Oregon State University

CS 534

Implementation Assignment 1

Students:
Trevor Fiez
Ben McCamish

Professor:
Dr. Xiaoli Fern

October 10, 2015

Problem I

First we can look at the gradient and loss for the training dataset, represented in Figures 1 and 2 respectively. This dataset is learning over the course of time, so eventually it should converge, most of the time. We were tasked with setting the alpha, or the learning rate, of the function. As you can see from the graphs, the loss and gradient completely diverge for values of alpha larger than 0.001 and converge at values less than that. The values shown here indicate that the optimal training rate would be $\alpha = 0.001$.

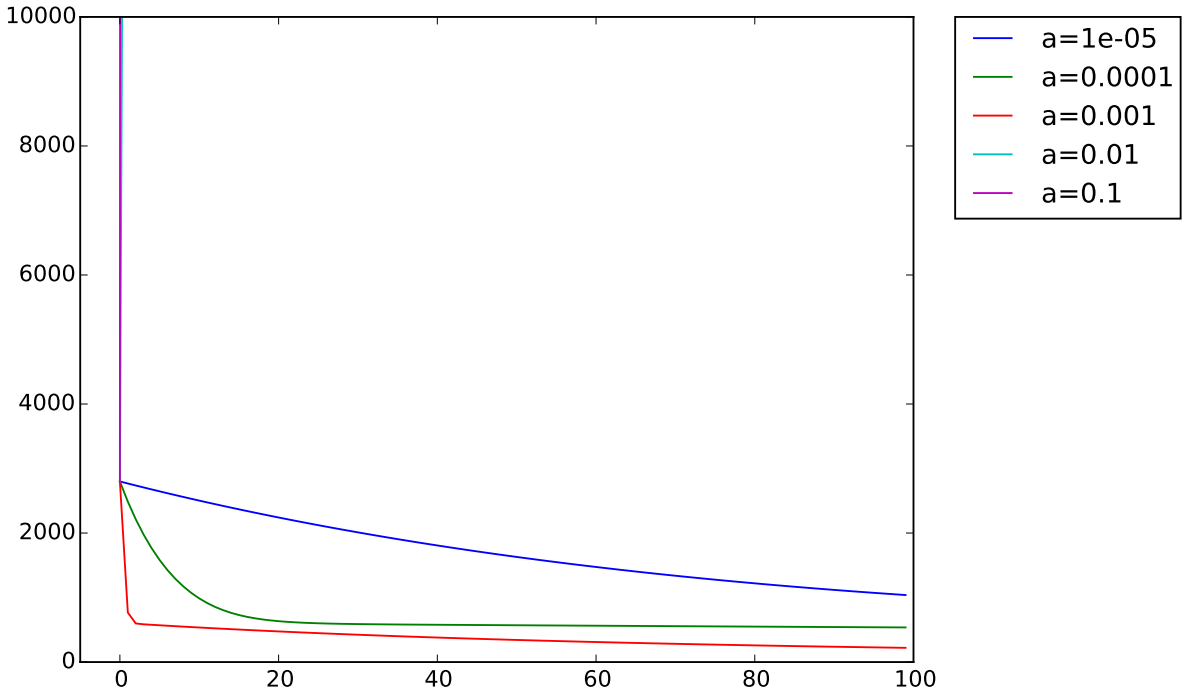

Figure 1: Gradient over time for training dataset.

Figures 3 and 4 illustrate the gradient and loss over the testing data set. It is interesting to note, for α values of 0.0001 and 0.00001, the loss and gradient fluctuate much more than when $\alpha = 0.001$.

Problem II

As expected, we see a concave curve in the Figure 5, indicating where we may experience overfitting or underfitting in the testing dataset. The training dataset has a steady increase in the SSE as the lambda increases. This may be due to the fact that gradient will be increasing since we are giving the weights too much importance, therefor ever increasing the amount of gradient we learn.

Problem III

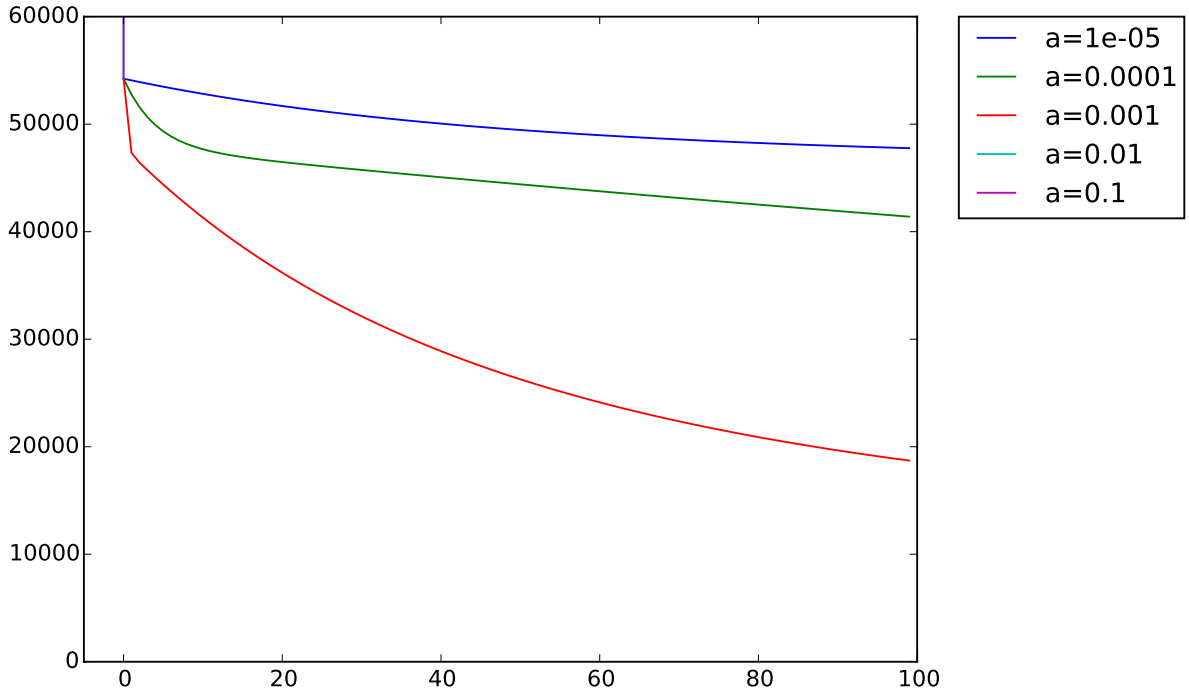

Figure 2: Loss over time for training dataset.

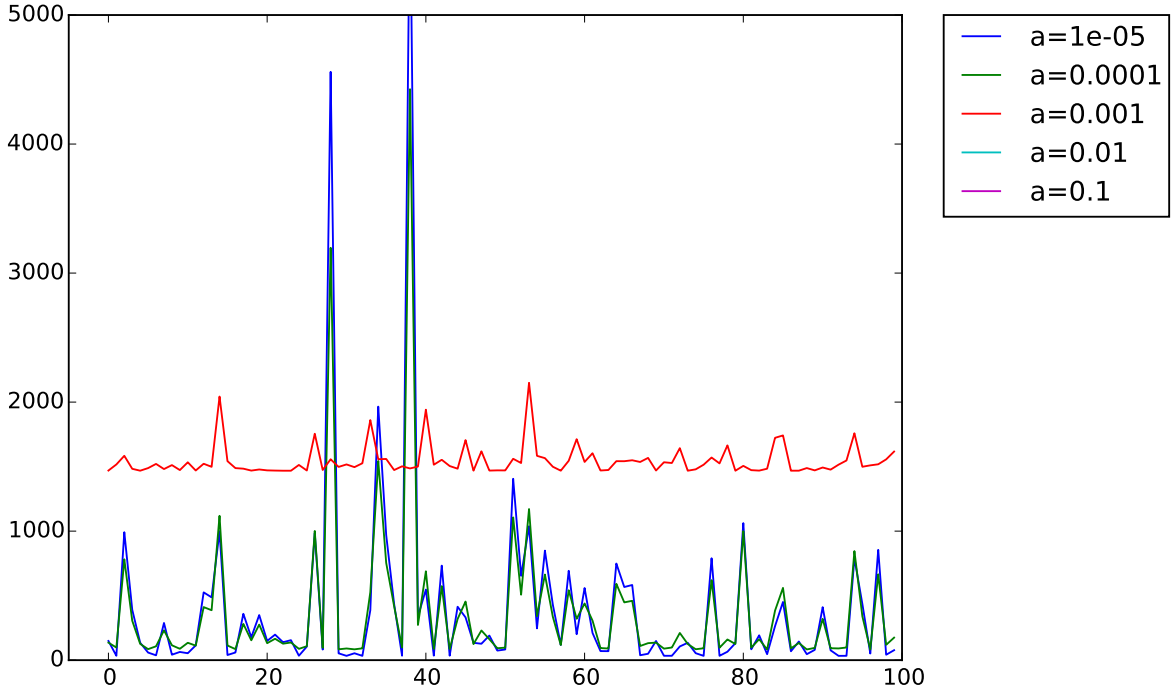

Figure 3: Loss over time for testing dataset.

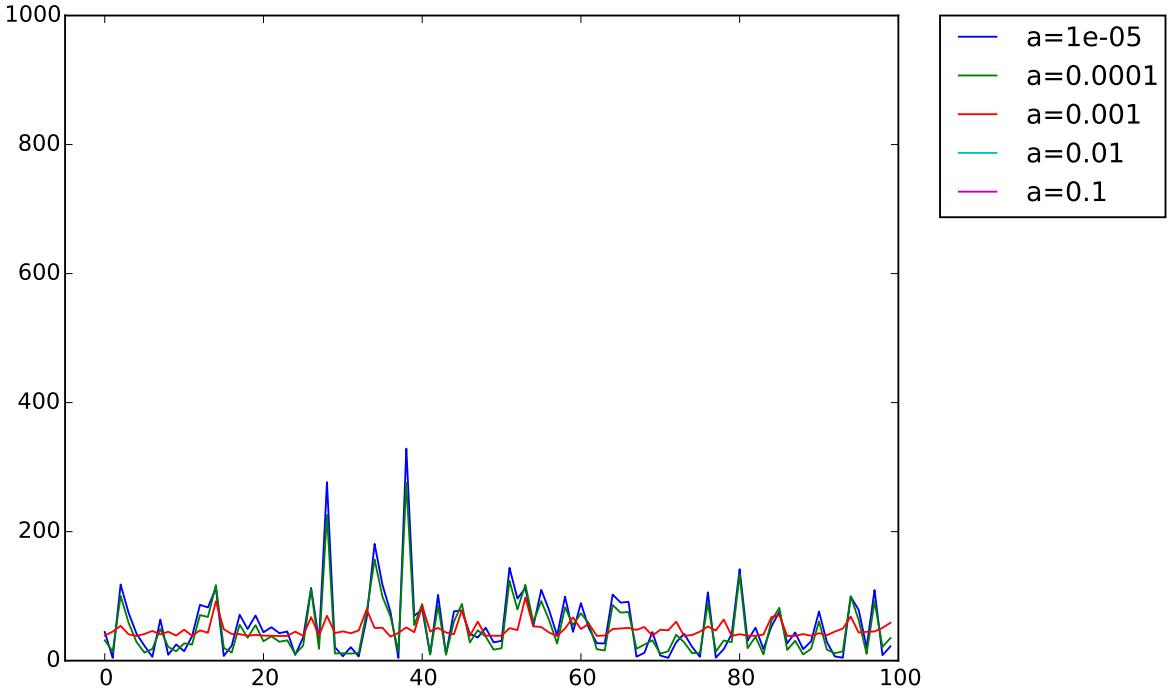

Figure 4: Gradient over time for testing dataset.

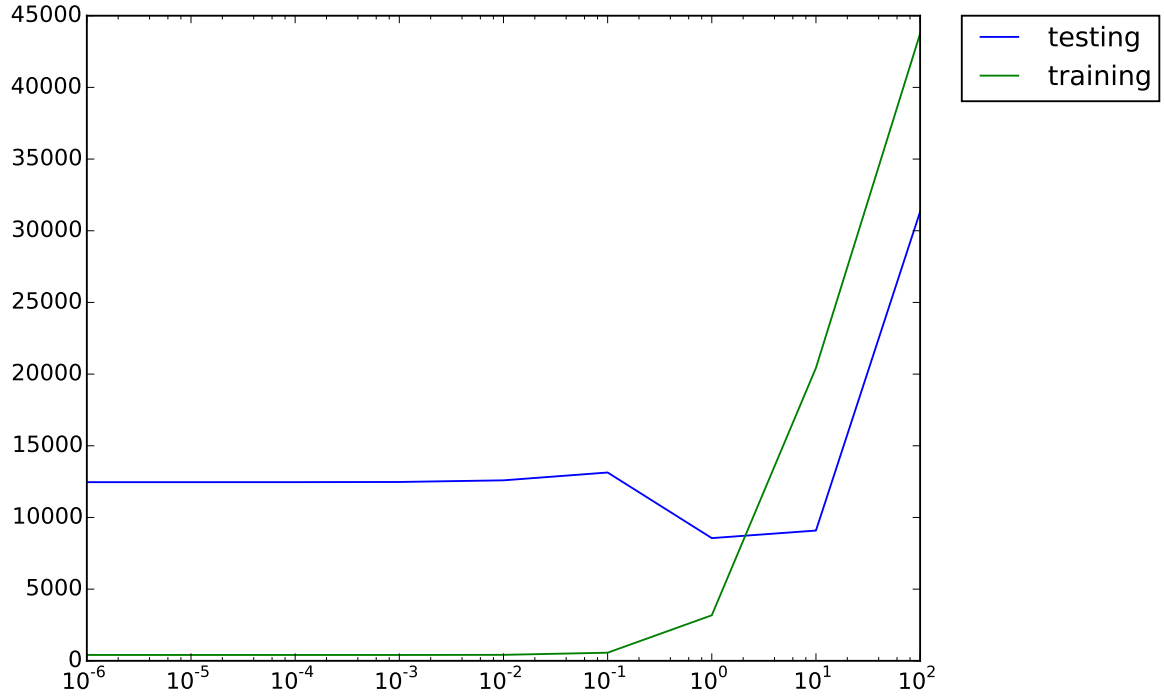

Figure 5: SSE with different lambda