Elijah Chou

Dr. Li Xiong

CS 334

10/15/2021

HW 3

1. (a): See selFeat.py for Q1.  
   (b): The following heatmap was produced using the q1b.py code included in the zip file. The PNG is saved in the same directory as “correlation\_heatmap.png”.  
   Chart

   Description automatically generated  
   (c) – (d): See selFeat.py for Q1. In addition, for part (d), the preprocessing step I did was to standardize all the data with a StandardScaler from the scikit-learn Python library.
2. (a) – (b): See lr.py and standardLR.py for Q2.
3. (a): See sgdLR.py for Q3.  
   (b): According to the plot generated below, it seems that the best learning rate is either 0.01 or 0.001. This is because both of these converge quite quickly compared to the other lower learning rates tested. The learning rate 1 was not tested or graphed due to overflow errors that occurred. The following graph was generated using the q3\_b.py code included in the zip file turned in on Canvas. The PNG file is saved as “trainmse\_vs\_epoch.png” in the same directory after running q3\_b.py.  
   Chart, line chart

   Description automatically generated  
     
   (c): The following graph was generated by the q3\_c.py code included in the turned-in zip file. It is saved as “train\_test\_mse\_vs\_epoch.png” in the same directory. The learning rate of 0.001 was used to train the model.  
   A picture containing chart

   Description automatically generated
4. (a): The following graph was produced by the q4.py file that was included in the zip file turned in for this homework assignment. It saves the graph as a PNG named “batch\_size\_effect.png” in the same directory.  
   Chart, scatter chart

   Description automatically generated  
   (b): By looking at the graph, the trade-off in terms of batch sizes is that the smaller the batch size, the more time it takes. However, it also shows that the smaller batch sizes also tend to have a much smaller MSE on both the training and test sets. Essentially the trade-off here is whether to choose a quicker runtime or a smaller MSE overall. Compared to the graphed closed form solution, it seems that the smaller batch sizes, while taking more time, could potentially lead to a lower training/test MSE.