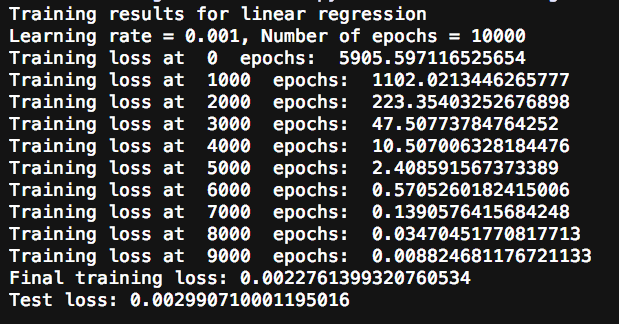
Homework 2

9/29

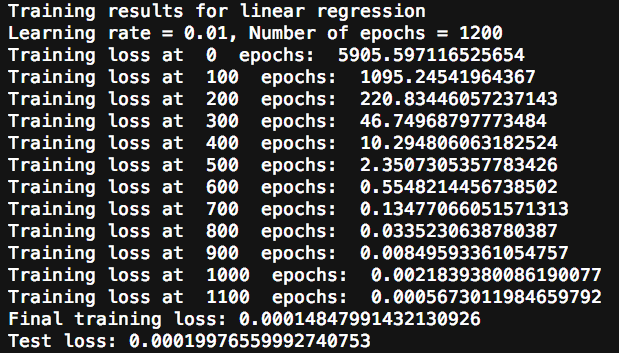
Chu Wang

The initial learning rate was 0.001 and number of epochs 100, which did not reach optimal result. I increased the number of epochs with learning rate = 0.001 and the testing loss converged to 0.0029 which is smaller than(1e-2) after 10000 epochs

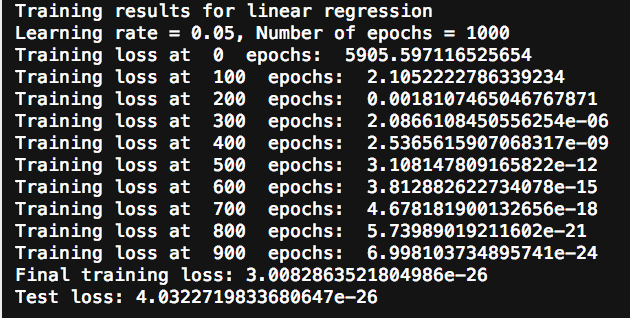


Then I changed the learning rate to 0.01, the convergence became much faster. After 1200 epochs, the testing loss went to 0.00019

Results after training with learning rate = 0.01:



Then I further increased the learning rate to 0.05 and as shown below, the loss reached 0.0018 after 200 epochs:



The relationship between learning rate and epochs

The smaller the learning rate, the slower the loss will converge to global minimum, which means more training epochs are needed to update the weights. With a larger learning rate, fewer training epochs will be needed