

Cuadrados mínimos

Supongamos que tenemos los siguientes datos de unas mediciones

$$x_1 = 15$$

Buscamos la mejor

$$x_2 = 12$$

constante que

$$x_3 = 16$$

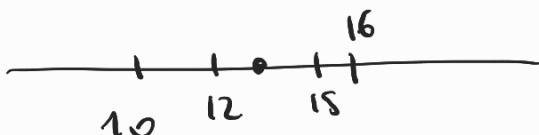
"resuma" o "represente"

$$x_4 = 10$$

a estos datos.

¿En qué sentido?

① Buscamos a:



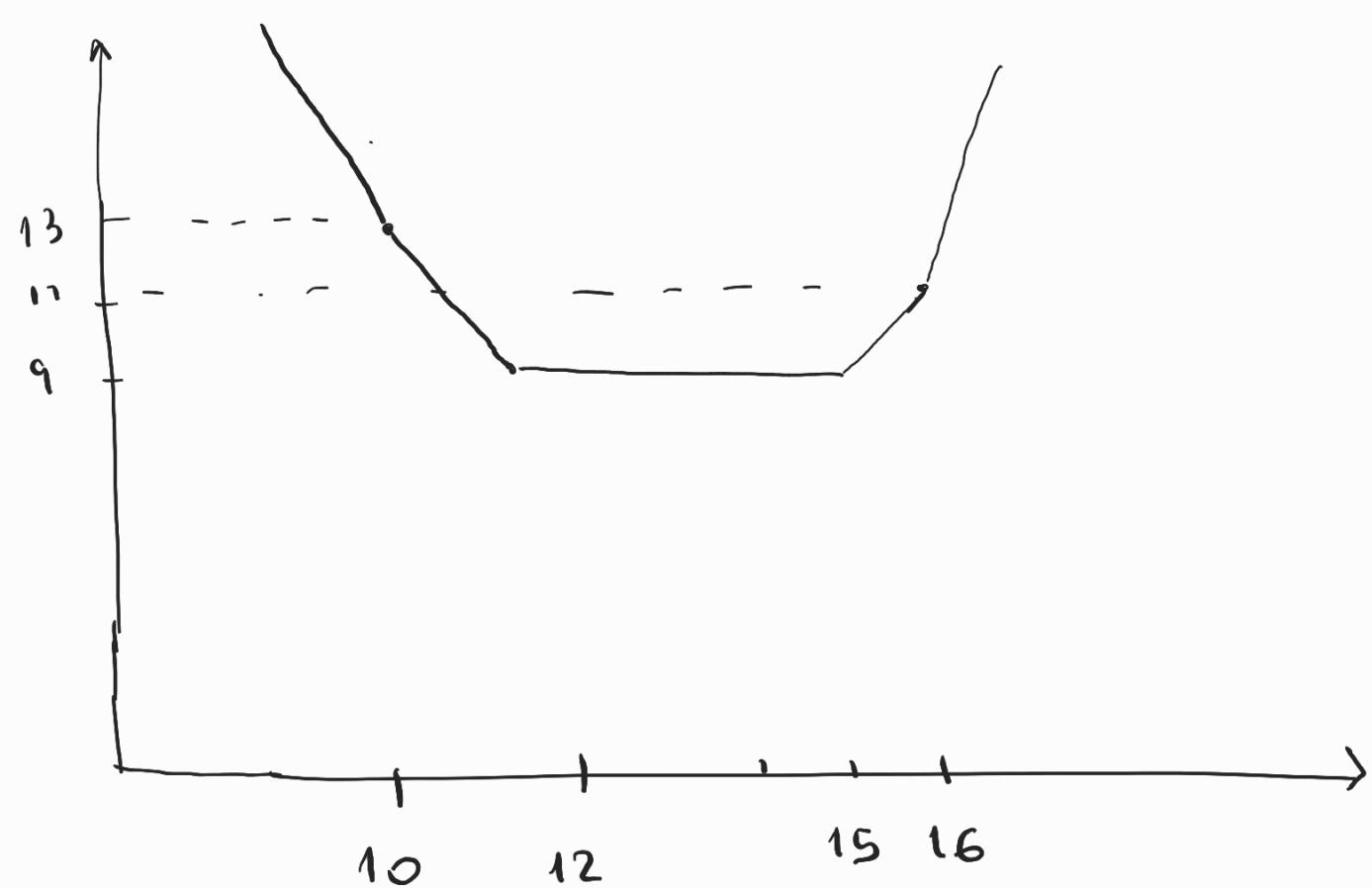
$|a - 15| + |a - 12| + |a - 16| + |a - 10| \rightarrow$ error que obtengo por usar "a" en lugar de mis datos.

sea mínimo.

Es decir $\min_a \sum_{i=1}^4 |a - x_i|$

$$F(a) = |a - 15| + |a - 12| + |a - 16| + |a - 10|$$

$$F(10) = 13, F(12) = 9, F(15) = 9, F(16) = 11$$



todos los $a \in [12, 15]$ son mínimos.

② Buscamos a:

$$(a-15)^2 + (a-12)^2 + (a-16)^2 + (a-10)^2 \rightarrow \text{error cuadrático.}$$

sea mínimo.

$$\text{Es decir } \min_a \sum_{i=1}^4 (a-x_i)^2$$

$$G(a) = (a-15)^2 + (a-12)^2 + (a-16)^2 + (a-10)^2$$

que es una cuadrática convexa \cup

Con lo cual el mínimo se alcanzará cuando a : $G'(a) = 0$

$$G'(a) = 2(a-15) + 2(a-12) + 2(a-16) + 2(a-10)$$

$$G'(a) = 0 \Leftrightarrow 4a - (15 + 12 + 16 + 10) = 0$$

$$\Leftrightarrow a = \frac{15 + 12 + 16 + 10}{4} = \frac{53}{4}$$

En general, dados los datos x_1, \dots, x_n el a que hace mínimos a

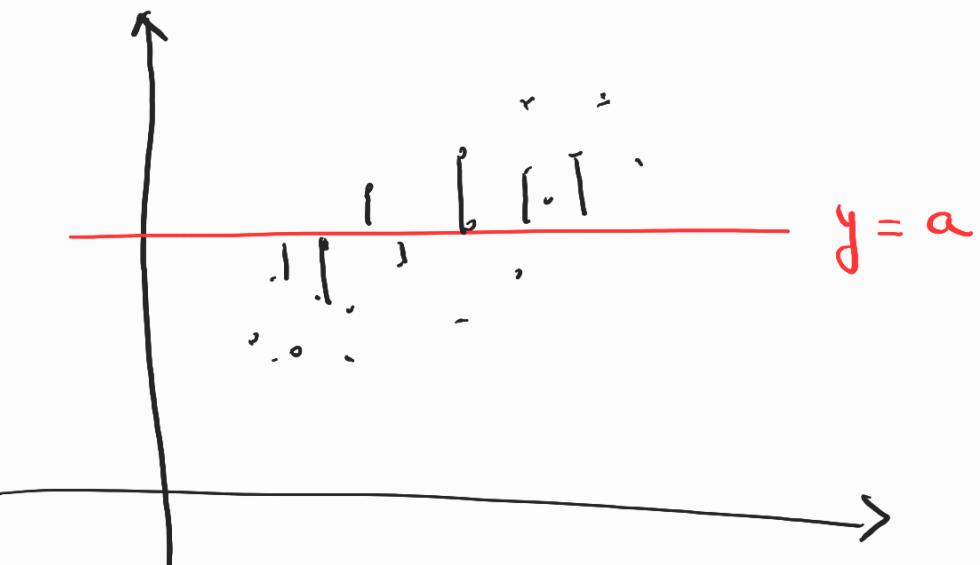
$$\sum_{i=1}^n (x_i - a)^2 \text{ es el promedio!}$$

$$a = \frac{\sum_{i=1}^n x_i}{n}$$

Si ahora tengo muchos datos en \mathbb{R}^2

$$(x_1, y_1) (x_2, y_2) \dots (x_n, y_n)$$

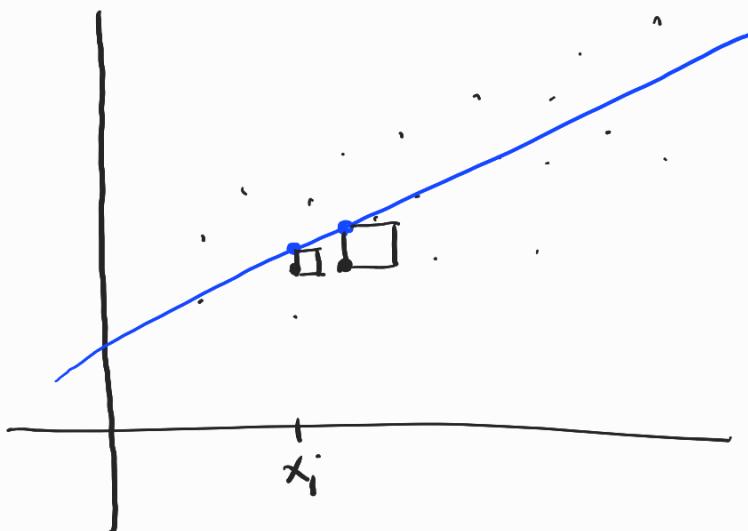
Busco la mejor constante por cuadrados mínimos



$$\sum_{i=1}^n (y_i - a)^2 \text{ sea min } a = \frac{1}{n} \sum y_i$$

Ahora buscamos la mejor recta

$$y = ax + b$$



tal que $\sum_{i=1}^n (ax_i + b - y_i)^2$ sea min.

Puedes derivar y usar A-L!

Más en general, en algunos casos tengo una tabla de datos y quiero encontrar

- Un polinomio de grado menor o igual a m que mejor aproxime a mis datos en el sentido de cuadrados mínimos.

En este caso, dada la familia de funciones $1, x, x^2, \dots, x^m$ busco las constantes a_0, a_1, \dots, a_m :

si $P(x) = a_0 + a_1 x + \dots + a_m x^m$

$$\sum_{i=1}^n (P(x_i) - y_i)^2 \text{ sea mím}$$

- Una combinación lineal de una familia de funciones dadas

ϕ_0, \dots, ϕ_m que mejor aproxime mis datos. Es decir, se busca

$$\min_{a_0, \dots, a_m} \sum (a_0 \phi_0(x_i) + \dots + a_m \phi_m(x_i) - f_i)^2$$

Volvamos al caso de la recta:

En un principio dados

$(x_1, y_1), \dots, (x_n, y_n)$ datos sería

ideal encontrar $L(x) = a_0 + a_1 x$

tales que $L(x_i) = y_i$.

(es claro que esto no va a ser normalmente el caso).

Si existieran a_0, a_1 serían solución de

$$a_0 + a_1 x_1 = y_1$$

$$a_0 + a_1 x_2 = y_2$$

:

$$a_0 + a_1 x_m = y_m$$

$$\rightsquigarrow \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Que normalmente no tendrá solución

si $m \geq n$ es más grande que 2.

Entonces nos preguntamos si habrá
en a_0, a_1 "mejores" como antes.

Llamamos $A = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{pmatrix}$ y $b = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$

Más en general,
el problema se puede pensar como
encontrar la solución $\bar{x} = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$

que hace que $\|A\bar{x} - b\|_2$

es mínimo.

Es decir, buscamos $\mathbf{I} \in \text{Im } A$

($\mathbf{I} = A\mathbf{x}$ para algún \mathbf{x})

donde $\text{Im } A$ es un subespacio de \mathbb{R}^m ,

que sea a distancia mínima deb.

Supongamos

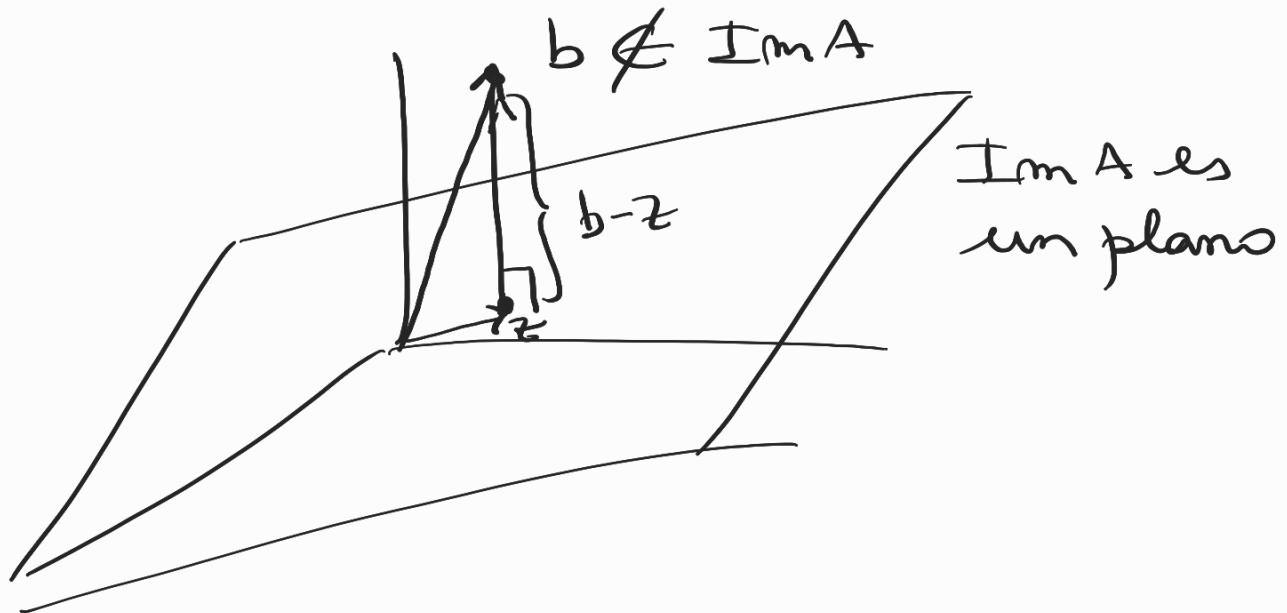
$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{3 \times 2} \quad y \quad b = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix}$$

$$\text{Im } A = \langle (1, 1, 1), (1, 2, 3) \rangle$$

Vemos que $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \notin \text{Im } A$ con
lo cual no existe $\begin{pmatrix} a_0 \\ a_1 \\ a_1 \end{pmatrix}: A\begin{pmatrix} a_0 \\ a_1 \\ a_1 \end{pmatrix} = b$

Pero buscamos $\begin{pmatrix} a_0 \\ a_1 \\ a_1 \end{pmatrix}: \|A\begin{pmatrix} a_0 \\ a_1 \\ a_1 \end{pmatrix} - b\|_2$

sea mínima.



$z \in S = \text{Im } A$ que tiene distancia mínima a b sabemos que es

$$z = P_S(b) \text{ que es igual}$$

que verifica que

$$b - z \perp \text{Im } A$$

es decir $b - z \in (\text{Im } A)^\perp$

$$(\text{Im } A)^\perp = \left\{ w \in \mathbb{R}^3 : \underbrace{\langle Ax, w \rangle}_{\text{vertical}} = \bar{0} \quad \forall x \in \mathbb{R}^2 \right\}$$

$$= \left\{ w \in \mathbb{R}^3 : (Ax)^T \cdot w = \bar{0} \quad \forall x \in \mathbb{R}^2 \right\}$$

$$= \left\{ w \in \mathbb{R}^3 : \underbrace{x^t A^t w = 0}_{x \in \mathbb{R}^2} \quad \forall x \in \mathbb{R}^2 \right\}$$

$$= \left\{ w \in \mathbb{R}^3 : A^t w = \bar{0} \right\}$$

$$= \text{Nú } A^t$$

buscamos $b - z \in \text{Nú } A^t$, es decir

$$A^t (b - z) = \bar{0}$$

Recapitulemos: Dados A, b ,

buscamos $z \in \text{Im } A$:

$$A^t (b - z) = \bar{0}.$$

Es decir, buscamos $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$:

$$A^t(b - A\begin{pmatrix} a \\ a_1 \end{pmatrix}) = \bar{0}$$

$$\Leftrightarrow A^t A \begin{pmatrix} a \\ a_1 \end{pmatrix} = A^t b$$

El $\begin{pmatrix} a \\ a_1 \end{pmatrix}$ que buscamos es solución de estas ecuaciones.

A estas ecuaciones se los llama **ecuaciones normales**.

En general se prueba que

Teorema: Dada $A \in \mathbb{R}^{m \times m}$, $b \in \mathbb{R}^m$
son equivalentes:

1) $a \in \mathbb{R}^m$ minimiza $\|Aa - b\|$

2) a es solución de $A^t A a = A^t b$

Además, si las columnas de A

son li, la solución a de $A^t A a = A^t b$

existe y es única. Notar $A^t A \in \mathbb{R}^{m \times m}$.

Dem: Se basa en que

$P_S: \mathbb{R}^m \rightarrow S$ satisface que :

•) $P_S(x)$ es el único elemento de S

que cumple $\langle x - P_S(x), w \rangle = 0$

$\forall w \in S$.

•) $P_S(x)$ es el único elemento de S

que satisface

$$\|x - P_S(x)\| = \min_{w \in S} \{\|x - w\|\}.$$

Veamos la última parte :

Si las columnas de A son li,

$A^t A \in \mathbb{R}^{m \times m}$ es inversible.

Veamos que es mono: Sea x :

$$A^t A x = \bar{0}$$

$$\Rightarrow \langle A^t A x, x \rangle = 0$$

$$\begin{aligned} &\Rightarrow 0 = (A^t A x)^t \cdot x = x^t A^t A x \\ &= (Ax)^t \cdot Ax = \langle Ax, Ax \rangle \\ &= \|Ax\|_2^2 \Rightarrow Ax = 0 \Rightarrow x = 0 \end{aligned}$$

A tiene
col. l.i.

$\Rightarrow A^t A$ es invertible.

OBS: Si $A^t A$ es invertible \Rightarrow

$A^t A$ resulta simétrica y definida positiva \Rightarrow podemos resolver el sistema lineal usando Cholesky.