ECE521: Assignment 1

Due on Monday, January 25, 2016

Davi Frossard

January 24, 2016

Task 1

Training Set Size (N)	**Validation Errors**
5	107
50	40
100	27
200	17
400	9
800	12

Using a larger training set increased the algorithm's accuracy since it makes data more diverse and harder to over-fit.

Task 2

Neighbors Considered (K)	Validation Errors
1	12
3	8
5	10
7	9
21	11
101	24
401	51

Smaller values of K tend to lead to over-fitting when the decision boundary for a variable is complex, that was not the case for this dataset, whose optimum K was 3; on the other hand an over-smoothed decision boundary (larger K) penalized accuracy.

Task 3

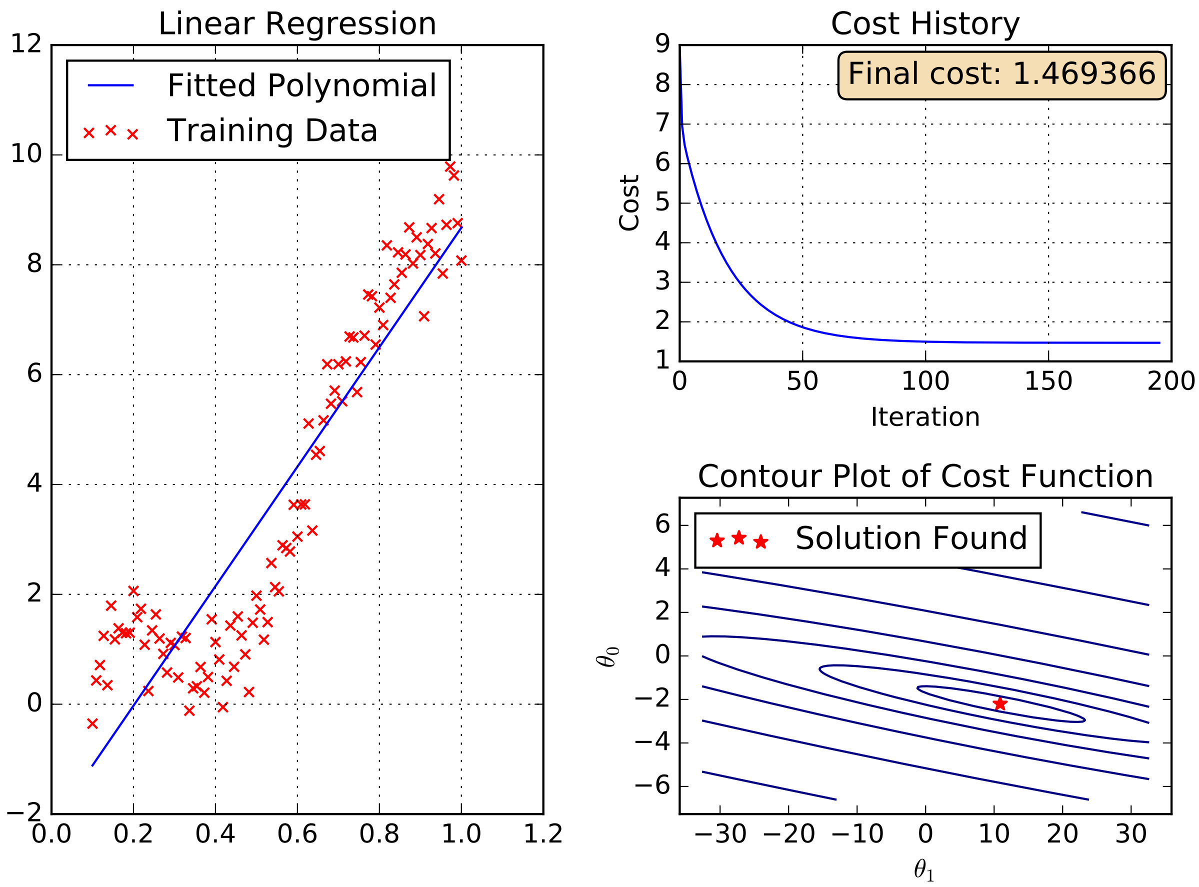

Figure 1: Linear regression of artificial dataset and cost plots

We first normalize the dataset by the maximum value to prevent overflows. Fitting the data to a first order polynomial with stochastic gradient descent yields the following model:

$$f(x) = -2.202718 + (10.878017) \cdot x^1$$

Task 4

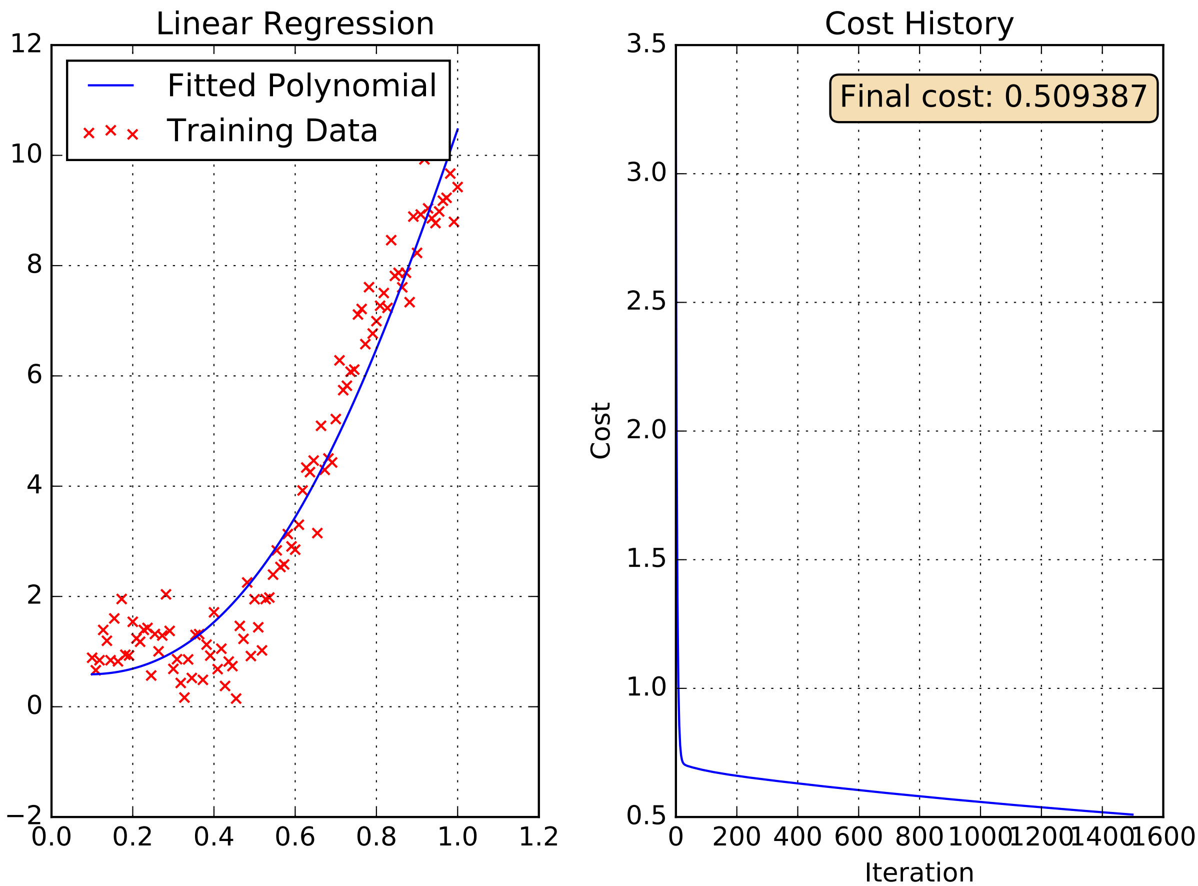

Figure 2: Multivariate regression of artificial dataset and cost plots

By using non linear parameters we give more malleability to the model since it can now make curves, which makes the final cost decrease considerably.

$$f(x) = 0.648815 + (-1.275800) \cdot x^1 + (6.138050) \cdot x^2 + (6.172230) \cdot x^3 + (2.059428) \cdot x^4 + (-3.271531) \cdot x^5$$

Task 5

Training Set Size (N)	Validation Errors
50	44
100	34
200	28
400	23
800	21

For the same reason as discussed in Task 1, larger values of N provide better results since it makes the dataset less prone to over-fitting.

Task 6

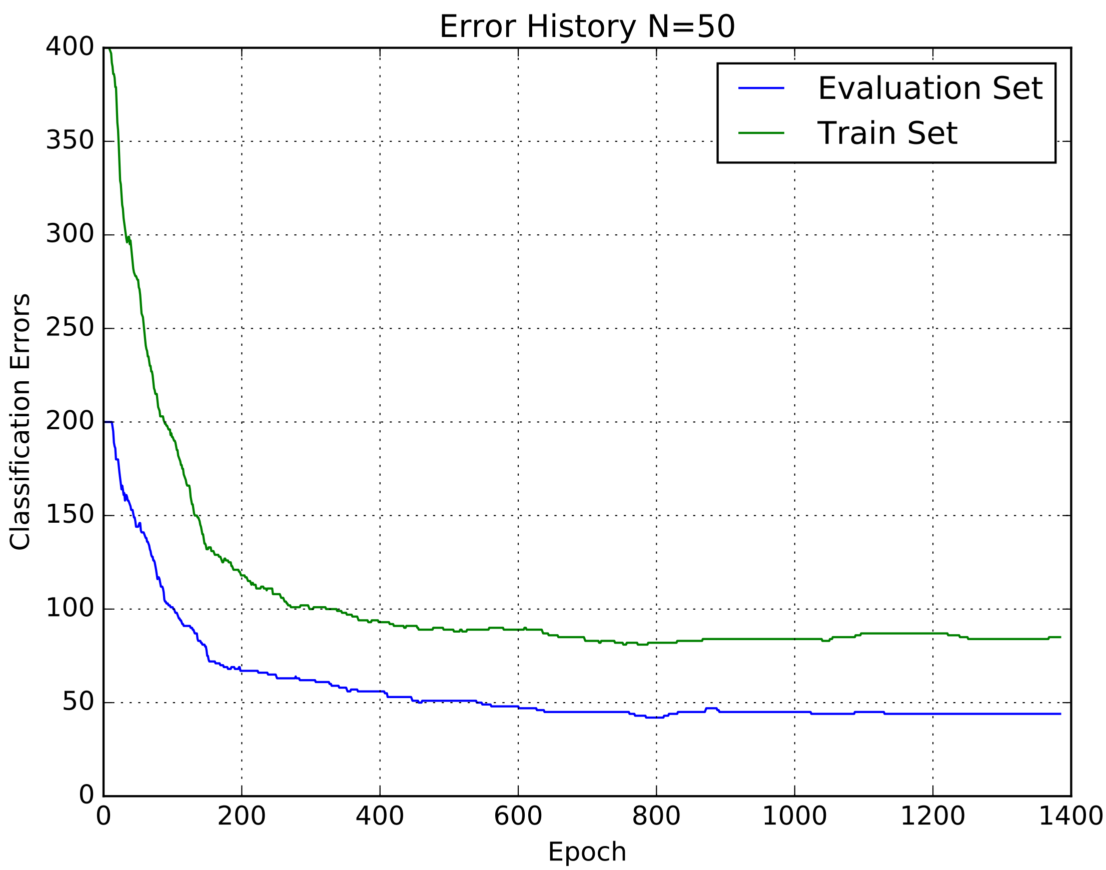

Figure 3: Classification errors by epoch

Task 7

λ	Validation Errors
0.0000	44
0.0001	44
0.0010	44
0.0100	44
0.1000	39
0.5000	200

A value of $\lambda = 0.1$ improves the model by decaying the weights and reducing over-fitting, improving the accuracy even with such a small dataset (50 points).