Project2: Implementing Polynomial Regression

Tamanna Baig

Problem Description:

The main objective of the project is to implement a polynomial regression function to predict a given student's Grade Point Average (GPA) based the number of minutes the student studies for in a week, along with the ounces of beer consumed by the student in that week.

The given dataset contains 300 records. Each line represents one student and contains 3 tab-separated values i.e., number of minutes spent studying/week, ounces of beer consumed/week and GPA scored per semester.

The python file when executed, will prompt the user to input 2 values: number of minutes spent studying/week and ounces of beer consumed/week. This will be prompted till the user enters 0 for both the values. The final output is the predicted GPA of the given students. Figure1 shows the initial plot from the given data.

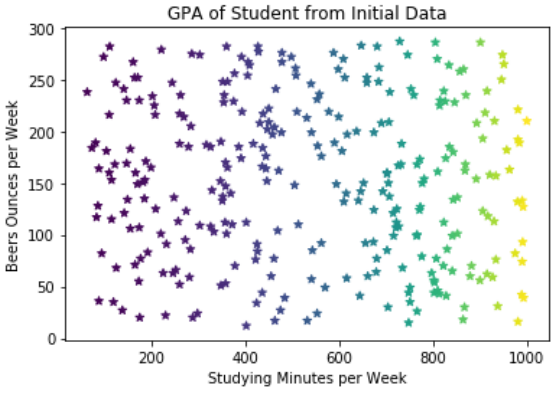

Figure 1. Initial plot of the data

Choosing Initial Values:

The initial plot (shown in Figure 2) was created by using iterations = 200,
alpha = 0.2,
initial weights = [1, 1, 1, 1, 1, 1]
Final J value = 0.0000036998573929704

On close observations, I gathered making iterations as 150, and alpha as 0.2 will give the algorithm enough time to learn and predict the correct GPA.

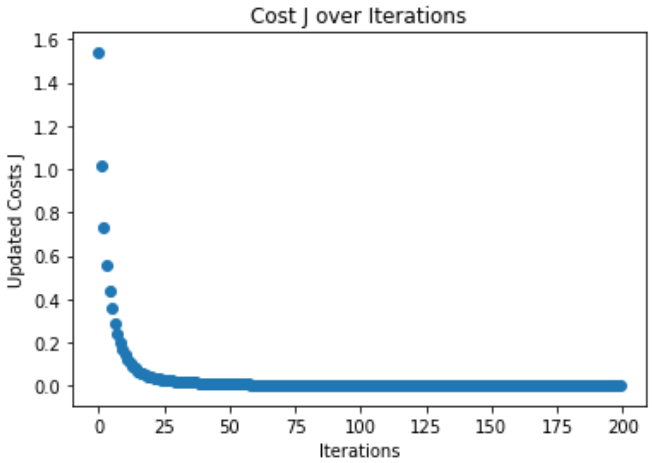

Figure 2. Initial plot with cost vs Iteration

Obtaining Final Values:

Figure 3 shows the final values of weights, J and errors generated by the learning algorithm on the training set. Figure 4 shows the Cost vs Iteration plot.

```
w0:  1.170186
w1:  1.171454
w2:  -0.000121
w3:  -3.8e-05
w4:  0.293568
w5:  0.000514
-----------------
J:  3.818490875011518e-06
-----------------
Mean Error:  0.008311
Median Error:  0.001817
```

Figure 3. Final Values after training

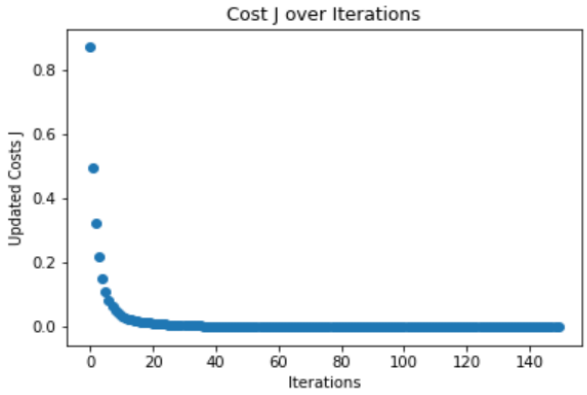

Figure 4. Final plot after training

Feature Scaling:

The two features are in minutes and ounces respectively, therefore, to make them both into a similar scale, **normalization** was applied. Figure 5 and 6 show values before and after normalization respectively.

	study_mins	ounces_beer	gpa
0	759	235	2.30
1	444	164	0.79
2	553	24	1.22
3	852	101	2.90
4	674	124	1.82

Figure 5. Values before normalization

	study_mins	ounces_beer	gpa
0	0.804626	1.063631	2.30
1	-0.359624	0.142682	0.79
2	0.043244	-1.673273	1.22
3	1.148357	-0.674497	2.90
4	0.490463	-0.376162	1.82

Figure 6. Values after normalization

Value of J on Test set:

The value of J on test set = 3.6981184920253093e-06