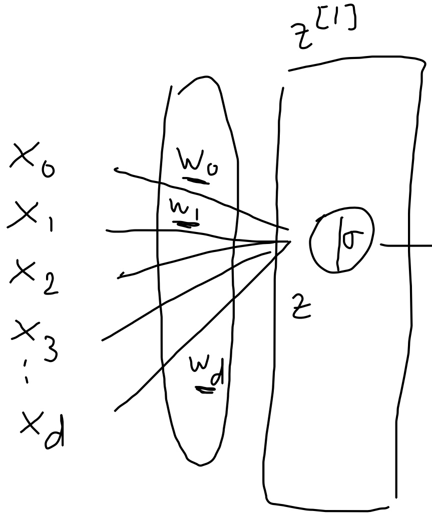
$z = w^T x$

$\frac{\partial z}{\partial w_j} = x_j$



$f^{[1]}$

$z = w^T x$

(x, y) $y \in \{0, 1\}$

$\rightarrow \sigma(z) = \underbrace{\sigma(w^T x)}$ $\rightarrow$ $\ell(f, y)$

$\boxed{f = \sigma(z)}$ $\quad f \in \mathbb{R}^1$

$\frac{\partial f}{\partial z} = \sigma'(z)$

$= \boxed{-y \log f - (1-y)\log(1-f)}$

$\boxed{\frac{\partial \ell}{\partial w_j}}$

$\left(\frac{\partial \ell}{\partial w_0}, \frac{\partial \ell}{\partial w_1}, \frac{\partial \ell}{\partial w_2}, \dots, \frac{\partial \ell}{\partial w_d}\right)$ = Gradiente $= \nabla_w \ell$

$\frac{\partial \ell}{\partial w_j} = \underbrace{\frac{\partial \ell}{\partial f}} \cdot \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial w_j}$ $\quad f(1-f)$

$= \left(-\frac{y}{f} + (1-y)\frac{1}{1-f}\right) \sigma'(z)\, x_j = \left(-y(1-\cancel{f}) + (1-\cancel{y})f\right) x_j = (f-y)x_j$

$= \boxed{(f-y)\cdot x}$

$\underbrace{(f-y)}_{\in \mathbb{R}} \cdot \underbrace{x}_{\in \mathbb{R}^{d+1}}$

Discesa del gradiente

$w \leftarrow w - \eta\, \nabla_w \ell$