Contents

List of Figures

List of Tables

Neural Nets Overview

This section covers the basic workings and theory behind a neural net. It is divided up into 3 sections; vocabulary, structure, and updating.

Vocabulary

This section goes over the nomenclature used in this chapter. Variables are all defined as well as other terminology.

- **Performance Function** ($g(x, \theta)$): The performance function defines how close the output of the neural net is to the expected value. This is used with a gradient descent to help train the net.
- **Gradient descent**: A mathematical procedure used to find the local minima based on a derivative. The derivative is subtracted from the total and then rerun again. Once the derivative equals zero, a minima has been located.
- **Step Size** (μ):a coefficient used to control how large the step size will be in the gradient descent function.
- **theta** (θ): a vector representing both the weights ($\mathbf{x}$) and the bias term b_0
- **A**: the result of a single layer neuron before the non-linear function is applied.
- **Z**: the result of applying the non-linear function to A

The nueral network structre is seen in the images below:

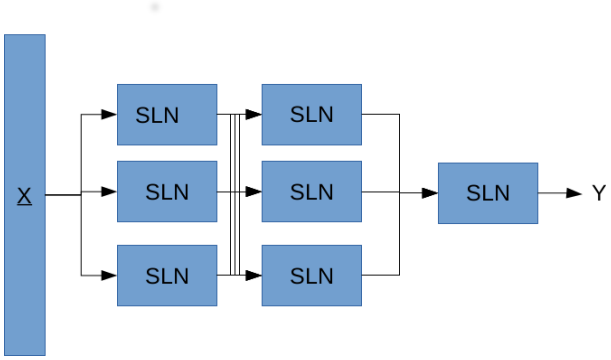

Figure 1: Multi Neuron Network

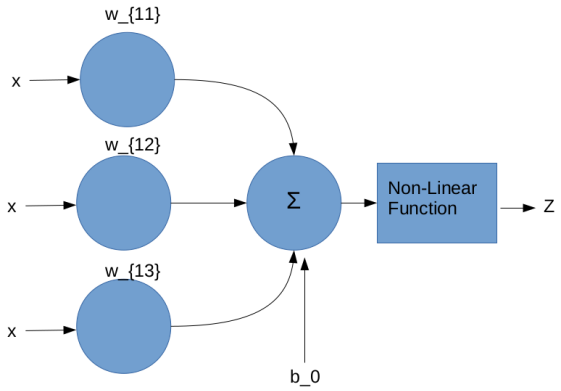

Figure 2: Single Layer Nueron

To solve for the minim using the gradient decent tequnique, the equation

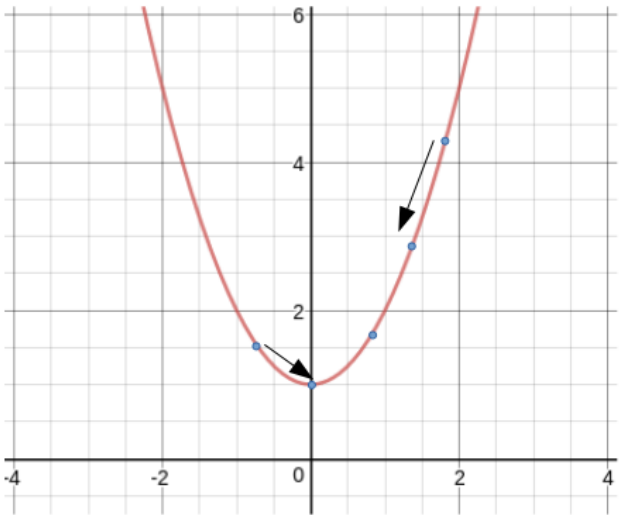

Figure 3: $J(\theta)$ Graph

$$\frac{1}{N}\sum_{i=1}^{N}(g(x+i,\theta)-y_i)^2 = J(\theta) \tag{1}$$

$$\int_{R^x} P(x)\cdot y(x) \tag{2}$$