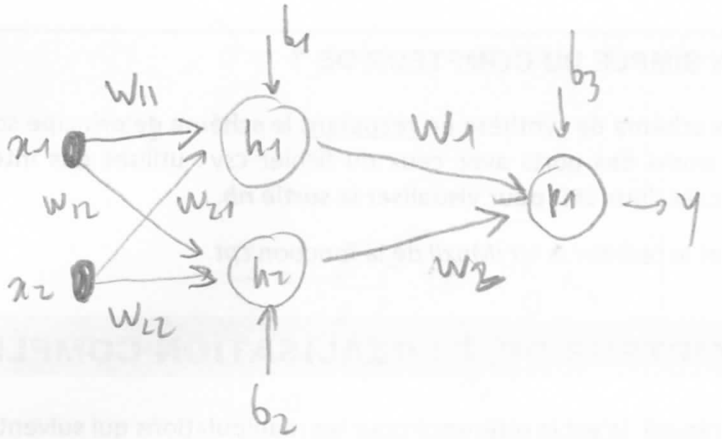

$$h_1 = b_1 + x_1 W_{11} + x_2 W_{21} = i_1 \qquad Z_1 = \frac{1}{1+e^{-i_1}}$$

$$h_2 = b_2 + x_1 W_{12} + x_2 W_{22} = i_2 \qquad Z_2 = \frac{1}{1+e^{-i_2}}$$

$$p_1 = b_3 + W_1 Z_1 + W_2 Z_2 = i_3 \qquad Z_3 = \frac{1}{1+e^{-i_3}}$$

$$C = \sum_k \left[-y^{(k)} \operatorname{Log} Z_3^{(k)} - (1-y^{(k)}) \operatorname{Log}(1-Z_3^{(k)}) \right]$$

Pour 1 échantillon

$$C(k) = -y \operatorname{Log} Z_3 - (1-y) \operatorname{Log}(1-Z_3)$$

$$\frac{\partial C}{\partial W_1} = \frac{\partial C}{\partial Z_3} \times \frac{\partial Z_3}{\partial W_1} = \underbrace{\left[-\frac{y}{Z_3} + \frac{(1-y)}{1-Z_3} \right]}_{=\frac{-y(1-Z_3)+Z_3(1-y)}{Z_3(1-Z_3)}} \times \overbrace{Z_3(1-Z_3)}^{Z_3'} \times Z_1$$

$$= \frac{(Z_3-y)}{Z_3(1-Z_3)}$$

$$= (Z_3-y) \times Z_1$$

$$\frac{\partial C}{\partial W_2} = (Z_3-y) \times Z_2$$

$$\frac{\partial C}{\partial b_3} = \frac{Z_3-y}{Z_3(1-Z_3)} \underbrace{\frac{\partial Z_3}{\partial b_3}}_{Z_3(1-Z_3)\times 1} = (Z_3-y)$$

Erreur issue de h_1: $$\frac{\partial C}{\partial Z_1} = \frac{Z_3-y}{Z_3(1-Z_3)} \overbrace{\frac{\partial Z_3}{\partial Z_1}}^{Z_3(1-Z_3)W_1} = (Z_3-y)W_1$$

$$\frac{\partial C}{\partial i_1} = \frac{\partial C}{\partial Z_1} \times \frac{\partial Z_1}{\partial i_1} = (Z_3-y) W_1 Z_1 (1-Z_1)$$

$$\frac{\partial C}{\partial w_{11}} = \frac{\partial C}{\partial i_1} \times \frac{\partial i_1}{\partial w_{11}} = (z_3 - y) w_1 z_1 (1 - z_1) x_1$$

Softmax

$$C = \sum_i \left[-y_i \, \mathrm{Log} \, z_i^{(3)} \right]$$

$$z_k^{(3)} = \frac{e^{s_k^{(3)}}}{\sum_{l \neq i} e^{s_l^{(3)}} + e^{s_i^{(3)}}}$$

$$\frac{\partial C}{\partial w_{ji}^{(2)}} = \frac{\partial C}{\partial s_i^{(3)}} \times \frac{\partial s_i^{(3)}}{\partial w_{ji}^{(3)}} \quad \rightarrow = z_j^{(2)}$$

$$= \sum_k \frac{\partial C}{\partial z_k^{(3)}} \times \frac{\partial z_k^{(3)}}{\partial s_i^{(3)}}$$

$$\frac{-y_k}{z_3^{(k)}}$$

$$\begin{cases} i = k \; : \; = \left\{ e^{s_i^{(3)}} \left[\sum_{l \neq i} e^{s_l^{(3)}} + e^{s_i^{(3)}} \right] - e^{s_i^{(3)}} \times e^{s_i^{(3)}} \right\} / (\Sigma)^2 \\ \qquad\quad = e^{s_i^{(3)}} \left[\sum_{l \neq i} \right] / \left(\sum_l \right)^2 = z_i^{(3)} \left(1 - z_i^{(3)} \right) \\ i \neq k \; : \; = \dfrac{-e^{s_k^{(3)}} e^{s_i^{(3)}}}{(\Sigma)^2} = -z_k^{(3)} \times z_i^{(3)} \end{cases}$$

$$= \frac{-y_i}{\cancel{z_i^{(3)}}} \times \cancel{z_i^{(3)}} \left(1 - z_i^{(3)} \right) + \sum_{k \neq i} \frac{-y_k}{\cancel{z_3^{(k)}}} \times \left(-\cancel{z_k^{(3)}} \times z_i^{(3)} \right)$$

$$= -y_i \left(1 - z_i^{(3)} \right) + \sum_{k \neq i} y_k z_i^{(3)}$$

$$= -y_i + \underbrace{\left(\sum_k y_k \right)}_{=1} z_i^{(3)} = z_i^{(3)} - y_i$$

$$\frac{\partial C}{\partial w_{ji}^{(2)}} = \left[z_i^{(3)} - y_i \right] z_j^{(2)}$$

$$S_k = \frac{e^{x_k}}{e^{x_0} + \dots + e^{x_{N-1}}} = \frac{e^{x_k}}{\sum_{\ell \neq k} e^{x_\ell} + e^{x_k}}$$

$$\frac{\partial S}{\partial x_\ell} = (\text{si } \ell = k)\ \frac{e^{x_k}\left[\sum_{\neq} + e^{x_k}\right] - e^{x_k} \cdot e^{x_k}}{(\Sigma)^2}$$

$$= \frac{e^{x_k}\left[\Sigma - e^{x_k}\right]}{\Sigma \times \Sigma} = S_k \times (1 - S_k)$$

$$= (\text{si } \ell \neq k) : \frac{e^{x_k} \times (-e^{x_\ell})}{\Sigma \times \Sigma} = S_k \times (-S_\ell).$$