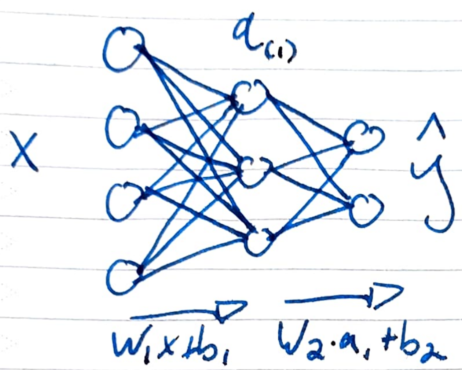
Back propagation Example

Network Architecture



input $\rightarrow$ 4×1

1 Hidden Layer $\rightarrow$ 3×1

output $\rightarrow$ 2×1

Sigmoid Activation $\rightarrow$ $\sigma(z) = \frac{1}{1+e^{-z}}$

Starting Weights and biases

$$W_1 = \begin{bmatrix} 0.2 & -0.1 & 0.4 & 0.1 \\ -0.3 & 0.2 & 0.1 & -0.2 \\ 0.1 & 0.3 & -0.2 & 0.2 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 0 \\ 0.1 \\ -0.1 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0.3 & -0.2 & 0.1 \\ -0.4 & 0.1 & 0.2 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 0.05 \\ -0.05 \end{bmatrix}$$

$$\text{input} = \begin{bmatrix} 1 \\ 0.5 \\ -1 \\ 2 \end{bmatrix} \qquad \text{True output} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Loss $\rightarrow$ MSE $\left(\frac{1}{2}\|\hat{y} - y\|^2\right)$

Forward Pass:

$$\hat{y} = \sigma(W_2 \, \underbrace{\sigma(W_1 x + b_1)}_{\text{first layer}} + b_2)$$

Layer 1 Pre-activation:

$$z_{(1)} = W_1 x + b_1$$

$$= \begin{bmatrix} -0.05 \\ -0.6 \\ 0.75 \end{bmatrix}$$

Layer 1 activation

$$a_1 = \sigma(z_1)$$

$$= \begin{bmatrix} 0.487 \\ 0.354 \\ 0.679 \end{bmatrix}$$

Layer 2 pre-activation

$$z_{(2)} = W_2 \cdot a_1 + b_2$$

$$= \begin{bmatrix} 0.1933 \\ -0.0737 \end{bmatrix}$$

Layer 2 activation

$$\hat{y} = \sigma(z_{(2)}) = \begin{bmatrix} 0.548 \\ 0.482 \end{bmatrix}$$

Loss: $L = \frac{1}{2}\|\hat{y} - y\|^2$

$= 0.21803$

Back Propogation:

Layer 2:

$$W_2' \rightarrow W_2 - \eta \frac{\partial L}{\partial W_2}$$

$$\frac{\partial L}{\partial W_2} = \frac{\partial L}{\partial a_2} \cdot \frac{\partial a_2}{\partial Z_2} \cdot \frac{\partial Z_2}{\partial W_2}$$

$$= \underbrace{(\hat{y} - y)}\circ \underbrace{\hat{y}\circ(1-\hat{y})} \; \underbrace{a_1}$$

$$= \delta_2 \cdot a_1^T$$

$$= \begin{bmatrix} -0.112 \\ 0.120 \end{bmatrix} \begin{bmatrix} 0.487 & 0.354 & 0.679 \end{bmatrix}$$

$$= \begin{bmatrix} -0.055 & -0.040 & -0.076 \\ 0.059 & 0.043 & 0.082 \end{bmatrix}$$

$$W_2' \rightarrow \begin{bmatrix} 0.306 & -0.196 & 0.108 \\ -0.406 & 0.096 & 0.192 \end{bmatrix}$$

$$b_2' \rightarrow b_2 - \eta \frac{\partial L}{\partial b_2}$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial b_2}$$

$$= (\hat{y} - y) \circ \hat{y}(1 - \hat{y}) \circ 1$$

$$= \delta^{(2)} \circ 1$$

$$= \begin{bmatrix} -0.112 \\ 0.120 \end{bmatrix} \rightarrow b' = \begin{bmatrix} 0.061 \\ -0.062 \end{bmatrix}$$

Layer 1:

$$W_1' \rightarrow W_1 - \eta \frac{\partial L}{\partial W_1}$$

$$\frac{\partial L}{\partial W_1} = \underbrace{\frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2}} \cdot \underbrace{\frac{\partial z_2}{\partial a_1}}_{a} \cdot \underbrace{\frac{\partial a_1}{\partial z_1}}_{b} \cdot \frac{\partial z_1}{\partial W_1}$$

$$= \delta^{(2)} \cdot W_2 \cdot a_1 \circ (1 - a_1) \cdot x$$

$$= ((W_2^T \cdot \delta^{(2)}) \circ a_1 \circ (1 - a_1))\} \cdot x^T$$

$$= \begin{bmatrix} -0.0204 & -0.01 & 0.020 & -0.041 \\ 0.008 & 0.004 & -0.008 & 0.016 \\ 0.002 & 0.001 & -0.003 & 0.007 \end{bmatrix}$$

$$\therefore W' = \begin{bmatrix} 0.202 & -0.099 & 0.400 & 0.104 \\ -0.304 & 0.200 & 0.101 & -0.202 \\ 0.100 & 0.300 & -0.200 & 0.200 \end{bmatrix}$$

$$b_1' = b_1 - \eta \frac{dL}{db_1}$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z_2} \cdot \frac{\partial z_2}{\partial a_1} \cdot \frac{\partial a_1}{\partial z_1} \cdot \frac{\partial z_1}{\partial b_1}$$

$$= \delta^{(2)} \cdot W_2 \circ a_1 \circ (1-a_1) \cdot 1$$

$$= (W_2^T \cdot \delta^{(2)}) \circ a_1 \circ (1-a_1)$$

$$= \begin{bmatrix} -0.020 \\ 0.008 \\ 0.003 \end{bmatrix}$$

$$\therefore b_1' = \begin{bmatrix} 0.002 \\ 0.099 \\ -0.100 \end{bmatrix}$$