2021/02/09

Exercise 3

Given an input $\mathbf{x} \in \mathbb{R}^2$, a weight vector $\mathbf{w} \in \mathbb{R}^2$ and a bias $b_0 \in \mathbb{R}$, draw the computational graph for the computation of the mean squared error $L = MSE(\hat{y}, y)$ of a prediction $\hat{y} = \sigma(\mathbf{w}^T\mathbf{x} + b_0)$ with respect to the true value y; σ is the sigmoid function.

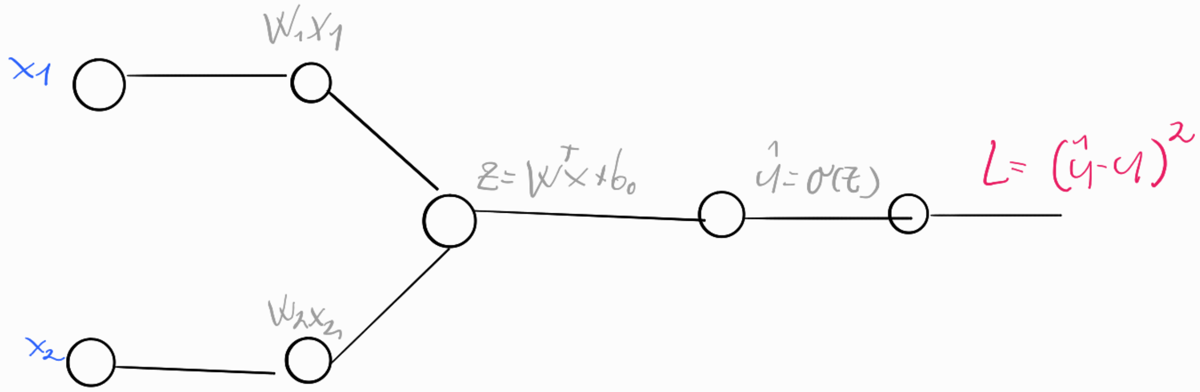
Consider the following values for $\mathbf{x}, \mathbf{w}, b_0$ and y:

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} \frac{4}{5} \\ \frac{7}{5} \end{bmatrix}, \quad b_0 = \frac{3}{5}, \quad y = 1, \tag{3}$$

and the MSE function $L = (\hat{y} - y)^2$. Report the correponding values on the computational graph.

Explain how the backpropagation method can be used to compute the gradient of L with respect to the inputs.

Draw the computational graph for the computation of the gradient and calculate the gradients values for each edge and node in the computational gradient graph.



$x_1 = 1$ $\quad Z = \frac{4}{5} + \frac{7}{5} + \frac{3}{5} = \frac{14}{5}$

$x_2 = 1$ $\quad \hat{y} = \sigma(\frac{14}{5}) \approx 0{,}94$

$w_1x_1 = \frac{4}{5} \cdot 1$ $\quad L = (0{,}94 - 1)^2 = -0{,}06^2 = 0{,}003$

$w_2x_2 = \frac{7}{5} \cdot 1$

Forward pass

Backpropagation $\quad$ define $\dot{v} = \frac{\partial L}{\partial v}$

$$\frac{\partial L}{\partial \hat{y}} = 2(\hat{y} - y) = 2(-0{,}06) = -0{,}12$$

$$\frac{\partial \hat{y}}{\partial z} = \sigma(z)(1 - \sigma(z)) = \hat{y}(1 - \hat{y}) = 0{,}05$$

$$\frac{\partial L}{\partial z} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z} = -0{,}12 \cdot 0{,}05 = -0{,}06$$

$$\frac{\partial z}{\partial w_1} = x_1 = 1 \longrightarrow \frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial w_1} = -0{,}006$$

$$\frac{\partial z}{\partial w_2} = x_2 = 1 \longrightarrow \frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial w_2} = -0{,}006$$

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

$$\frac{\partial L}{\partial \hat{y}} = 2(\hat{y} - y)$$

$$\hat{y} = \sigma(z)$$

$$\frac{\partial \hat{y}}{\partial z} = \sigma(z)\,(1 - \sigma(z))$$

$$z = w^T x + b_0$$

$$\frac{\partial z}{\partial w_1} = x_1 \qquad \frac{\partial z}{\partial w_2} = x_2 \qquad \frac{\partial z}{\partial b_0} = 1$$

w_1 $\quad$ w_2 $\quad$ b_0