Neural Networks

A simple example

Henrik Lund Mortensen

September 7, 2017

Introduction

$$\text{softmax}(\vec{x})_i = \frac{e^{\vec{x}_i}}{\sum_k e^{\vec{x}_k}} \tag{1}$$

$$\text{Err}(\hat{\vec{y}}, \vec{y}) = \frac{-1}{N} \sum_{n \in N} \sum_{i \in C} \vec{y}_{n,i} \log \hat{\vec{y}}_{n,i} \tag{2}$$

Backpropagation

$$\frac{\partial L}{\partial \vec{b}^{(i)}} = \frac{\partial L}{\partial \vec{b}^{(i+1)}} \cdot W^{T(i+1)} \circ (1 - \vec{z^{(i)}} \circ \vec{z^{(i)}}) \tag{3}$$

$$\frac{\partial L}{\partial \vec{b}^{(N_{\text{layer}})}} = \frac{\hat{\vec{y}} - \vec{y}}{N_{\text{samples}}} \tag{4}$$

$$\frac{\partial L}{\partial \vec{W}^{(i)}} = \begin{cases} \vec{z}^{(i-1)} \otimes \frac{dL}{db^{(i)}} & \text{if } i \geq 2 \\ \vec{x} \otimes \frac{dL}{db^{(i)}} & \text{if } i = 1 \end{cases} \tag{5}$$