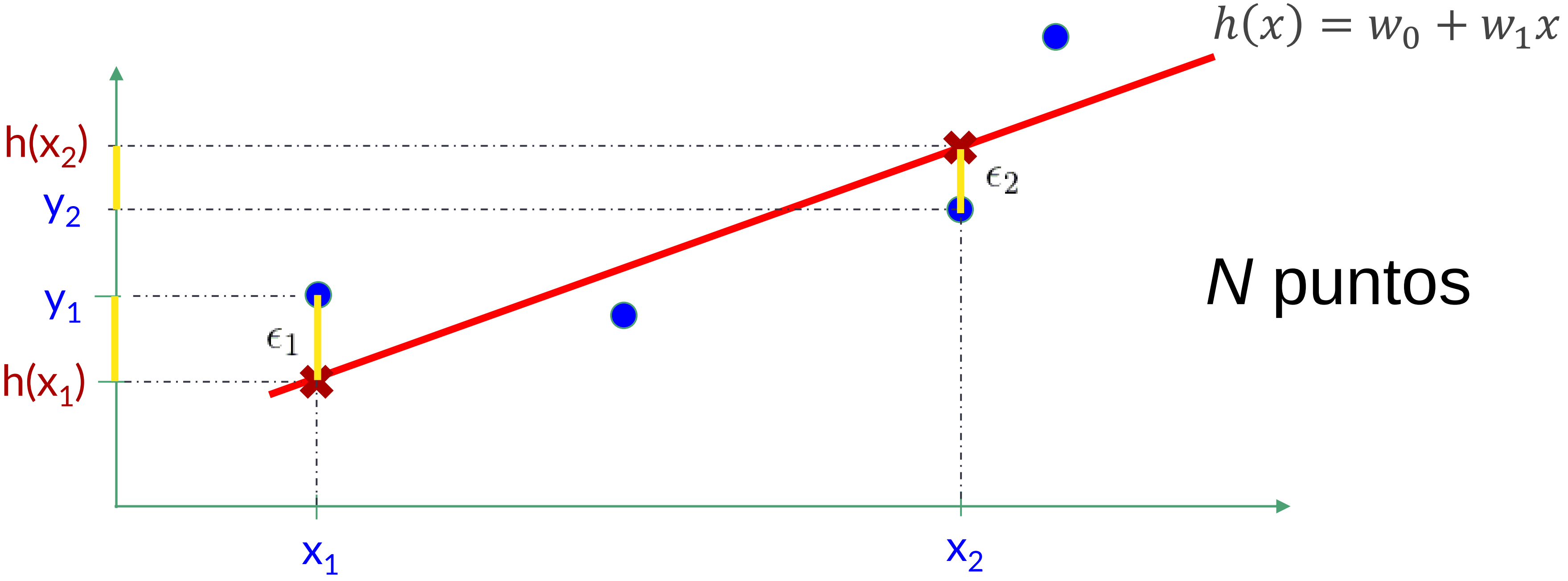
Mínimos cuadrados



Se busca la recta que minimice $$\sum_n \epsilon_n^2 = \epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2 + \epsilon_4^2 + \epsilon_5^2 + \epsilon_6^2 + \cdots + \epsilon_N^2$$

Es decir, se minimiza la distancia entre la recta y todos los puntos, para obtener los coeficientes *a* y *b*

Mínimos cuadrados



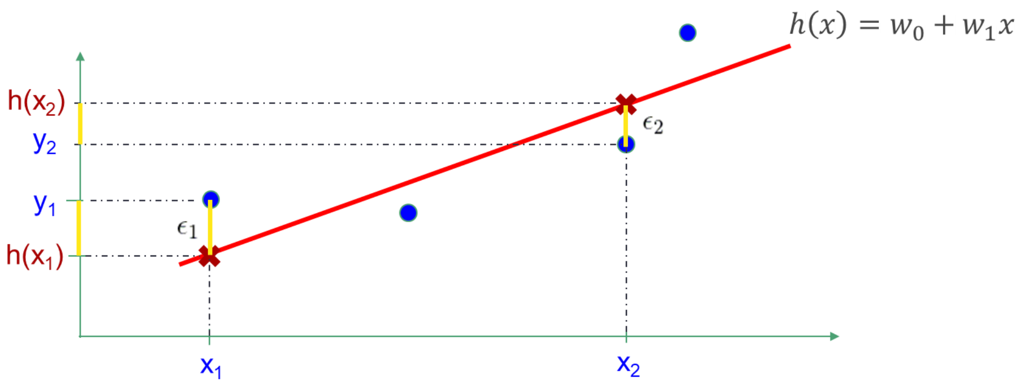
N puntos

$$\epsilon_1^2 = (h(x_1) - y_1)^2$$
$$\epsilon_2 = (h(x_2) - y_2)^2$$

$$\Rightarrow E = \frac{1}{N}\sum_{n=1}^{N}(h(x_n) - y_n))^2$$

Función de error
(error cuadrático medio)

Mínimos cuadrados



- **Datos:** $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$

$$\begin{aligned} h(x_1) &= \boxed{w_0 + w_1 x_1} \\ h(x_2) &= w_0 + w_1 x_2 \\ &\vdots \\ h(x_n) &= w_0 + w_1 x_n \end{aligned} \Rightarrow \underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} w_0 \\ w_1 \end{bmatrix}}_{w}$$

- **Errores cuadráticos:**

$$\begin{aligned} \epsilon_1^2 &= (w_0 + w_1 x_1 - y_1)^2 \\ \epsilon_2^2 &= (w_0 + w_1 x_2 - y_2)^2 \\ &\vdots \\ \epsilon_n^2 &= (w_0 + w_1 x_n - y_n)^2 \end{aligned} \Rightarrow \left(\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right)^2$$

$$|X \quad w - y|^2$$

Mínimos cuadrados

Minimizar $E_{in}(\mathbf{w}) = \frac{1}{N}|X\mathbf{w} - \mathbf{y}|^2$

$$\nabla E_{in}(\mathbf{w}) = \frac{2}{N}X^T(X\mathbf{w} - \mathbf{y}) = \mathbf{0}$$

$$X^TX\mathbf{w} - X^T\mathbf{y} = 0$$

$$X^TX\mathbf{w} = X^T\mathbf{y}$$

$$(X^TX)^{-1}X^TX\mathbf{w} = (X^TX)^{-1}X^T\mathbf{y}$$

$$I \cdot \mathbf{w} = (X^TX)^{-1}X^T\mathbf{y}$$

$$\Rightarrow \boxed{\mathbf{w} = (X^TX)^{-1}X^T\mathbf{y}}$$

$$n \times k \ \cdot k \times n = n \times n$$

¿Por qué no puedo hacer esto?

$$\nabla E_{in}(\mathbf{w}) = \frac{2}{N}X^T(X\mathbf{w} - \mathbf{y}) = \mathbf{0}$$

$$X^TX\mathbf{w} - X^T\mathbf{y} = 0$$

$$X^TX\mathbf{w} = X^T\mathbf{y}$$

$$(X^T)^{-1}X^TX\mathbf{w} = (X^T)^{-1}X^T\mathbf{y}$$

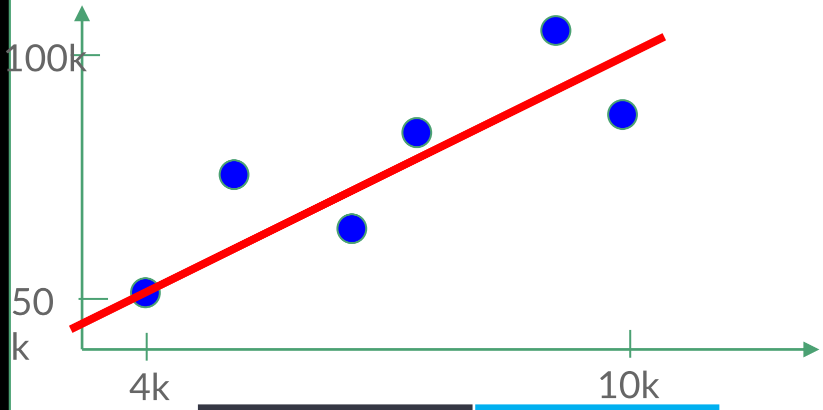
$$X \cdot \mathbf{w} = \mathbf{y}$$

$$X^{-1}X \cdot \mathbf{w} = X^{-1}\mathbf{y}$$

$$\Rightarrow \mathbf{w} = X^{-1}\mathbf{y}$$

Porque solo tienen inversa las matrices **cuadradas**. Normalmente X no va a ser cuadrada, mientras que (X^TX) siempre lo es

Mínimos cuadrados



Publicidad x_1	Ventas y
4	50
5	75
6	60
7	80
9	110
10	85

$$\boldsymbol{w} = (X^T X)^{-1} X^T \boldsymbol{y}$$

$$X = \begin{bmatrix} 1 & 4 \\ 1 & 5 \\ 1 & 6 \\ 1 & 7 \\ 1 & 9 \\ 1 & 10 \end{bmatrix}, y = \begin{bmatrix} 50 \\ 75 \\ 60 \\ 80 \\ 110 \\ 85 \end{bmatrix}$$

$$X^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 5 & 6 & 7 & 9 & 10 \end{bmatrix}$$

$$\Rightarrow \mathbf{w} = (X^T X)^{-1} X^T \boldsymbol{y} = \begin{bmatrix} 27.85 \\ 7.14 \end{bmatrix}$$

Ventas = 27.85 + 7.14 * Publicidad

THE BRIDGE DIGITAL TALENT ACCELERATOR

¡Gracias!

Contacto: Rafael Zambrano