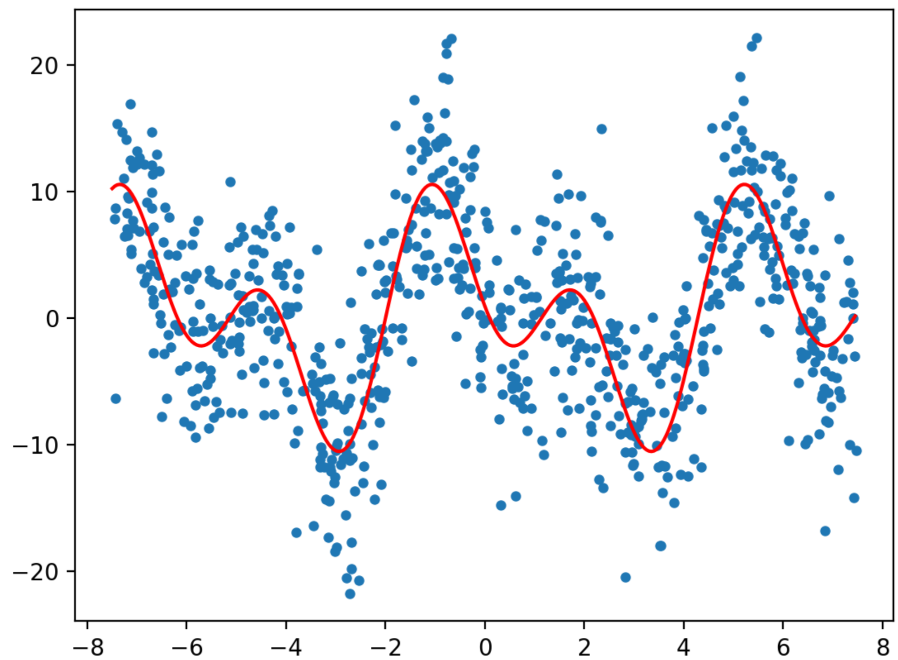
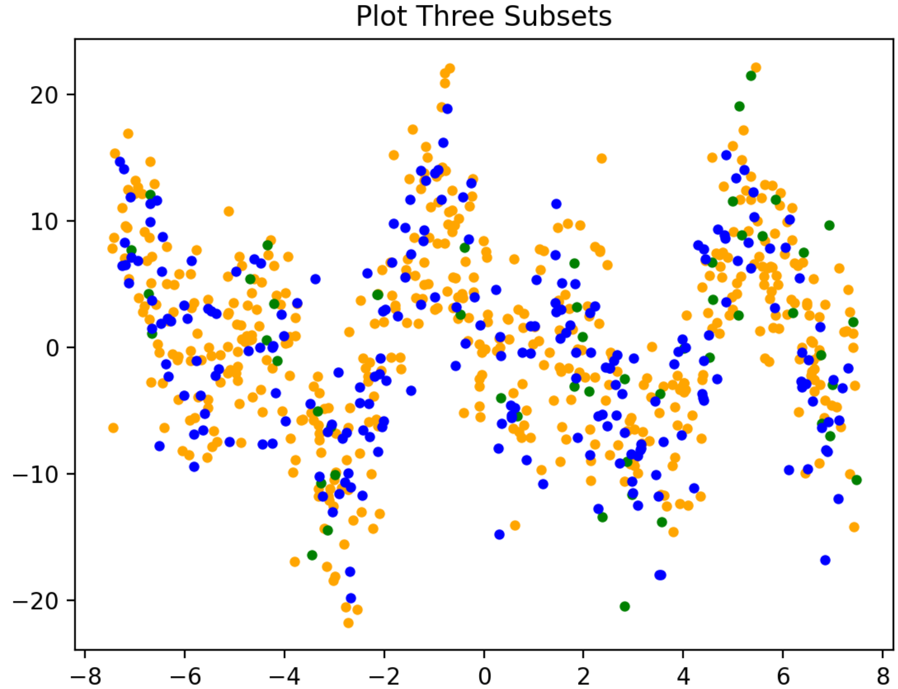
Assignment 1

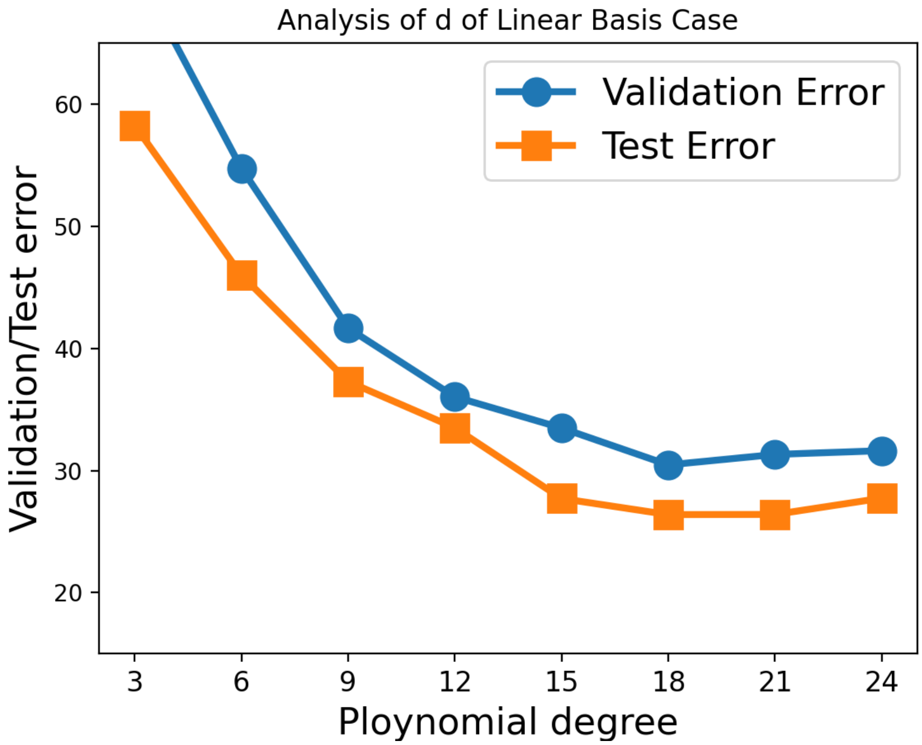
Images

Plot the Data

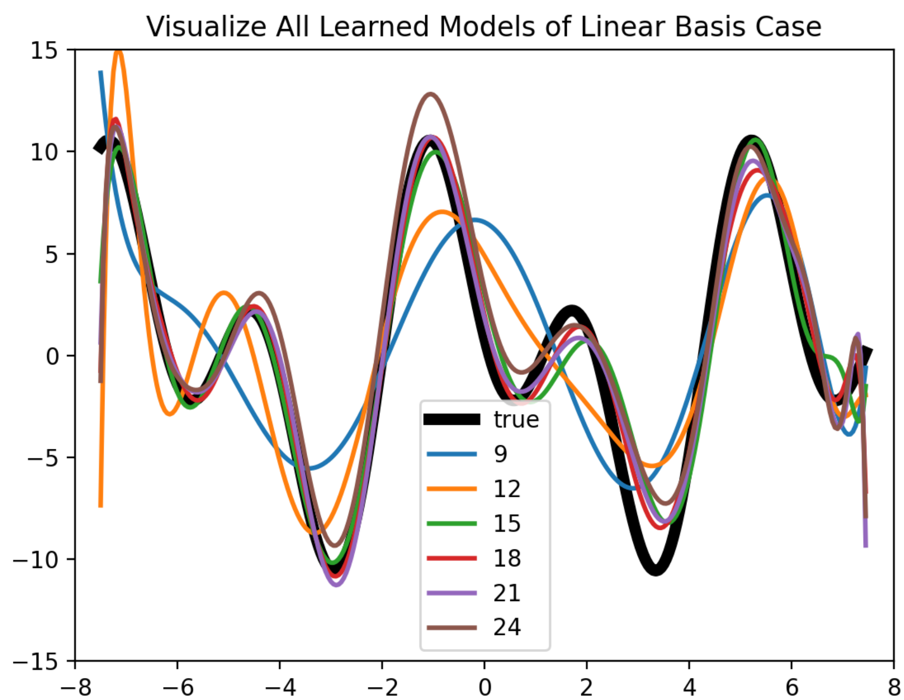


Plot Three Subsets

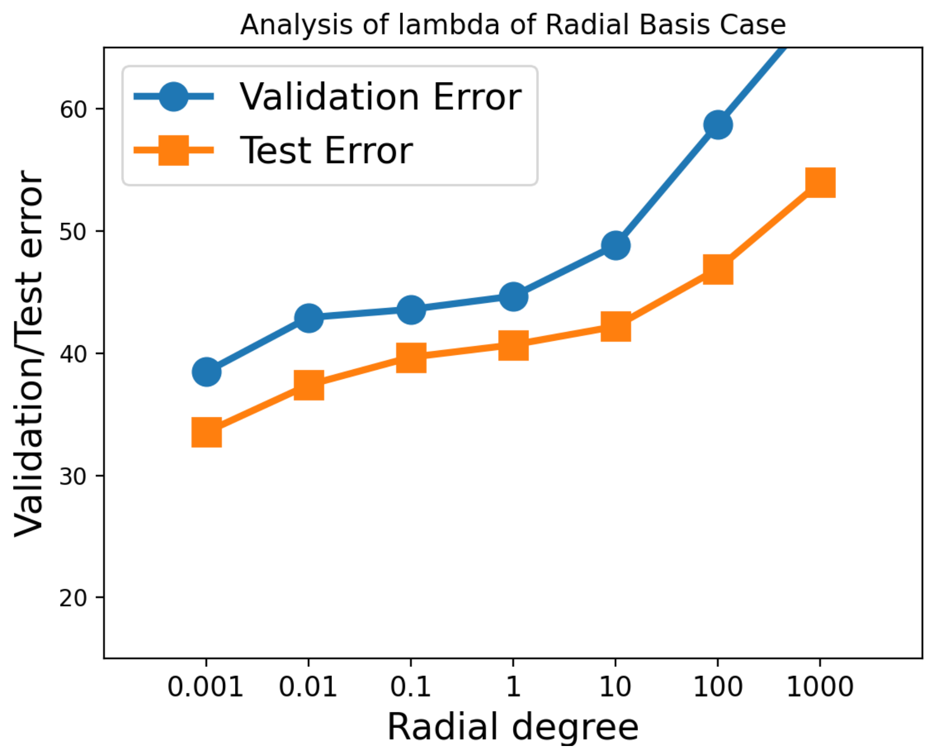


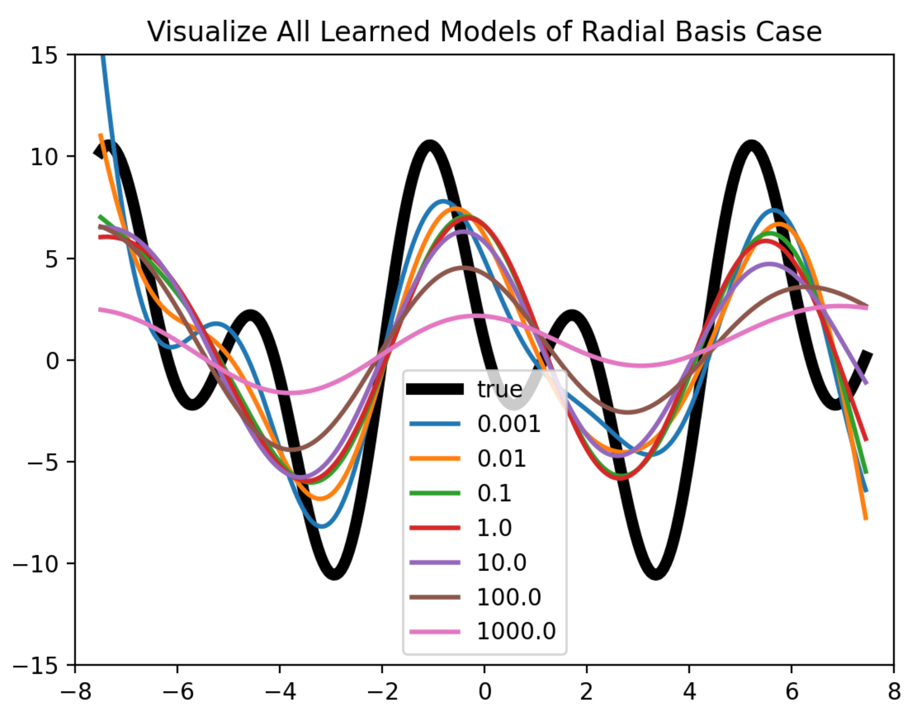


Analysis of d of Linear Basis Case
Validation Error
Test Error
Validation/Test error
Ploynomial degree
20
30
40
50
60
3
6
9
12
15
18
21
24



Visualize All Learned Models of Linear Basis Case
true
9
12
15
18
21
24
15
10
5
0
−5
−10
−15
−8
−6
−4
−2
0
2
4
6
8

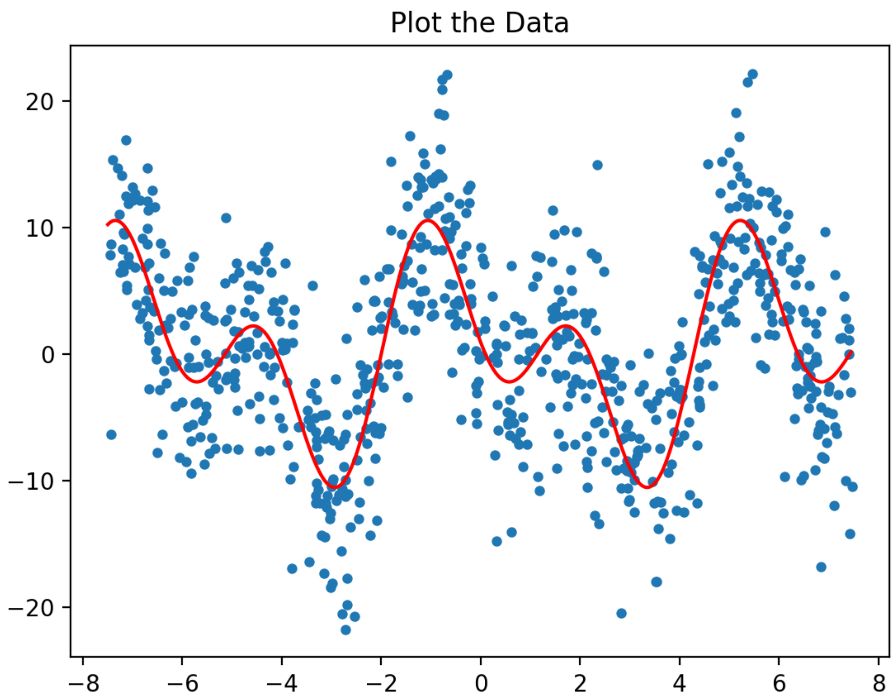

Analysis of lambda of Radial Basis Case
Validation Error
Test Error
Validation/Test error
Radial degree
0.001
0.01
0.1
1
10
100
1000


Visualize All Learned Models of Radial Basis Case
true
0.001
0.01
0.1
1.0
10.0
100.0
1000.0

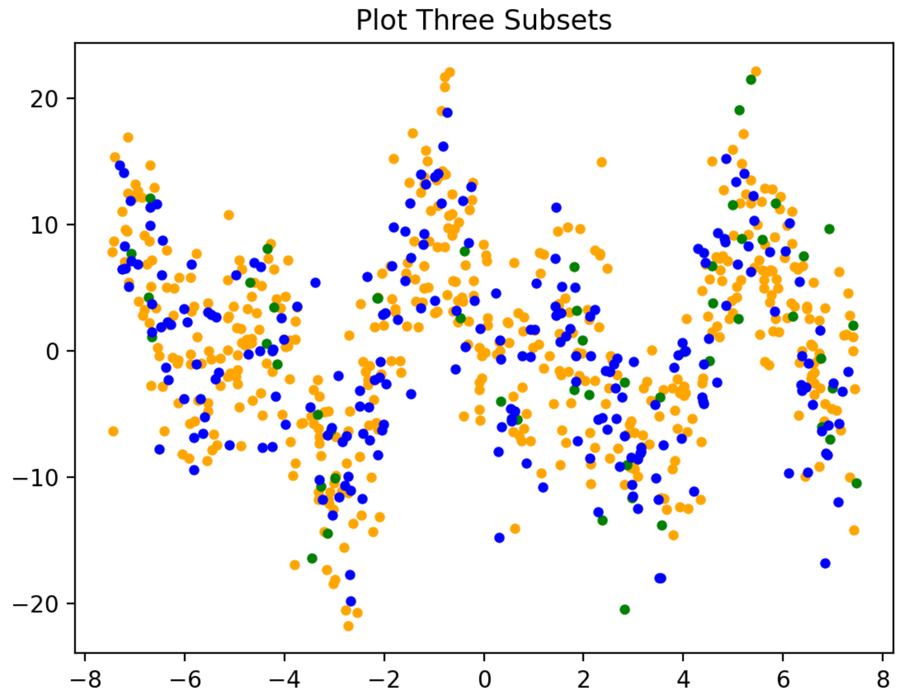
P.S.

I write the following document to simply discuss the non-programming part of Assignment 1.

I also add the pictures required with some comments.



With numpy.random module, I got a set of data points randomly distributed around the true function.



Then, I used train_test_split in sklearn.model_selectioon to split the data points into Training set, Validation set and Test set.

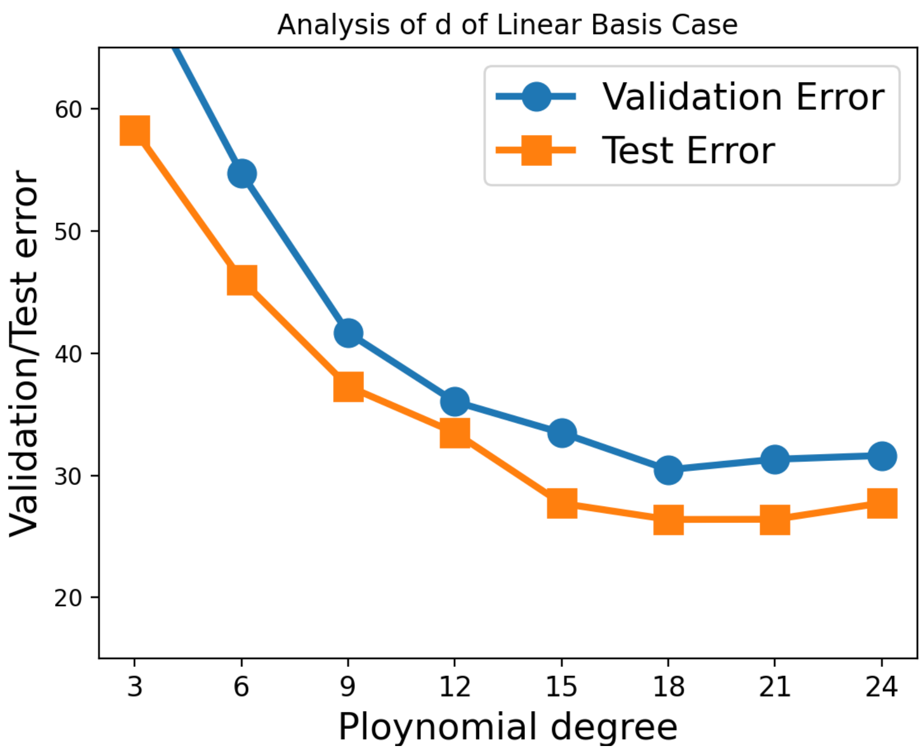
1. ** Regression with Polynomial Basis Functions**

a) Completed in code.

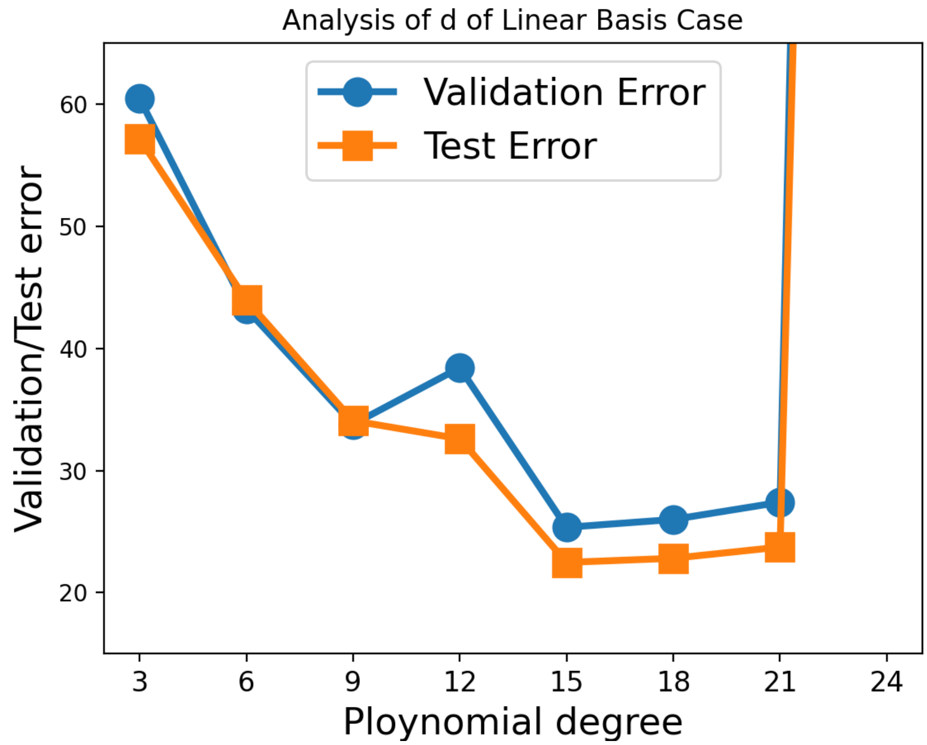
b) Completed in code.

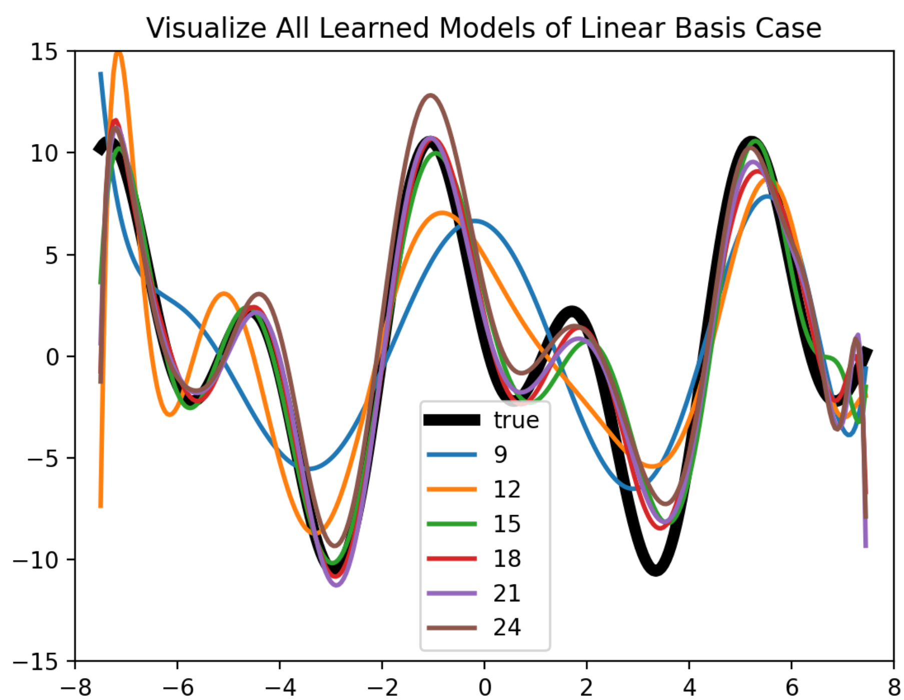
c) Completed in code.

d) Which choice of d do you expect will generalize best?



From the image above, d = 18 will get a lowest validation error. Actually, if we run the code some more times, we can get different images. But the lowest point is always around d = 18 which usually 15 if not 18. Below is another result I got.


Analysis of d of Linear Basis Case
Validation Error
Test Error
Validation/Test error
Ploynomial degree
20
30
40
50
60
3
6
9
12
15
18
21
24

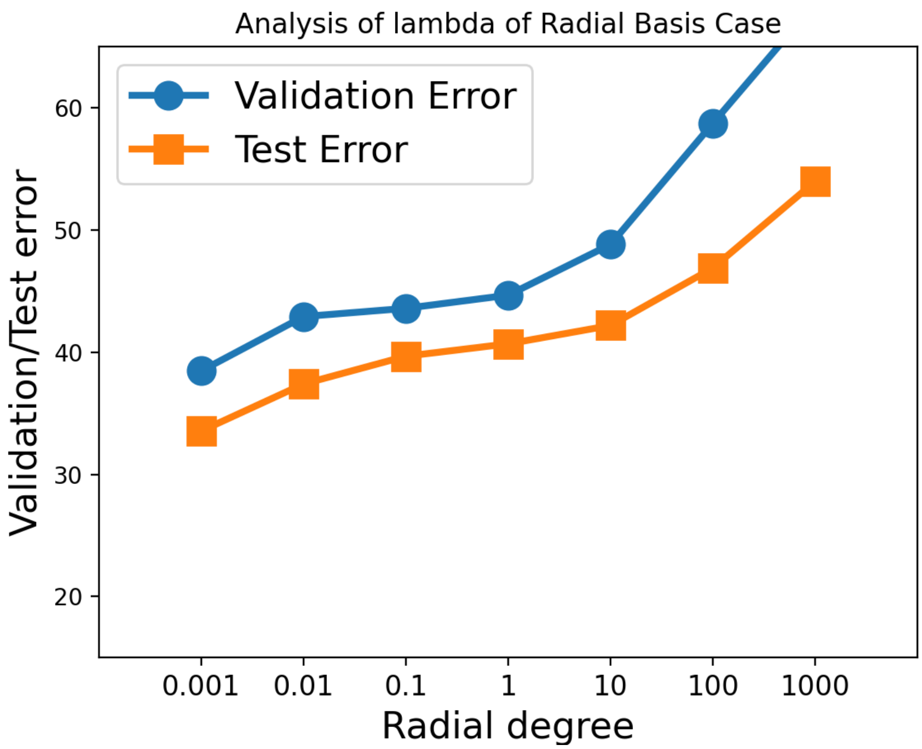

Visualize All Learned Models of Linear Basis Case
true
9
12
15
18
21
24
15
10
5
0
−5
−10
−15
−8
−6
−4
−2
0
2
4
6
8

2. ** Regression with Radial Basis Functions**

a) Completed in code.

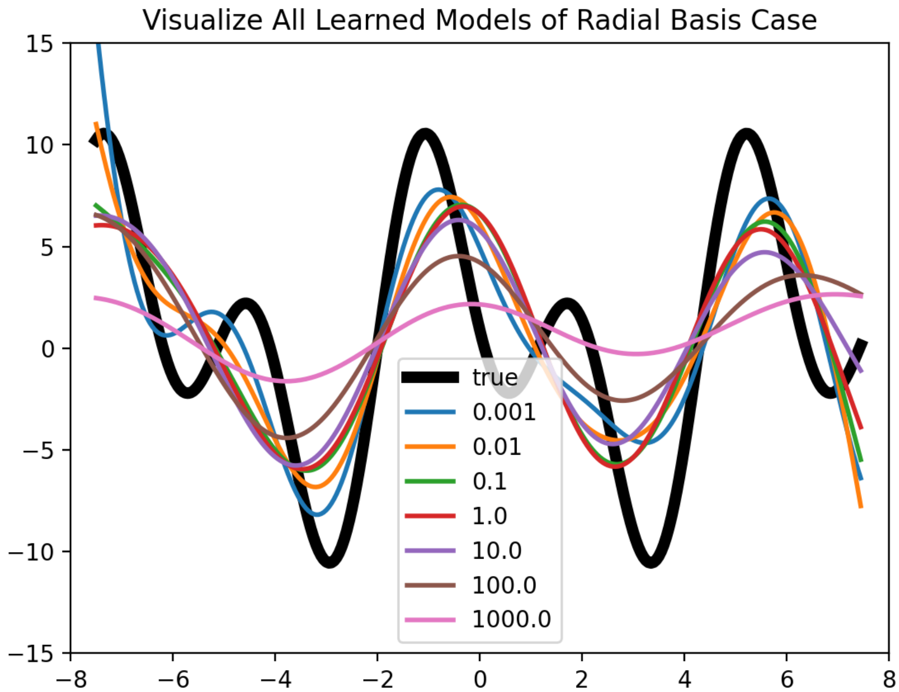
b) Completed in code.

c) What are some ideal values of **λ**?



From the image above, **λ** = 0.001 will be a good choice. I compressed the **λ** axis logarithmically to get a more intuitive vision. The Validation Error will decrease while **λ** decreases. So I would guess we can get lower Validation Error if we use a smaller **λ**. Even though this may cause huge calculation workload. It can be seen that **λ** and Validation Error are basically in a logarithmic relationship.

d) How does the linearity of the model change with **λ**?



While the $\boldsymbol{\lambda}$ decreases, the models tend to be more and more nonlinear. At the same time, the models tend to be more and more close to true function.