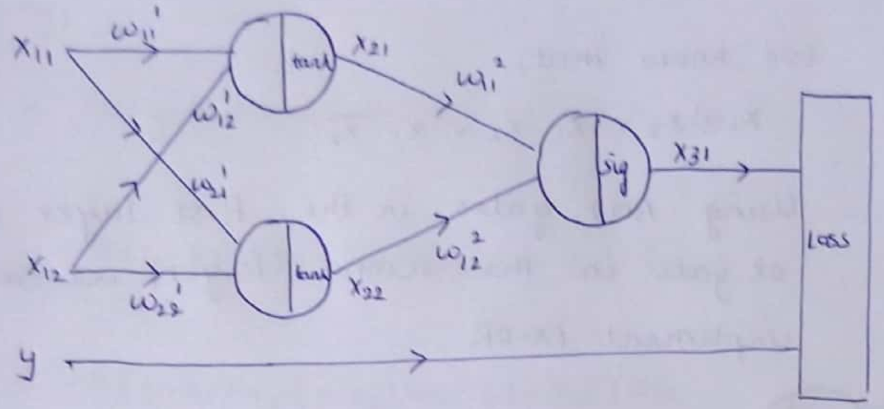
20171104 - Pset 18 - Q1



$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}} \qquad \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\frac{\partial}{\partial x} \text{sig}(x) = \text{sig}(x)(1-\text{sig}(x)) \qquad \frac{\partial}{\partial x}\tanh(x) = 1 - \tanh^2(x)$$

$$\bar{x}_2 = \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = \tanh\left(\begin{bmatrix} w_{11}^1 & w_{12}^1 \\ w_{21}^1 & w_{22}^1 \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix}\right) = \tanh\left(\begin{bmatrix} w_{11}^1 x_{11} + w_{12}^1 x_{12} \\ w_{21}^1 x_{11} + w_{22}^1 x_{12} \end{bmatrix}\right)$$

$$\bar{x}_3 = [x_{31}] = \text{sig}\left([w_{11}^2 \quad w_{12}^2]\begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}\right) = \text{sig}(w_{11}^2 x_{21} + w_{12}^2 x_{22})$$

$$\text{MSE loss} = \Sigma (y - x_3)^2 = E$$

$$\frac{\partial E}{\partial w_{11}^2} = \frac{\partial E}{\partial x_3} \cdot \frac{\partial x_3}{\partial w_{11}^2}$$

Now, $\frac{\partial E}{\partial x_3} = 2(y - x_3)(-1)$ and $\frac{\partial x_3}{\partial w_{11}^2} = x_3(1-x_3)(x_{21})$

since $x_3 = \text{sig}(w_{11}^2 x_{21} + w_{12}^2 x_{22})$

$$\therefore \frac{\partial E}{\partial w_{11}^2} = -2 x_3 x_{21} (1 - x_3)(y - x_3)$$

Similarly, $\frac{\partial E}{\partial w_{12}^2} = -2 x_3 x_{22} (1 - x_3)(y - x_3)$

$$\frac{\partial E}{\partial w_{11}^1} = \frac{\partial E}{\partial x_{21}} \cdot \frac{\partial x_{21}}{\partial w_{11}^1} = \frac{\partial E}{\partial x_3} \cdot \frac{\partial x_3}{\partial x_{21}} \cdot \frac{\partial x_{21}}{\partial w_{11}^1}$$

Now, $\frac{\partial E}{\partial x_3} = 2(y - x_3)(-1)$; $\frac{\partial x_3}{\partial x_{21}} = x_3(1-x_3) w_{11}^2$; $\frac{\partial x_{21}}{\partial w_{11}^1} = (1 - x_{21}^2) x_{11}$

$$\therefore \frac{\partial E}{\partial w_{11}^1} = -2(y - x_3) x_3 (1 - x_3) w_{11}^2 (1 - x_{21}^2) x_{11}$$

Similarly,

$$\frac{\partial E}{\partial \omega_{12}^1} = -2(y-x_3)\, x_3(1-x_3)\,\omega_{11}^2\,(1-x_{21}^2)\, x_{12}$$

$$\frac{\partial E}{\partial \omega_{21}^1} = -2(y-x_3)\, x_3(1-x_3)\,\omega_{12}^2\,(1-x_{22}^2)\, x_{11}$$

$$\frac{\partial E}{\partial \omega_{22}^1} = -2(y-x_3)\, x_3(1-x_3)\,\omega_{12}^2\,(1-x_{22}^2)\, x_{12}$$

Updates:

$$\omega_{11}^1 \leftarrow \omega_{11}^1 + 2\eta\, x_3\, x_{11}\, \omega_{11}^2\, (y-x_3)(1-x_3)(1-x_{21}^2)$$

$$\omega_{12}^1 \leftarrow \omega_{12}^1 + 2\eta\, x_3\, x_{12}\, \omega_{11}^2\, (y-x_3)(1-x_3)(1-x_{21}^2)$$

$$\omega_{21}^1 \leftarrow \omega_{21}^1 + 2\eta\, x_3\, x_{11}\, \omega_{12}^2\, (y-x_3)(1-x_3)(1-x_{22}^2)$$

$$\omega_{22}^1 \leftarrow \omega_{22}^1 + 2\eta\, x_3\, x_{12}\, \omega_{12}^2\, (y-x_3)(1-x_3)(1-x_{22}^2)$$

$$\omega_{11}^2 \leftarrow \omega_{11}^2 + 2\eta\, x_3\, x_{21}\, (y-x_3)(1-x_3)$$

$$\omega_{12}^2 \leftarrow \omega_{12}^2 + 2\eta\, x_3\, x_{22}\, (y-x_3)(1-x_3)$$