Analysis of Stochastic Gradient Descent on Linear Regression Models

Author: Joseph Cauthen
Homework 1 for CS 4990

Sep 21st, 2019

1 Analysis of Performance for Stochastic Gradient Descent Linear Regression Models

The performance of each model in exercise 4 changes along with the batch size. These changes take place in one of two ways, runtime and loss reduction per epoch. For this analysis, I have prepared a chart that analyzes each of these metrics as the batch size increases from 1 to the size of the dataset. I initialized the weight vector of each model to a fixed number, 5, and trained each model over 100 epochs with a learning rate of 1E-3.

1.1 Runtime

The runtime decreased as the batch size increased. This makes sense as the training algorithm must conduct fewer gradient calculations and alterations as the batch size grows.
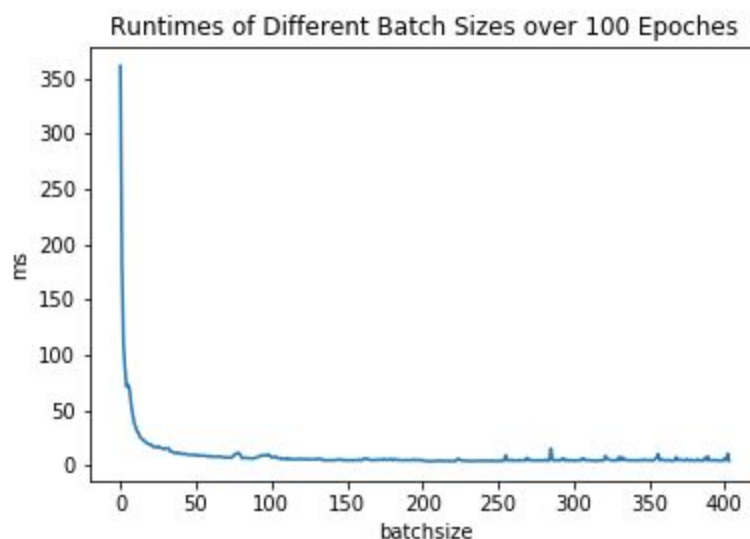

Figure 1: Graph visualizing the runtime reduction as batch sizes increase

1.2 Loss Reduction per Epoch

The loss reduction per epoch went up as the batch size grew. Smaller batch sizes means that the algorithm recalculates the weight vector more times per epoch. As long as the batch adequately reflects the overall data, the weights move in a favorable direction multiple times per epoch. Larger batch sizes reduce the total weight calculations and therefore reduce the accuracy improvement over epochs.
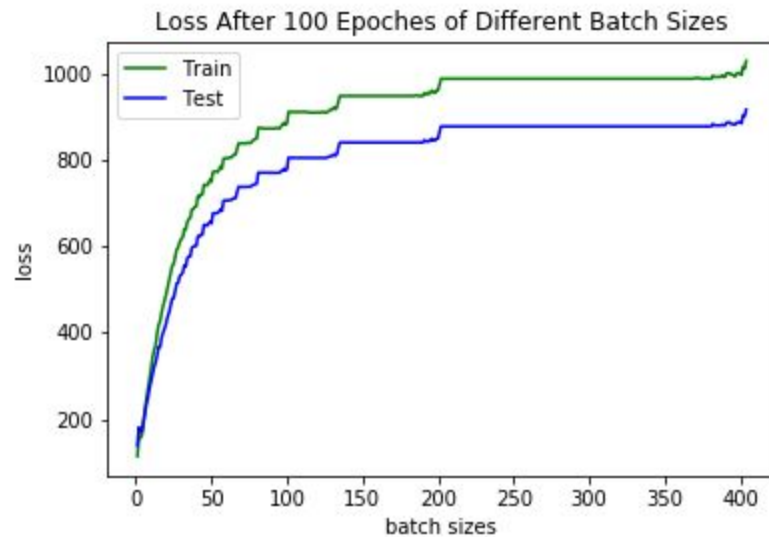

Figure 2: Graph of final loss values after 100 epochs as batch size increases