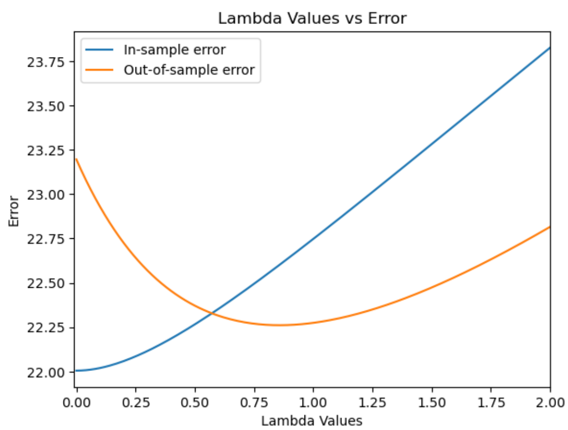
Prog4 Report - CSCD 484

Nate Wilson

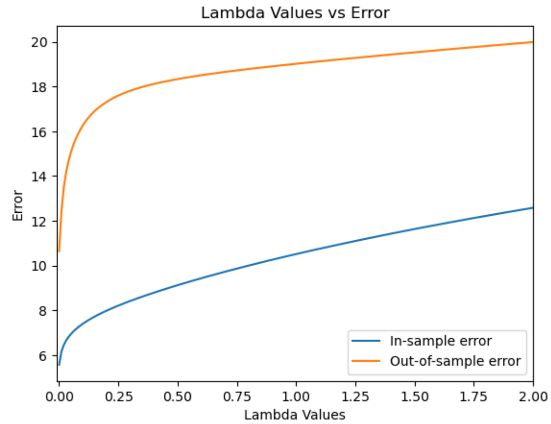
February 19, 2024

Closed Form Parameter Tuning

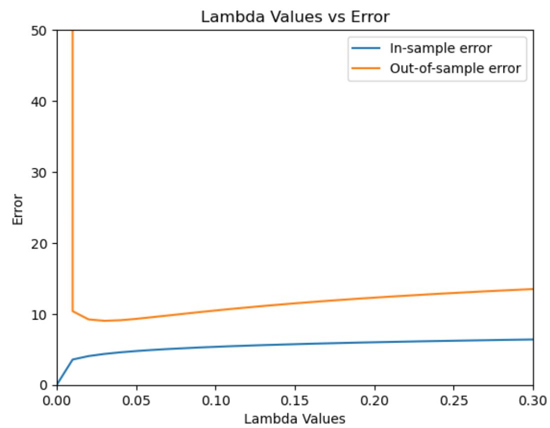
For each graph below, the degree space increases by 1, while iterating through different lambda values. 201 total lambda values are checked on each graph with a step size of 1/100 between the interval $[0, 2]$. Some graphs may not show the complete interval as they have been zoomed in to provide a better view of the curves in the main area of interest.



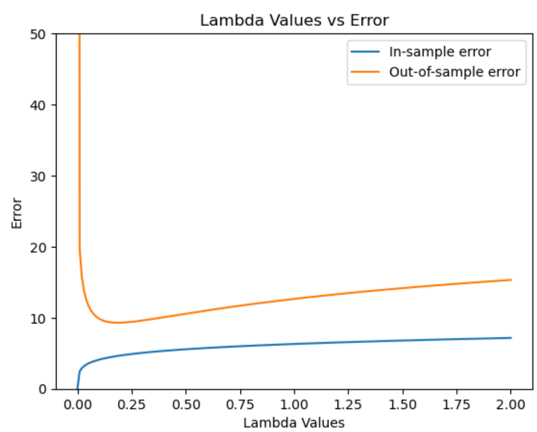
Lambda vs Error in Degree Space 1



Lambda vs Error in Degree Space 2



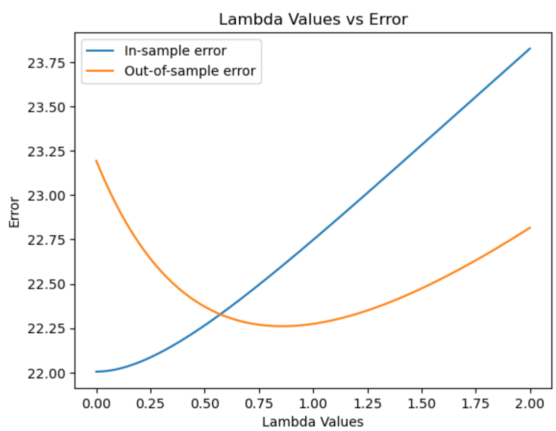
Lambda vs Error in Degree Space 3



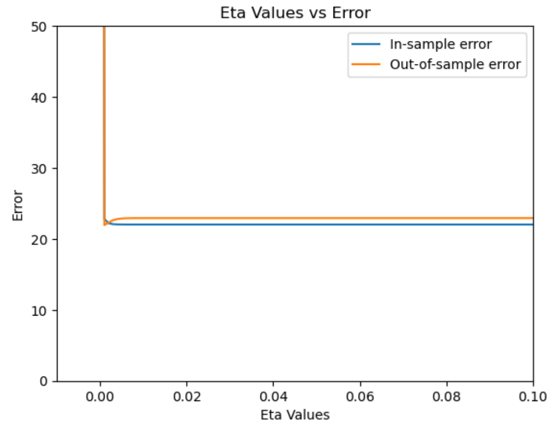
Lambda vs Error in Degree Space 4

Gradient Descent Parameter Tuning

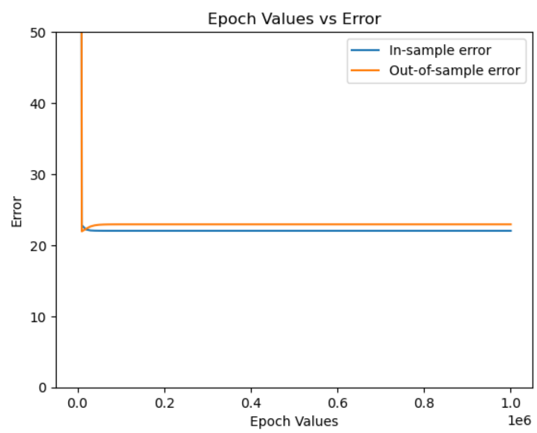
Gradient Descent parameter tuning was a bit more difficult due to the number of parameters that can be changed. Since we have 4 tunable parameters, we would need a 4D graph (5D if you include error) to be able to plot each variable against the other. That being said, we can test each parameter against the error in a respective z-space while holding all the other variables constant to get an idea of what is happening. We start in degree space 1 and plot lambda vs error, eta vs error, and epochs vs error, then switch to degree space 2, 3, and 4 repeating the same pattern. When tuning the parameters, all other variables that weren't being changed were held constant with the values, $\lambda = 0.1$, $\eta = 0.01$, and epochs = 100000 in the first 3 degree spaces. For λ, the step size is the same at 1/100 on the interval $[0, 2]$. η had a step size of 1/1000 on the interval $[0, 0.1]$, and epochs had a step size of $10,000$ on $[0, 1000000]$. Due to the time required to be able to run the model with the given parameters in degree space 4, some intervals had to be lowered in order to get a result in a somewhat time friendly manner. These intervals were lowered to $[0, 0.5]$ for λ, and $[0, 500000]$ for epochs, while η had the same interval. Ideally, we would want to be able to test all degree spaces with the same constant values but we can still gain some insight into whether or not the model is proficient in the 4th degree space using this testing. In a perfect world, we would also want to be able to test all variables against each other at the same time, which would give us the most ideal model.



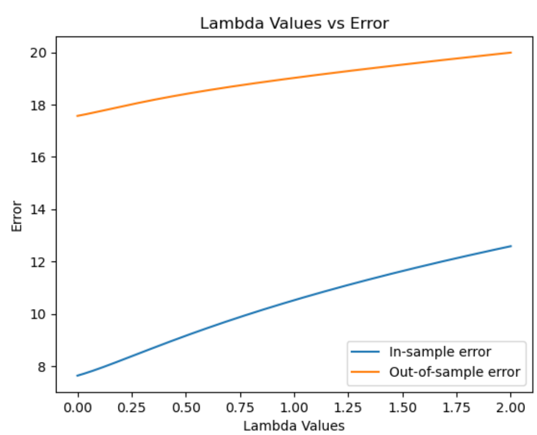
Lambda vs Error in Degree Space 1



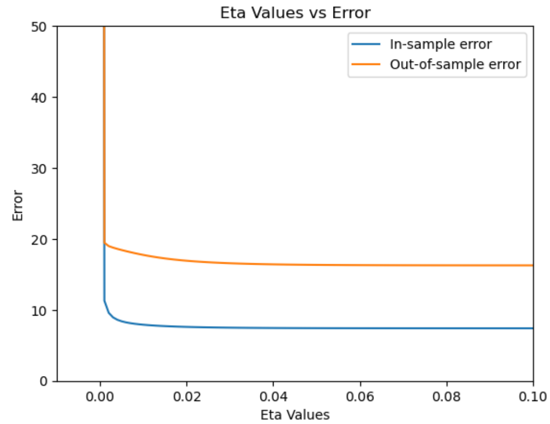
Eta vs Error in Degree Space 1



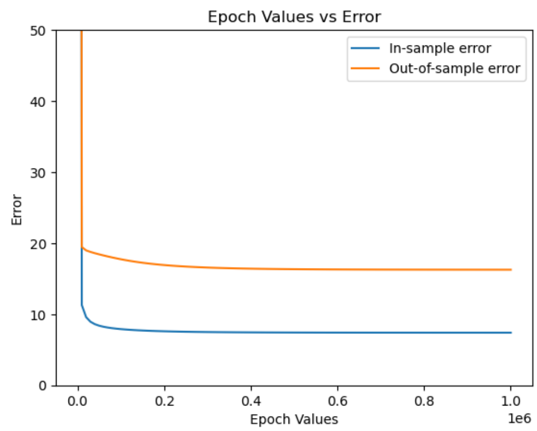
Epochs vs Error in Degree Space 1



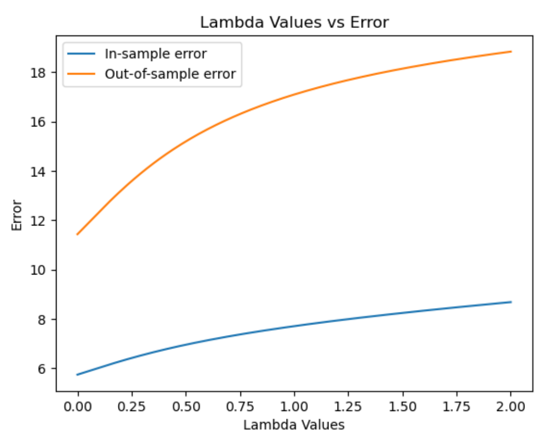
Lambda vs Error in Degree Space 2



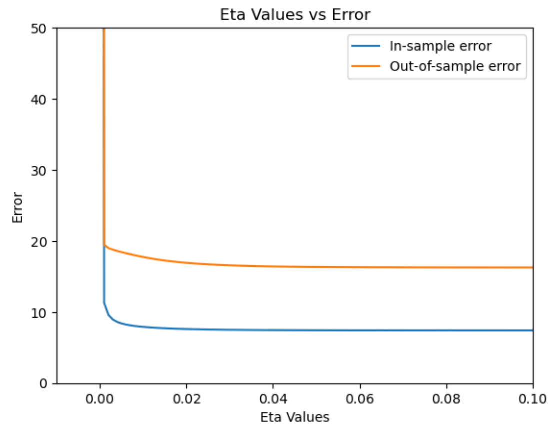
Eta vs Error in Degree Space 2



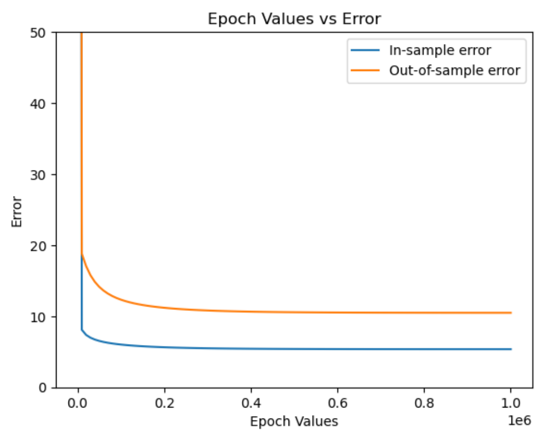
Epochs vs Error in Degree Space 2



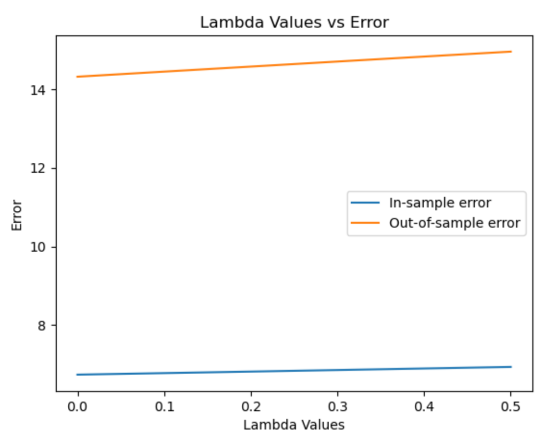
Lambda vs Error in Degree Space 3



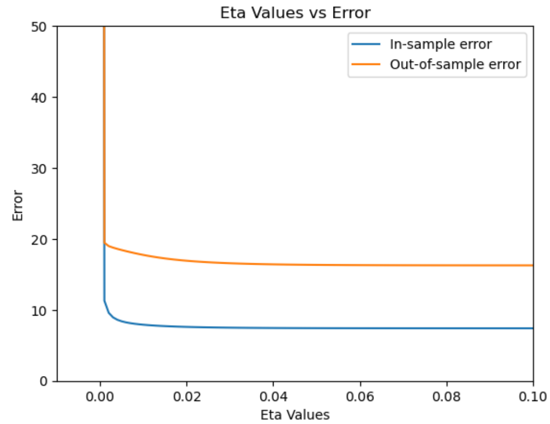
Eta vs Error in Degree Space 3



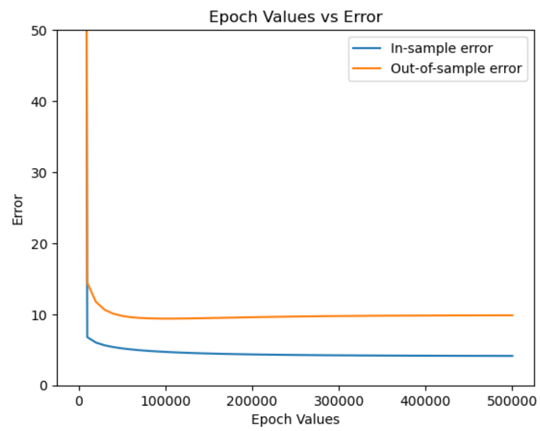
Epochs vs Error in Degree Space 3



Lambda vs Error in Degree Space 4



Eta vs Error in Degree Space 4



Epochs vs Error in Degree Space 4

Model Selection

For the closed form method, we would want to select a model that has $\lambda \approx 0.03$ in degree space 3. For the gradient descent method, we would choose a model that has $\lambda \approx 0.25$, $\eta \approx 0.001$, and epochs ≈ 1000000 in degree space 3. These models minimize the out-of-sample error while also keeping the in-sample errors minimal to prevent overfitting or underfitting of the data. (The actual w vectors for each model were not included as they have 560 features)