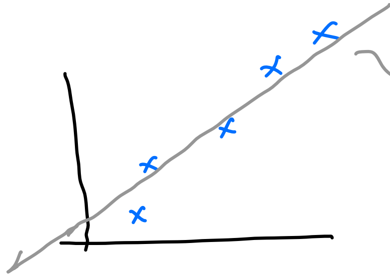
clase pasada



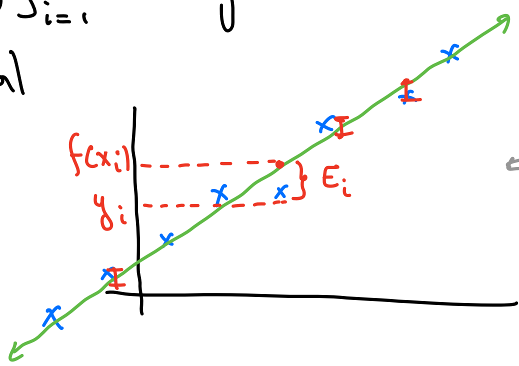
$(x_i, y_i) \to \boxed{y = \underset{=}{m}\,x + \underset{=}{b}}$

modelo

Reg. lineal es un metodo para encontrar los parametros m y b usando los datos (x_i, y_i)

Reg. Lineal (recordar)

Idea: sea $\{(x_i, y_i)\}_{i=1}^{n}$ un conjunto de datos t. q tiene un distribucion con tendencia lineal



Tendencia lineal (datos satisfacen)

$f(x_i) = m\,x_i + b$ ✓

para todo $i = 1, \dots, n.$ (supuesto)

$f(x)$: La recta que aproximadamente describe la dist. de los datos.

Error $\rightarrow$ $E_i = \left(y_i - f(x_i)\right)^2$ en el punto i

Sumando todos lo errores tenemos

Error total $\rightarrow$ $E_{Total} = \sum_{i=1}^{n} E_i = \sum_{i=1}^{n} \left(y_i - \underbrace{f(x_i)}_{\text{polinomio lineal}}\right)^2$

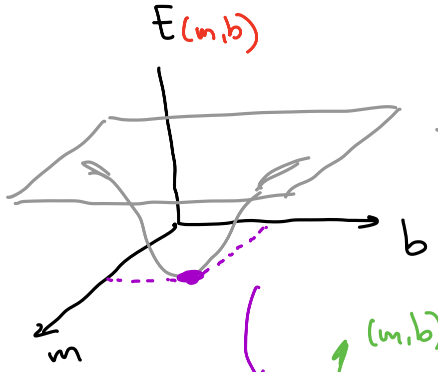
$\Rightarrow E_{Total} = \sum_{i=1}^{n} \left(\underset{=}{y_i} - \overbrace{\underset{?}{m}\,\underset{=}{x_i} - \underset{?}{b}}\right)^2$ ✓

son conocidos

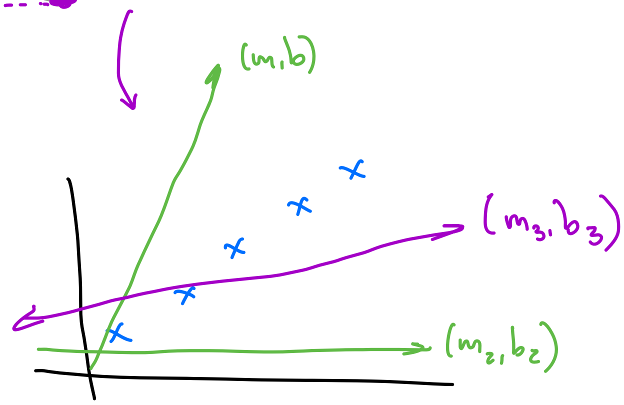
$$\boxed{E(m,b)} \rightarrow \text{depende de dos parametros}$$

$\Updownarrow$

se puede ver como función de 2 variables!



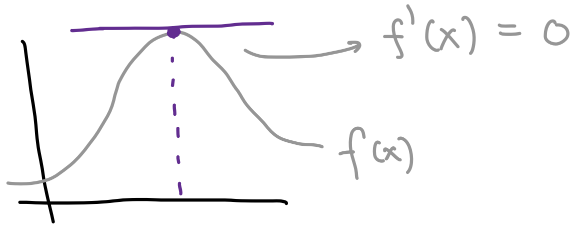
→ superficie para el c. datos que tiene (x_i, y_i)



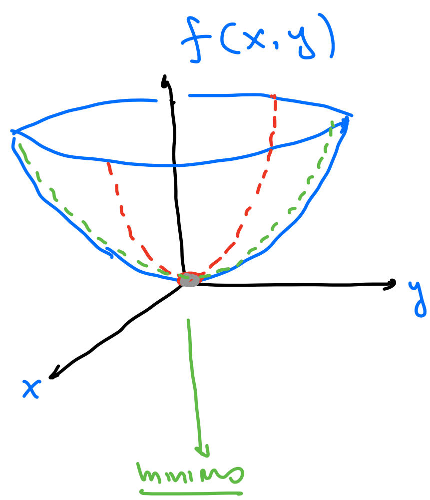
Idea clave: Para que valores de m y b, la función $E(m,b)$ alcanza el valor mínimo.

$\Downarrow$

¿ Como hago esto ?



$\Rightarrow$



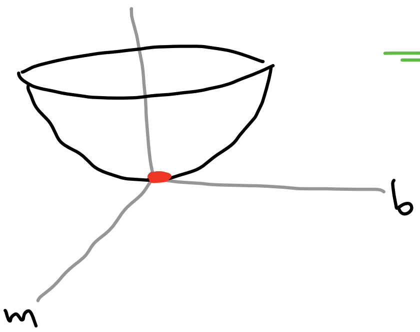
$$\frac{\partial f(x,y)}{\partial x} = 0 \quad \text{y} \quad \frac{\partial f(x,y)}{\partial y} = 0$$

y constante x constante

¿ Como uso esto para nuestro problema ?

En nuesto caso

$E(m,b)$



$\Longrightarrow$ El mínimo lo encontramos para

$$\underbrace{\frac{\partial E(m,b)}{\partial m} = 0}_{\text{me da 1 ec.}} \quad y \quad \underbrace{\frac{\partial E(m,b)}{\partial b} = 0}_{\text{me da 2 ec.}}$$

Ec 1:

$$\frac{\partial}{\partial m} E(m,b) = \frac{\partial}{\partial m}\left(\sum_{i=1}^{n} (y_i - m x_i - b)^2 \right)$$

$$= \sum_{i=1}^{n} \frac{\partial}{\partial m} (y_i - m x_i - b)^2 = \sum_{i=1}^{n} 2 (y_i - m x_i - b) \cdot (-x_i)$$

Regla de cadena

$$0 = \sum_{i=1}^{n} 2 (y_i - m x_i - b)(-x_i)$$

$$\underbrace{\sum_{i=1}^{n} y_i x_i}_{D} = m \underbrace{\sum_{i=1}^{n} x_i^2}_{E} + b \underbrace{\sum_{i=1}^{n} x_i}_{B}$$

$$\boxed{D = mE + bB} \quad (1)$$

Ec 2

$$\frac{\partial}{\partial b} E(m,b) = \frac{\partial}{\partial b}\left(\sum_{i=1}^{n} (y_i - m x_i - b)^2 \right)$$

$$\frac{\partial}{\partial b} E(m,b) = \sum_{i=1}^{n} 2 (y_i - m x_i - b)(-1)$$

$$0 = \sum_{i=1}^{n} (y_i - m x_i - b)$$

Nota:
(A, B, C, D, E)
Son numeros que vienen de los datos

$$\underbrace{\sum_{i=1}^{n} y_i}_{A} = m \underbrace{\sum_{i=1}^{n} x_i}_{B} + b \underbrace{\sum_{i=1}^{n} 1}_{C}$$

$$\boxed{A = m B + b \cdot C} \quad (2)$$

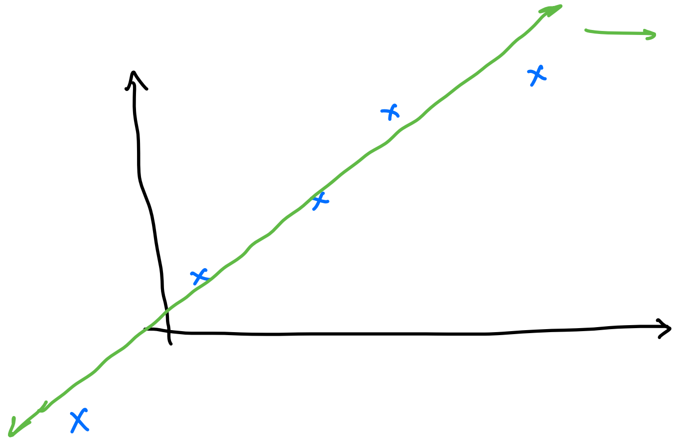
Usando alguno de los metodos de sustitucion, igualacion o reduccion se resuelve el sistema de (1) y (2) simultaneamente. lo que arroja:

$$m = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} (x_i)^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \quad (1)$$

$$b = \frac{\sum y_i \sum (x_i)^2 - (\sum x_i y_i)(\sum x_i)}{n \sum (x_i)^2 - (\sum x_i)^2} \quad (2)$$

Entonces para m y b encontrados, tenemos que

$$E = \sum_{i=1}^{n} (y_i - \overbrace{f(x_i)}^{m x_i + b})^2$$ es minimo!



$f(x) = mx + b$