- Regresión lineal. → Problemas de clasificación

→ Parto de un conjunto de datos, para encontrar el mejor modelo.

Aprendizaje Supervisado: problemas de regresión.

n° muestra

$(x^i, y^i) \longrightarrow x^{(i)} = (x^i_1, x^i_2, x^i_3 \ldots) \mid i = 1 \ldots m$

atributo

$y^i \equiv$ continuo

EJ: Precio Vivienda.

- Vamos a suponer que existe una relación lineal entre x e y.

→ Hipótesis: $h_w(x) = \overbrace{w_0}^{x_0 \cdot w_0} + x_1 \cdot w_1 + x_2 \cdot w_2 + \ldots x_n \cdot w_n$

↳ w = Vector de pesos

$$\boxed{h_w(x) = \sum_k x_k \cdot w_k}$$

- ¿Y cómo calculamos los pesos?

Error en el ejemplo i por $h_{\bar{w}}(\bar{x})$.

→ Esto en un punto ejemplo.

Definir una función $\underbrace{J(\vec{w}) = \frac{1}{2m} \cdot \sum_i \left(y^i \cdot h_w(x^i)\right)^2}$ para encontrar el error que comete $\bar{w}$.

"Función de Costo"

↳ para establecer un valor x y w que se acerque a lo esperado → Objetivo (que se acerque).

↳ Min/Max $J(\vec{w})$

"Coeficiente de Correlación"

Ésta se llama "error cuadrático medio" ECM

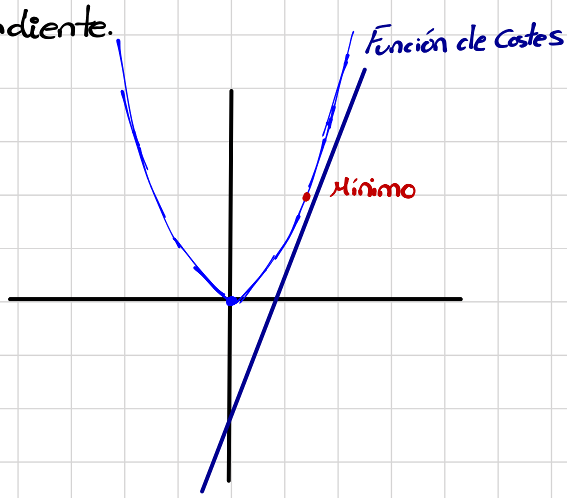
- Una vez tenemos nuestra ecuación de coste, es importante saber cuando nos acercamos a los valores con un mínimo de error.

Para ello, introducimos el término de gradiante descendiente.

Método del gradiante descendiente.

- Pretende encontrar un mínimo de la función.

$$\frac{dJ(\bar{w})}{d\bar{w}} = \left(\frac{\partial J(\bar{w})}{\partial w_0}, \ldots, \frac{\partial J(\bar{w})}{\partial w_n}\right)$$



Ejemplo

DataSet

x_i	y
1	1
3	2
4	5

$\alpha = 0{,}1 \to$ Valor alpha (valor cualquiera)

$w = (\phi, 0{,}25) \to$ Vector Inicial

w_0 w_1

1) Hipótesis inicial $\to \left[h_{\vec{w}}(\vec{x}) = \phi + 0{,}25 x_i \right]$

- Gradiante Decreciente $\to$

x_0	x_1	y	$\hat{y}$		error cometido $y - \hat{y}$
1	1	1	0,25	$\to h_{\vec{w}}(\vec{x}) = 0 \cdot x_0 + 0{,}25 \cdot 1$	0,75
1	3	2	0,75	$\to$ " $= 0 \cdot x_0 + 0{,}25 \cdot 3$ (0,75)	1,25
1	4	5	1	$\to$ " $= 0 \cdot x_0 + 0{,}25 \cdot 4$	4

$\delta_0 =$ error gradiante eje x_0

error gradiante eje $x_1 = \delta_1$

$(y - \hat{y}) \cdot x_0$	$(y - \hat{y}) \cdot x_1$
0,75	0,75
1,25	3,75
4	16

$\delta_0 = 6/3 = 2$

$\delta_1 = \frac{20{,}5}{3} = 6.83$

$w_0 \leftarrow w_0 + \alpha \cdot \delta_0 = \phi + 0{,}1 \cdot 2 = 0{,}2$

$w_1 \leftarrow w_1 + \alpha \cdot \delta_1 = 0{,}25 + 0{,}1 \cdot 6{,}83 = 0{,}933$

2) Nueva Hipótesis

$\hookrightarrow h_{\vec{w}}(\vec{x}) = 0{,}2 \cdot x_0 + 0{,}933 \cdot x_1$ ($x_0 \to 1$)

$\hookrightarrow$ (Volvemos a hacer lo mismo).