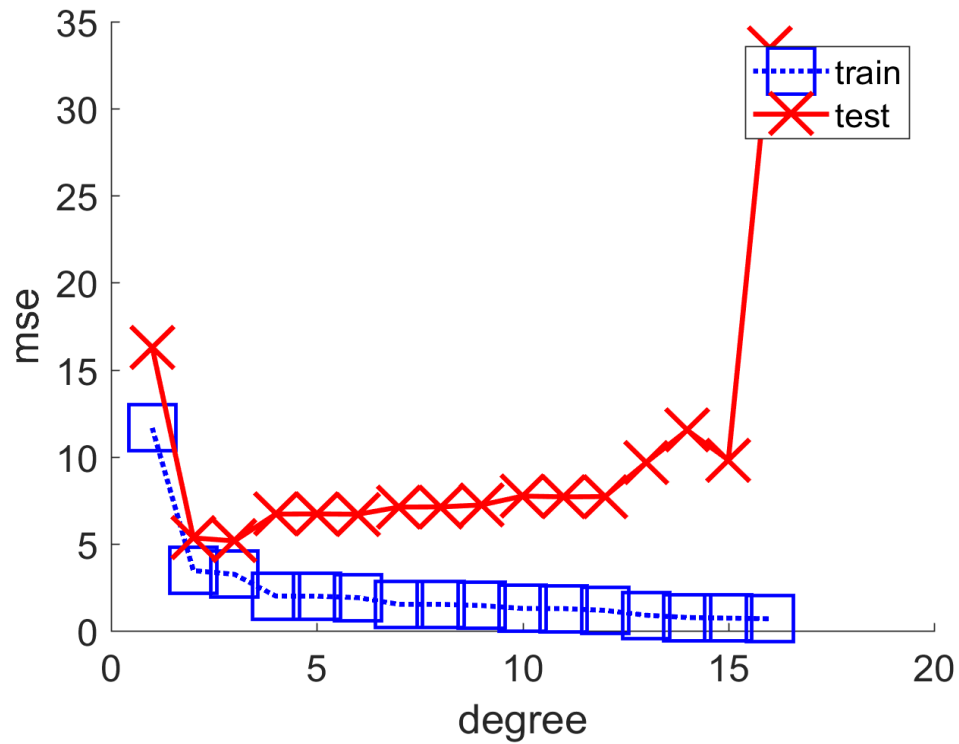


## Homework Week 2

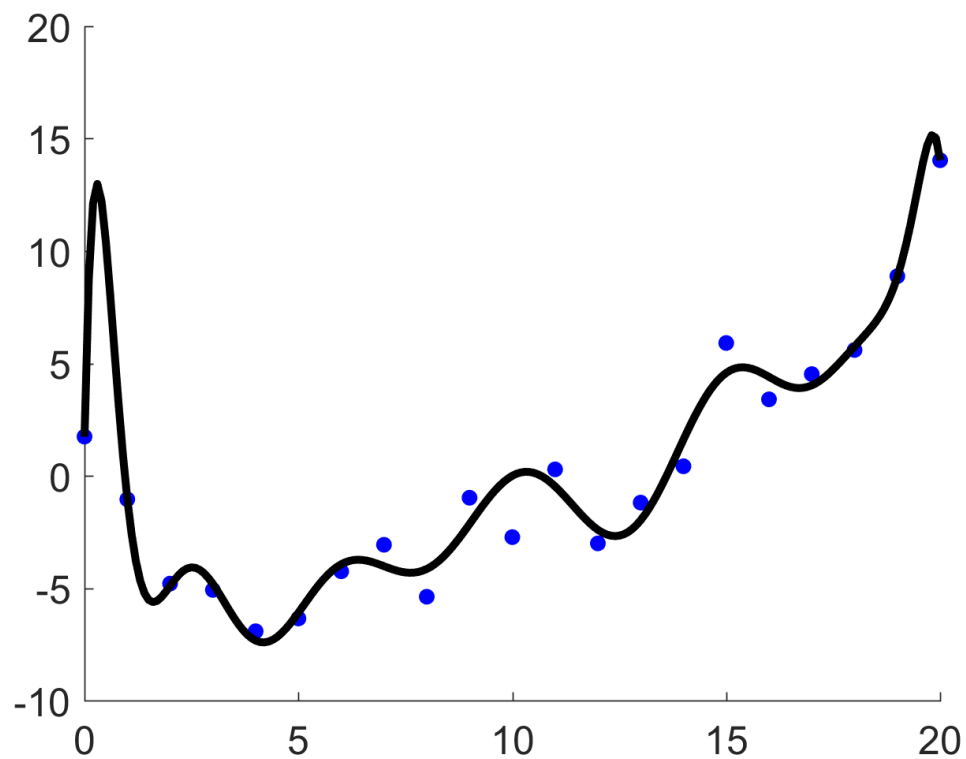
### Question 1

b) (i)



(ii) As we can see from the figure above, for the training set, the MSE is initially high when the degree of polynomial is one. But, as the degree of the polynomial increases, the MSE initially decreases. This is because the curve starts fitting the data better. But, for the test set, MSE is high for degree 1, then it decreases for degree 2 and 3 polynomials but thereafter, it increases gradually with the increase in degree. . There is a sharp increase in the MSE for degree 15 and greater polynomials which could be because there is possibly an overfit of the curve to the training data after a certain increase in the polynomial degree.

c) (i)

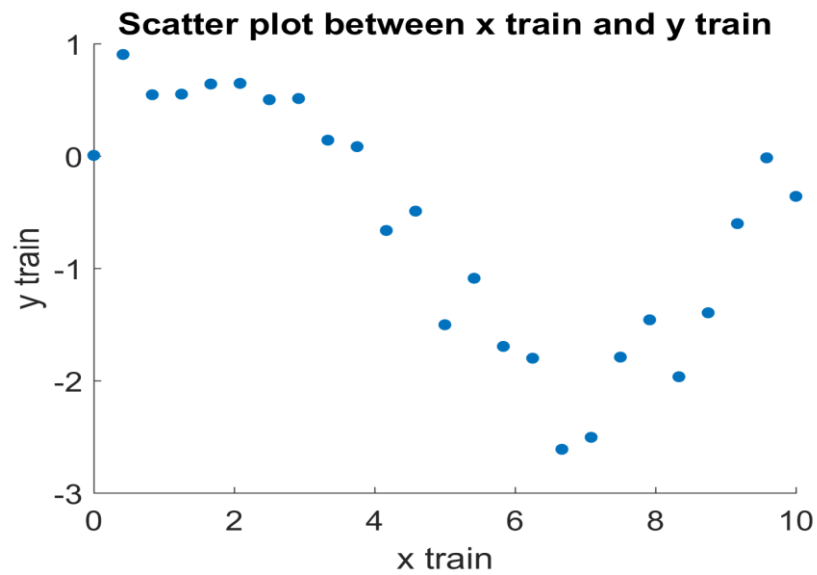


(ii) The degree of the polynomial is 14. It isn't changing during the demo.

(iii) The variable  $\lambda$  controls the effect of regularization. Increasing the value of the regularizer makes the curve fit more exactly to the training data point. This means that with an increase in the value of  $\lambda$ , MSE also increases. This can lead to overfitting on training data point and thus decrease in generalization when dealing with new test data points.

## Question 2

c)



d) The weight vectors are:

Weight\_1dim = [0.5579; -0.2349]

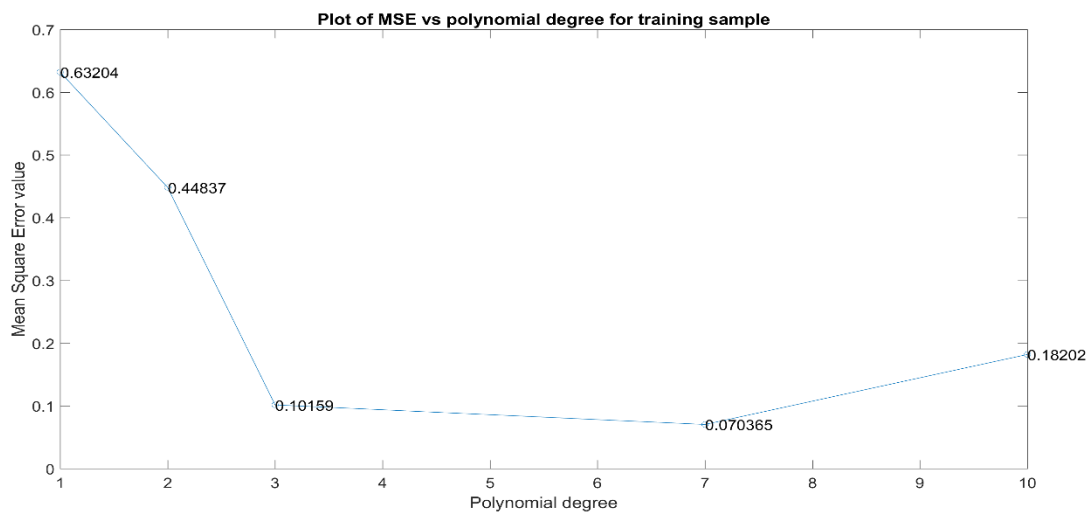
Weight\_2dim = [1.4077; -0.7670; 0.0532]

Weight\_3dim = [0.1830; 0.8720; -0.3650; 0.0279]

Weight\_7dim = [0.2809; 0.3639; 0.1329; -0.1979; 0.0625; -0.0105; 0.0009; -0.0000]

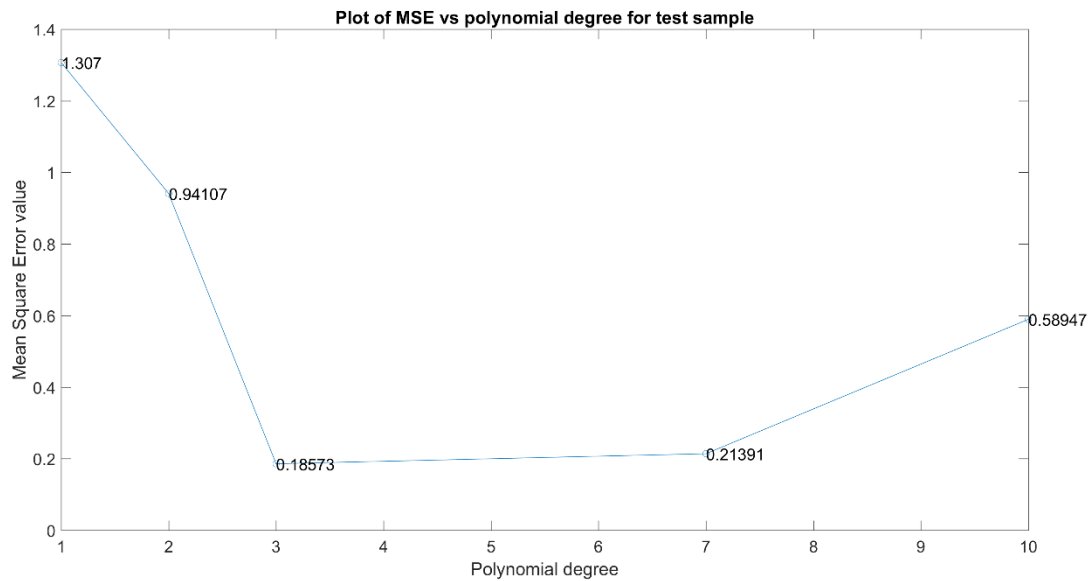
Weight\_10dim =  $1.0 \times 10^{-3} \times [0.0002; 0.0007; 0.0022; 0.0069; 0.0180; 0.0341; 0.0176; -0.1079; 0.0306; -0.0030; 0.0001]$

e)



Based on the training sample MSE only, polynomial degree 7 seems to be the best model.

f)



Degrees of polynomial	MSE
1	1.3070
2	0.9411
3	0.1857
7	0.2139
10	0.5895

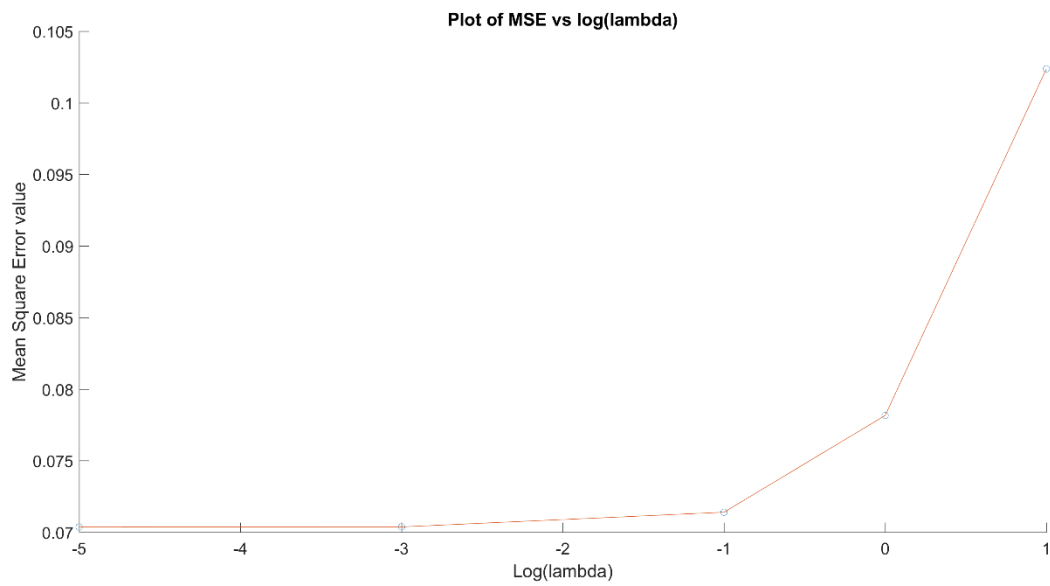
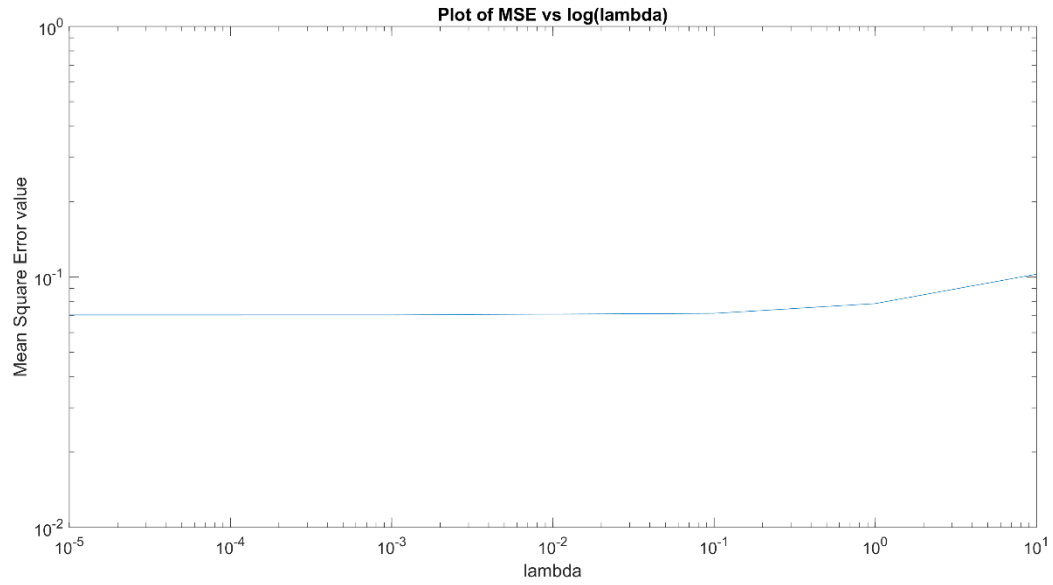
Based on the test sample MSE only, polynomial degree 3 seems to be the best model.

g)

Value of lambda	Value of weight
$10^{-5}$	[0.2809, 0.3639, 0.1329, -0.1979, 0.0625, -0.0105, 0.0009, 0.0000]
$10^{-3}$	[0.2810, 0.3636, 0.1328, -0.1976, 0.0624, -0.01050, 0.0009, 0.0000]
$10^{-1}$	[0.2911, 0.3350, 0.1230, -0.1756, 0.0536, -0.0090, 0.0008, 0.0000]
1	[0.2833, 0.2512, 0.0684, -0.0768, 0.0141, -0.0021, 0.0003, 0.0000]
10	[0.1306, 0.1073, 0.0777, 0.0203, -0.0310, 0.0060, -0.0004, 0.0000]

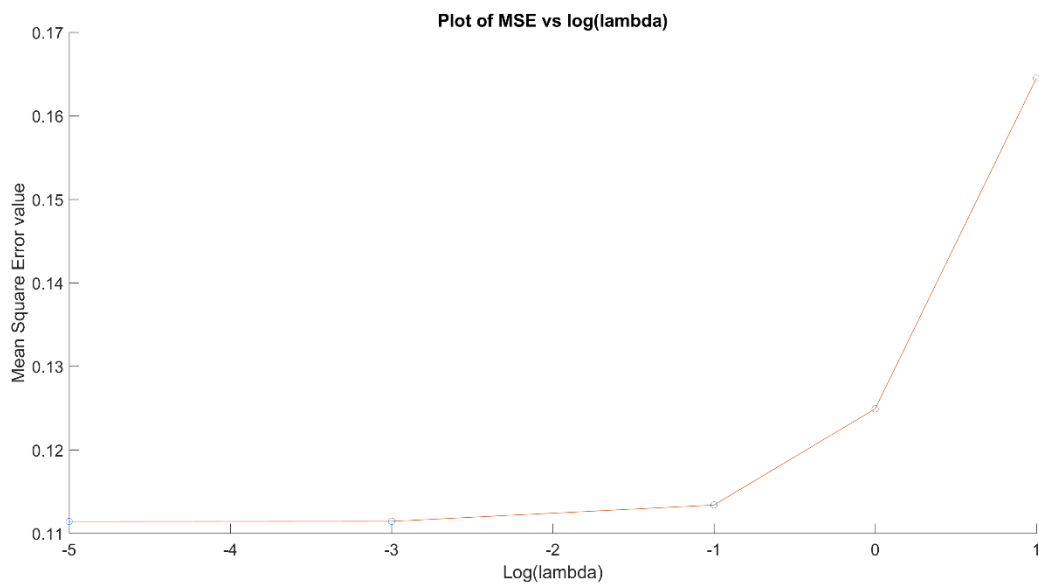
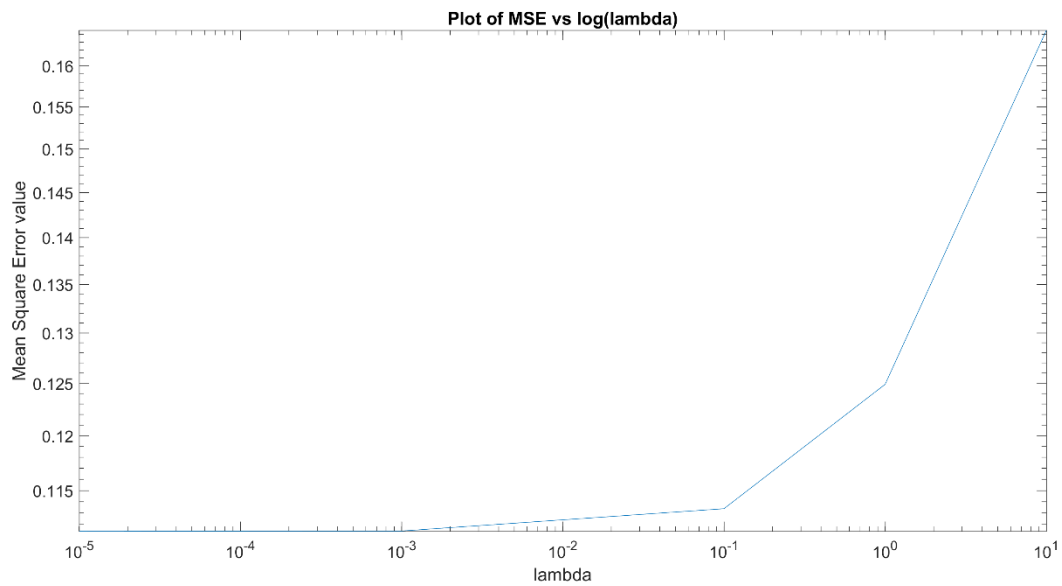
h)

Train MSE = [0.0704, 0.0704, 0.0714, 0.0782, 0.1024]



From the Graph, it can be seen that as the value of lambda increases, the MSE increases too for the training data. The change is more rapid when lambda changes from 1 to 10.

Test MSE = [0.1114, 0.1114, 0.1134, 0.1249, 0.1645]



From the Graph, as the value of lambda increases, the MSE increases too for the test data. The change is more rapid when lambda changes from 1 to 10.