Calc

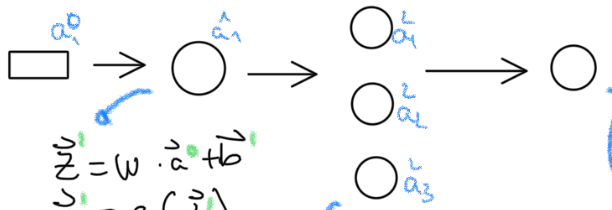
$\vec{a}^0 = \vec{X} = [1] \qquad \vec{y} = [5]$

$g(z) = \text{ReLU} = \max(z, 0)$

$g'(z) = \text{ReLU}' \quad \begin{matrix} 1 \Leftarrow z>0 \\ 0 \Leftarrow \text{else} \end{matrix}$

$C = (\vec{a}^3 - y)^2$

$C' = 2(\vec{a}^3 - y)$



$\vec{z}^1 = w \cdot \vec{a}^0 + \vec{b}^1$

$\vec{a}^1 = g(\vec{z}^1)$

$\vec{z}^2 = W^2 \cdot \vec{a}^1 + \vec{b}^2$

$\vec{a}^2 = g(\vec{z}^2)$

$\vec{z}^3 = W^3 \cdot \vec{a}^2 + \vec{b}^3$

$\vec{a}^3 = g(\vec{z}^3)$

Forward Feed

$W^1 = [2]$

$b^1 = [3]$

$a_1 = 5 = W^1 \cdot a^0_1 + b^1$

Backprop

$W^2 = \begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix} \quad \vec{b}^2 = \begin{bmatrix} 11 \\ 13 \\ 17 \end{bmatrix}$

$W^3 = [19\ 27\ 31] \quad \vec{b}^3 = [37]$

$\vec{a}^1 = [5]$

$$f_{NN}([1]) = g\left(W^3 \begin{bmatrix} g(5 \cdot g(2 \cdot 1 + 3) + 11) \\ g(7 \cdot g(2 \cdot 1 + 3) + 13) \\ g(9 \cdot g(2 \cdot 1 + 3) + 17) \end{bmatrix} + \vec{b}^3 \right)$$

$$= g\left(W^3 \begin{array}{cc} \begin{bmatrix} 36 \\ 48 \\ 62 \end{bmatrix} & \begin{matrix} a^2_1 \\ a^2_2 \\ a^2_3 \end{matrix} \end{array} + \vec{b}^3 \right)$$

$$= g(19 \cdot 36 + 27 \cdot 48 + 31 \cdot 62 + 37) = 3939 = \vec{a}^3$$

LAYER 3

$$\frac{\partial C}{\partial \vec{a}_1^3} = 2(\vec{a}-y) = 7868$$

$$\frac{\partial C}{\partial W_{mn}^3} = \frac{\partial C}{\partial \vec{a}_m^3} \frac{\partial a_m^3}{\partial \vec{z}_m^3} \frac{\partial z_m^3}{\partial w_{mn}^3} =$$

$$\frac{\partial C}{\partial W_{11}^3} = 2(\vec{a}_1^3 - \vec{y}_1)\, g'(z_1^3) \cdot \vec{a}_1^2 = 283243$$

$$\frac{\partial C}{\partial b_1^3} = 2(\vec{a}_1^3 - \vec{y}_1)\, g'(z_1^3) = 7868$$

$W_{11}^2 \to a_1^1 \quad W_{21}^2 \to a_2^1$

$w^2 \quad W_{12}^2 \to a_1^1 \quad W_{22} \to a_2^1$

$$\begin{bmatrix} w_{11} & w_{21} \\ w_{12} & w_{22} \end{bmatrix} \begin{bmatrix} a_1^1 \\ a_2^1 \end{bmatrix} = \begin{bmatrix} w_{11} \cdot a_1^1 + w_{12} \cdot a_2^1 \\ w_{12} \cdot a_1^1 + w_{22} \cdot a_2^1 \end{bmatrix} = \begin{bmatrix} a_1^2 \\ a_2^2 \end{bmatrix}$$

$$\frac{\partial C}{\partial W_{12}^3} = 2(\vec{a}_1^3 - \vec{y}_1)\, g'(z_1^3) \cdot \vec{a}_2^2 = 377664$$

$$\frac{\partial C}{\partial W_{13}^3} = 2(\vec{a}_1^3 - \vec{y}_1)\, g'(z_1^3) \cdot \vec{a}_3^2 = 487816$$

LAYER 2

CACHE POS=1

$$\frac{\partial C}{\partial \vec{a}_1^2} = \underbrace{\frac{\partial C}{\partial a_1^3} \frac{\partial a_1^3}{\partial z_1^3}}_{7868} \frac{\partial z_1^3}{\partial a_1^2} \overset{19}{=} 149492$$

$$\frac{\partial C}{\partial b_1^2} = 7868$$

$$\frac{\partial C}{\partial \vec{a}_2^2} = \frac{\partial C}{\partial a_1^3} \frac{\partial a_1^3}{\partial z_1^3} \frac{\partial z_1^3}{\partial a_2^2} \overset{27}{=} 212436$$

$$\frac{\partial C}{\partial w_{11}^2} = \frac{\partial C}{\partial a_1^2} \frac{\partial a_1^2}{\partial z_1^2} \frac{\partial z_1^2}{\partial w_{11}^2} = 7868 \cdot a_1^1 =$$

$$\frac{\partial C}{\partial \vec{a}_3^2} = \frac{\partial C}{\partial a_1^3} \frac{\partial a_1^3}{\partial z_1^3} \frac{\partial z_1^3}{\partial a_3^2} \overset{31}{=} 243908$$

$$\frac{\partial C}{\partial w_{12}^2} = \frac{\partial C}{\partial a_1^2} \frac{\partial a_1^2}{\partial z_1^2} \frac{\partial z_1^2}{\partial w_{11}^2} = 7868 \cdot a_2^1 =$$

Error δ

$$\frac{\partial C}{\partial w_{13}^2} = \frac{\partial C}{\partial a_1^2} \frac{\partial a_1^2}{\partial z_1^2} \frac{\partial z_1^2}{\partial w_{11}^2} = 7868 \cdot a_3^1 =$$

LAYER 1

$$\frac{\partial C}{\partial \vec{a}_1^1} = \sum_{m=1}^{3} \frac{\partial C}{\partial a_m^2} \frac{\partial a_m^2}{\partial z_m^2} \frac{\partial z_m^2}{\partial a_1^1} = 149492 \cdot 5 + 212436 \cdot 7 + 243908 \cdot 9 = 4429684$$

$$\frac{\partial C}{\partial W_{11}^1} = \frac{\partial C}{\partial a_1^1} \frac{\partial a_1^1}{\partial z_1^1} \frac{\partial z_1^1}{\partial w_{11}^1} = 3029180 \cdot 1 \cdot X_1 = 4429684$$

$$\frac{\partial C}{\partial b_1^1} = \frac{\partial C}{\partial a_1^1} \frac{\partial a_1^1}{\partial z_1^1} \frac{\partial z_1^1}{\partial b_1^1} = 4429684$$

$$\frac{\partial C}{\partial a_k^L} = 2(a_k^L - y_k) \qquad \frac{\partial C}{\partial w_{mn}^L} = \frac{\partial C}{\partial a_m^L} \frac{\partial a_m^L}{\partial z_m^L} \frac{\partial z_m^L}{\partial w_{mn}^L} = 2(a_m^L - y_m)\, \sigma'(z_m^L)\, a_n^{L-1}$$

$$\frac{\partial C}{\partial b_m^L} = \frac{\partial C}{\partial a_m^L}\frac{\partial a_m}{\partial z_m^L}\frac{\partial z_m}{\partial b_m^L} = 2(a_m^L - y_m)\,\sigma'(z_m^L)$$

Last Layer

$L=3$

Layer $L-1$ with $L = 2, \dots N$

CACHE LOOKUP

$$\frac{\partial C}{\partial a_n^{L-1}} = \sum_{m=1}^{N_L} \frac{\partial C}{\partial a_m^L}\frac{\partial a_m^L}{\partial z_m^L}\frac{\partial z_m^L}{\partial a_N^{L-1}} = \sum_{m=1}^{N_L} \boxed{\frac{\partial C}{\partial a_m^L}}\,\sigma'(z_m^L)\,W_{mn}^L$$

$$\frac{\partial C}{\partial W_{mn}^{L-1}} = \frac{\partial C}{\partial a_m^{L-1}}\frac{\partial a_m^{L-1}}{\partial z_m^{L-1}}\frac{\partial z_m^{L-1}}{\partial W_{mn}^{L-1}} = \boxed{\frac{\partial C}{\partial a_m^{L-1}}}\,\sigma'(z_m^{L-1})\,a_n^{L-2}$$

Not the same as above

$$\frac{\partial C}{\partial b_m^{L-1}} = \frac{\partial C}{\partial a_m^{L-1}}\frac{\partial a_m^{L-1}}{\partial z_m^{L-1}}\frac{\partial z_m^{L-1}}{\partial b_m^{L-1}} = \boxed{\frac{\partial C}{\partial a_m^{L-1}}}\,\sigma'(z_m^{L-1})$$