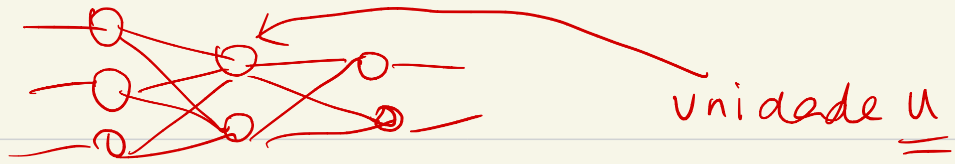


Atualização dos pesos por unidade $\underline{u}$:

$$w_u = w_u - \eta \nabla J(w_u)$$

$$\nabla J(w_u) = \left[\frac{\partial J}{\partial w_{u,0}}, \frac{\partial J}{\partial w_{u,1}}, \ldots, \frac{\partial J}{\partial w_{u,n}}\right]$$

Por entrada $\underline{i}$ da unidade $\underline{u}$:

$x_{u,i} \rightarrow$ input i of unit u

$w_{u,i} \rightarrow$ weight of input i of unit u

$$w_{u,i} = w_{u,i} - \eta \frac{\partial J}{\partial w_{u,i}}$$

Vamos usar a regra da cadeia:

$$\frac{\partial J}{\partial w_{u,i}} = \frac{\partial J}{\partial sum_u} \times \frac{\partial sum_u}{\partial w_{u,i}}, \text{ where}$$

$$sum_u = \sum_i w_{u,i} x_{u,i}$$

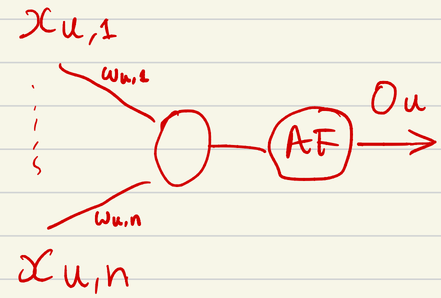
$$\frac{\partial sum_u}{\partial w_{u,i}} = \frac{\partial}{\partial w_{u,i}} \sum_i w_{u,i} \cdot x_{u,i} = x_{u,i}$$

$$\frac{\partial J}{\partial w_{u,i}} = \frac{\partial J}{\partial sum_u} \times x_{u,i}$$

falta calcular esta derivada parcial

Para a camada de saída:

$$\frac{\partial J}{\partial soma_u} = \underbrace{\frac{\partial J}{\partial O_u}}_{①} \cdot \underbrace{\frac{\partial O_u}{\partial soma_u}}_{②}$$



$$\frac{\partial J}{\partial O_u} = \frac{\partial}{\partial O_u} \frac{1}{2} \sum_k (t_k - O_k)^2$$

cada unidade da camada de saída

$$\frac{\partial J}{\partial O_u} = \frac{\partial}{\partial O_u} \frac{1}{2}\left((t_1 - O_1)^2 + (t_2 - O_2)^2 + \ldots + (t_n - O_n)^2\right)$$

① $$\frac{\partial J}{\partial O_u} = \frac{\partial}{\partial O_u} \frac{1}{2} (t_u - O_u)^2 = -(t_u - O_u)$$

Assumindo $O_u(soma_u) = \dfrac{1}{1 + e^{-soma_u}}$

função logística

② $$\frac{\partial O_u}{\partial soma_u} = O(soma_u)(1 - O(soma_u))$$

Atualização dos pesos da camada de saída

$$W_{u,i} = W_{u,i} + \eta \underbrace{\underbrace{\underbrace{(t_u - O_u)}_{\frac{\partial J}{\partial O_u}} \underbrace{O_u(1-O_u)}_{\frac{\partial O_u}{\partial soma_u}}}_{\frac{\partial J}{\partial soma_u}} \; \underbrace{x_{u,i}}_{\frac{\partial soma_u}{\partial w_{u,i}}}}_{\frac{\partial J}{\partial w_{u,i}}}$$

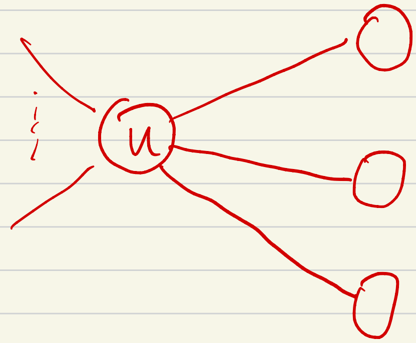
$$W_{u,i} = W_{u,i} + \eta \, \underbrace{\delta_u}_{-\frac{\partial J}{\partial soma_u}} \, x_{u,i}$$

termo de erro da unidade u

$$\delta_u = (t_u - O_u) O_u (1 - O_u)$$

Para uma unidade na camada escondida:

$$\frac{\partial J}{\partial soma_u}$$



a saída de u apenas influencia as unidades da camada seguinte

$$\frac{\partial J}{\partial soma_u} = \sum_{k \in seguinte(u)} \frac{\partial J}{\partial soma_k} \cdot \frac{\partial soma_k}{\partial soma_u}$$

$$\frac{\partial J}{\partial soma_u} = \sum_{k} -\delta_k \cdot \frac{\partial soma_k}{\partial soma_u}$$

erro da unidade k da camada seguinte

$$\frac{\partial soma_k}{\partial soma_u} = \frac{\partial soma_k}{\partial O_u} \cdot \frac{\partial O_u}{\partial soma_u}$$

$$soma_k = w_{k,1} x_{k,1} + \dots + w_{k,u} \underbrace{x_{k,u}}_{O_u} + \dots$$

$$\frac{\partial\, soma_k}{\partial\, soma_u} = w_{k,u} \frac{\partial O_u}{\partial\, soma_u}$$

$$\hookrightarrow O_u(1-O_u)$$

logo

$$\frac{\partial J}{\partial\, soma_u} = \sum_{k \in seguinte(u)} -\delta_k \cdot w_{k,u} \cdot O_u(1-O_u)$$

$$\delta_u = \frac{-\partial J}{\partial\, soma_u} = O_u(1-O_u)\sum_k \delta_k\, w_{k,u}$$

e

$$w_{u,i} = w_{u,i} + \eta\, \delta_u\, x_{u,i}$$