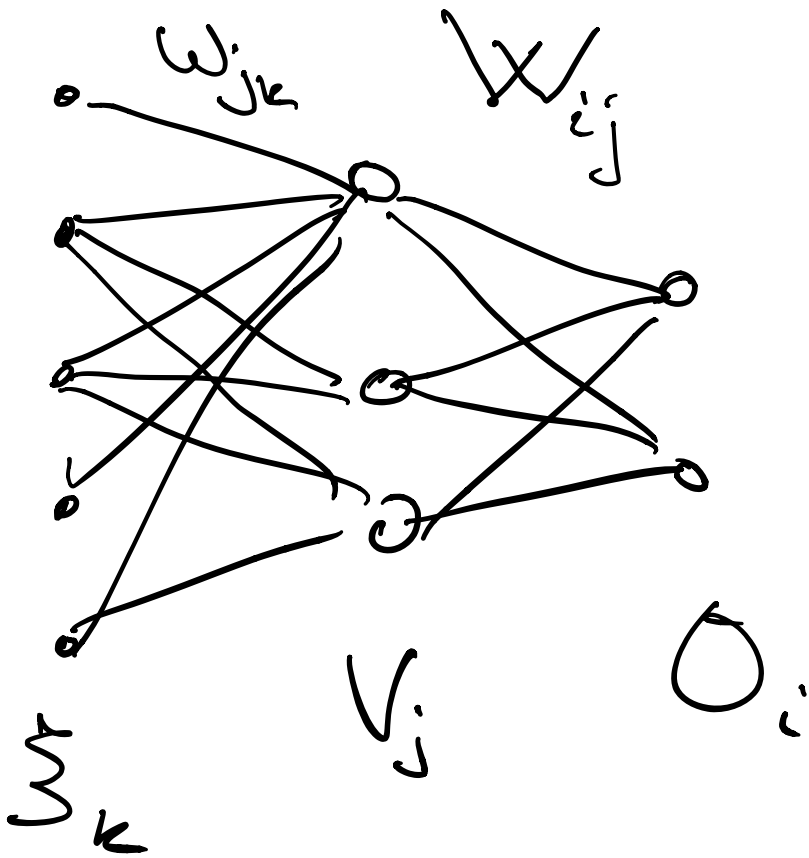


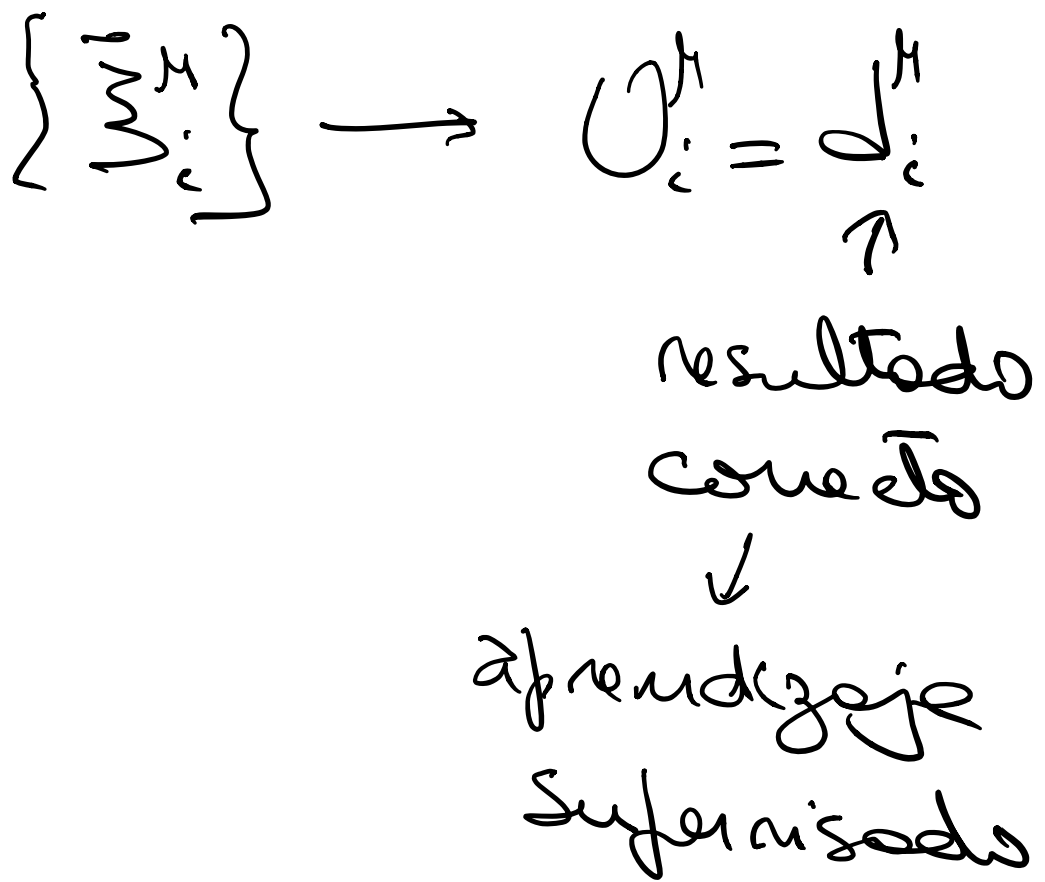
11/oct/18

Repaso

Perceptron con una capa oculta



Tenemos un conjunto de asociaciones correctas



De los resultados correctos
usaré algunos para entrenar
y otros para testear

1º paso

en la capa de entrada
presento un ejemplo
conocido $\left\{ \sum_{k=1}^M \right\} M=1$

Actualizar los neurones
de la capa oculta

La j hace lo siguiente

calcula

$$h_j^H = \sum_k w_{jk} \sum_k^H$$

y luego

$$V_j^H = g(h_j^H)$$

$$= g\left(\sum_k w_{jk} \sum_k^H\right)$$

• Conociendo todos los V_j
ahora calcula los

$$\begin{aligned}
 O_i^{\mu} &= g(h_i^{\mu}) \\
 &= g\left(\sum_j W_{ij} U_j\right) \\
 &= g\left(\sum_j W_{ij} g\left(\sum_k W_{jk} \sum_{\mu} U_k^{\mu}\right)\right)
 \end{aligned}$$

Tengo todas las medidas
 O_i^{μ}

• Calcular el error

$$\begin{aligned}
 E(w, W) &= \frac{1}{2} \sum_{\mu} \sum_i [d_i^{\mu} - O_i^{\mu}]^2 \\
 &= \frac{1}{2} \sum_{\mu} \sum_i \left[d_i^{\mu} - g\left(\sum_j W_{ij} g\left(\sum_k W_{jk} \sum_{\mu} U_k^{\mu}\right)\right) \right]^2
 \end{aligned}$$

• Modificar los W_{ij} cuando

"descenso por el gradiente"

$$\Delta W_{ij} = -\eta \frac{\partial E}{\partial W_{ij}} \quad \checkmark \quad \checkmark$$
$$= \eta \sum_{\mu} \delta_i^{\mu} V_j^{\mu}$$

donde $\delta_i^{\mu} = g'(h_i^{\mu}) [d_i^{\mu} - O_i^{\mu}]$

• Modificar los w_{jk}

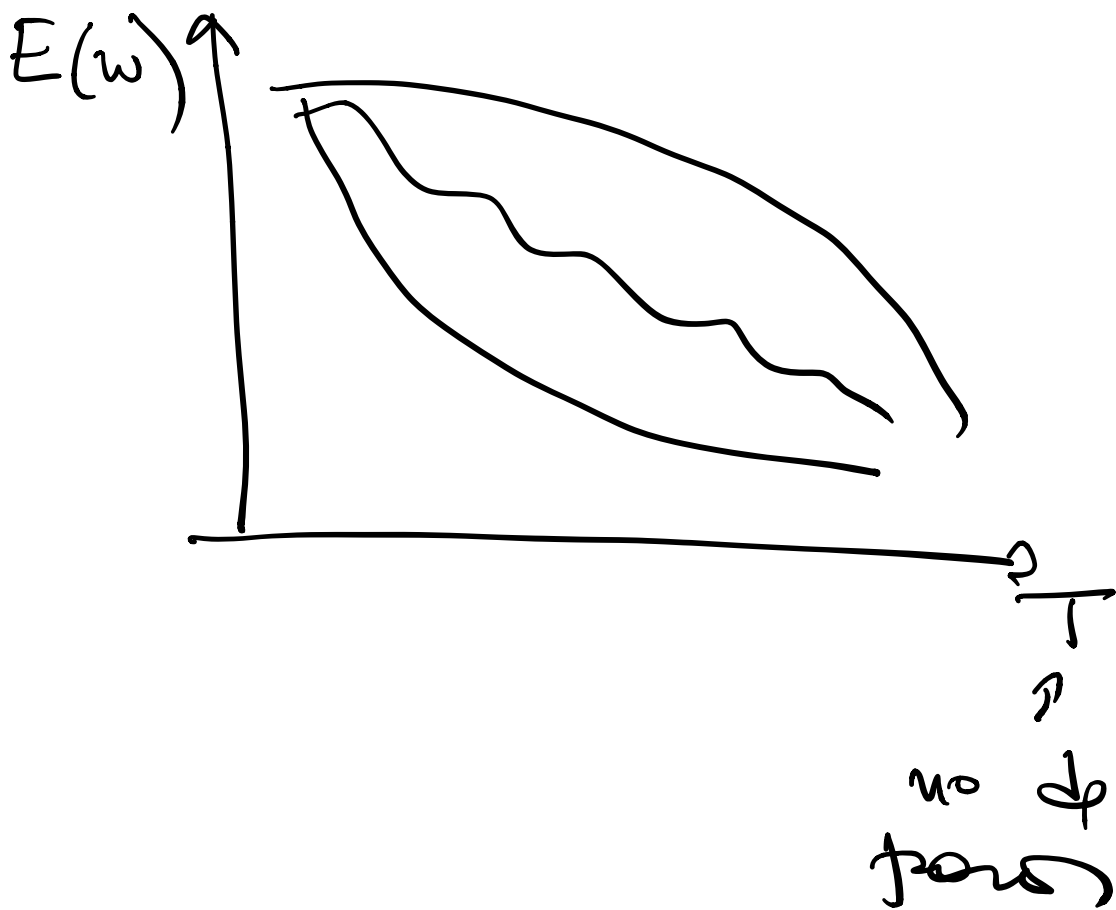
$$\Delta w_{jk} = -\eta \frac{\partial E}{\partial w_{jk}} = -\eta \sum_{\mu} \frac{\partial E}{\partial V_j^{\mu}} \frac{\partial V_j^{\mu}}{\partial w_{jk}}$$

$$= \eta \sum_{\mu} \delta_j^{\mu} \sum_k^{\mu}$$

donde

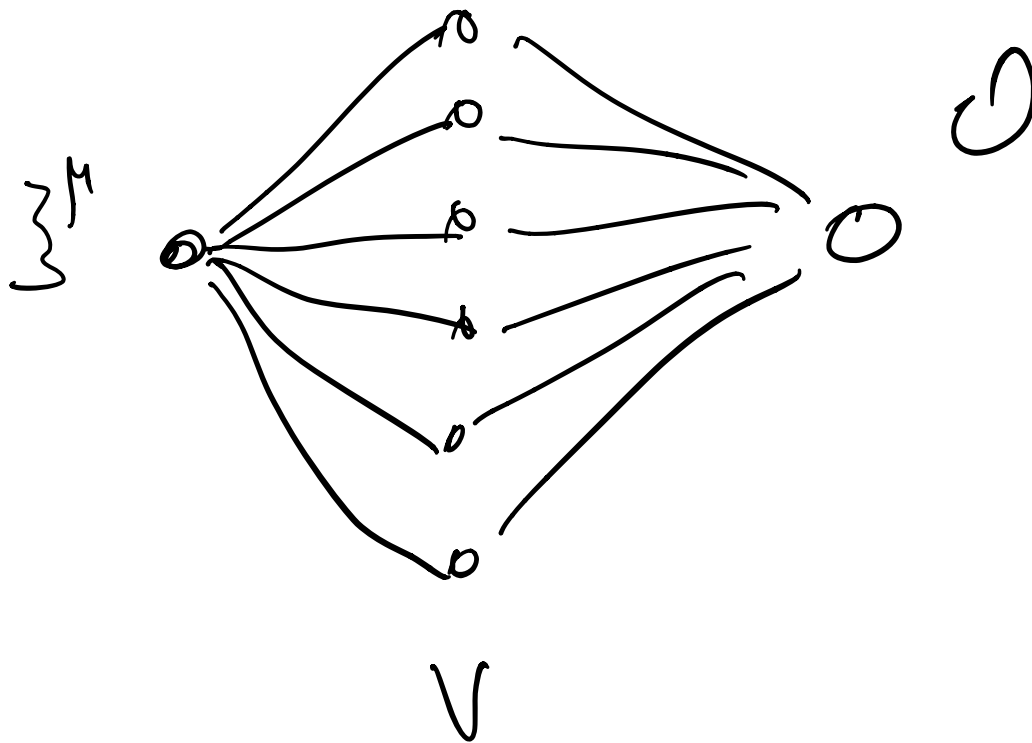
$$\delta_j^{\mu} = g'(h_j^{\mu}) \sum_i W_{ij} \delta_i$$

Una vez \hat{g} mostre todos los ejemplos del conjunto de entrenamiento, amplíezlo nuevamente

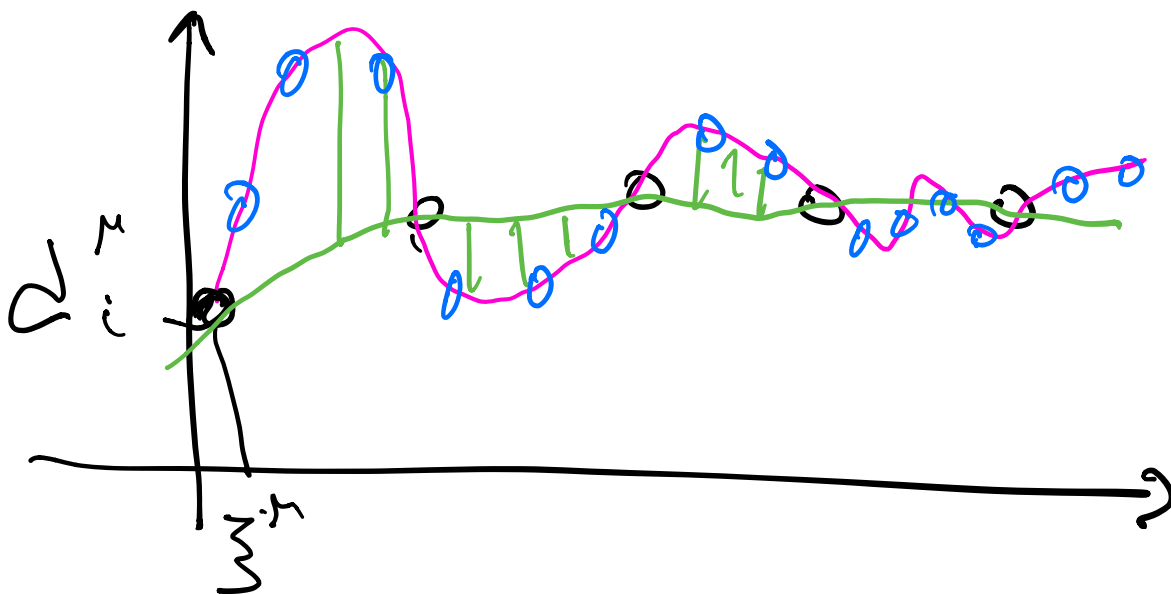


1 para: mostre todos los ejemplos

Ejemplo

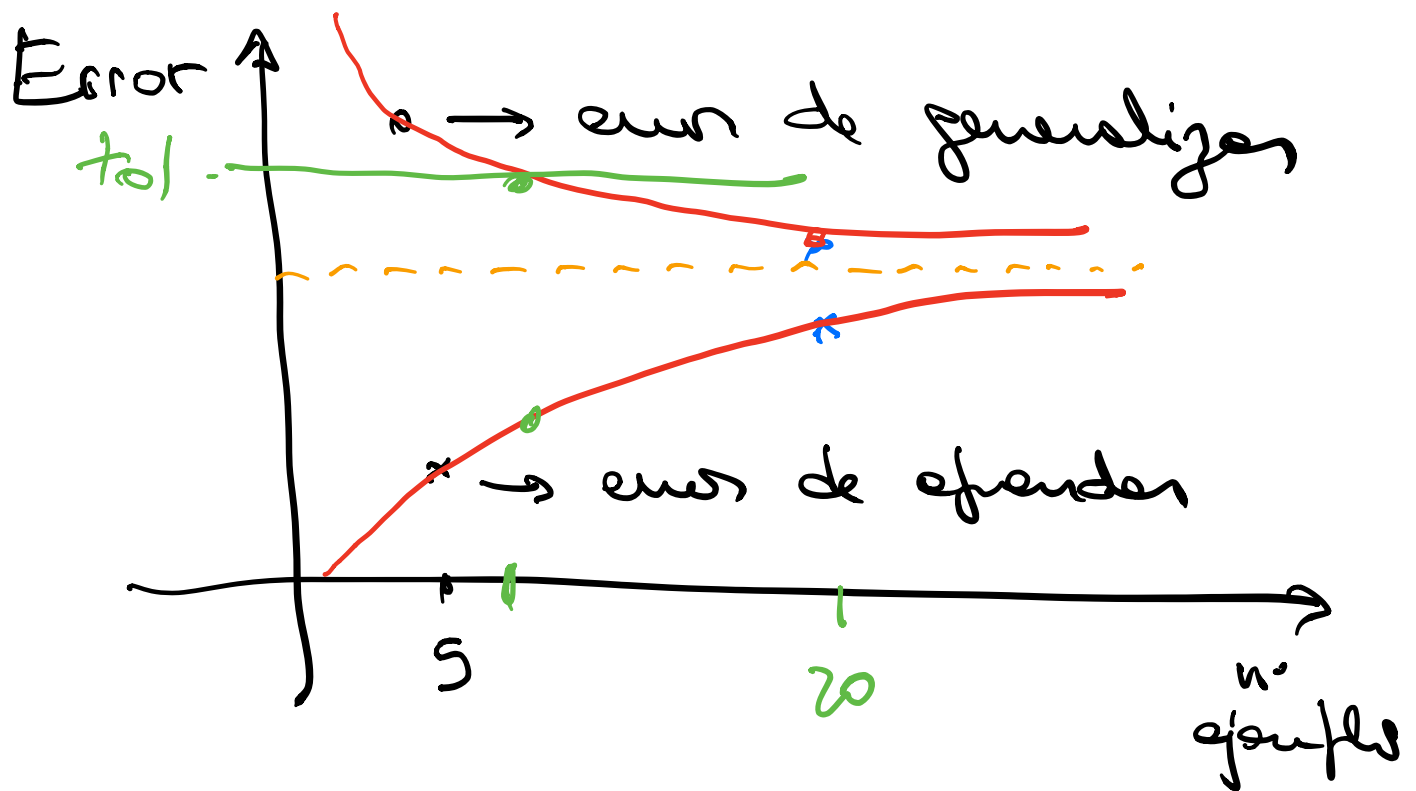


Tarea: aproximar una función

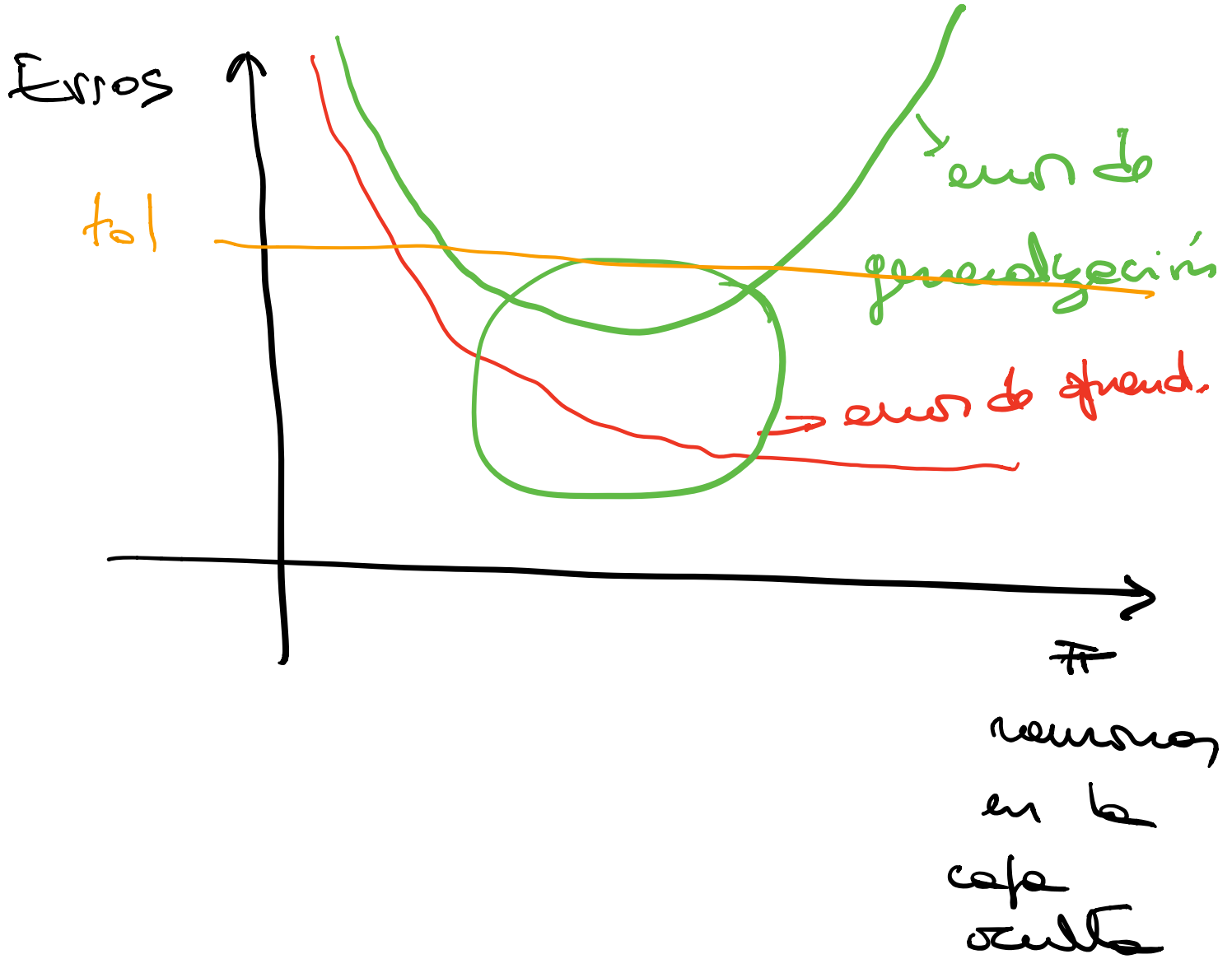


Si tengo cinco datos

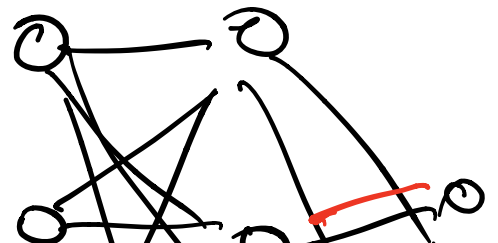
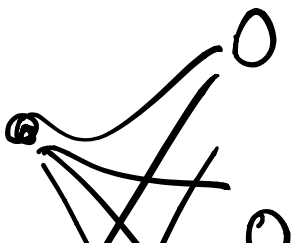
Si el n° de ejemplos es pequeño aprendo bien pero generalizo mal

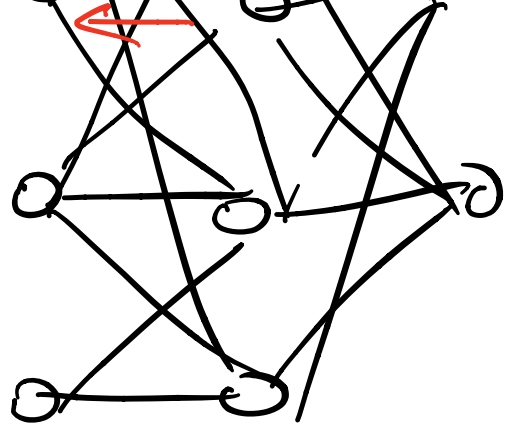
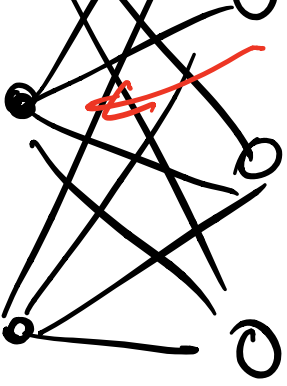


¿Cuántos neuronas ocultas uso?



Ahora tenemos más
 capas





$$\sum^k V_{j,l}$$

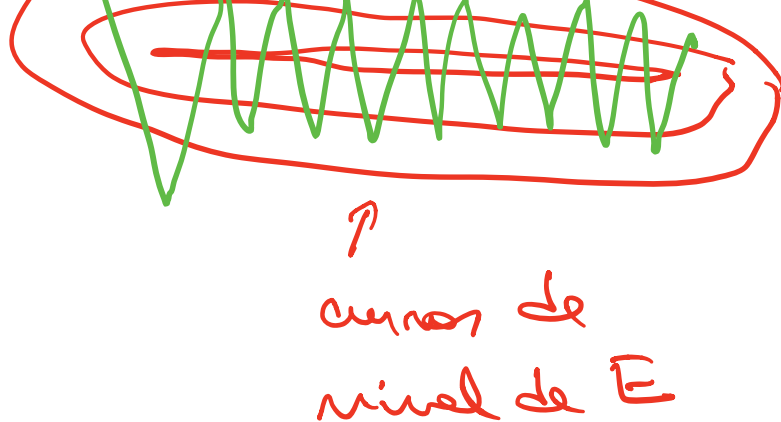
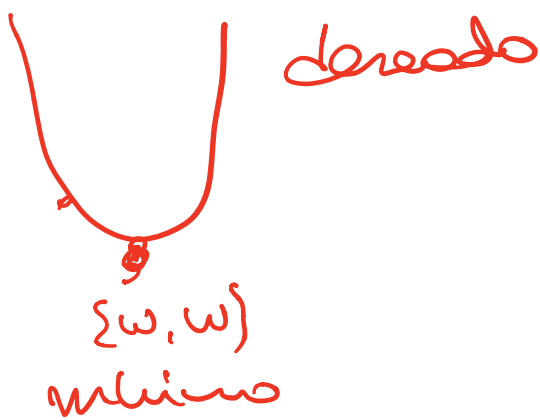
$$V_{j,l-2} \quad V_{j,l-1} \quad 0$$

$$\Delta W_{ij}^m = \eta \delta_i V_{j,m-1}$$

case $\rightarrow \delta_i = g'(h_i) \sum_j W_{ji}^{m+1} \delta_j^{m+1}$

Esquema del error





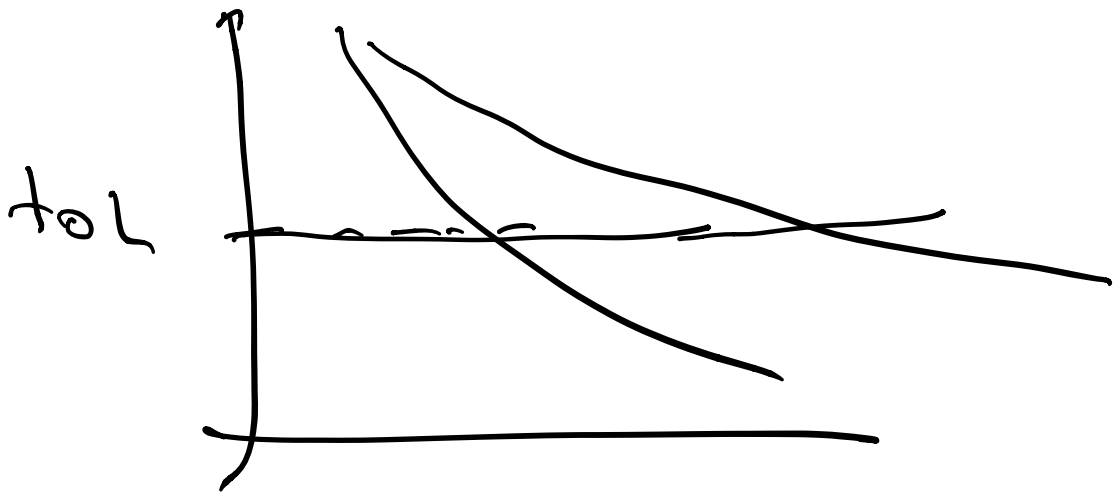
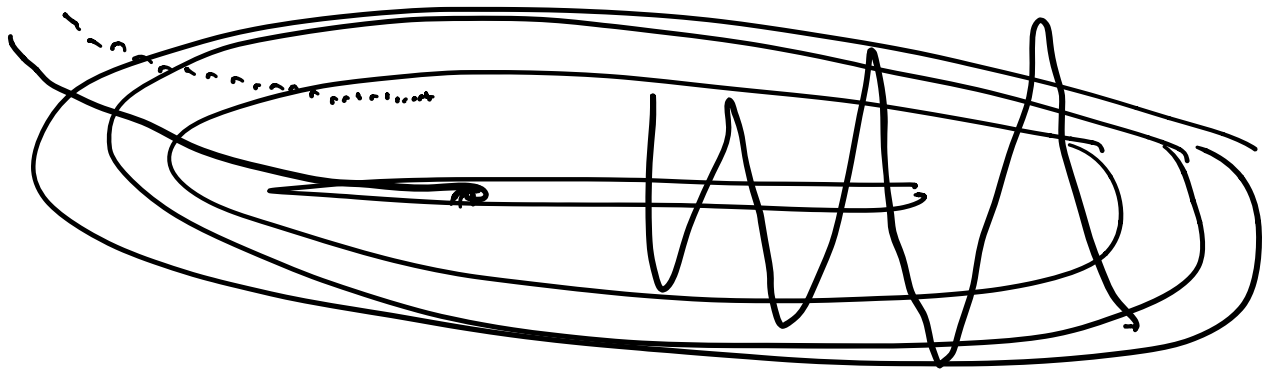
Método del momento

$$\Delta \omega_{p_f} (t+1) = - \frac{\partial E}{\partial \omega_{p_f}} + \alpha \Delta \omega_{p_f} (t)$$

$$\Sigma: \frac{\partial E}{\partial \omega_{p_f}} \approx 0$$

perfecto

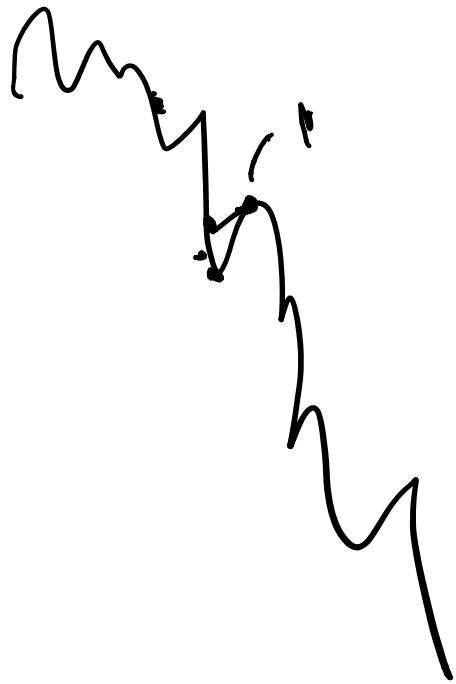
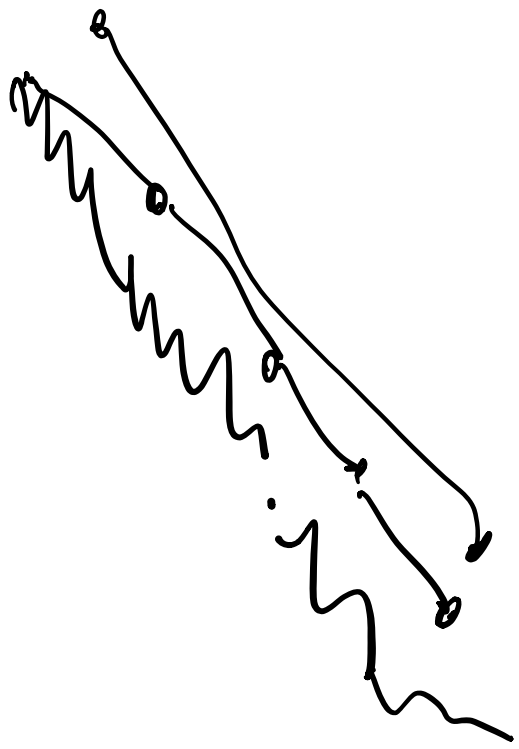
$$\Delta \omega_{p_f} (t) \approx \frac{1}{(1-\alpha)} \frac{\partial E}{\partial \omega_{p_f}}$$



Méthode de paramètres adaptatifs

$$\gamma(t+1) = \gamma(t) + \Delta\gamma$$

$$\Delta\gamma = \begin{cases} +a & \text{Si } \Delta E < 0 \text{ durant } k \text{ pers} \\ -b\gamma(t) & \text{Si } \Delta E > 0 \end{cases}$$



Calculo h_i

bajo con pred.

$$\frac{1}{1 + e^{-\frac{h}{T}}}$$

sobre con pred.

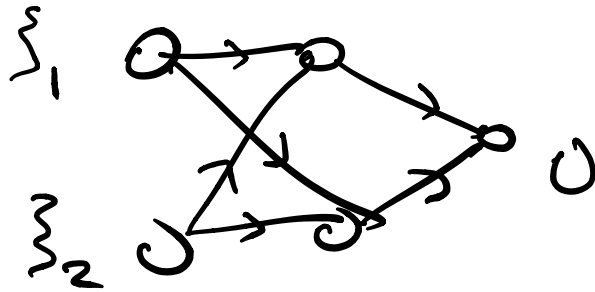
$$1 - \frac{1}{1 + e^{-\frac{h}{T}}}$$

Ejemplos

El XOR anda muy mal

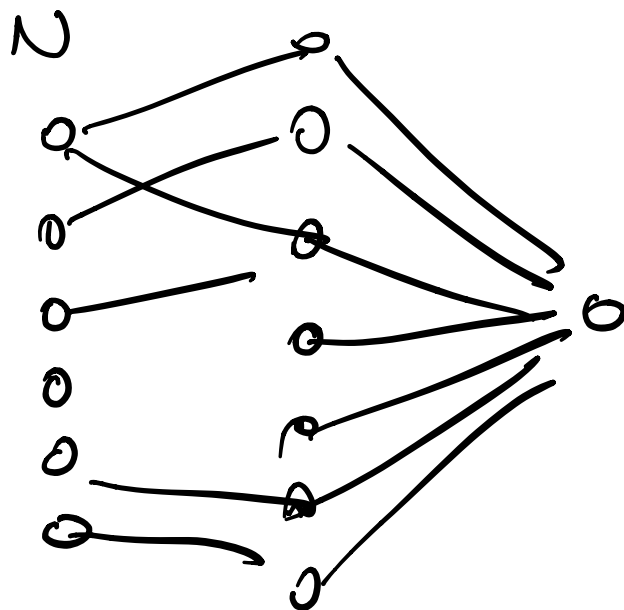
con un perceptron simple

Puedo usar capa oculta



XOR : función paridad

Se puede generalizar

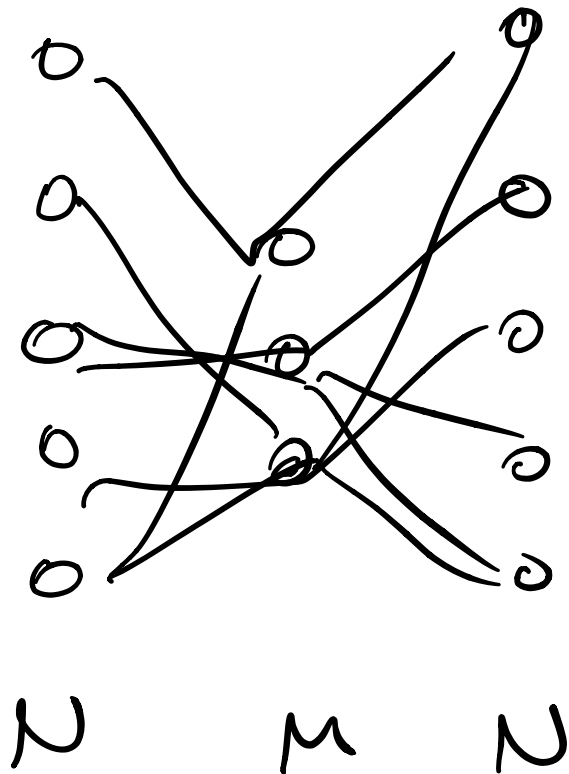


M

Si tengo un no impar de +1
en la entrada $O = 0$

Si tengo un no par de +1
 $O = +1$

Auto encoder



P

$$O_k^M = \sum_k^M$$

