

Backpropagation

Output Layer (Out)

$$\frac{\partial C_0}{\partial w_{jk}^{out}} = \left(\frac{\partial C_0}{\partial a_j^{out}} \right) * \left(\frac{\partial a_j^{out}}{\partial z_j^{out}} \right) * \left(\frac{\partial z_j^{out}}{\partial w_{jk}^{out}} \right)$$

Output gradients:

$$= 2(a_j^{out} - y_j) * (\sigma'(z_j^{out})) * (a_k^{H1}) = 7.1644$$

Layer H1

$$\begin{aligned} \left(\frac{\partial C_0}{\partial a_k^{H1}} \right) &= \sum_{j=0}^{n_L-1} \left(\frac{\partial C_0}{\partial a_j^{out}} \right) * \left(\frac{\partial a_j^{out}}{\partial z_j^{out}} \right) * \left(\frac{\partial z_j^{out}}{\partial a_k^{H1}} \right) \\ &= \sum_{j=0}^{n_{out}-1} 2(a_j^{out} - y_j) * (\sigma'(z_j^{out})) * (w_{jk}^{out}) \end{aligned}$$

H1 gradient:

$$\frac{\partial C_0}{\partial w_{jk}^{H1}} = \left(\frac{\partial C_0}{\partial a_j^{H1}} \right) * \left(\frac{\partial a_j^{H1}}{\partial z_j^{H1}} \right) * \left(\frac{\partial z_j^{H1}}{\partial w_{jk}^{H1}} \right) = 7.196$$

Layer H0

$$\begin{aligned} \left(\frac{\partial C_0}{\partial a_k^{H0}} \right) &= \sum_{j=0}^{n_{H1}-1} \left(\frac{\partial C_0}{\partial a_j^{H1}} \right) * \left(\frac{\partial a_j^{H1}}{\partial z_j^{H1}} \right) * \left(\frac{\partial z_j^{H1}}{\partial a_k^{H0}} \right) \\ &= \sum_{j=0}^{n_{H1}-1} \left(\sum_{j=0}^{n_{out}-1} \left(\left(\frac{\partial C_0}{\partial a_j^{out}} \right) * \left(\frac{\partial a_j^{out}}{\partial z_j^{out}} \right) * \left(\frac{\partial z_j^{out}}{\partial a_k^{H1}} \right) \right) * (\sigma'(z_j^{H1})) * (w_{jk}^{H1}) \right) \end{aligned}$$

H0 gradient

$$\frac{\partial C_0}{\partial w_{jk}^{H0}} = \left(\frac{\partial C_0}{\partial a_k^{H0}} \right) * \left(\frac{\partial a_k^{H0}}{\partial z_j^{H0}} \right) * \left(\frac{\partial z_j^{H0}}{\partial w_{jk}^{H0}} \right) = 0.01755125$$

H0

- Input: .5
- Weight: .1
- Bias: 2
- $z: (.5 * .1) + 2 = 2.05$
- $a: \text{ReLU}(2.05) = 2.05$

Output

- $dC/da = 2(a - y) = 2(3.51025 - 0) = 7.0205$
- $da/dz = \text{relu}'(z) = 1$
- $dz/dw = a(H1) = 1.0205$
- Product = $7.0205 * 1 * 1.0205 = 7.1644$

H1

- Input: 2.05
- Weight: .01
- Bias: 1
- $z: (2.05 * .01) + 1 = 1.0205$
- $a: \text{ReLU}(1.0205) = 1.0205$

H1

- $dC/da(\text{Out}) = 2(a - y) = 7.0205$
- $da(\text{Out})/dz(\text{Out}) = \text{relu}'(z(\text{Out})) = 1$
- $dz(\text{Out})/da(H1) = \text{weight}(\text{Out}) = .5$
- $dC/da(H1) = \text{sum}(\text{products}) = 3.51025$
- $da(H1)/dz(H1) = \text{relu}'(z(H1)) = 1$
- $dz(H1)/dw(H1) = a(H0) = 2.05$
- $dC/dw(H1) = 3.51025 * 1 * 2.05 = 7.196$

Output

- Input: 1.0205
- Weight: .5
- Bias: 3
- $z: (1.0205 * .5) + 3 = 3.51025$
- $a: \text{ReLU}(3.51025) = 3.51025$

H0

- $dC/da(H1) = 3.51025$
- $da(H1)/dz(H1) = 1$
- $dz(H1)/da(H0) = \text{weight}(H1) = .01$
- $dC/da(H0) = \text{sum}(\text{products}) = .0351025$
- $da(H0)/dz(H0) = \text{relu}'(z(H0)) = 1$
- $dz(H0)/dw(H0) = \text{input} = .5$
- $dC/dw(H0) = .0351025 * 1 * .5 = .01755125$