Week 9 Homework Submission

Deep Learning

Iuliia Skobleva
Matriculation Number 03723809

December 15, 2019

Deep Learning

Problem 1

The goal is to prove that the identity holds:

$$y = \log \sum_{i=1}^{N} e^{x_i} = a + \log \sum_{i=1}^{N} e^{x_i - a} \tag{1}$$

where $a = \max_i x_i$. We can write:

$$\begin{aligned} y &= a + \log \sum_{i=1}^{N} e^{x_i} e^{-a} = \\ &= a + \log e^{-a} \sum_{i=1}^{N} e^{x_i} = \\ &= a + \log e^{-a} + \log \sum_{i=1}^{N} e^{x_i} = \\ &= a - a + \log \sum_{i=1}^{N} e^{x_i} = \\ &= \log \sum_{i=1}^{N} e^{x_i} \end{aligned} \tag{2}$$

Problem 2

We can compute the output of the softmax function, $\pi_i = \frac{e^{x_i}}{\sum_{i=1}^{N} e^{x_i}}$ with $a = \max_i x_i$, in a numerically stable way. We can show that the following identity holds:

$$\frac{e^{x_i}}{\sum_{i=1}^{N} e^{x_i}} = \frac{e^{x_i - a}}{\sum_{i=1}^{N} e^{x_i - a}} \tag{3}$$

We can write:

$$\begin{aligned} \frac{e^{x_i - a}}{\sum_{i=1}^{N} e^{x_i - a}} &= \frac{e^{x_i} e^{-a}}{\sum_{i=1}^{N} e^{x_i} e^{-a}} = \\ &= \frac{e^{-a} e^{x_i}}{e^{-a} \sum_{i=1}^{N} e^{x_i}} = \\ &= \frac{e^{x_i}}{\sum_{i=1}^{N} e^{x_i}} \end{aligned} \tag{4}$$