Bryan Faryadi (801178567) - Homework 1

Github repository: https://github.com/bfaryadi/ECGR4105_hw1_bf

1.1 Linear models (4500 iterations, 0.01 training rate)
X1: [w0, w1] = [5.92767927, -2.03823168]
X2: [0.73630934, 0.55751073]
X3: [2.8713337, -0.52044792]

Explanations for these answers can be found by reading the comments in the code.

1.2 Graphs shown in python notebook

1.3 X1 has the lowest loss for explaining the output by a large margin. It settles around 0.985, compared to 3.6 and 3.63 for X2 and X3 respectively.

1.4 Based on my training observations for X1, final loss goes down when training rate or iterations increase. However, the effect of a higher training rate becomes insignificant with enough iterations. For example, at 1500 iterations, a training rate of 0.1 had a lower final loss of roughly 0.005 compared to a training rate of 0.01. However, at 4500 iterations, the difference was only 0.00000001. Results are similar for X2 and X3.

2.1 [w0, w1, w2, w3] = [4.15118728, -1.8394291, 0.72473856, -0.09513266]

2.2 Graph shown in python notebook

2.3 At 4500 iterations, the difference in final loss between a 0.01 and 0.1 training rate is ~0.00018. This is greater than the difference for just X1, but still insignificant. Using 1500 iterations, however, the difference between 0.01 and 0.1 is ~0.072, which is much more significant. We can see that the training rate matters more with a higher number of explanatory variables, and that many more iterations are required in order to begin converging toward a lower final loss value.

2.4 Using the model described in 2.1, here are the predictions:

Prediction for (1, 1, 1): Y = 2.9413640816645295
Prediction for (2, 0, 4): Y = 0.09179843075891547
Prediction for (3, 2, 1): Y = -0.012755552822328653