Assignment 1

ELL784

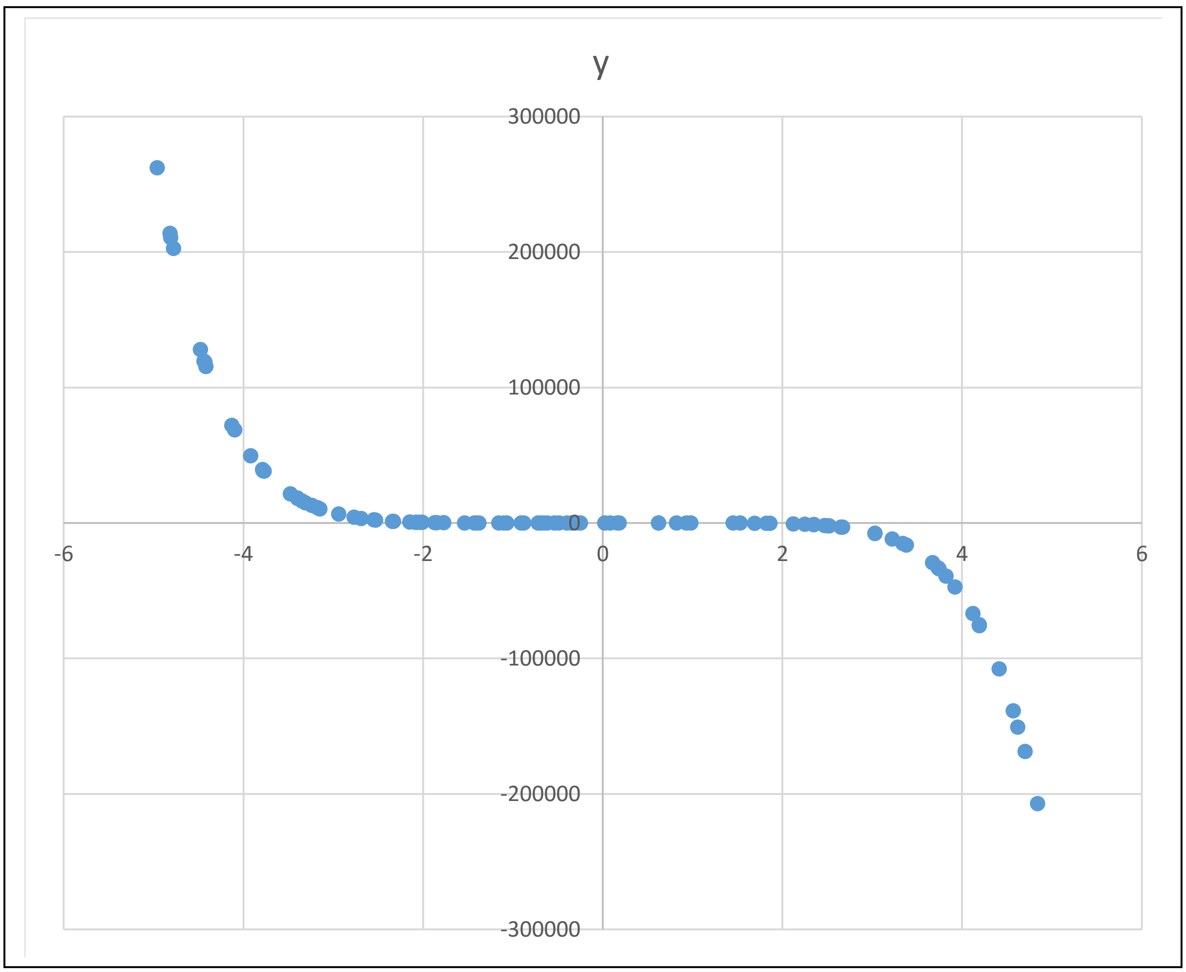
Submitted By: Harshdeep Gupta

Entry No- 2013MT60597

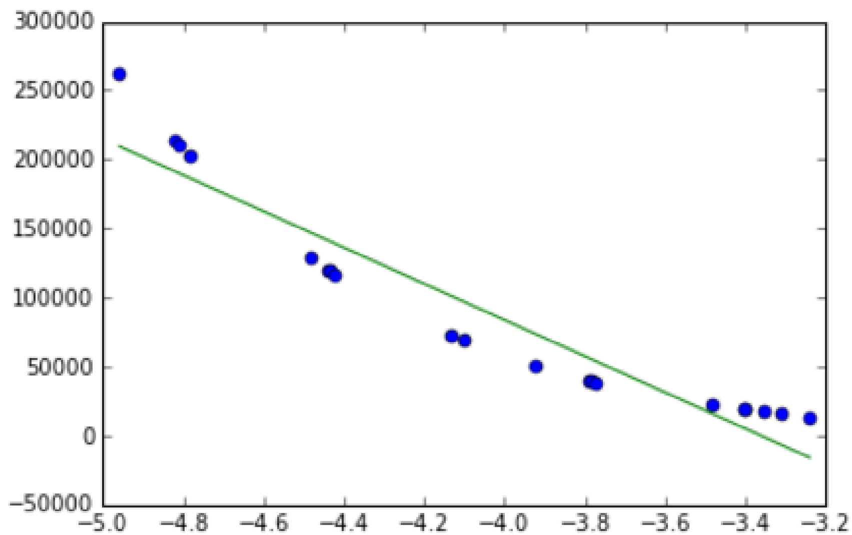
Part 1

The Data:-

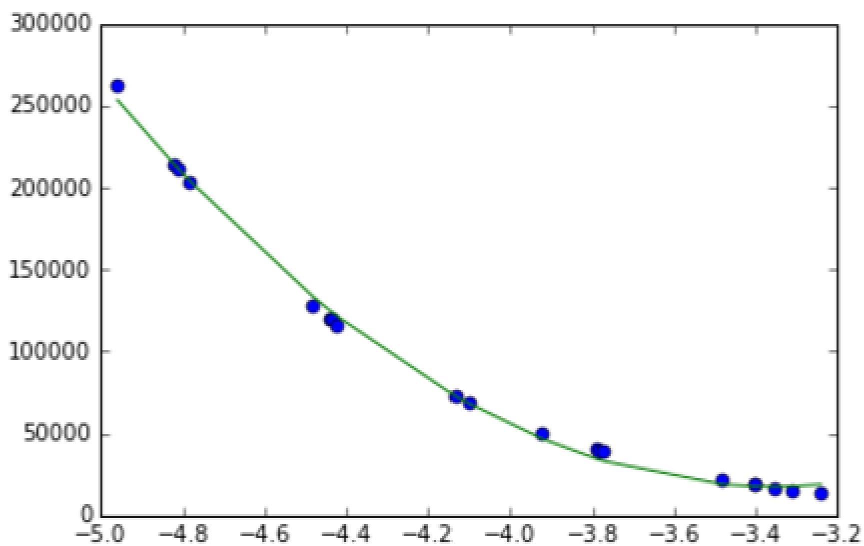
The plot of the data as given:

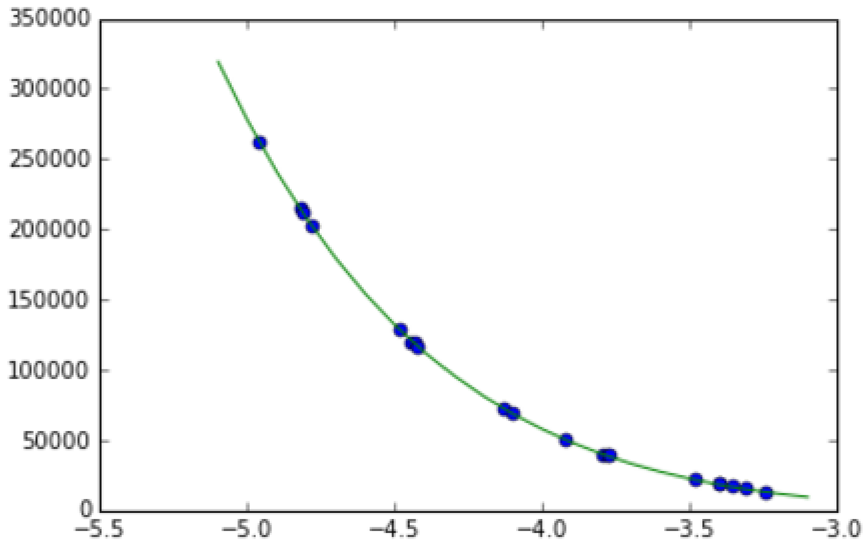


- Using the First twenty points only
    - **Regression with Linear Basis Functions, no regularization**
    - Theta = [-131101.94166835 -440682.09031097]
    - RMS error = 103740.46909601474



    - **Regression with Quadratic Basis Functions, no regularization**
    - Theta = [92093.62943895   618787.04644468 1056948.65695317]
    - RMS error = 17797.548636733027



- **Regression with X^5 Basis Functions, no regularization**
- Theta = [-1559.67288411 -22514.55413373 -142887.30678472 -475289.15934071
- -812414.81955637 -565601.85879763]
- RMS error = 1560.3124785126483
- 

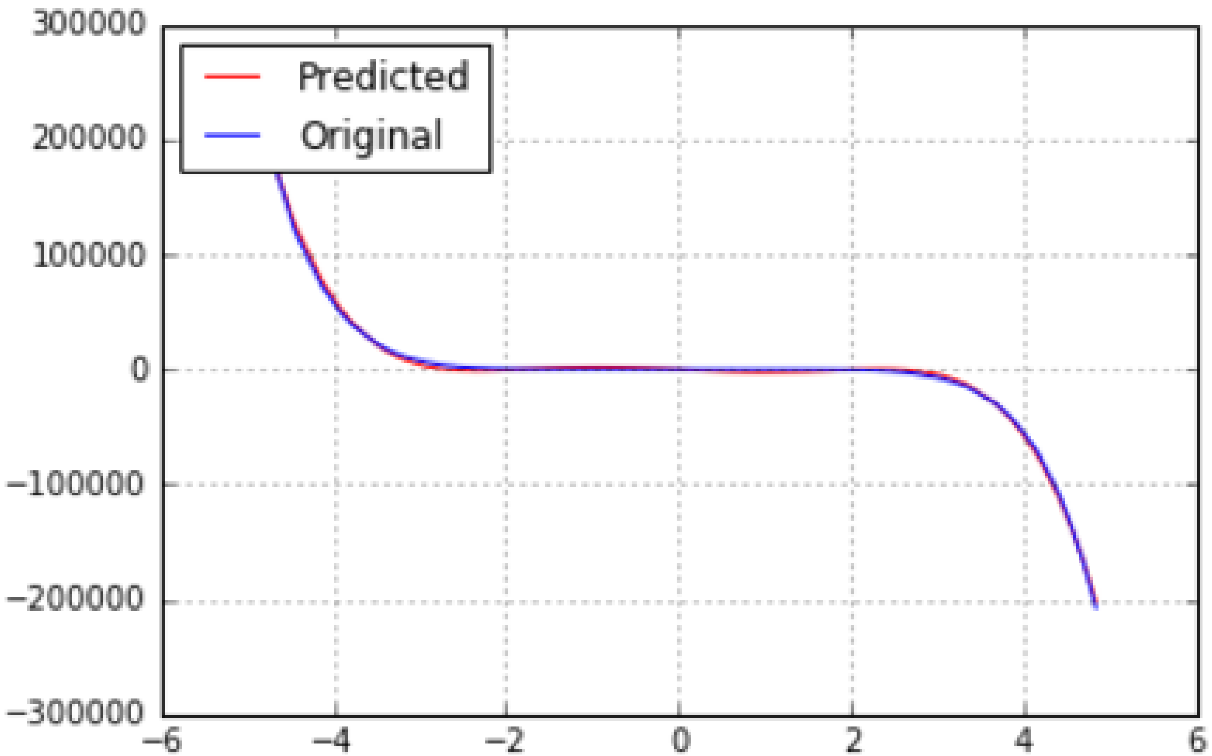
- Estimation of Beta:

- **Using all the 100 points in the dataset**
  - For obtaining optimal results, this is what I did:
    - Ran regression for all polynomials with degree 1 to 10
    - For each polynomial , put alpha from 0 to 30, adding 3 to each previous value to generate 10 values of alpha
- *Result*

For the given data, I found that a fifth order polynomial, with regularization coefficient set to 10 works pretty well (gives the least cross validation error). The parameters of my model are [a0,a1,…,a5]:

[-117.57479698 -3575.76286303 -199.62516982 1448.89434622   22.37773588 -134.19724626].

The final fit looks like:



Estimation of Beta

$$\frac{1}{\beta_{\mathrm{ML}}} = \frac{1}{N}\sum_{n=1}^{N}\{y(x_n, \mathbf{w}_{\mathrm{ML}}) - t_n\}^2 .$$

Therefore, using the above model, and plugging the values we get our beta as :-

0.08457689868069149

Part 2:

Here, we use a multivariate regression model. The parameters which are important for obtaining low errors are:

- Suitable choice of basis functions
- Choosing a suitable value of Regularization parameter (α)

- To choose a suitable α, we need to apply proper cross validation techniques and use that alpha that gives the least error according to some error function.
- **Results**
- In my final model I use upto quadratic basis functions of all the variables to capture nonlinear dependencies on the data. However I don't multiply two different variables with one another.
- For finding a suitable value of α:
    - I used 10-fold cross validation
    - Scanned for values of from 10^-15 to 1, and using that one which gives the lowest cross validation error.
- When I tried to include cubic basis functions into my model, I saw that my error on test data increased, which shows that my model was over fitting my training data. So I switched back to using quadratic basis functions.
- During cross validation, I found that setting alpha too low (<10^-14) or too high (>10^-12) increased the amount of error. The lowest amount of cross validation error that I got was when alpha was set to 10^-13.

So the best results are obtained using a quadratic model with ridge shrinkage. During the time of writing this report , my rank of leaderboard is seven, my best rank being five.0