Name: <Your Name>
Github Username: <Your Github username>
Purdue Username: <Your PUID>
Instructor: <Ghasemi/Pare/Qiu/Sahay>

Problem1_writeup.

(1) Estimated Functions:

$\widehat{y_1}(x) = a_1 x + b$ *(write numerical values for a_i's and b's)*

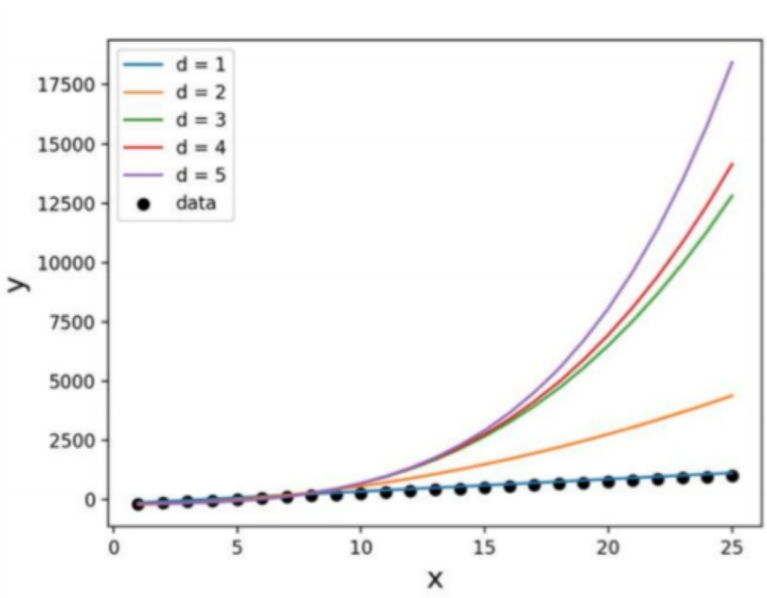
$\widehat{y_2}(x) = a_2 x^2 + a_1 x + b$

$\widehat{y_3}(x) = a_3 x^3 + a_2 x^2 + a_1 x + b$

$\widehat{y_4}(x) = a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + b$

$\widehat{y_5}(x) = a_5 x^5 + a_4 x^4 + a_3 x^3 + a_2 x^2 + a_1 x + b$

(2) Data Visualization: *(insert plot obtained from data in poly.txt)*



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

(Discuss relationship of data and insert numerical value of c calculated from best regression)

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, $\widehat{y_1(x)}$, and the data in the plot above.

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

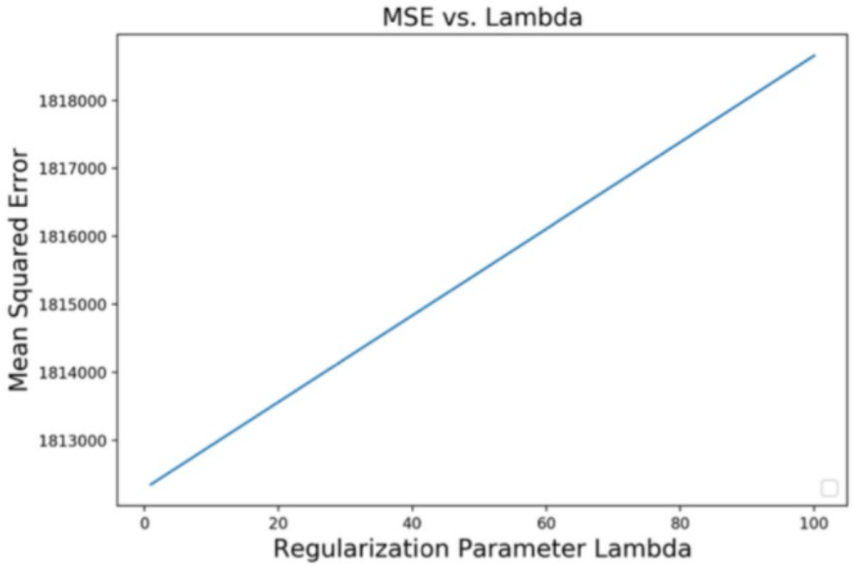
Sample answer:

If we measured a new data point, x=3, the corresponding predicted value would be, $\widehat{y_1(3)}$= c.

Problem2_writeup.

(1) <u>Plot the mean squared error as a function of lambda in Ridge Regression:</u>

*(Insert plot obtained by completing the **main** function)*



(2) Find best lambda:

Sample answer:

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)*

$$\hat{y}(x) = a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 + a_6 x_6 + b$$

(4) Draw a prediction plot using Google data

(Note that the plot below is not the solution)