Project Proposal, EENG 515, Fall 2019

Alexander Vidal

April 12, 2020

i. **Overview and problem statement:** Logistic regression, a predictive statistics technique now common in machine learning, seeks to model a binary explanatory variable using a logistic function. This technique is an optimization in the sense that a loss function will be minimized for a training set of data to condition for future predictions. Since logistic regression is useful in predicting image context, in this project, we seek to use numerical optimization techniques to accurately predict the context of an image presented given a training set of image data.

ii. **Data:**

In our project, we plan to utilize the classic "MNIST" dataset, which contains a set of 70,000 small images of handwritten characters. Associated with each image is the "correct" character which we can use to verify the performance of our model. Since there are multiple character types, we will use a multinomial logistic regression to explain each type of character presented in the data. The performance of our model will be directly related to how well the logistic regression for each explanatory variable (each character) explains the true character from the image data.

iii. **Approach:**

To train our model, we will use a logistic loss function with three optimization techniques for comparison. The logistic loss function is (generally) defined as:

$$L(f(\vec{x}), y) = \log(1 + e^{-yf(\vec{x})})$$

This function is convex and continuous, so gradient descent methods can be utilized. However, it is somewhat sensitive to outliers in the data.

We will perform the optimization using the following numerical optimization techniques:

(a) Gradient/Steepest Descent
(b) Stochastic Gradient Descent (SGD)
(c) BFGS

In addition, we will compare the performance of the algorithms in terms of:

(a) Solution performance
(b) CPU time
(c) Storage complexity
(d) Rate of convergence
(e) Iteration complexity