Bryan Faryadi (801178567) - Homework 2

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

Problem 1

In both 1a and 1b, the model could not converge, and instead approached infinity after few iterations.

Problem 2

2a. The training results for normalized vs. standardized data were close, and there was not a clear winner. The standardized data resulted in a lower cost for the training data, but a higher cost for the validation data. The normalized data has its training and validations costs closer to equivalent, so I would be inclined to choose that as the best training method.

2b. Similar to 2a, the standardized data had a lower training cost and higher validation cost. However, the final costs for training and validation sets were much lower than for 2a across the board. This time, the standardized data had a smaller gap between the training and validation costs than the normalized data, which is the opposite of what I observed in 2a. Because of that, I would choose standardization as the better scaling method here.

Clearly, both normalization and standardization result in much better training processes, as the costs did not even converge in problem 1.

Problem 3

For both 3a and 3b, adding parameters penalization majorly increased the final cost for both the normalized and standardized datasets. In the case of the standardized examples, the validation loss converged in a very strange manner. The training loss looks normal, but the validation loss dips down and then up before converging while increasing instead of while decreasing. I am not sure why this is happening, but the final loss is still not radically different from that of the training set.