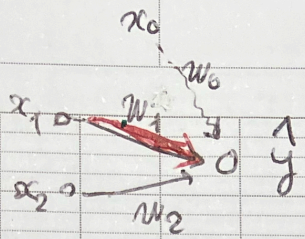
Deep : notes   oct. 2022

1) pas de couche cachée :

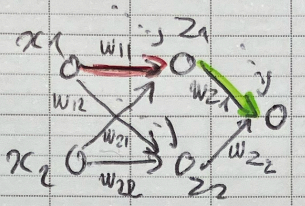


si f est linéaire.

$$\Delta w_1 = -\frac{\partial E}{\partial w_1} = -\frac{\partial E}{\partial \hat y}\cdot\frac{\partial \hat y}{\partial w_1} = (y-\hat y)\,x_1$$

idem pour Δw_2

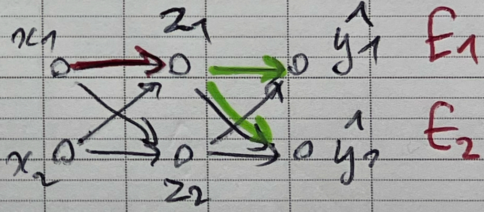
2) une couche cachée :



si f est linéaire.

$$\Delta w_{11} = -\frac{\partial E}{\partial w_1^1} = -\underbrace{\frac{\partial E}{\partial z_1}}_{\underbrace{\frac{\partial E}{\partial \hat y}}_{-(y-\hat y)}\cdot\underbrace{\frac{\partial \hat y}{\partial z_1}}_{w_{21}}}\;\underbrace{\frac{\partial z_1}{\partial w_{11}}}_{x_1} = (y-\hat y)\,w_{21}\,x_1$$

3) avec deux neurones de sortie;



si f est linéaire :

$$\Delta w_{11} = -\left(\frac{\partial E_1}{\partial w_{11}} + \frac{\partial E_2}{\partial w_{11}}\right)$$

Le calcul est le même que (2) :
$$\begin{cases} -\dfrac{\partial E_1}{\partial w_{11}} = (y_1-\hat y_1)\,w_{2_{11}}\,x_1 \\[2ex] -\dfrac{\partial E_2}{w_{11}} = (y_2-\hat y_2)\,w_{2_{12}}\,x_1 \end{cases}$$

et donc :

$$\Delta w_{11} = \left[(y_1-\hat y_1)\,w_{2_{11}} + (y_2-\hat y_2)\,w_{2_{12}}\right]x_1$$