Exercise: Perceptron and Multi-Layer Perceptron

Due: Optional

Problem 1 (Sigmoid function)

Let $\sigma(x) = \frac{1}{1+\exp(-x)}$. Show that

$$\sigma'(x) = \sigma(x)\big(1 - \sigma(x)\big).$$

Solution 1

Let $f(x) = (1+x)^{-1}$ and $g(x) = \exp(-x)$. Then we have $\sigma(x) = f(g(x))$. From calculus, we know

$$f'(x) = -(1+x)^{-2} \quad \text{and} \quad g'(x) = -\exp(-x).$$

According to the chain rule, we know

$$\begin{aligned}
\sigma'(x) &= f'(g(x))g'(x) = -\big(1+g(x)\big)^{-2}\big(-\exp(-x)\big) \\
&= \frac{1}{\big(1+\exp(-x)\big)^2}\exp(-x) = \frac{\exp(-x)}{1+\exp(-x)}\frac{1}{1+\exp(-x)} \\
&= \sigma(x)(1-\sigma(x)).
\end{aligned}$$

Problem 2 (Multi-Layer Perceptron)

Consider a fully-connected MLP with 5 layers: 1 input layer, 1 output layer and 3 hidden layers. Assume the input layer has 6 nodes, the three hidden layers have 6, 8, 10 nodes respectively, and the output layer has 3 nodes. Compute the number of trainable parameters.

Solution 2

According to the definition of MLPs, the trainable parameters include the weights and bias.

- Weight parameters include 4 matrices: $\mathbf{W}^2, \mathbf{W}^3, \mathbf{W}^4, \mathbf{W}^5$. The size of these matrices are as follows

$$\mathbf{W}^2 \in \mathbb{R}^{6\times 6}, \quad \mathbf{W}^3 \in \mathbb{R}^{8\times 6}, \quad \mathbf{W}^4 \in \mathbb{R}^{10\times 8} \quad \mathbf{W}^5 \in \mathbb{R}^{3\times 10}$$

- Bias parameters include 4 vectors: $\mathbf{b}^2, \mathbf{b}^3, \mathbf{b}^4, \mathbf{b}^5$. The size of these matrices are as follows

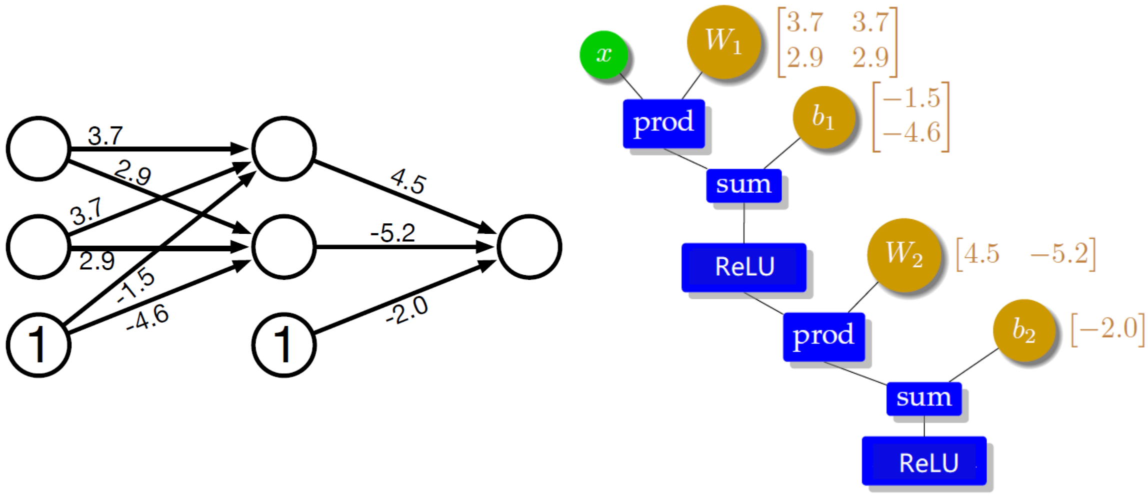
$$\mathbf{b}^2 \in \mathbb{R}^{6}, \quad \mathbf{b}^3 \in \mathbb{R}^{8}, \quad \mathbf{b}^4 \in \mathbb{R}^{10}, \quad \mathbf{b}^5 \in \mathbb{R}^{3}$$

Therefore, the total number of parameters are

$$\underbrace{6*6+6*8+8*10+10*3}_{\text{weight parameters}} + \underbrace{6+8+10+3}_{\text{bias parameters}} = 194+27 = 221.$$

Problem 3 (Forward Propagation)

Consider the following MLP with three layers. Let the input vector be $\mathbf{x} = (2, -1)^\top$. Apply the forward propagation to compute the final output.



Solution 3

The solution can be found in the following figure