

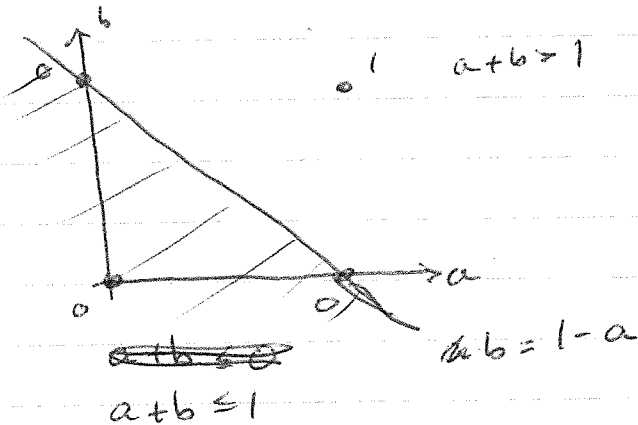
# Math et Tech Exercices - Réseaux de neurones.

#1. a)  $\hat{y}(a,b) = h(a+b-1)$

$$h(a+b-1) = 1 \Leftrightarrow a+b-1 > 0 \Leftrightarrow a+b > 1 \Leftrightarrow a=1 \text{ et } b=1$$

$$h(a+b-1) = 0 \Leftrightarrow a+b-1 \leq 0 \Leftrightarrow a+b \leq 1 \Leftrightarrow \left( a = \begin{smallmatrix} 0 \\ \text{ou} \\ 1 \end{smallmatrix} \text{ et } b = 0 \right)$$

$$\left( a = 0 \text{ et } b = \begin{smallmatrix} 0 \\ \text{ou} \\ 1 \end{smallmatrix} \right)$$

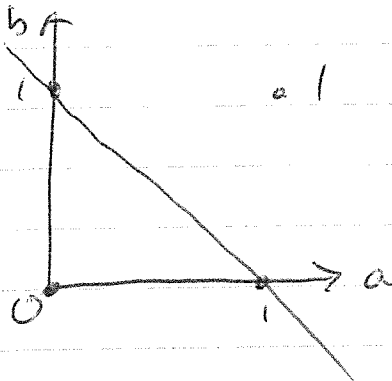


b)  $\hat{y}(a,b) = h(2a+2b-1)$

$$h(2a+2b-1) = 1 \Leftrightarrow 2a+2b-1 > 0 \Leftrightarrow a+b > \frac{1}{2} \Leftrightarrow \left( a = 1 \text{ et } b = \begin{smallmatrix} 0 \\ \text{ou} \\ 1 \end{smallmatrix} \right)$$

$$\left( a = \begin{smallmatrix} 0 \\ \text{ou} \\ 1 \end{smallmatrix} \text{ et } b = 1 \right)$$

$$h(2a+2b-1) = 0 \Leftrightarrow 2a+2b-1 \leq 0 \Leftrightarrow a+b \leq \frac{1}{2} \Leftrightarrow a=0 \text{ et } b=0$$



#1c). Soit  $\hat{y}_1 = h(2a+2b-1)$  la sortie du neurone 1

$\hat{y}_2 = h(-a-b+2)$  la sortie du neurone 2.

Alors

$$\hat{y}(a,b) = h(\hat{y}_1 + \hat{y}_2 - 1)$$

On a donc  $\hat{y}(a,b) = 1 \Leftrightarrow h(\hat{y}_1 + \hat{y}_2 - 1) = 1$

$$\Leftrightarrow \hat{y}_1 + \hat{y}_2 - 1 > 0$$

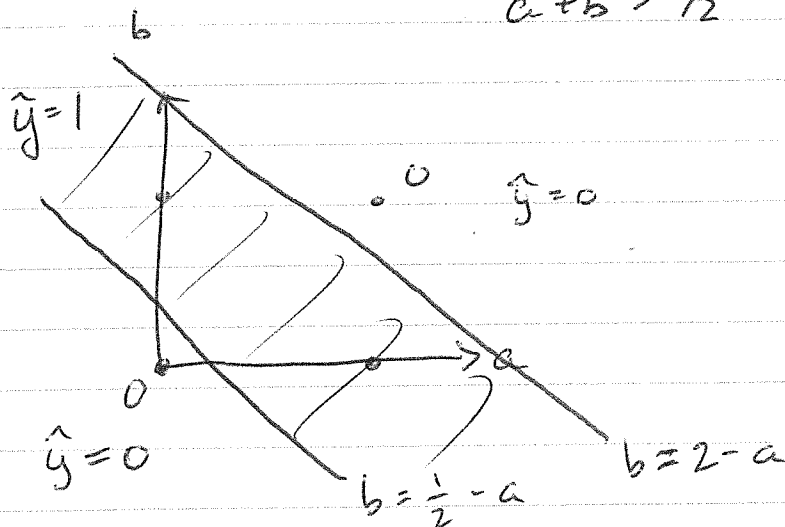
$$\Leftrightarrow \hat{y}_1 + \hat{y}_2 > 1$$

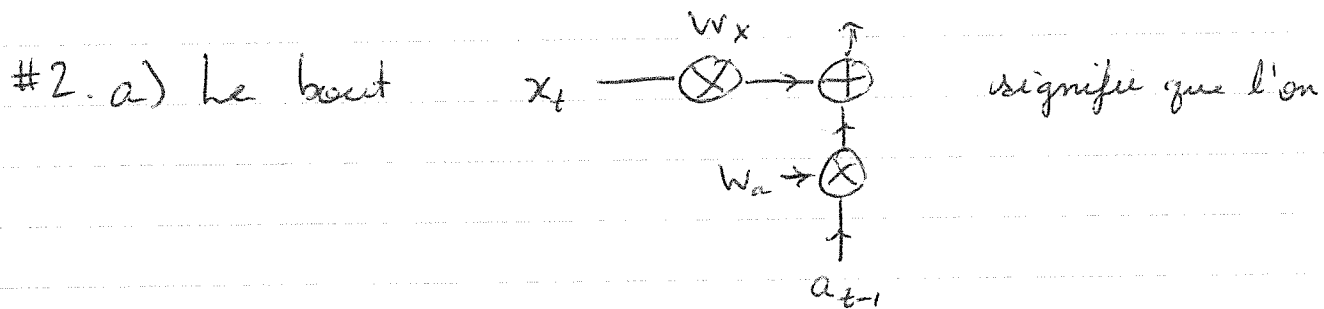
$$\Leftrightarrow \hat{y}_1 = 1 \text{ et } \hat{y}_2 = 1$$

$$\Leftrightarrow h(2a+2b-1) = 1 \text{ et } h(-a-b+2) = 1$$

$$\Leftrightarrow 2a+2b-1 > 0 \text{ et } -a-b+2 > 0$$

$$\Leftrightarrow a+b > \frac{1}{2} \text{ et } a+b < 2$$





fera  $w_x x_t + w_a a_{t-1}$ . Ensuite, la sortie du prochain état sera  $g_1(w_x x_t + w_a a_{t-1}) = a_t$ , donc il faut pouvoir additionner  $w_x x_t + w_a a_{t-1}$  et faut que le vecteur qui en résulte ait une la même taille de  $a_t$ .

$$g_1 \left( \begin{matrix} w_x & x_t & + & w_a & a_{t-1} \end{matrix} \right) = a_t$$

$\begin{matrix} m \times n & n \times 1 & & m \times m & m \times 1 \end{matrix}$ 
 $\begin{matrix} & & & & m \times 1 \end{matrix}$

Il faut que  $w_x$  ait la dimension  $m \times n$  et  $w_a$ ,  $m \times m$ .

On a ensuite  $y_t = g_2(w_y a_t)$ . Il faut que  $w_y$

$\begin{matrix} 1 \times 1 & & 1 \times m & m \times 1 \end{matrix}$

soit de taille  $1 \times m$ .

b) En suivant le schéma, on trouve

$$y_t = g_2(w_y a_t)$$

$$a_t = g_1(w_x x_t + w_a a_{t-1})$$

$$c). \frac{\partial}{\partial w_{x,u}} \mathcal{L}(\hat{y}_t, \hat{y}_t) = 2(y_t - \hat{y}_t) \frac{\partial \hat{y}_t}{\partial w_{x,u}}$$

$$= 2(y_t - \hat{y}_t) \frac{\partial}{\partial w_{x,11}} (g_2(w_y a_t))$$

$$\frac{\partial}{\partial w_{x,11}} g_2(w_y a_t) = g_2'(w_y a_t) w_y \frac{\partial a_t}{\partial w_{x,11}}$$

$$\frac{\partial a_t}{\partial w_{x,11}} = \frac{\partial}{\partial w_{x,11}} (g_1(w_x x_t + w_a a_{t-1}))$$

$$= g_1'(w_x x_t + w_a a_{t-1}) \left( \frac{\partial w_x}{\partial w_{x,11}} x_t + w_a \frac{\partial a_{t-1}}{\partial w_{x,11}} \right)$$

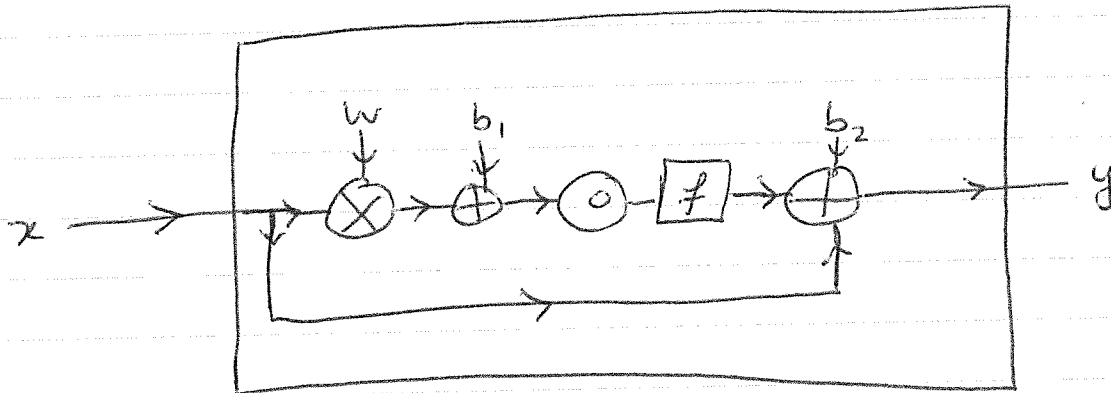
$$\frac{\partial w_x}{\partial w_{x,11}} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \\ \vdots & & & \\ 0 & & & 0 \end{pmatrix}_{m \times n} \Rightarrow \frac{\partial w_x}{\partial w_{x,11}} x_t = \begin{pmatrix} x_{t,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Si on combine, on arrive à

$$\frac{\partial \mathcal{L}}{\partial w_{x,11}} = 2(y_t - \hat{y}) g_2'(w_y a_t) w_y g_1'(w_x x_t + w_a a_{t-1}) \cdot \left( \begin{pmatrix} x_{t,1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + w_a \frac{\partial a_{t-1}}{\partial w_{x,11}} \right)$$

(En fait,  $x_t$  dépend probablement de  $w_x$ , mais on a supposé que non pour simplifier les calculs.)

#3. On pourrait faire quelque chose comme ceci



#4. a)  $\hat{y}_1$  est un vecteur de deux composantes : la première est la sortie du neurone 1 et la deuxième, la sortie du neurone 2. On peut donc écrire

$$\hat{y}_1 = \begin{pmatrix} f(w_1 x) \\ f(w_2 x) \end{pmatrix}$$

La deuxième couche reçoit  $\hat{y}_1$  comme entrée. On a

$$\hat{y}_2 = \begin{pmatrix} f(w_3 \hat{y}_1) \\ f(w_4 \hat{y}_1) \\ f(w_5 \hat{y}_1) \end{pmatrix}$$

Enfin, la dernière couche reçoit  $\hat{y}_2$  comme entrée. La sortie finale est

$$\hat{y}_3 = f(w_6 \hat{y}_2)$$

b)  $\frac{\partial \mathcal{L}}{\partial w_1} = 2(y - \hat{y}_3) \frac{\partial \hat{y}_3}{\partial w_1}$

$$\frac{\partial \hat{y}_3}{\partial w_1} = f'(w_6 \hat{y}_2) \cdot w_6 \frac{\partial \hat{y}_2}{\partial w_1}$$

$$\frac{\partial \hat{y}_2}{\partial w_1} = \begin{pmatrix} f'(w_3 \hat{y}_1) w_3 \frac{\partial \hat{y}_1}{\partial w_1} \\ f'(w_4 \hat{y}_1) w_4 \frac{\partial \hat{y}_1}{\partial w_1} \\ f'(w_5 \hat{y}_1) w_5 \frac{\partial \hat{y}_1}{\partial w_1} \end{pmatrix}$$

$$\frac{\partial \hat{y}_1}{\partial w_1} = \begin{pmatrix} f'(w_1 x) x \\ 0 \end{pmatrix} \Rightarrow w_i \frac{\partial \hat{y}_1}{\partial w_1} = w_i f'(w_1 x) x$$

$$\Rightarrow \frac{\partial \hat{y}_2}{\partial w_1} = \begin{pmatrix} f'(w_3 \hat{y}_1) w_{31} \\ f'(w_4 \hat{y}_1) w_{41} \\ f'(w_5 \hat{y}_1) w_{51} \end{pmatrix} f'(w_1 x) x$$

$$\Rightarrow w_6 \frac{\partial \hat{y}_2}{\partial w_1} = \left[ w_{61} f'(w_3 \hat{y}_1) w_{31} + w_{62} f'(w_4 \hat{y}_1) w_{41} + w_{63} f'(w_5 \hat{y}_1) w_{51} \right] f'(w_1 x) x$$

$$\Rightarrow \frac{\partial \mathcal{L}}{\partial w_1} = 2(y - \hat{y}_3) f'(w_6 \hat{y}_2) \left[ w_{61} f'(w_3 \hat{y}_1) w_{31} + w_{62} f'(w_4 \hat{y}_1) w_{41} + w_{63} f'(w_5 \hat{y}_1) w_{51} \right] f'(w_1 x) x$$

comme voulu.