

Gradient Calculation Example

Simple Network: $y = \sigma(wx + b)$

Computed Gradients

Given Values

Input
 $x = 2$

Weight
 $w = 0.5$

Bias
 $b = 1$

Target
 $y^* = 1$

Learning Rate
 $\eta = 0.1$

Final Gradients

$$\frac{\partial L}{\partial w} = \frac{\partial L}{\partial z} \cdot x \quad -0.05$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial z} \quad -0.025$$

FORWARD Step 1: Linear Combination

$$z = wx + b = 0.5 \times 2 + 1$$

$$z = 2$$

FORWARD Step 2: Activation

$$\hat{y} = \sigma(z) = 1 / (1 + e^{-z})$$

$$\hat{y} = 0.881$$

FORWARD Step 3: Loss

$$L = (\hat{y} - y^*)^2 = (0.881 - 1)^2$$

$L = 0.014$

BACKWARD Step 4: $\partial L/\partial \hat{y}$

$$\partial L/\partial \hat{y} = 2(\hat{y} - y^*) = 2(0.881 - 1)$$

$$\partial L/\partial \hat{y} = -0.238$$

BACKWARD Step 5: $\partial L/\partial z$

$$\partial L/\partial z = \partial L/\partial \hat{y} \cdot \sigma'(z) = -0.238 \times 0.105$$

$$\partial L/\partial z = -0.025$$

BACKWARD Step 6: $\partial L/\partial w$ & $\partial L/\partial b$

$$\partial L/\partial w = \partial L/\partial z \cdot x = -0.025 \times 2$$

$$\partial L/\partial w = -0.05, \partial L/\partial b = -0.025$$

Parameter Update ($\eta = 0.1$)

$$w_{\text{new}} = 0.5 - 0.1 \times (-0.05) = 0.505$$

$$b_{\text{new}} = 1.0 - 0.1 \times (-0.025) = 1.0025$$

