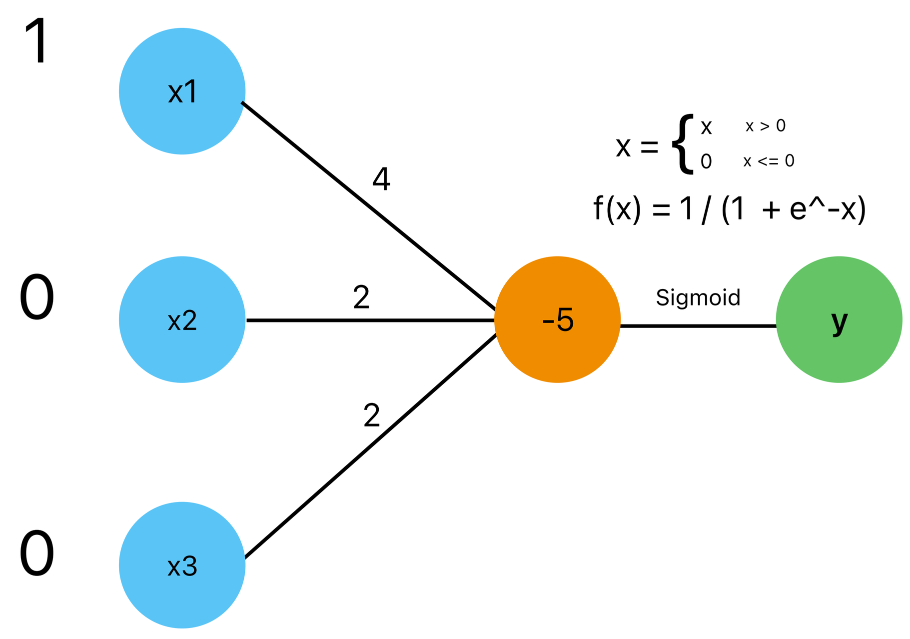
Ax + By + Cz

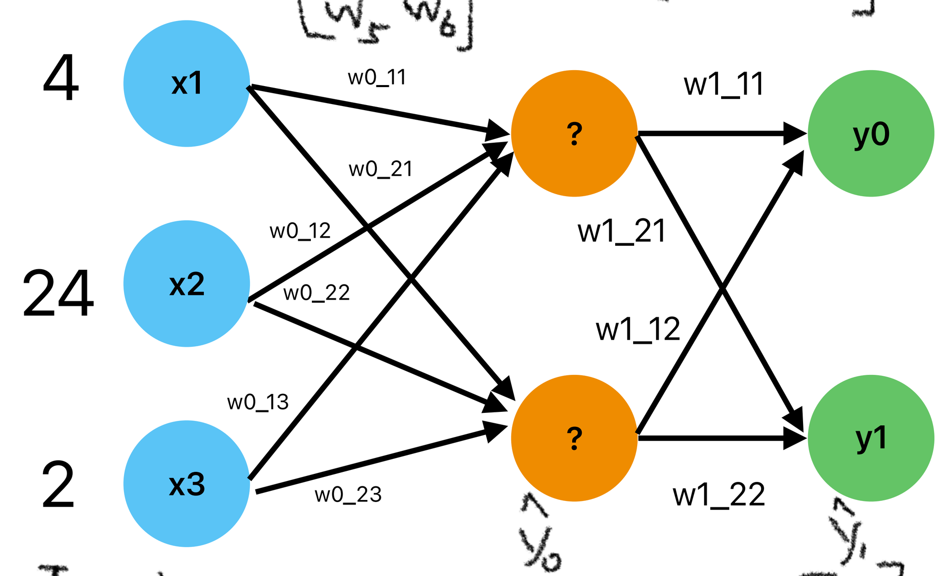
$x = \begin{cases} x & x > 0 \\ 0 & x <= 0 \end{cases}$

f(x) = 1 / (1 + e^-x)



$$W0 = \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \\ w_5 & w_6 \end{bmatrix} \qquad W1 = \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix}$$

02/24/2022



$$\text{Input} = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \qquad \hat{y}_0 = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \hat{y}_1 = \begin{bmatrix} a \\ b \end{bmatrix}$$

Loss

$$\begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{2}(\hat{y}_1 - \text{target})^2 = \begin{bmatrix} (a-\hat{a})^2 \\ (b-\hat{b})^2 \end{bmatrix}$$

$$\hat{y}_0 = \text{ReLU}(\text{Input} \otimes W0 + \text{bias}_0)$$

$$\text{sum} = \text{Input} \otimes W0 = \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} \otimes \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \\ w_5 & w_6 \end{bmatrix} = \begin{bmatrix} 3w_1 + 24w_3 + 2w_5 \\ 3w_2 + 24w_4 + 2w_6 \end{bmatrix}$$

$$z_1 = \text{sum} + \text{bias}_0 = \begin{bmatrix} 3w_1 + 24w_3 + 2w_5 + b \\ 3w_2 + 24w_4 + 2w_6 + b \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\hat{y}_1 = \text{ReLU}(\hat{y}_0 \otimes W1 + \text{bias}_1)$$

$$\hat{y}_0 \otimes W1 = \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} w_1 & w_2 \\ w_3 & w_4 \end{bmatrix} = \begin{bmatrix} xw_1 + yw_3 \\ xw_2 + yw_4 \end{bmatrix}$$

$$z_2 = \hat{y} \otimes w_1 + \text{bias}_1 = \begin{bmatrix} xw_1 + yw_3 + b \\ xw_2 + yw_4 + b \end{bmatrix} \simeq \begin{bmatrix} a \\ b \end{bmatrix}$$

$$\text{target} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix}$$

$$\frac{d}{d\,\text{loss}} = \hat{y}_1 - \text{target} = \begin{bmatrix} a \\ b \end{bmatrix} - \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix}$$

$$\text{ReLU} = \begin{cases} x & x > 0 \\ 0 & x <= 0 \end{cases} \qquad \text{ReLU}'(\hat{y}_1) = \begin{bmatrix} a' \\ b' \end{bmatrix}$$

$$\text{ReLU}' = \begin{cases} 1 & x > 0 \\ 0 & x <= 0 \end{cases} \qquad \text{ReLU}'(\hat{y}_0) = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$\partial\text{Loss} \otimes \text{ReLU}'(\hat{y}_1) = \begin{bmatrix} c \\ d \end{bmatrix} \otimes \begin{bmatrix} a' \\ b' \end{bmatrix} = [ca' + db'] = L'_1$$

$$\partial\text{Loss} \otimes \text{ReLU}'(\hat{y}_0) = \begin{bmatrix} c \\ d \end{bmatrix} \otimes \begin{bmatrix} x' \\ y' \end{bmatrix} = [cx' + dy'] = L'_0$$

$$\frac{\partial L}{\partial w_1} = L'_1 \otimes \hat{y}_0 = [ca' + db'] \otimes \begin{bmatrix} xw_1 + yw_3 \\ xw_2 + yw_4 \end{bmatrix} = \begin{bmatrix} (xw_1 + yw_3)(ca' + db') \\ (xw_2 + yw_4)(ca' + db') \end{bmatrix} = \begin{bmatrix} \hat{w}_{11} \\ \hat{w}_{12} \end{bmatrix}$$

$$\frac{\partial L}{\partial w_0} = L'_0 \otimes \text{Input} = [ca' + db'] \otimes \begin{bmatrix} 3 \\ 24 \\ 2 \end{bmatrix} = \begin{bmatrix} 3(ca' + db') \\ 24(ca' + db') \\ 2(ca' + db') \end{bmatrix}$$