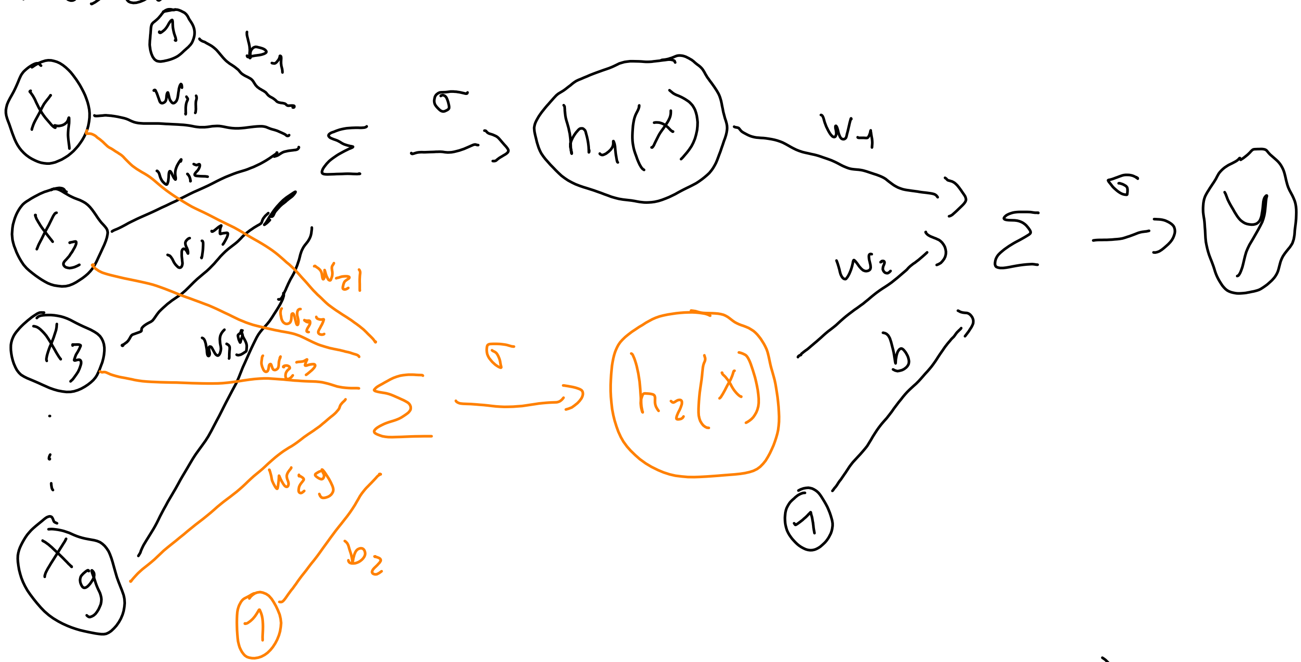
Réseau de neurones du TPS :



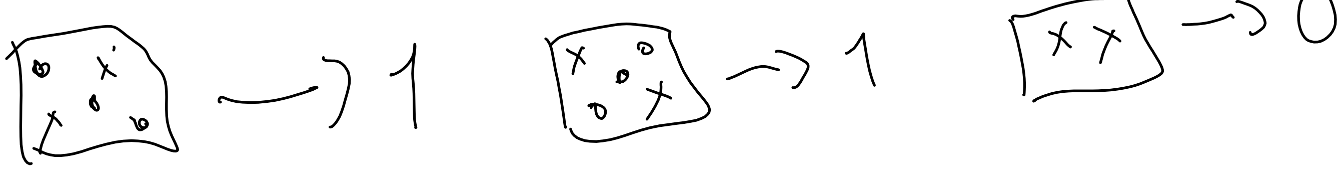
$$R_W(X) = \sigma\big(b + w_1 h_1(X) + w_2 h_2(X)\big)$$

$$= \sigma\Big(b + w_1\,\sigma\big(b_1 + w_{11}X_1 + w_{12}X_2 + \dots + w_{19}X_9\big) + w_2\,\sigma\big(b_2 + w_{21}X_1 + w_{22}X_2 + \dots + w_{29}X_9\big)\Big)$$

$$= \sigma\big(b + (w_1 \;\; w_2)\cdot H\big)$$

où $H = \begin{pmatrix} h_1(X) \\ h_2(X) \end{pmatrix} = \begin{pmatrix} \sigma(b_1 + W_1 \cdot X) \\ \sigma(b_2 + W_2 \cdot X) \end{pmatrix}$

Pb : trouver $b, w_1, w_2, \dots,$ (23 paramètres)



Calcul du gradient de f :

$$f(W) = \frac{1}{2} \sum_J \left(\sigma(s_J) - y_J\right)^2$$

- $\sigma'(x) = \sigma(x) \times (1 - \sigma(x))$
- $(u(v(x)))' = u'(v) \times v'$

$$(u^2)' = 2 \times u \times u'$$

$$f(W) = \frac{1}{2} \sum_J \left\{ \sigma\Big(\underbrace{b + w_1 \sigma(h_{1J}) + w_2 \sigma(h_{2J})}_{s_J}\Big) - y_J \right\}^2$$

$$\frac{\partial f}{\partial b}(W) = \frac{1}{2} \sum_J 2\left(\sigma(s_J) - y_J\right) \times \frac{\partial}{\partial b}\left(\sigma\left(b + w_1 \sigma(h_{1J}) + w_2 \sigma(h_{2J})\right) - y_J\right)$$

$$= \frac{1}{\cancel{2}} \sum_J \cancel{2}\left(\sigma(s_J) - y_J\right) \sigma(s_J) \times (1 - \sigma(s_J)) \times \frac{\partial}{\partial b}\left(b + w_1 \sigma(h_{1J}) + w_2 \sigma(h_{2J})\right)$$

$$= \sum_J \left(\sigma(s_J) - y_J\right) \sigma(s_J) \times (1 - \sigma(s_J)) \times 1$$

$$\frac{\partial f}{\partial w_1}(W) = \frac{1}{2} \sum_J 2\left(\sigma(s_J) - y_J\right) \times \frac{\partial}{\partial w_1}\left[\sigma(s_J) - y_J\right]$$

$$= \sum_J \left(\sigma(s_J) - y_J\right) \times \sigma(s_J) \times (1 - \sigma(s_J)) \times \frac{\partial}{\partial w_1}\left[s_J\right]$$

$$= \sum_J \left(\sigma(s_J) - y_J\right) \times \sigma(s_J) \times (1 - \sigma(s_J)) \times \underbrace{\frac{\partial}{\partial w_1}\left[b + w_1 \sigma(h_{1J}) + w_2 \sigma(h_{2J})\right]}_{\sigma(h_{1J})}$$

$$\frac{\partial f}{\partial b_1}(W) = \sum_J \left(\sigma(s_J) - y_J\right) \sigma(s_J) (1 - \sigma(s_J)) \times \underbrace{\frac{\partial}{\partial b_1}\left[b + w_1 \sigma(b_1 + W \cdot X) + w_2 \sigma(h_{2J})\right]}_{w_1 \times \sigma(h_{1J}) \times (1 - \sigma(h_{1J})) \times \underbrace{\frac{\partial}{\partial b_1}\left[b_1 + W \cdot X\right]}_{1}}$$

$$= \sum_J \left(\sigma(s_J) - y_J\right) \sigma(s_J) (1 - \sigma(s_J)) w_1 \sigma(h_{1J}) \times (1 - \sigma(h_{1J})) \times 1$$