

Name: <Your Name>  
 Github Username: <Your Github username>  
 Purdue Username: <Your PUID>  
 Instructor: < Ding/Qiu/Carrillo >

## Problem 1.

### (1) Estimated Functions:

$\hat{y}_1(x) = a_1x + b$  (Write numerical values for  $a_i$ 's and  $b$ 's )

$\hat{y}_2(x) = a_2x^2 + a_1x + b$

$\hat{y}_3(x) = a_3x^3 + a_2x^2 + a_1x + b$

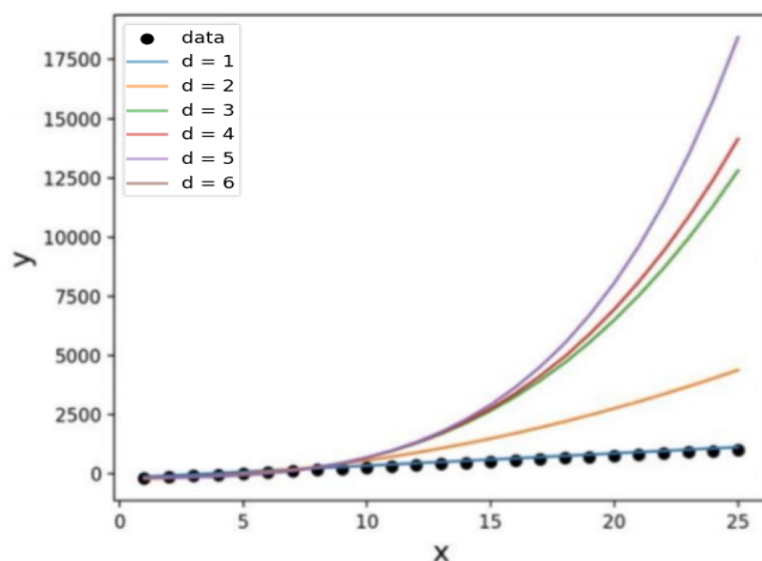
$\hat{y}_4(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + b$

$\hat{y}_5(x) = a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + b$

$\hat{y}_6(x) = a_6x^6 + a_5x^5 + a_4x^4 + a_3x^3 + a_2x^2 + a_1x + b$

### (2) Data Visualization:

(Insert plot obtained from data in poly.txt. Note that the plot below is not the solution)



### (3) What degree polynomial does the relationship seem to follow? Please explain your answer.

(Discuss relationship of data and state a numerical value for the best-fitting polynomial degree)

Sample answer:

The data seems to best follow a first order polynomial (i.e., a line) which can be seen from the low error between the estimated regression function,  $\hat{y}_1(x)$  and the data in the plot above.

**(4) If we measured a new data point,  $x = 4$ , what would be the predicted value of  $y$ , based on the polynomial identified as the best fit in Question (3)?**

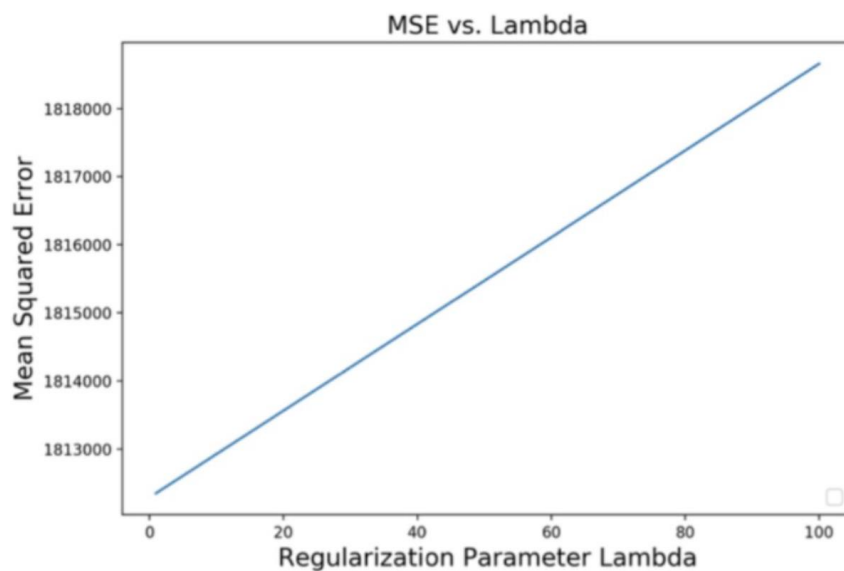
Sample answer:

If we measured a new data point,  $x = 4$ , the corresponding predicted value would be  $\hat{y}_1(4)$ .

## Problem 2.

### (1) Plot the mean squared error as a function of lambda in Ridge Regression:

(Insert plot obtained by completing the *main* function. Note that the plot below is not the solution)



### (2) Find best lambda:

Sample answer: *(insert numerical values for c and d)*

Based on the range of Lambda values tested, the best lambda value is  $c$ , which yields an MSE of  $d$  as shown on the plot above.

### (3) Find equation of the best fitted model:

*(Insert numerical values for  $a_i$ 's and  $b$ )*

$$\hat{y}(x) = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + b$$

**(4) Plot the predicted stock prices and actual stock prices using Google data**

*(Note that the plot below is not the solution)*

