

CSCD 484

Hammed Gafar

Subproject 5

Methods:

For this project, the house-price dataset was used, and both the training and validation datasets were normalized using the normalization in `utils.py` in the code provided in the assignment.

Different combinations of hyperparameters were used to train the model, and the hyperparameter combinations that significantly reduced validation error was selected as the best.

For the Closed form, the hyperparameters are

- The degree of Z-space
- λ

For Gradient Descent, the hyperparameters are

- The degree of Z-space
- λ
- η (learning rate)
- The number of epochs

An automated script was used to test which hyperparameter combination significantly reduced the validation dataset error. The various hyperparameter combinations, sample error, and validation error were output in a CSV file for ease of visualization.

The closed form hyperparameter combinations are:

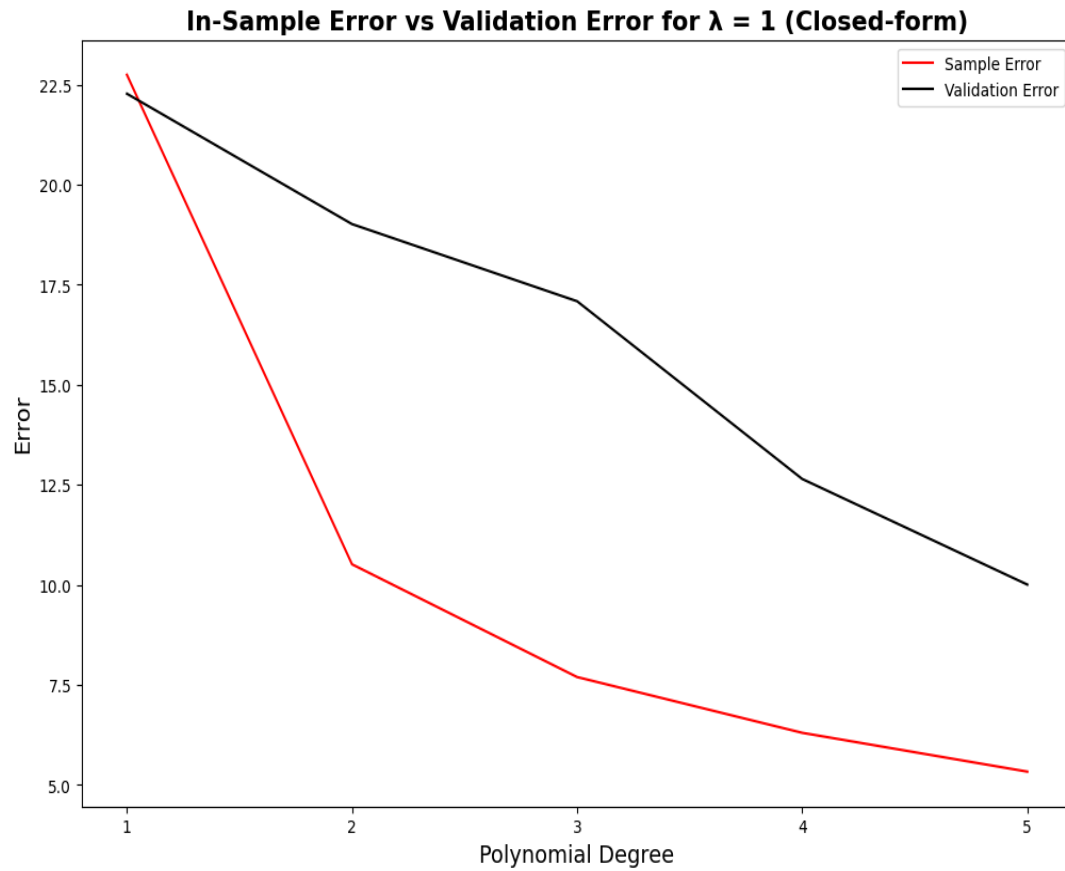
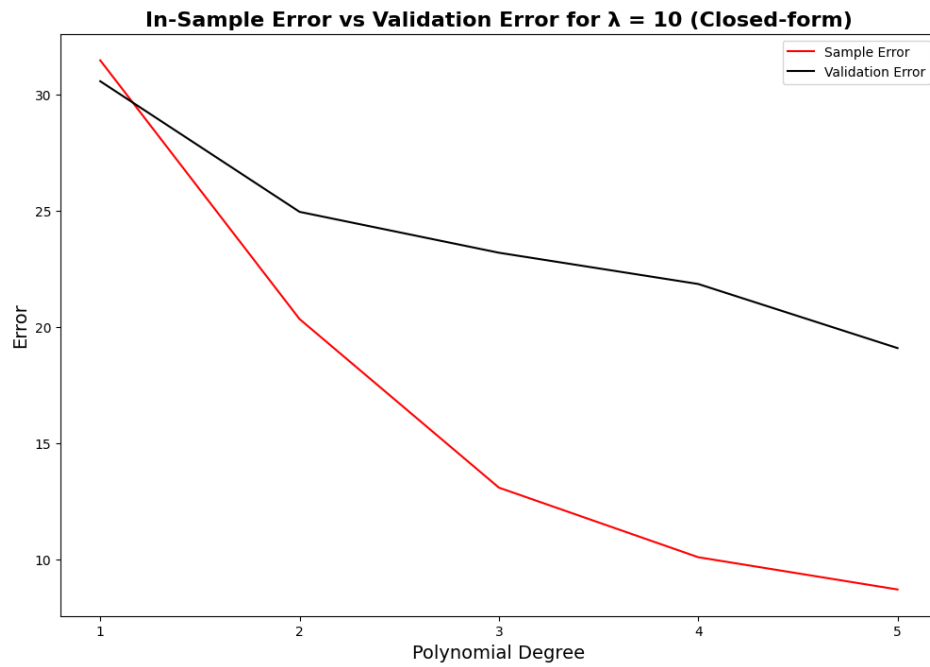
- Z-transform Degree: 1, 2, 3, 4, 5
- λ : 10, 1, 0.1, 0.01, 0.001

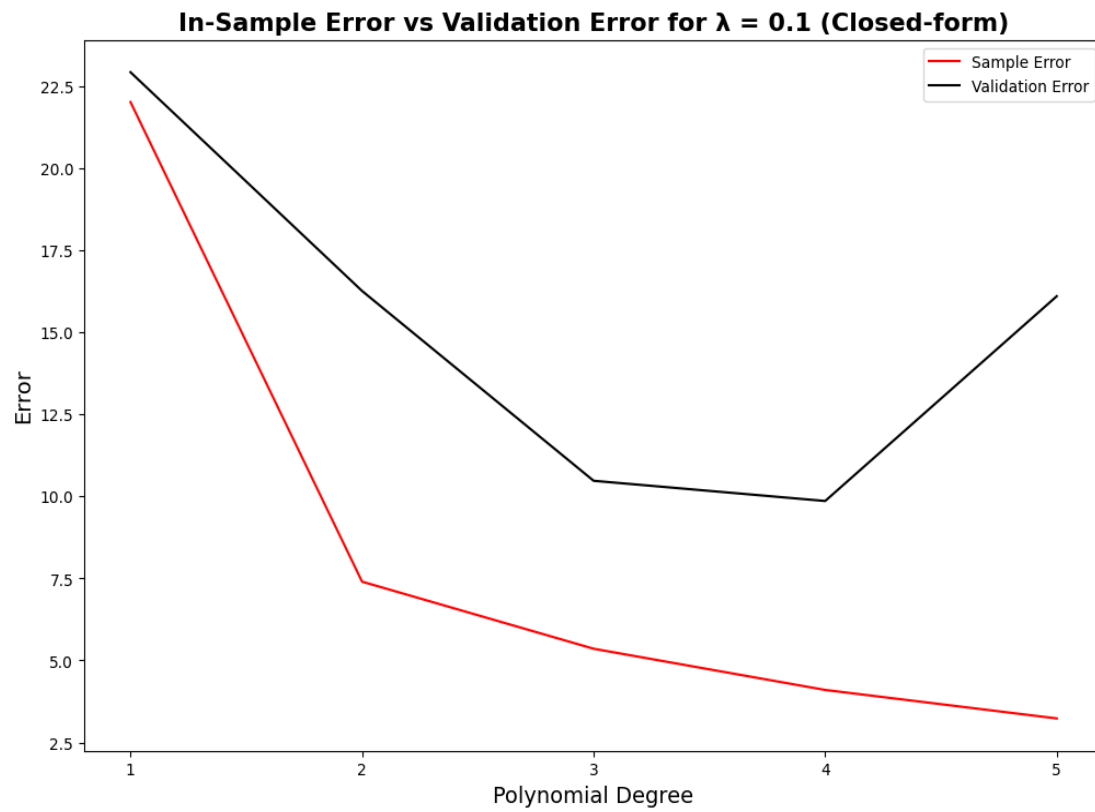
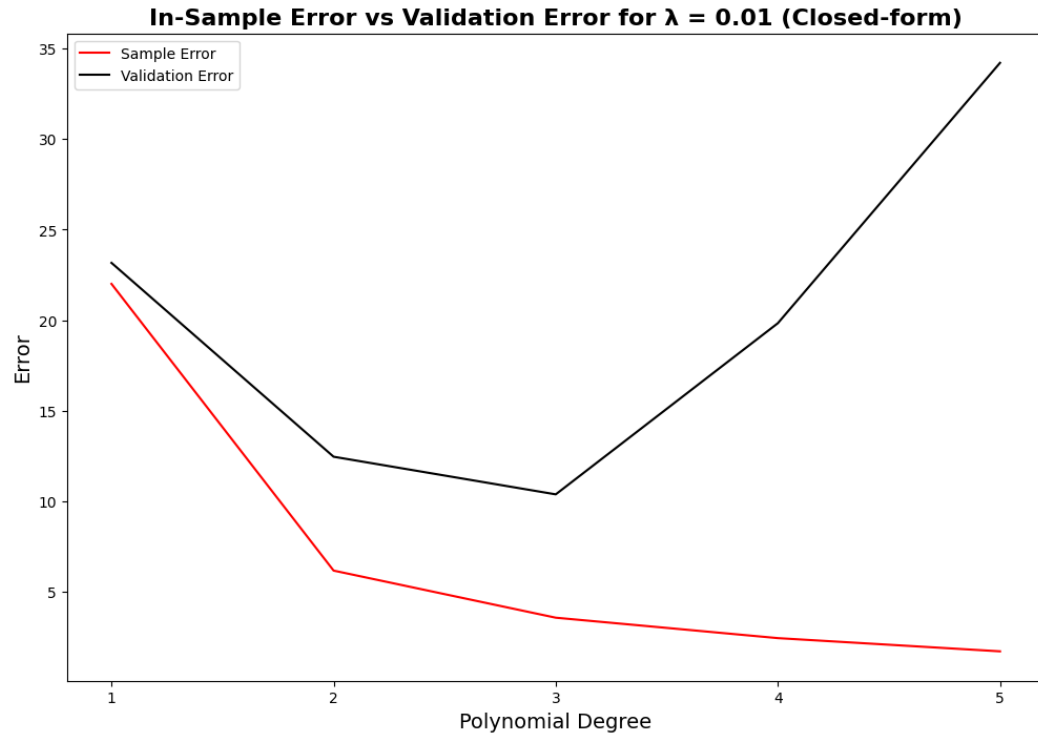
The gradient hyperparameter combinations are:

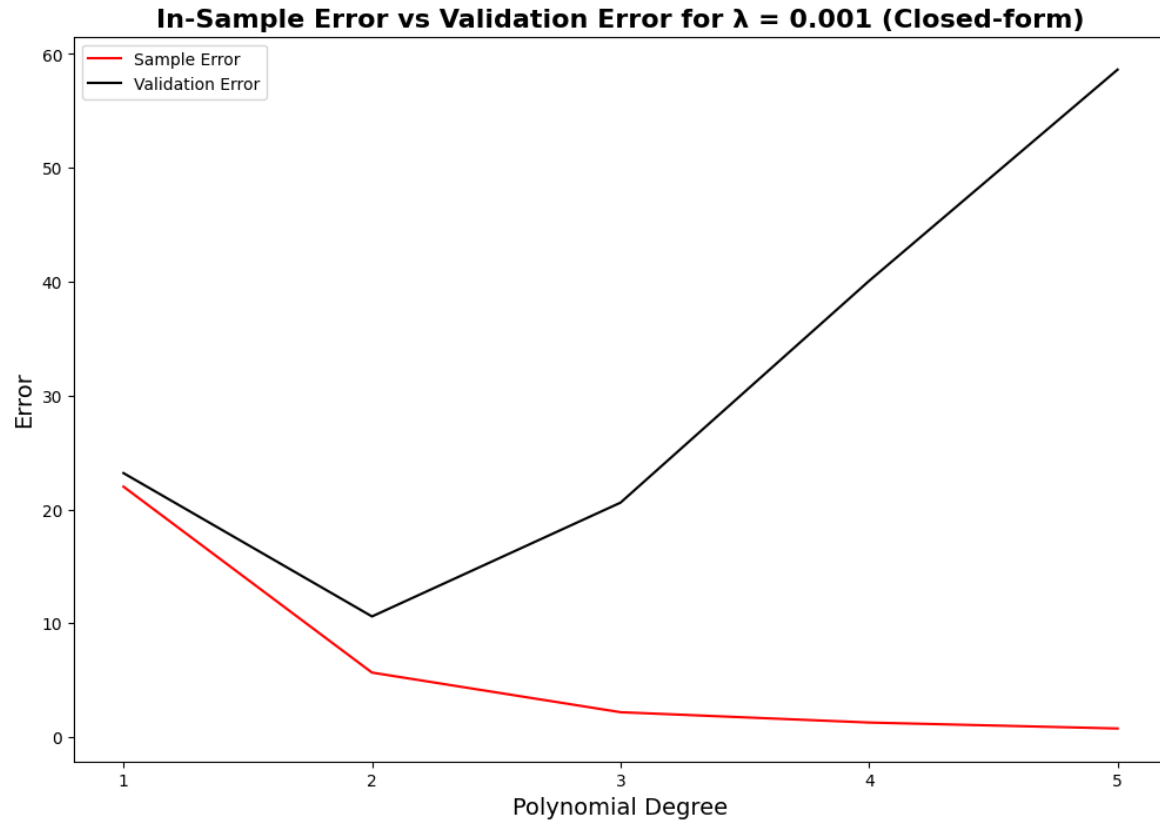
- λ : 10, 1, 0.1, 0.01, 0.001
- Degree: 1, 2, 3, 4
- Learning rates: 0.01, 0.001
- Epochs: 1000

Note: I did not test a wide range of values due to hardware limits.

CLOSED-FORM





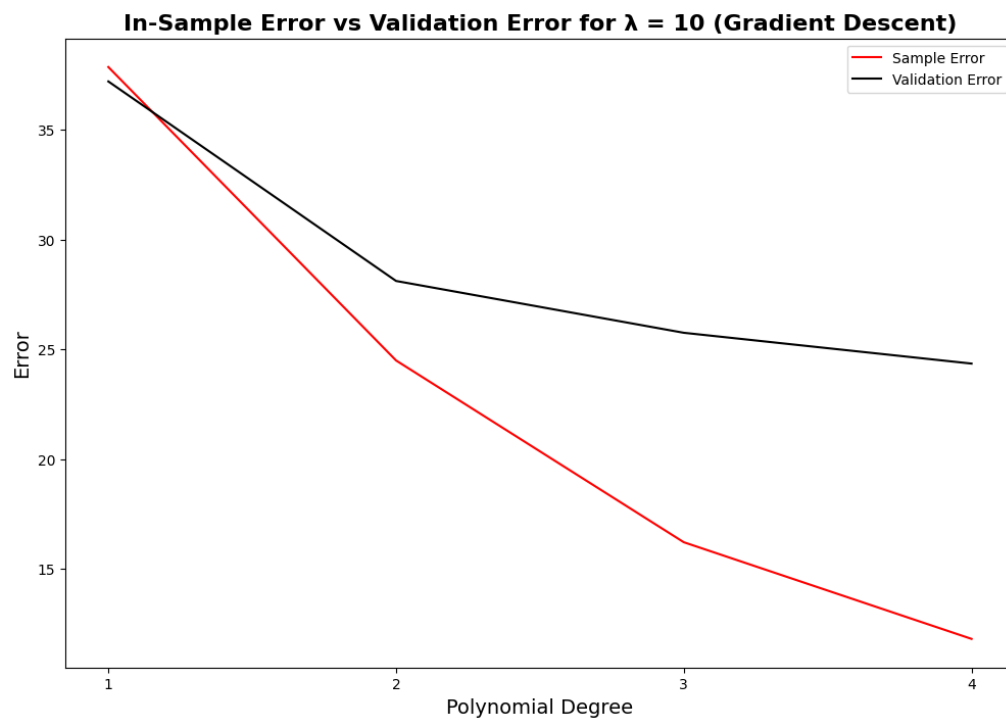
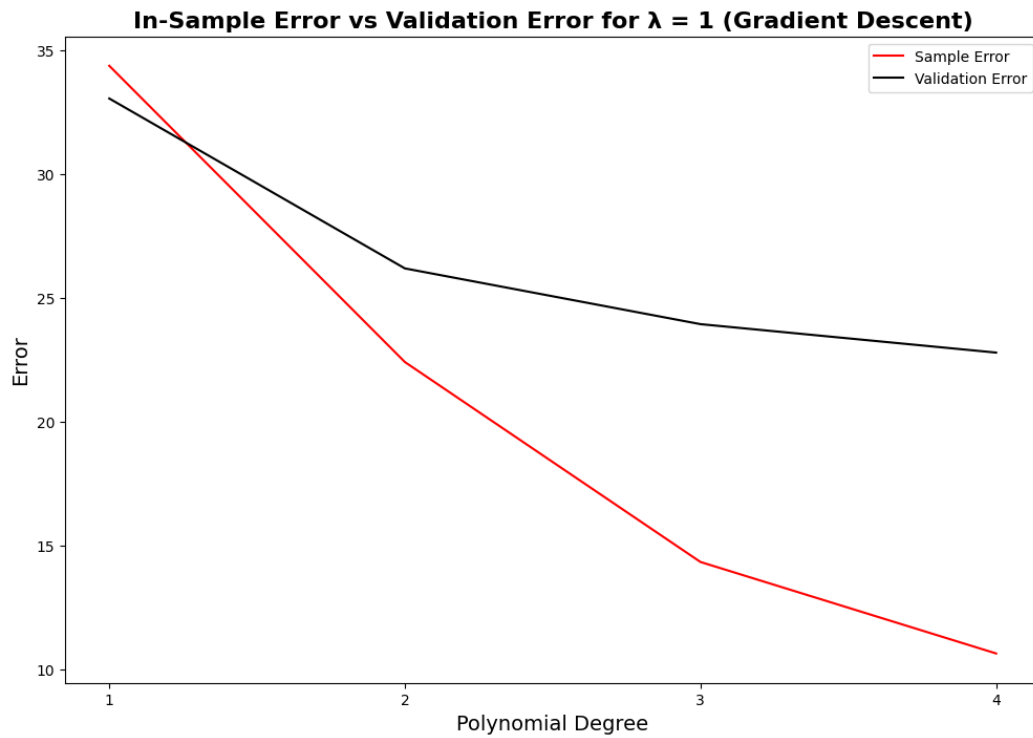


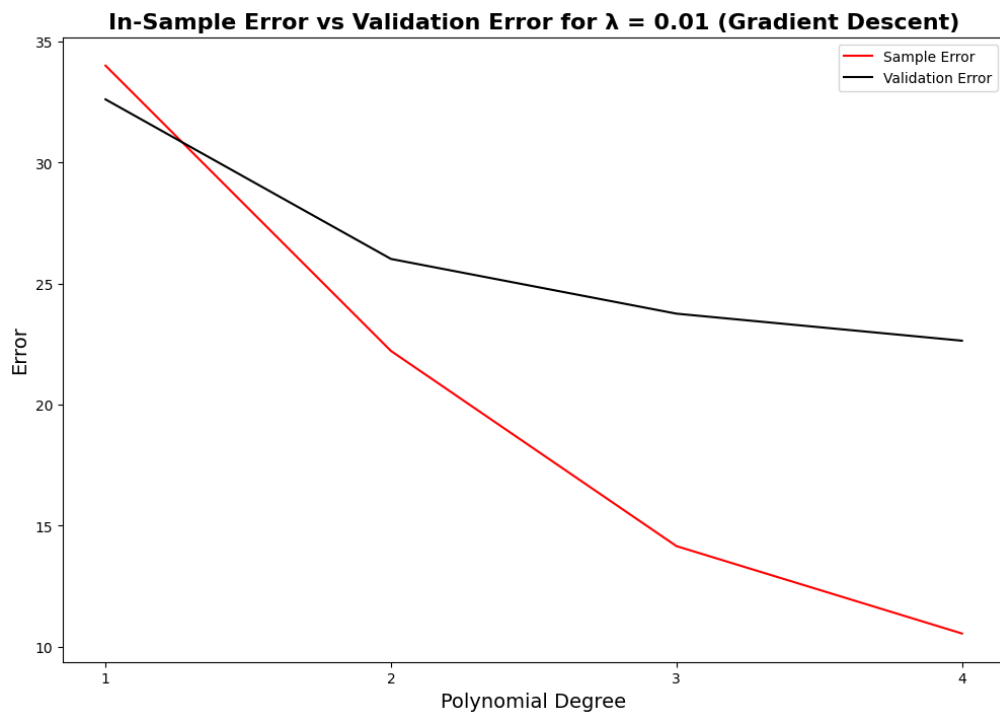
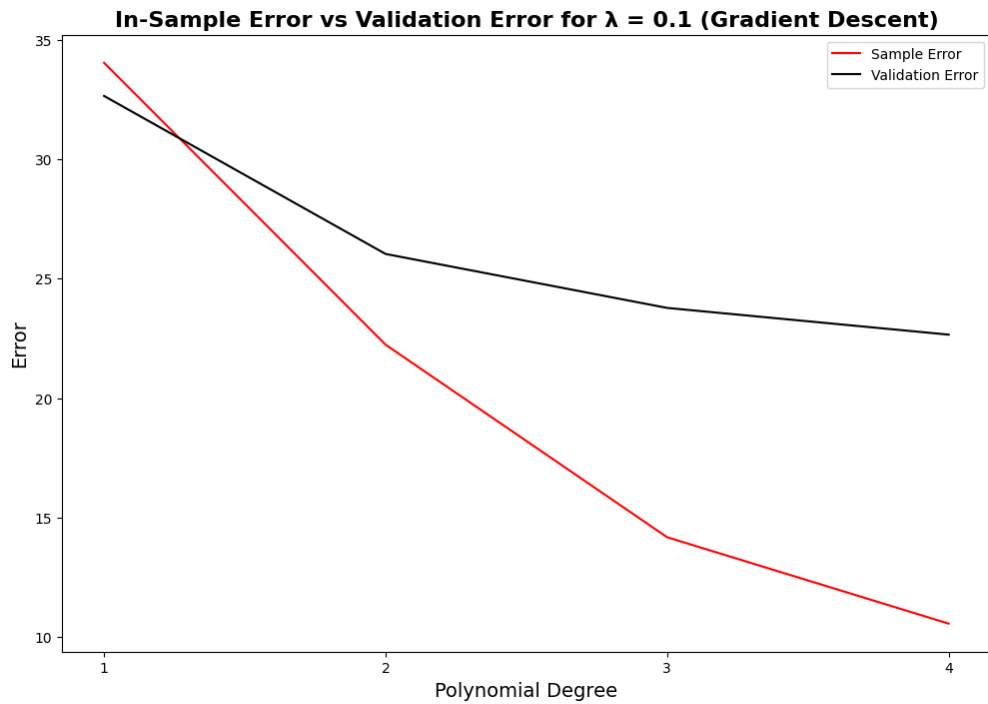
Based on the observation of the figures above. As λ approaches 0, the validation error went up at a lower degree so the optimal λ is between 10 and 1. Observing both plots ($\lambda = 1$ and 10), $\lambda=1$ seems to be the best because at $\lambda=10$ the validation error does not seem to decrease below 20 compared to $\lambda=1$ which decreases the validation error below 12.5. Thus, $\lambda=1$ can further reduce the validation error

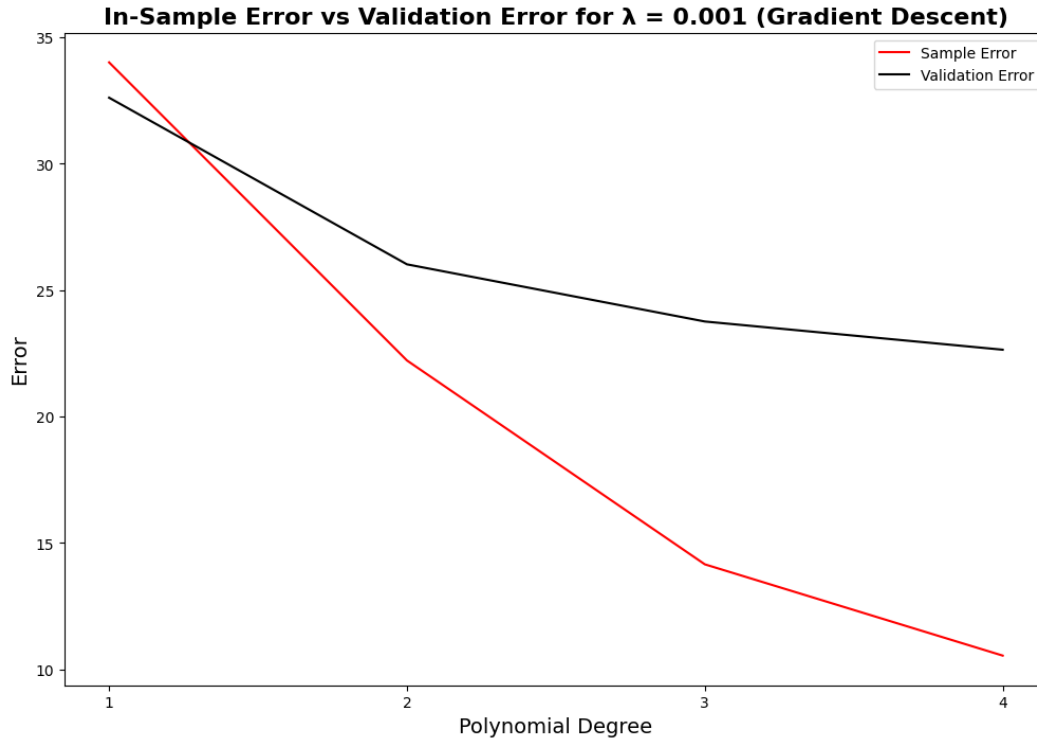
I am unable to go beyond degree 5 due to hardware limits. In conclusion, the best hyperparameters for the closed form based on the house-price dataset are:

$\lambda = 1$ and Degree = 5

GRADIENT DESCENT







Based on the observation of the output CSV file, the epoch used was 1000 which indicates I need a longer walk to reach the minimum because of this, the final learning rate considered for further analysis is 0.01. Keeping that in mind, Observing the figures above lambda does not seem to significantly affect the validation error, and transforming the features to a higher dimension reduced the error. In conclusion, the best hyperparameters for the gradient descent form based on the housing dataset are:

Lamda = 0.1, learning rate = 0.01, degree = 4, epoch = 1000