Cuadrados mínimos:

Supongamos que tenemos los siguientes datos de unas mediciones

$x_1 = 15$

$x_2 = 12$

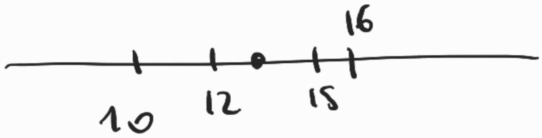
$x_3 = 16$

$x_4 = 10$

Buscamos la mejor constante que "resuma" o "represente" a estos datos.

¿En qué sentido?

(1) Buscamos a:



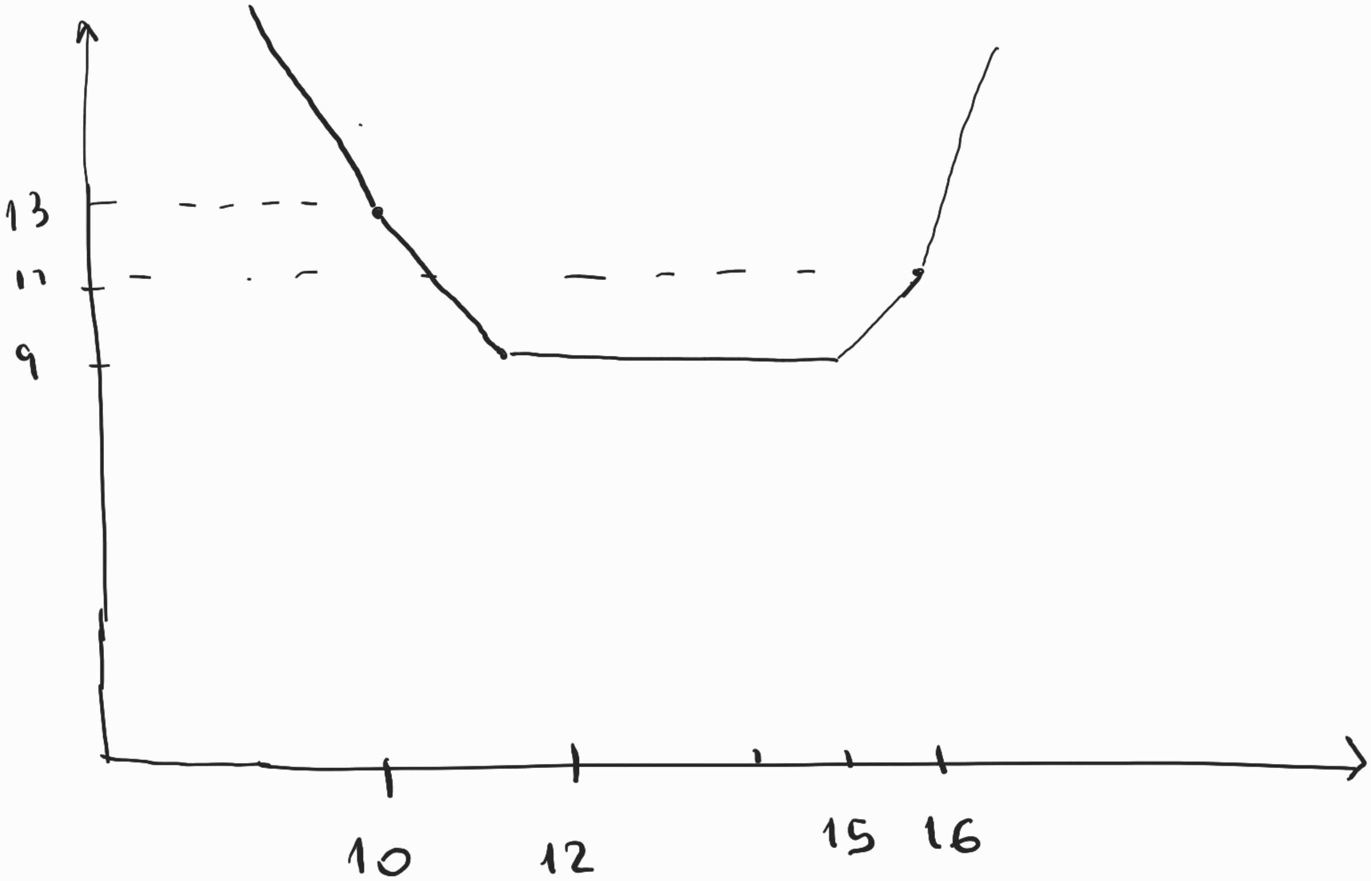
$$|a-15| + |a-12| + |a-16| + |a-10| \to$$ error que obtengo por usar "a" en lugar de mis datos.

sea mínimo.

Es decir $\min\limits_{a} \sum\limits_{i=1}^{4} |a - x_i|$

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

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



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

(2) Buscamos a:

$(a-15)^2 + (a-12)^2 + (a-16)^2 + (a-10)^2$ → error cuadrático.

sea mínimo.

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 ∪

Con lo cual el mínimo se alcanzará

cuando a: $G'(a)=0$

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

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

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

En general, dados los datos $x_1, \dots, x_n$

el a que hace mínimos a

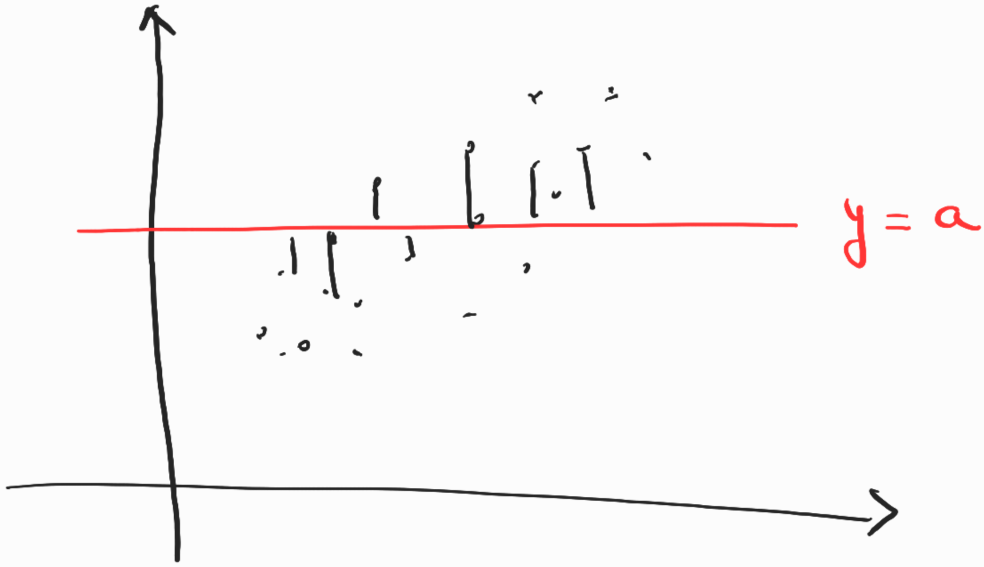
$$\sum_{i=1}^{n} (x_i - a)^2 \text{ es el promedio!}$$

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

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

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

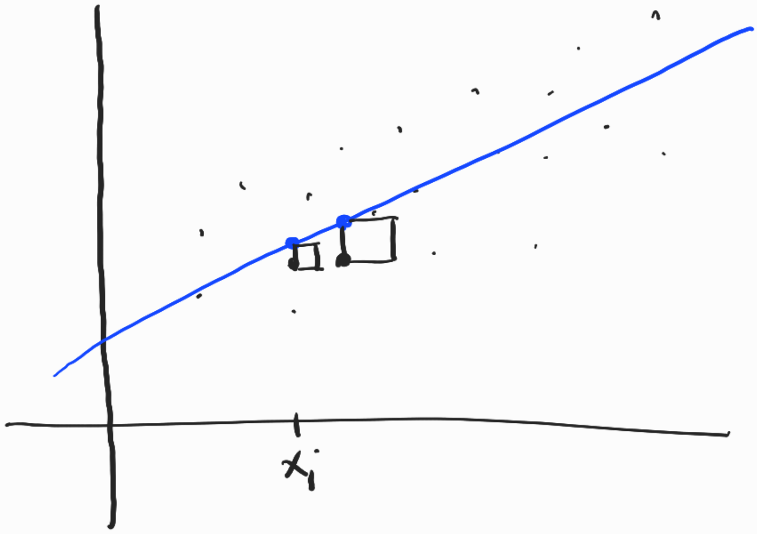
Busco la mejor constante por cuadrados mínimos



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

Ahora buscamos la mejor recta

$y = ax + b$



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

Puedo derivar y $\boxed{\text{usar A.L!}}$

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

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

- Una combinación lineal de una familia de funciones dadas

$\phi_0, \ldots, \phi_m$ que mejor aproxime mi datos. Es decir, se busca

$$\min_{a_0,\ldots,a_m} \sum \left(a_0\phi_0(x_i)+\cdots+a_m\phi_m(x_i)-y_i\right)^2$$

Volvamos al caso de la recta:

En un principio dados $(x_1,y_1) \ldots (x_n,y_n)$ datos sería ideal encontrar $L(x)=a_0+a_1x$ 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

$$a_0 + a_1 x_1 = y_1$$
$$a_0 + a_1 x_2 = y_2$$
$$\vdots$$
$$a_0 + a_1 x_n = y_n$$

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

Que normalmente no tendrá solución si n es más grande que 2.

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

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

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

que hace que $\| A \underline{X} - b \|_2$ es mínimo.

Es decir, buscamos $\underline{Y} \in \operatorname{Im} A$

$(\underline{Y} = A\underline{X}$ para algún $\underline{X})$

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

que esté a distancia mínima de b.

Supongamos

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

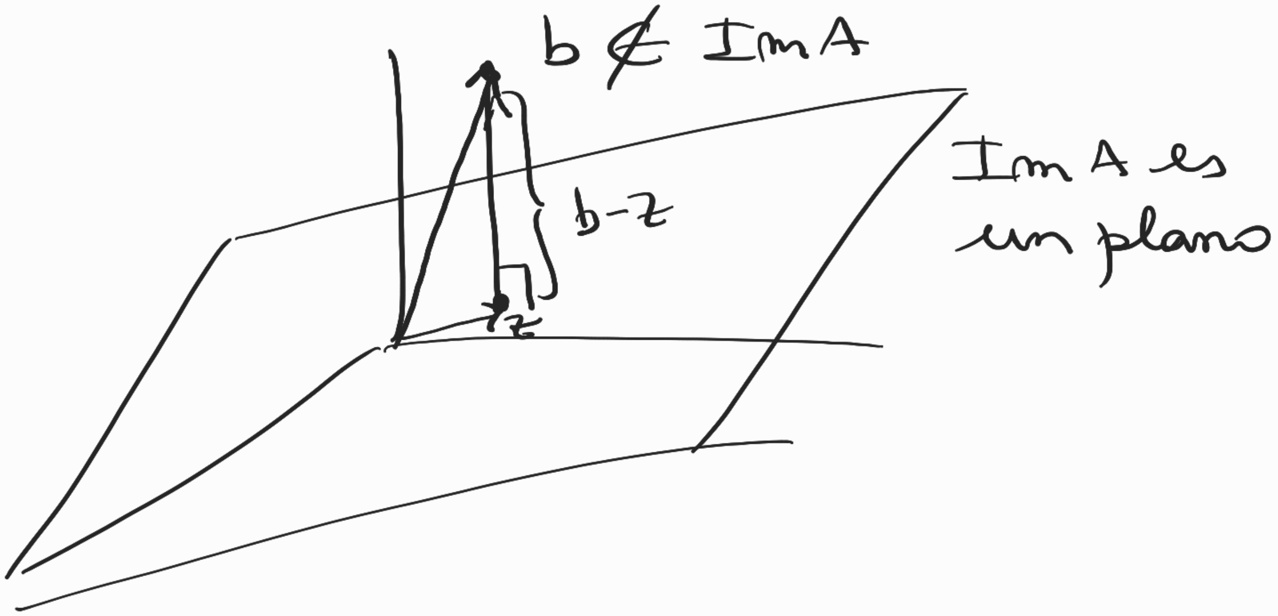
$$\operatorname{Im} A = \langle (1,1,1), (1,2,3) \rangle$$

Vemos que $\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} \notin \operatorname{Im} A$ con

lo cual no existe $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} : A\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = b$

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

sea mínima.



$z \in S = \operatorname{Im} A$ que tiene distancia

mínima a b sabemos que es

$z = P_S(b)$ que es aquel

que verifica que

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

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

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

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

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

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

$$= \text{Nu } A^t$$

luego $b - z \in \text{Nu } A^t$, 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 $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$:

$$A^t\left(b - A\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}\right) = \bar{0}$$

$$\Longleftrightarrow \quad A^t A \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = A^t b$$

El $\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$ que buscamos es soluc. de estas ecuaciones.

A estas ecuaciones se las llama $\boxed{\text{ecuaciones normales}}$.

En general se prueba que

<u>Teorema</u>: Dada $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$ 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 $A^t A \in \mathbb{R}^{m \times m}$.

<u>Dem</u>: Se basa en que

$P_S : \mathbb{R}^m \to S$ 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:

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

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

$$\Rightarrow 0 = (A^t A x)^t \cdot x = x^t A^t \cdot A x$$

$$= (Ax)^t \cdot Ax = \langle Ax, Ax \rangle$$

$$= \|Ax\|_2^2 \Rightarrow Ax = 0 \underset{\substack{A \text{ tiene} \\ \text{col. l.i}}}{\Rightarrow} x = 0$$

$\Rightarrow A^t A$ es inversible.

<u>OBS</u>: Si $A^t A$ es inversible $\Rightarrow$ $A^t A$ resulta simétrica y definida positiva $\Rightarrow$ podemos resolver el sistema lineal usando Cholesky.