

Name: Roshan Sundar
GitHub Username: rmsundar
Purdue Username: 0032402630
Instructor: Ghasemi

Problem1_writeup.

(1) Estimated Functions:

$$y_1(x) = 29.0587x + 92.7676$$

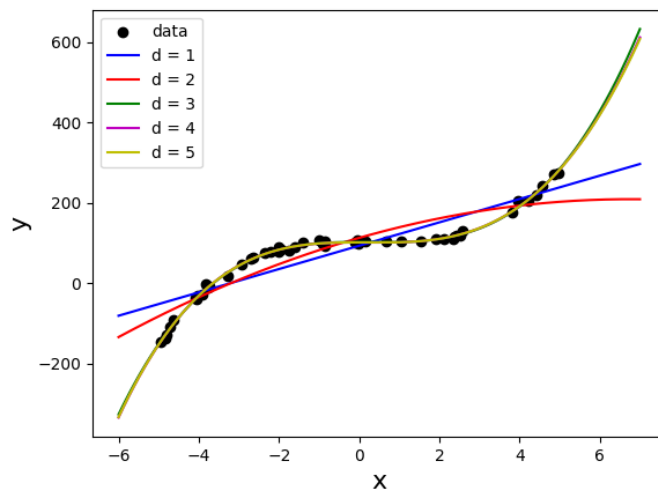
$$y_2(x) = -2.1111x^2 + 28.5066x + 112.3148$$

$$y_3(x) = 1.7574x^3 - 1.4324x^2 - 0.3307x + 101.8661$$

$$y_4(x) = -0.0152x^4 + 1.7541x^3 - 1.0821x^2 - 0.2558x + 100.9145$$

$$y_5(x) = -0.0004x^5 - 0.0154x^4 + 1.7668x^3 - 1.0743x^2 - 0.3227x + 100.8874$$

(2) Data Visualization:



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

Sample answer:

The data seems to match $d = 3, 4, 5$. In following Occam's Razor, the simplest model that matches should be chosen. $d=3$, or $y_3(x)$, is the model that matches this criterion.

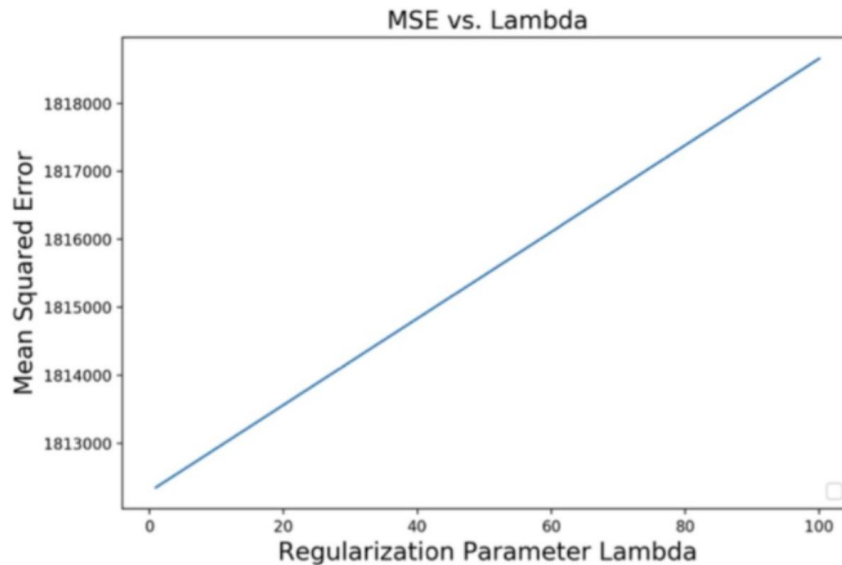
(4) If we measured a new data point, $x = 3$, what would be the predicted value of y , based on the polynomial identified as the best fit in Question (3)?

$$y_3(x) = 135.4328$$

Problem2_writeup.

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

(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)

