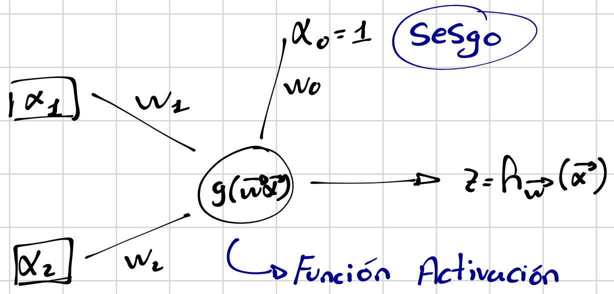
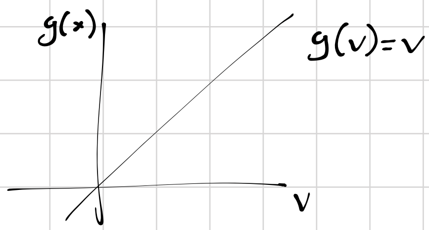


Variables entradas

a) Función: Activación Lineal

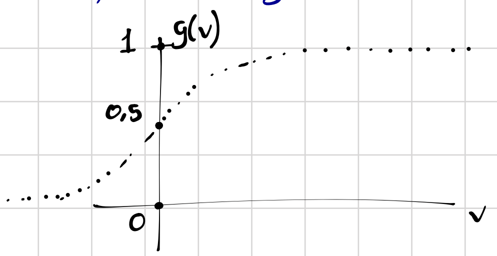


$$\underbrace{\sum_{i=0}^{2} w_i \cdot x_i}$$

$$g\left(\sum_{i}^{n} w_i \cdot x_i\right) \Rightarrow h_{\vec{w}}(\vec{x}) =$$

$$= w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_i x_i$$

b) F. Act.: logística



$$\underbrace{g(v) = \frac{1}{1 + e^{-v}}}$$

$$h_{\vec{w}}(\vec{x}) = \frac{1}{1 + e^{-\underbrace{(w_0 + w_1 \cdot x_1 + w_2 \cdot x_2)}_{=0 \rightarrow \text{Frontera de decisión}}}}$$

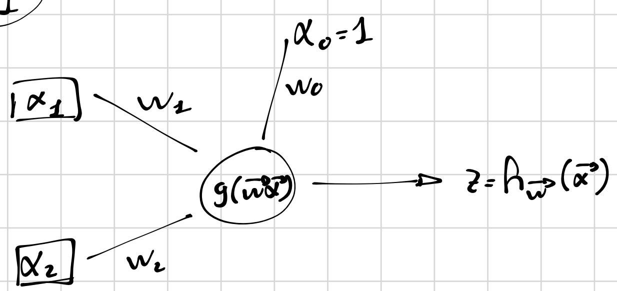
probabilidad de que x pertenezca a la clase (problemas de clasificación)

Frontera de decisión

$\underbrace{h_{\vec{w}}(\vec{x}) \geq 0,5}$ Pertenece a la clase

$\underbrace{h_{\vec{w}}(\vec{x}) < 0,5}$ No pertenece a la clase

1



$g(v) = v$

$\vec{w} = (1, -1, 0{,}5)$ — w_0, w_1, w_2

x_1	x_2	Y
1	1	1
2	2	3
3	5	4

a) Salida de la red y el ECM (error cuadrático medio).

x_1	x_2	Y	z	(Y−z)	$(Y-z)^2$
1	1	1	0,5	0,5	0,25
2	2	3	0	3	9
3	5	4	0,5	3,5	12,25

$\Sigma = 21.5 \rightarrow ECM = \frac{21{,}5}{3} = 7{,}1\hat{6}$

b) Aplicar el algoritmo del gradiante $(\alpha = 0,1)$

· Función de pérdida → ECM

$(y-z)\cdot x_0$	$(y-z)\cdot x_1$	$(y-z)\cdot x_2$
0,5	0,5	0,5
3	6	6
3,5	5	17,5
$\underbrace{}_{7}$	$\underbrace{}_{17}$	$\underbrace{}_{24}$

$d_0 = 7/3$
$d_1 = 17/3$
$d_2 = 24/3$

$w_0 \leftarrow w_0 + \alpha \cdot d_0 = 1,24$
$w_1 \leftarrow w_1 + \alpha \cdot d_1 = -0,44$
$w_2 \leftarrow w_2 + \alpha \cdot d_2 = 1,3$

Nueva hipótesis: $h_{\vec{w}}(\vec{x}^0) = 1,24 - 0,44 x_1 + 1,3 \cdot x_2$

c) Calcular ECM de nuevo

x_1	x_2	y	z	$(y-z)$	$(y-z)^2$
1	1	1	2,1	−1,1	1,21
1	2	3	2,96	0,04	0,0016
3	5	4	6,42	−2,42	5,88

$\Sigma = 7,0916 \rightarrow ECM = \frac{7,0916}{3} = 2,36$

Efectivamente se ha reducido

(3) $g(v) = \frac{1}{1+e^{-v}} = \sigma(v)$

$\vec{w} = (\underset{w_0}{-1}, \underset{w_1}{1,5}, \underset{w_2}{-1})$

x_1	x_2	y
2	1	0
4	2	0
4	4	1

a) Salida de la red y entropía cruzada binaria.

Ejemplo hecho en logaritmo neperiano.
→ Se puede cualquier log. dependiendo de la unidad de medida.
$\log_2 \ldots \rightarrow$ bits

$$ECB = -\frac{1}{m} \sum_i \left(y_i \cdot \log \overset{z}{\sigma(\vec{w} \cdot \vec{x}_i^0)} + (1-y_i) \cdot \log\left(1 - \sigma(\vec{w} \cdot \vec{x}_i^0)\right) \right)$$

valores dataset

x_1	x_2	y	$\vec{w} \cdot \vec{x}$	z	$y \cdot \log(z)$	$(1-y)\cdot \log(1-z)$
2	1	0	1	0,73	0	−1,31
4	2	0	3	0,95	0	−2,99
4	4	1	1	0,73	−0,31	0

$\frac{1}{1+e^{-v}} = \frac{1}{1+e^{-1}} = 0,73$

$v = -1 \cdot 1 + 2 \cdot 1.5 + 1 \cdot (-1) = -1+3-1 = 1$

$-0,31 + -4,3$

$\Sigma = -4,61 \rightarrow -\left(\frac{-4,61}{3}\right) = 1,53$

→ Verosimilitud

$e^{-4,61} = 0,0099$

b) Aplicar algoritmo del gradiente. $(\alpha = 0,1)$

- Función de pérdida = ECB

x_1	x_2	y	z	$(y-z)$	$(y-z)^2$	$(y-z)\cdot x_0$	$(y-z)\cdot x_1$	$(y-z)\cdot x_2$
1	1	1	2,1	−1,1	1,21	−0,73	−1,46	−0,73
1	2	3	2,96	0,04	0,0016	−0,95	−3,8	−1,9
3	5	4	6,42	−2,42	5,88	0,27	1,08	1,08
						−1,41	−4,18	−1,55

$d_0 = \frac{-1,41}{3} = -0,47$ $\qquad w_0 \leftarrow w_0 + \alpha \cdot d_0 = -1,047$

$d_1 = \frac{-4,18}{3} = -1,47$ $\qquad w_1 \leftarrow w_1 + \alpha \cdot d_1 = 1,361$

$d_2 = \frac{-1,55}{3} = -0,51$ $\qquad w_2 \leftarrow w_2 \cdot \alpha \cdot d_2 = -1,051$

c) (volver a repetir).

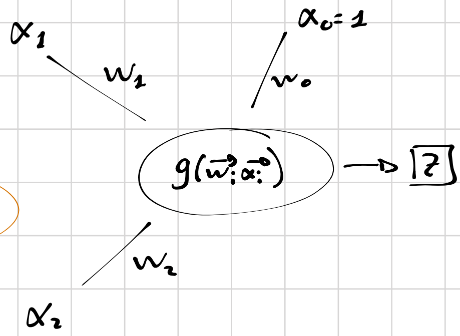
④ Calcular Error Cuadrático Medio / F. Pérdida → "ECM" → Aplicar gradiente $(\alpha = 0,1)$.

Salida Red $\qquad w = (w_0, w_1, w_2) = (1, 1, 1)$

x_1	x_2	y	z	$(y-z)$	$(y-z)^2$
0	1	1	2	−1	1
1	2	2	4	−2	4
1	2	3	4	−1	1

$\boxed{\Sigma = 6} \rightarrow ECM = \frac{6}{3} = 2$

Error Cuadrático Medio



Función de Activación: lineal

$$\boxed{g\left(\sum_{i=0}^{n} w_i \cdot x_i\right) \Rightarrow h_{\vec{w}}(\vec{x}) = w_0 + w_1\cdot x_1 + w_2\cdot x_2 + \ldots + w_i x_i}$$

- $(1\cdot 1) + (0\cdot 1) + (1\cdot 1) = 2$
- $(1\cdot 1) + (1\cdot 1) + (2\cdot 1) = 4$
- $(1\cdot 1) + (1\cdot 1) + (2\cdot 1) = 4$

$(y-z)\cdot x_0$	$(y-z)\cdot x_1$	$(y-z)\cdot x_2$
−1	0	−1
−2	−2	−4
−1	−1	−2
−4	−3	−7

Función Pérdida = "ECM" $(\alpha = 0,1)$

$d_0 = -4/3 \qquad w_0 \leftarrow w_0 + \alpha_0 \cdot d_0 = \boxed{0,8\hat{6}}$

$d_1 = -3/3 \qquad w_1 \leftarrow w_1 + \alpha_1 \cdot d_1 = \boxed{0,9}$

$d_2 = -7/3 \qquad w_2 \leftarrow w_2 + \alpha_2 \cdot d_2 = \boxed{0,7\hat{6}}$

Nueva Hipótesis $= h_{\vec{w}}(\vec{x}) = 0,86 + 0,9\cdot x_1 + 0,76\cdot x_2$

⑤ $W = (w_0, w_1, w_2) = (1, -1, 0.5)$

Entropía Cruzada Binaria

$\hookrightarrow -\frac{1}{m} \cdot \sum_i \left(y_i \cdot \log\left(\underbrace{\vec{w}_i \cdot \vec{x}_i}_{z}\right) + (1-y_i) \cdot \left(\log\left(1 - (\vec{w}_i \cdot \vec{x}_i)\right)\right)\right)$

neperiano

	x_1	x_2	y	$w_i \cdot x_i$	z	$y \cdot \log z$	$(1-y) \cdot \log(1-z)$
a	2	1	0	-0,5	0,37	0	-0,46
b	3	3	0	-0,5	0,37	0	-0,46
c	-1	2	1	3	0,95	-0,051	0

$-0{,}051 + (-0{,}92) = \left\{\sum = -0{,}971\right\} \rightarrow \underbrace{-\left(\frac{-0{,}971}{3}\right)}_{0{,}3236}$

Verosimilitud $= e^{(EBC)} = e^{0{,}3236} = \underline{1{,}38}$

· Hipótesis Inicial: $h_{\vec{w}}(\vec{x}) = 1 \cdot x_0 - 1 \cdot x_1 + 0{,}5 \cdot x_2$

$a = w_0 x_0 + w_1 x_1 + w_2 x_2 = 1 \cdot 1 + (-1 \cdot 2) + (0{,}5 \cdot 1) = 1 - 2 + 0{,}5 = -0{,}5$

$b = 1 \cdot 1 + (-1 \cdot 3) + (0{,}5 \cdot 3) = 1 - 3 + 1{,}5 = -0{,}5$

$c = 1 \cdot 1 + (-1 \quad -1)) + (0{,}5 \cdot 2) = 1 + 1 + 1 = 3$

· Como es un caso de entropía cruzada binaria, su función de activación será logística.

$$g(v) = \frac{1}{1 + e^{-v}} \qquad \left[v = \vec{w}_i \cdot \vec{x}_i \right]$$

$$g(a) = \frac{1}{1 + e^{-(-0{,}5)}} = \underline{0{,}37}$$

$$g(b) = \frac{1}{1 + e^{-(-0{,}5)}} = \underline{0{,}37}$$

$$g(c) = \frac{1}{1 + e^{-(3)}} = \underline{0{,}95}$$

⑥ Iteración del algoritmo del gradiente $(\alpha = 0{,}1) \rightarrow$ Función Pérdida: Ent. Cruz. Binaria.

	x_1	x_2	y	z	$(y-z)$	$(y-z) \cdot x_0$	$(y-z) \cdot x_1$	$(y-z) \cdot x_2$
a	2	1	0	0,37	-0,37	-0,37	-0,74	-0,37
b	3	3	0	0,37	-0,37	-0,37	-1,11	-1,11
c	-1	2	1	0,95	0,05	0,05	-0,05	0,1
						-0,69	(-1,9)	(-1,38)

Nueva Hipótesis Inicial

$h_{\vec{w}}(\vec{x}) = 0{,}977 \cdot x_0 - 1{,}063 \cdot x_1 + 0{,}454 \cdot x_2$

$\delta_0 = -0{,}69/3 = -0{,}23$

$\delta_1 = -1{,}9/3 = -0{,}63$

$\delta_2 = -1{,}38/3 = -0{,}46$

$w_0 \leftarrow w_0 + \alpha \cdot \delta_0 = 1 + 0{,}1 \cdot (-0{,}23) = 0{,}977$

$w_1 \leftarrow w_1 + \alpha \cdot \delta_1 = -1 + 0{,}1 \cdot (-0{,}63) = -1{,}063$

$w_2 \leftarrow w_2 + \alpha \cdot \delta_2 = 0{,}5 + 0{,}1 \cdot (-0{,}46) = 0{,}454$