Backpropagation

Parameters w_1, w_2, b

Step 1: linear transformation $\sum = x_1 w_1 + x_2 w_2 + b$ $w_1 \approx 0.1$

Step 3: Estimation error (loss) $\mathcal{L} = \frac{1}{2}(\hat{y} - y)^2$

Loss \mathcal{L}

Input X = (1, 2)

 $w_1^{new} = w_1 - \eta \frac{\partial \mathcal{L}}{\partial w_1}$

Step 2: Non-linear Activation (e.g., ReLU) $\hat{y} = \sigma(\Sigma) = \max(0, \Sigma)$

 $\frac{\partial \mathcal{L}}{\partial \Sigma} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial \Sigma} = \begin{cases} \hat{y} - y, & if \Sigma > 0 \\ 0, & if \Sigma \le 0 \end{cases} \qquad \frac{\partial \mathcal{L}}{\partial \hat{y}} = \hat{y} - y$

 $w_2^{new} = w_2 - \eta \frac{\partial \mathcal{L}}{\partial w_2} \quad \frac{\partial \mathcal{L}}{\partial w_1} = \frac{\partial \mathcal{L}}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial \Sigma} \cdot \frac{\partial \Sigma}{\partial w_1} = \frac{\partial \mathcal{L}}{\partial \Sigma} \cdot x_1$ $b^{new} = b - \eta \frac{\partial \mathcal{L}}{\partial b} \qquad \frac{\partial \mathcal{L}}{\partial w_2} = \frac{\partial \mathcal{L}}{\partial \Sigma} \cdot x_2 \qquad \frac{\partial \mathcal{L}}{\partial b} = \frac{\partial \mathcal{L}}{\partial \Sigma} \cdot 1$

Assume $\eta = 0.1$