

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

```
Training results for linear regression
Learning rate = 0.001, Number of epochs = 10000
Training loss at 0 epochs: 5905.597116525654
Training loss at 1000 epochs: 1102.0213446265777
Training loss at 2000 epochs: 223.35403252676898
Training loss at 3000 epochs: 47.50773784764252
Training loss at 4000 epochs: 10.507006328184476
Training loss at 5000 epochs: 2.408591567373389
Training loss at 6000 epochs: 0.5705260182415006
Training loss at 7000 epochs: 0.1390576415684248
Training loss at 8000 epochs: 0.03470451770817713
Training loss at 9000 epochs: 0.008824681176721133
Final training loss: 0.0022761399320760534
Test loss: 0.002990710001195016
```

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:

```
Training results for linear regression
Learning rate = 0.01, Number of epochs = 1200
Training loss at 0 epochs: 5905.597116525654
Training loss at 100 epochs: 1095.24541964367
Training loss at 200 epochs: 220.83446057237143
Training loss at 300 epochs: 46.74968797773484
Training loss at 400 epochs: 10.294806063182524
Training loss at 500 epochs: 2.3507305357783426
Training loss at 600 epochs: 0.5548214456738502
Training loss at 700 epochs: 0.13477066051571313
Training loss at 800 epochs: 0.0335230638780387
Training loss at 900 epochs: 0.00849593361054757
Training loss at 1000 epochs: 0.0021839380086190077
Training loss at 1100 epochs: 0.0005673011984659792
Final training loss: 0.00014847991432130926
Test loss: 0.00019976559992740753
```

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

```
Training results for linear regression
Learning rate = 0.05, Number of epochs = 1000
Training loss at 0 epochs: 5905.597116525654
Training loss at 100 epochs: 2.1052222786339234
Training loss at 200 epochs: 0.0018107465046767871
Training loss at 300 epochs: 2.0866108450556254e-06
Training loss at 400 epochs: 2.5365615907068317e-09
Training loss at 500 epochs: 3.108147809165822e-12
Training loss at 600 epochs: 3.812882622734078e-15
Training loss at 700 epochs: 4.678181900132656e-18
Training loss at 800 epochs: 5.73989019211602e-21
Training loss at 900 epochs: 6.998103734895741e-24
Final training loss: 3.0082863521804986e-26
Test loss: 4.0322719833680647e-26
```

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