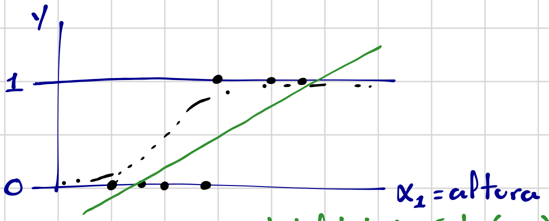
Aprendizaje Supervisado: Regresión logística (para problemas de clasificación, en nuestro caso, binaria). → Problemas de regresión



clase = ¿jugador baloncesto?

→ probabilidad de éxito ($y=1$).

$$h_{\vec{w}}(\vec{x}) = \underbrace{p[y=1/\vec{x}]}_{[0,1]} = \underbrace{\vec{x}\cdot\vec{w}}_{\mathbb{R}}$$

Por tanto, la idea sería transformar la regresión logística en regresión lineal.

↳ $\underbrace{[0,1] \rightarrow \mathbb{R}}$

Para ello, aplicamos una función, denominada logit → $\text{logit}\,[y=1/\vec{x}] = \vec{x}\cdot\vec{w}$

↳ $\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$

Por tanto, nuestra hipótesis quedaría como:

$$h_{\vec{w}}(\vec{x}) = p[y=1/\vec{x}] = \text{logit}^{-1}(v)$$

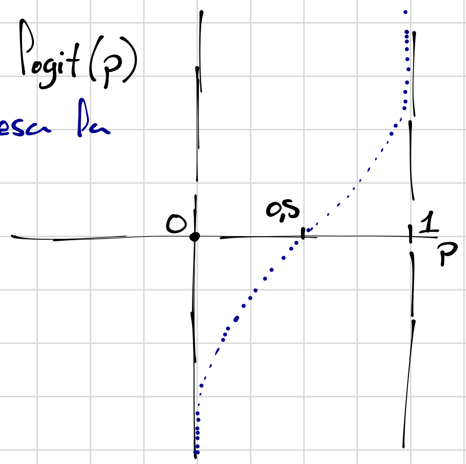
$$\text{logit}^{-1}(v) = O(v) = \frac{1}{1+e^{-v}}$$

logit(p)

Nos interesa la inversa

"Función logística"

Para obtener la probabilidad



- Dada la introducción:

$(\vec{x}_i, y_i) \quad i = 1 \ldots\ldots m$

$y_i \in \{\phi, 1\}$ ↳ pertenencia

- Modelo → $h_{\vec{w}}(\vec{x}) = \frac{1}{1+e^{\vec{w}\cdot\vec{x}}}$

↳ Elección de $\vec{w}$: Maximizar la verosimilitud de la hipótesis.

- Supongamos un ejemplo cualquiera (x_i, y_i)

$$p[y=y_i/\vec{x}_i] = O(\vec{w}\cdot\vec{x}_i)^{y_i} \cdot [1-O(\vec{w}\cdot\vec{x}_i)]^{(1-y_i)}$$

$$MAX\ V(\vec{w}) = \prod_i O(\vec{w}\cdot\vec{x}_i)^{y_i} \cdot (1-O(\vec{w}\cdot\vec{x}_i))^{(1-y_i)}$$

↳ Verosimilitud: probabilidad de que el modelo de esos datos. ??

$$\text{MIN } J(\vec{w}) = -\frac{1}{m}\cdot\sum y_i\cdot\log\left(\sigma(\vec{w}^T\cdot\vec{x}_i)\right) + (1-y_i)\cdot\log\left(1-\sigma(\vec{w}^T\cdot\vec{x}_i)\right)$$

→ Verosimilitud (likelyhood).

Conocida como
entropía cruzada binaria (proveniente de la "teoría de la información" de Shanon).

Al aplicar la derivada ahora con sumas, la ecuación resultante sería:

$$\frac{1}{m}\cdot\sum_i\left(y_i - h_{\vec{w}}(\vec{x}_i)\right)\cdot x_{i_k} \longrightarrow \text{Gradiante Decreciente}$$

Dataset

x	y
1	0
2	0
5	1

$\alpha = 0,5$ (tasa de probabilidad).

w inicial (peso) $= (\underset{w_0}{-2}, \underset{w_1}{1})$

1) $\hat{y} = h_{\vec{w}}(x) = 1/1+e^{-(-2x_0+x_1)}$

x_0	x_1	y	$\hat{y}$	$(y-\hat{y})$	$(y-\hat{y})\cdot x_0$	$(y-\hat{y})\cdot x_1$
1	1	0	0,269	-0269	-0,269	-0,269
1	2	0	0,5	-0,5	-0,5	-1
1	5	1	0,953	0,047	0,047	0,235
				Σ	-0,722	-1,034

$\delta_0 = \frac{-0,722}{3} = -0,241$ $\qquad w_0 \leftarrow w_0 + \alpha\delta_0 = -2+0,5(-0,241) = \underline{-2,12}$

$\delta_1 = \frac{-1,034}{3} = -0,345$ $\qquad w_1 \leftarrow w_1 + \alpha\cdot\delta_1 = 1+0,5(-0,345) = \underline{0,828}$

Por tanto, mi nueva $\hat{y}$ sería:

$$\hat{y} = h_{\vec{w}}(x) = \frac{1}{1+e^{-(-2,12\cdot x_0+0,828\cdot x_1)}}$$

↓

Volvemos a Iterar