Assignment 6: Introduction to Neural Networks and Backpropagation

Machine Learning

Fall 2019

Learning Objectives

- Gain some familiarity with some of the key ideas in machine learning.
- Review of mathematical concepts we will be using in the beginning part of this course.
- Familiarize yourself with computational tools for machine learning.
- Learn linear regression using a "top-down" approach.

HMC Multivariable Chain Rule Page

Hidden layer to loss function.

$$e = -y \ln z - (1-y)\ln(1-z) \tag{1}$$
$$\frac{de}{dz} = -y\frac{1}{z} - (1-y)\frac{1}{1-z} \tag{2}$$
$$z = \sigma(s) \tag{3}$$
$$\frac{de}{s} = \frac{dz}{s}\frac{de}{dz} \tag{4}$$
$$= \sigma(s)(1-\sigma(s))\frac{de}{dz} \tag{5}$$
$$s = \mathbf{w}^\top \mathbf{h} \tag{6}$$
$$\nabla_h e = \frac{de}{ds}\nabla_{\mathbf{h}} s \tag{7}$$
$$= \frac{de}{ds}\mathbf{w} \tag{8}$$

Input layer to hidden layer.

$$s_1 = \mathbf{w_1}^\top \mathbf{x} \tag{9}$$
$$h_1 = \sigma(s_1) \tag{10}$$
$$\nabla_{\mathbf{w_1}} e = \frac{de}{ds_1}\nabla_{\mathbf{w_1}} s_1 \tag{11}$$
$$= \frac{de}{ds_1}\mathbf{x} \tag{12}$$
$$\frac{de}{s_1} = \frac{dh_1}{s_1}\frac{de}{dh_1} \tag{13}$$
$$= \sigma(s_1)(1-\sigma(s_1))\frac{de}{dh_1} \tag{14}$$