Bryan Faryadi (801178567) - Homework 2

Github repository: [https://github.com/bfaryadi/ECGR4105\_hw2\_bf](https://github.com/bfaryadi/ECGR4105_hw1_bf)

Problem 1

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

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 every example in problem 3, introducing a parameters penalty increased the training cost but decreased the validation cost compared to problem 2. This indicates that overfitting is successfully being prevented via the penalty. The difference is very slight in all cases, but it is worth nothing that the penalty had a greater effect on the normalized data than the standardized data.