

Homework#6

Yuhang Peng

5.10

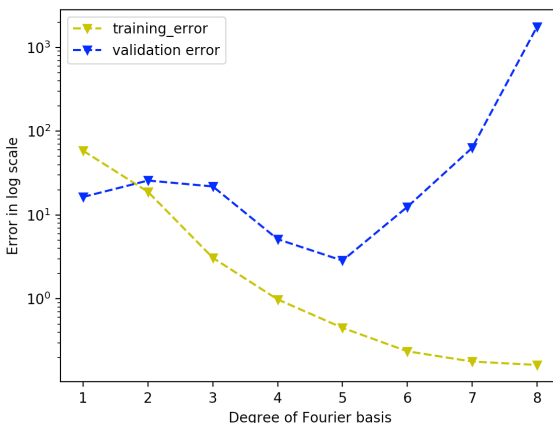
Eric: From the result plot, $D = 5$ provided the best fit. This model provided the smallest testing error. Next step, he need to train again on the entire dataset based on this model and observe the result.

Stanley: The $D = 3$, $D = 8$ can be selected be the model. However, the under-fitting and over-fitting should be considered at here. One suggestion is k-fold cross-validation, it may take a little bit longer, but the result is much better. After this, compare the result again, select the best one, and trained on the entire dataset on the selected model. If the result doesn't change compare with hold out cross-validation, just select the $D = 8$, because the training error is smaller. Then, train again on the entire dataset.

Kyle: From the plot, it doesn't show the minimal error point in the testing error curve. The training model(Degree) is not enough, add more model (15-20) and observe the result again.

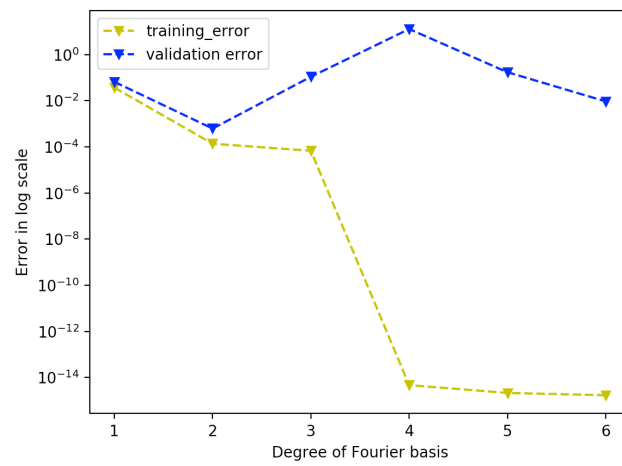
Kenneth: From the result plot, $D = 6$ provided the best fit. This model provided the smallest testing error. Next step, he need to train again on the entire dataset based on this model and observe the result.

5.11



Because each time, the testing data will be random selected, the figure will be a little bit different. After compare several results, it showed that the $D = 5$ is the best model that fit to the data.

5.12



From the plot, it is easy to find that $D = 2$ is the best model that fit the dataset.