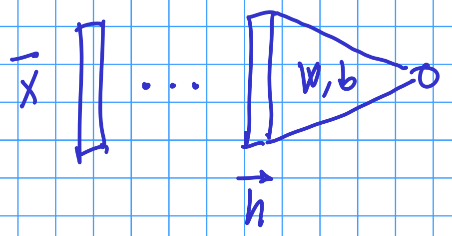
CLASIFICACIÓN 2 CLASES:

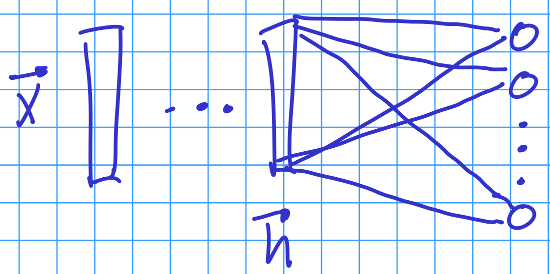


$z = W \cdot \vec{h} + b$, $y = \sigma(z) = \text{PROB}(t = 1 / x)$

LOSS $\rightarrow$ CROSS ENTROPY ...

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

K CLASES:



$$\left.\begin{aligned} z_1 &= W_1 \vec{h} + b_1 \\ z_2 &= W_2 \vec{h} + b_2 \\ &\vdots \\ z_k &= W_k \vec{h} + b_k \end{aligned}\right\}$$

$$\downarrow \qquad \downarrow \qquad \downarrow$$

$$\vec{z} \qquad W \qquad \vec{b}$$

$\vec{z} = W \vec{h} + \vec{b}$ $\qquad$ $\vec{y} = \text{softmax}(\vec{z})$

$$y_1 = \frac{e^{z_1}}{e^{z_1} + e^{z_2} + \cdots + e^{z_k}}$$

$$y_2 = \frac{e^{z_2}}{e^{z_1} + e^{z_2} + \cdots + e^{z_k}}$$

$$y_j = \frac{e^{z_j}}{\sum_{j=1}^{k} e^{z_j}}$$

$y_j \rightarrow \text{PROB}(t = t_j | \vec{x})$

LOSS → CROSS-ENTROPY:

MAX. prob. de las clases reales.

N puntos → $\ell = \prod_{i=1}^{N} P(t = t_i \mid \vec{x}_i)$

$$L = \log \ell = \sum_{i=1}^{N} \log P(t = t_i \mid \vec{x}_i) = \sum_{i=1}^{N} \log \begin{cases} y_1 & \text{si } t_i = 1 \\ y_2 & \text{si } t_i = 2 \\ \vdots & \\ y_k & \text{si } t_i = k \end{cases} =$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{K} t_{ij} \cdot \log y_{ij}$$

one-hot

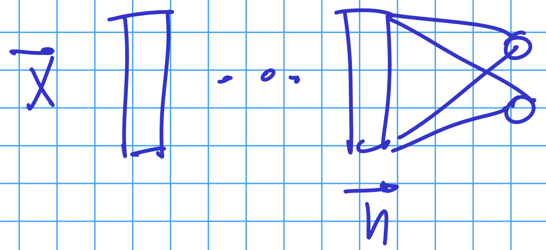
EJEMPLO:

$\vec{x}$, $t = 3$, $k = 4$ clases

$$t = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \qquad y = \begin{bmatrix} 0.1 \\ 0.05 \\ 0.8 \\ 0.05 \end{bmatrix}$$

$\log 0.8$

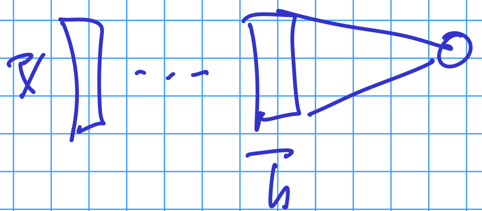
¿SOFTMAX CON 2 CLASES?



$$z_1 \rightarrow y_1 = \frac{e^{z_1}}{e^{z_1}+e^{z_2}}$$

$$z_2 \rightarrow y_2 = \frac{e^{z_2}}{e^{z_1}+e^{z_2}}$$

SIGMOIDE CON 1 NEURONA SALIDA



$$z_1 \rightarrow y_1 = \frac{e^{z_1}}{e^{z_1}+1} = \frac{e^{z_1}}{e^{z_1}+e^{z_2}}$$

$$z_2 = 0$$

$$y_2 = \frac{e^{z_2}}{e^{z_1}+e^{z_2}} = \frac{1}{e^{z_1}+1} = 1-y_1$$