Modelo $P(y=1|x) = \sigma(z)$

Con $z = w_0 + w_1 x_1 + w_2 x_2$

y $\sigma(t) = \dfrac{1}{1+e^{-t}}$

$$\mathcal{L}(w) = \dfrac{1}{m} \sum_{i=1}^{m} -y_i \, Ln(p_i) - (1-p_i) \ln(1-p_i) + \dfrac{\lambda}{2m}(w_1^2 + w_2^2)$$

x_1 x_2

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 2 \\ 2 & 1 \\ 0 & 2 \\ 2 & 2 \end{pmatrix} \qquad y = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \qquad m=6 \qquad \text{pero } X = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

$$w_0 = \begin{pmatrix} w_0^0 \\ w_1^0 \\ w_2^0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \lambda = 1 \quad \eta = 1$$

gradiente a calcular

$$grad = \begin{pmatrix} \dfrac{d\mathcal{L}}{dw_0} \\[4pt] \dfrac{d\mathcal{L}}{dw_1} \\[4pt] \dfrac{d\mathcal{L}}{dw_2} \end{pmatrix} \qquad \text{ó} \quad grad = X^{\top}(\sigma(z) - y)$$

pero $z = Xw$

$$\Rightarrow \boxed{grad = X^{\top}(\sigma(Xw) - y)}$$

para cada iteración

de modo que calculamos y reemplazamos.

i	x_{1i}	x_{2i}	y_i	$z_i = w_0 + w_1 x_{1i} + w_2 y_{2i}$	$\hat{p}_i = \delta(z_i)$
	0	1	0	0	0.50
	1	0	0	1	0.731
	1	2	1	1	0.731
	2	1	1	2	0.88
	0	2	1	0	0.50
	2	2	0	2	0.88

$$\text{grad} = X^t \, w_0 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 2 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1.223 \\ 1.985 \\ 0.605 \end{pmatrix}$$

$$\frac{d\mathcal{L}}{d w_0} = 1.223$$

$$\frac{d\mathcal{L}}{w_1} = 1.985$$

$$\frac{d\mathcal{L}}{w_2} = 0.605$$

$m = 6$

$$w_0^1 = 0 - 1 \times \frac{1.223}{6} = -0.2039$$

$$w_1^1 = 1 - \frac{1 \times 1.985}{6} = 0.5025$$

$$w_2^1 = 0 - 1 \times 0.605 = -0.10175$$

```
iteracion

 i1
X
[[1 0 1]
 [1 1 0]
 [1 1 2]
 [1 2 1]
 [1 0 2]
 [1 2 2]]
X.T
[[1 1 1 1 1 1]
 [0 1 1 2 0 2]
 [1 0 2 1 2 2]]
theta
[0 1 0]
y
[0 0 1 1 1 0]
z:
[0 1 1 2 0 2]
h_theta, p:
[0.5        0.73105858 0.73105858 0.88079708 0.5        0.88079708]
grad:
[1.22371131 1.98530547 0.60450839]
theta = theta - alpha * grad:
[-0.20395189  0.50244909 -0.1007514 ]
iteracion

 i2
X
[[1 0 1]
 [1 1 0]
 [1 1 2]
 [1 2 1]
 [1 0 2]
 [1 2 2]]
X.T
[[1 1 1 1 1 1]
 [0 1 1 2 0 2]
 [1 0 2 1 2 2]]
theta
[-0.20395189  0.50244909 -0.1007514 ]
y
[0 0 1 1 1 0]
z:
[-0.30470328  0.2984972   0.09699441  0.70019489 -0.40545468  0.59944349]
h_theta, p:
[0.42440813 0.5740751  0.52422961 0.66823098 0.4000025  0.64552898]
grad:
[ 0.2364753   0.72582463 -0.76783871]
theta = theta - alpha * grad:
[-0.20937246  0.2977368   0.04401362]
```

|: array([-0.20937246, 0.2977368 , 0.044013621])