Regresión Logística

x_1	x_2	x_m	y
x_{11}	x_{21}	x_{1m}	y_1
x_{n1}	x_{n2}	x_{nm}	y_n

$y_i \in \{0,1\}$

$\vec{x}_i \rightarrow y_i$: 1 o 0

Máxima Verosimilitud

$$\max_w \prod_{i=1}^{n} f_{y_i/x_i}(y_i/x_i)$$

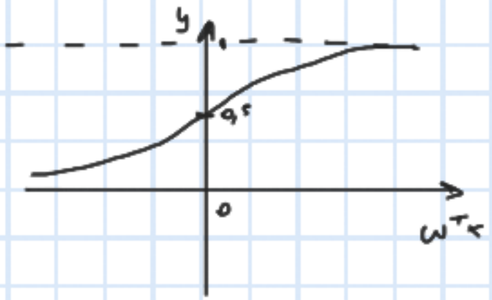
$y_i/x_i \rightarrow B(p_i)$

$B(p_i)$?

$$P_i = \frac{1}{1+e^{-w^T x_i}}$$

$$\sigma(w^T x) = \frac{1}{1+e^{-w^T x}}$$

$w^T x = w_0 x_0 + w_1 x_1 \dots$



$$P_w(y_i = 1 / \vec{x}_i = x_i) = \frac{1}{1+e^{-w^T x_i}}$$

$$P_w(y_i = 0 / \vec{x}_i = x_i) = 1 - \frac{1}{1+e^{-w^T x_i}}$$

Desarrollo

$$\max_w \prod_{i=1}^{n} P_w(y_i = y_i / \vec{x}_i = x_i)$$

$$B \rightarrow p_i^{y_i} (1-p_i)^{1-y_i}$$

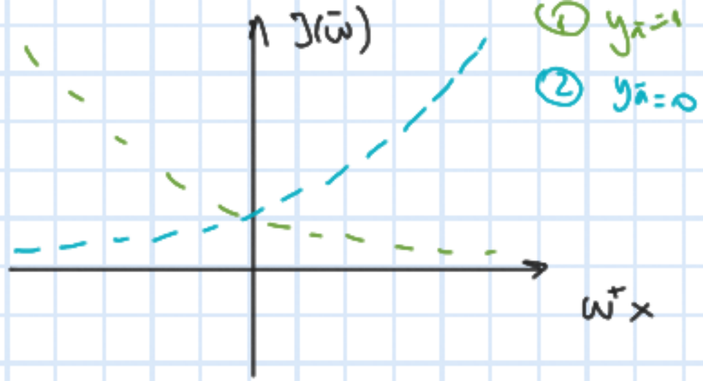
$$\max_w \sum_{i=1}^{n} \ln\left(P_w(y_i = y_i / \vec{x}_i = x_i)\right)$$

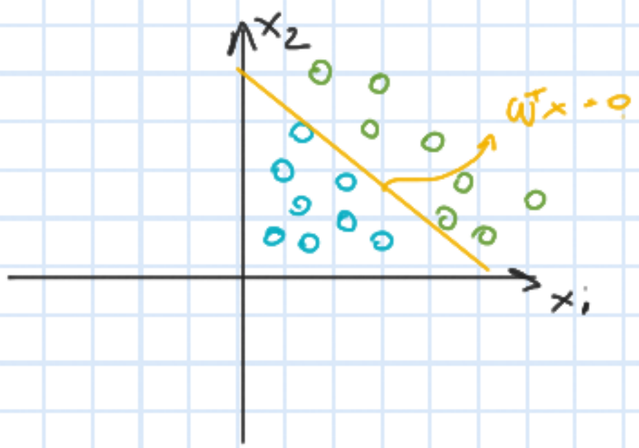
$$\max_w \sum_{i=1}^{n} \ln\left(\sigma(w^T x)^{y_i} (1-\sigma(w^T x))^{1-y_i}\right)$$

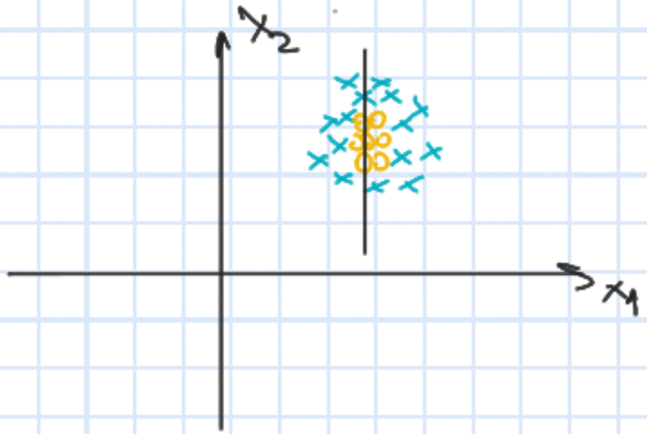
$$\max_w \sum_{i=1}^{n} \left[y_i \underbrace{\ln(\sigma(w^T x))}_{①} + (1-y_i) \underbrace{\ln(1-\sigma(w^T x))}_{②} \right]$$

$$\min_w \sum_{i=1}^{n} \left[-y_i ① - (1-y_i) ② \right]$$

$$J(w) = \underbrace{\frac{1}{n} \sum_{i=1}^{n} \left[-y_i ① - (1-y_i) ② \right]}_{\text{Binary Cross Entropy}}$$


J(w̄)
① yᵢ=1
② yᵢ=0
wᵀx


x₂
wᵀx = 0
x₁


x₂
x₁

Cálculo

$$\bar{\nabla}_w(J(w)) = \bar{0} \qquad \begin{pmatrix} \partial/\partial w_1 J(w) \\ \partial/\partial w_2 J(w) \\ \vdots \\ \partial/\partial w_m J(w) \end{pmatrix} = \bar{0}$$

$\bar{w}^T x_i = h$

$\nabla(h) = z$

$j = 1, \dots, m$

$$\frac{\partial}{\partial w_j}\left[-y_i \ln \underbrace{\left(\frac{1}{1+e^{-\overbrace{w^T x}^{h}}}\right)}_{z} - (1-y_i)\ln\left(1 - \frac{1}{1+e^{-w^T x}}\right)\right] = \frac{\partial}{\partial w_j} (A)$$

$$\frac{\partial A}{\partial w_j} = \frac{\partial A}{\partial z}\,\frac{\partial z}{\partial h}\,\frac{\partial h}{\partial w}$$

$$\frac{\partial A}{\partial z} = \frac{-y_i}{z} \cdot \frac{(1-y)}{1-z} = \frac{z - y_i}{z(1-z)}$$

$$\frac{\partial z}{\partial h} = z(1-z)$$

$$\frac{\partial h}{w} = x_i$$

$$\nabla(A_i) = \underbrace{(z_i - y_i)}_{\text{error}} x_i$$

$$= \left(\frac{1}{1+e^{-w^T x_i}} - y_i\right) x_i$$

$\Downarrow$

GD

SGD

MB

$$\bar{w} \leftarrow \bar{w} - \alpha \cdot \frac{1}{n} \underbrace{\sum_{i=1}^{n} \left(\frac{1}{1+e^{-w^T x_i}} - y_i\right) \bar{x}_i}_{\nabla J(w)}$$

(1) GD, SGD, MB optimizan $\bar{w}$

(2) Predicciones $\hat{y}$ → $p \geq 0,5 : 1$

$p < 0,5 : 0$

(3) Métricas

Matriz Confusión

PREDICCIÓN	VALORES VERDADEROS: VALOR T	VALORES VERDADEROS: VALOR F
PT	TRUE POSITIVE	FALSE POSITIVO
PF	FALSO NEGATIVO	TRUE NEGATIVO

$$A = \frac{TP + TN}{FP + FN + TP + TN}$$

$$\frac{TP}{TP + FP} = P$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

MSE

- ROC
- AUC
- ↓
- DL