Name/Roll No: Himanshu Ruhela 2018IMT-039
Course: IMT
Course Code: ITIT - 4107

Assignment Number: 4
Deadline for Submission: 23:59 18th October 2021

Aim

Suppose you are the CEO of a restaurant franchise and are considering different cities for opening a new outlet. The chain already has trucks in various cities and you have data for profits and populations from the cities. You would like to use this data to help you select which city to expand to next.

The file ex1data1.txt contains the dataset for our linear regression problem. The first column is the population of a city and the second column is the profit of a food truck in that city. A negative value for profit indicates a loss.

1. Use a scatter plot to visualize the data, since it has only two properties to plot (profit and population).

2. Consider a simple linear model with two parameters and one input variable and mean square error cost function to implement the gradient descent algorithm to find the intercepts. Assume a suitable terminating condition.

3. Plot the model alongside the scatterplot to show the fit model.

4. Perform steps 1,2,3 in batch mode for varying values of alpha, learning rate and plot the results.

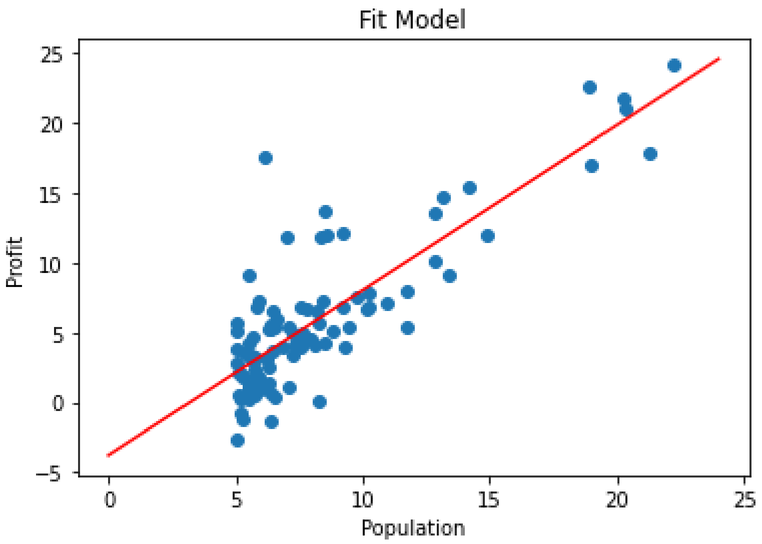
5. For each of the experiments performed above in steps 1,2,3,4 with varying learning rates visualize the cost function as a contour plot as well as plot the values of parameters to visualize the stepwise traversion of the parameters on this contour plot.

Procedure/Algorithm

We will implement linear regression with one variable and fit the linear regression parameters θ to our dataset using gradient descent. The objective of linear regression is to minimize the cost function.

The data is plotted and shown below –

Results



Learning Rate: 0.01

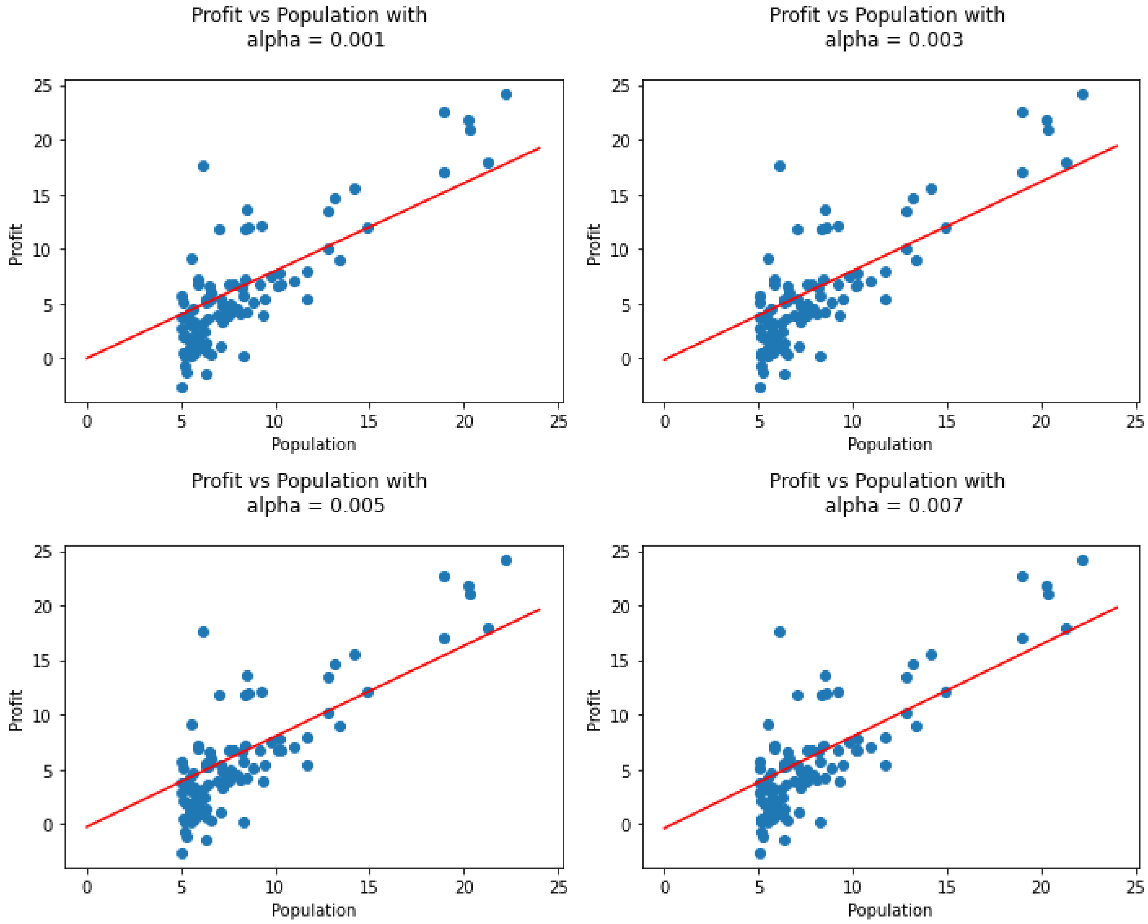
Initial Cost: 32.072733877455676

Final Cost: 4.4780276098799705

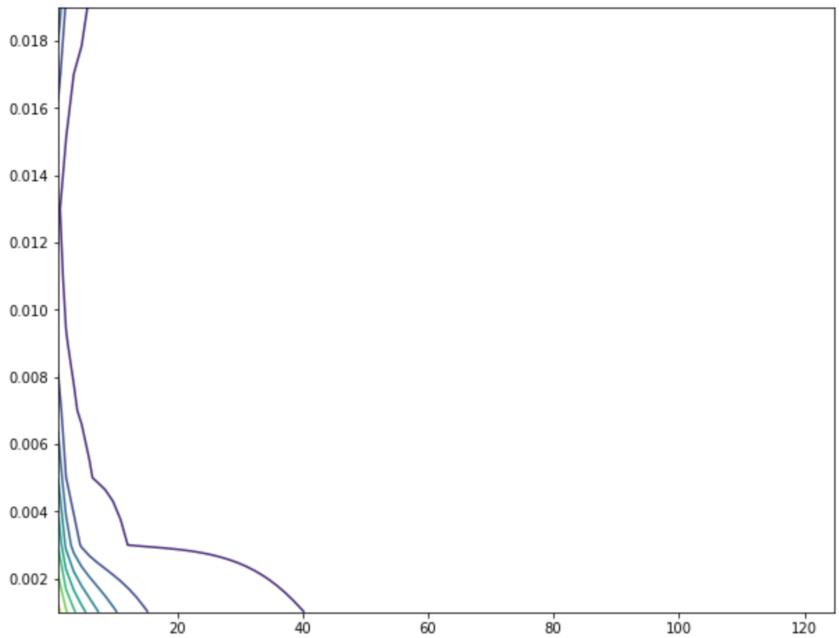
Theta0: 1.1822127747573767

Theta1: -3.78806857272311

Some of the graphs plotting different learning rates:



Contour Plot:



Jupyter Notebook:

https://github.com/verdantfire/ITIT-4103-2021/blob/main/Assignment%204/2018IMT-039_ML_Assignment_4.ipynb

Repository:

https://github.com/verdantfire/ITIT-4103-2021/tree/main/Assignment%204

Inference

The choice of correct learning rate is very important as it ensures that Gradient Descent converges in a reasonable time:

If we choose α to be very large, Gradient Descent can overshoot the minimum. It may fail to converge or even diverge.

If we choose α to be very small, Gradient Descent will take small steps to reach local minima and will take a longer time to reach minima.

We plot the contour to show how $J(\theta)$ varies with changes in $\theta 0$ and $\theta 1$. The cost function $J(\theta)$ is bowl-shaped and has a global minimum. This minimum is the optimal point for $\theta 0$ and $\theta 1$, and each step of gradient descent moves closer to this point.