

Name: ID Number:

1 Linear Regression

1. Randomly generate 50 samples with seed number = 10.

$$X = \text{np.random.rand}(n)$$
$$y = 4 + 5X + \text{np.random.rand}(n)$$



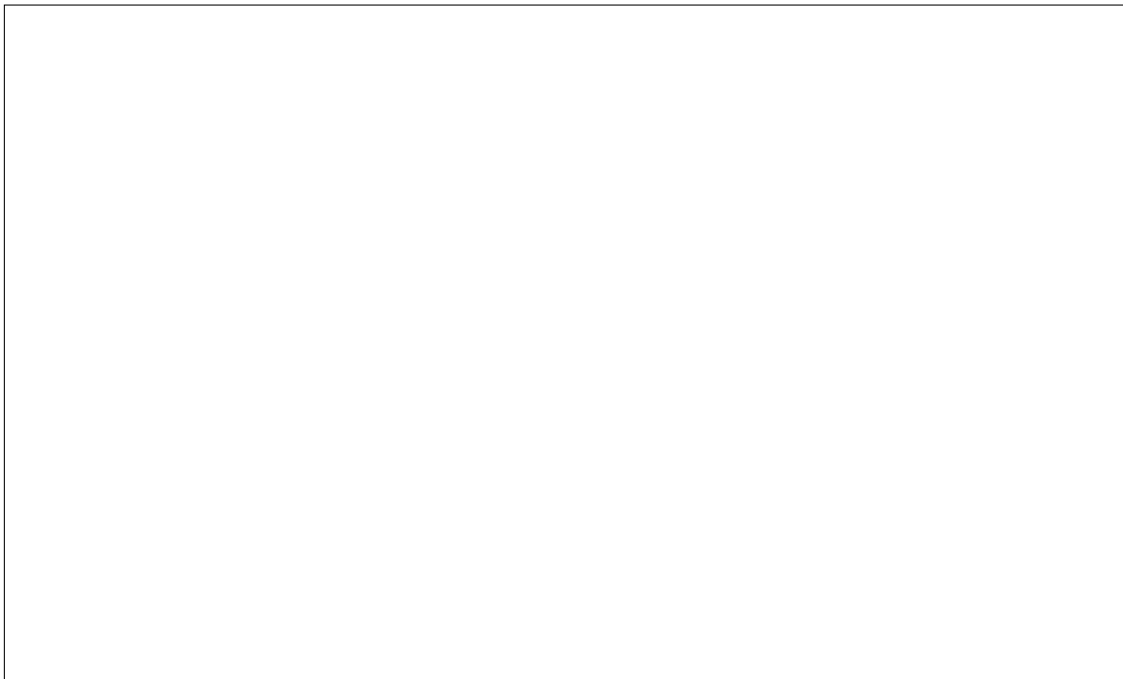
2. Plot scatter graph of the generated samples.



3. Write a function to calculate a cost function of Linear Regression—Mean Squared Error (J)—in the following format:

$$J = \text{computeCostFunction}(X, y, \theta)$$

where θ is a weight value.



4. Write a function to find the optimal theta of Linear Regression by using Gradient Descent Algorithm in the following format:

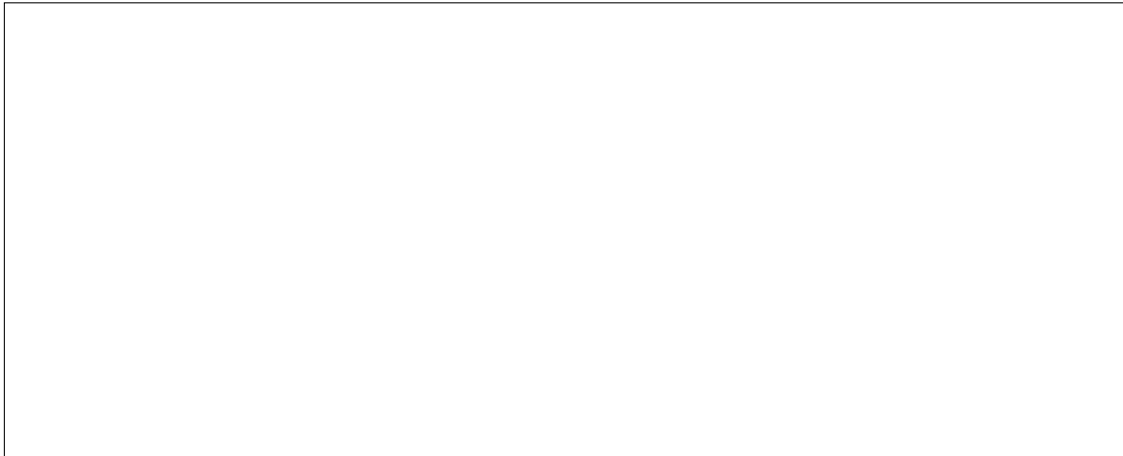
$$[\theta, J_{History}] = \text{gradientDescent}(X, y, \theta, \alpha, n_{iteration})$$

where α is a learning rate, $n_{iteration}$ is the number of iteration, and $J_{History}$ is a vector that store J for every iteration.

5. Demonstrate the equation of Linear Regression by using Gradient Descent Algorithm. The number of iterations ($n_{iteration}$) is 100 and the Learning Rate (α) is 0.1.

6. Plot the convergence curve of the algorithm and report what you see. (Hint: Cost function value -vs- iteration number)

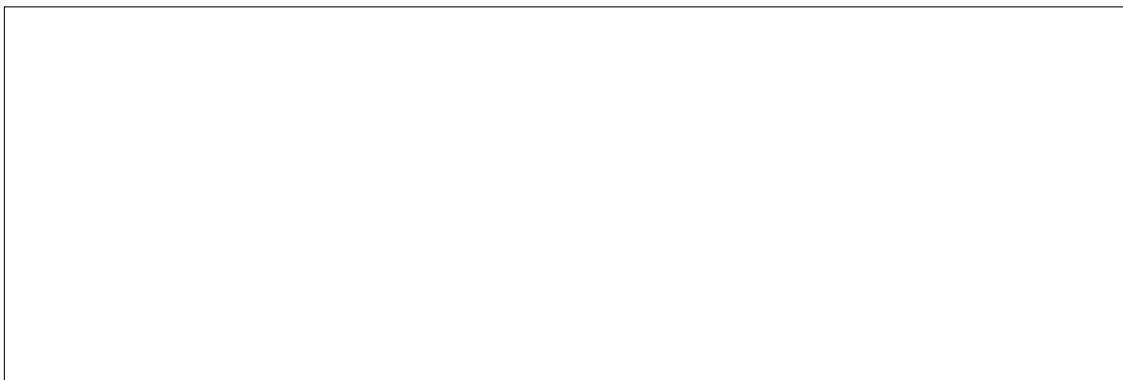
7. Demonstrate how the gradient descent work by simply plot the gradient of each iteration in J .



8. Plot the linear equation on the graph created in Question 2.



9. Change the Learning rate α to 0.01 and 0.001, respectively, and observe the results.



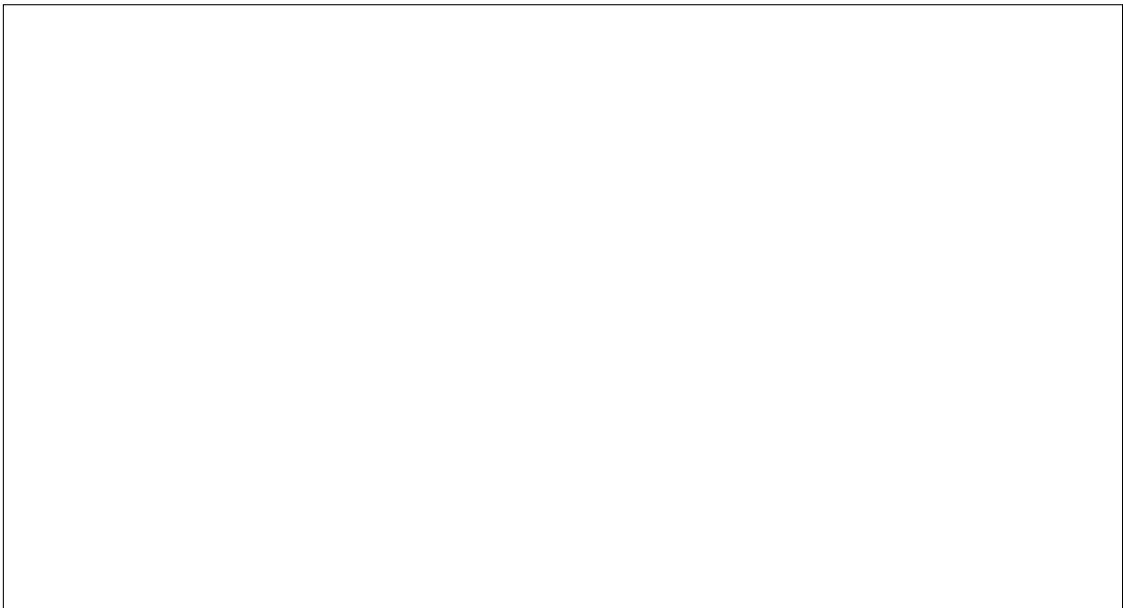
10. Plot the convergence curves corresponds to Question 9.



11. Observe the gradient of each α .



12. Change the number of iterations to 1000 and observe the result.



2 Linear Regression—Multi-Variable

Each residence in Boston is described by 13 features ($I_1 - I_{13}$). The task is to employ the linear regression algorithm to predict the value of owner-occupied home in Boston (O_{14}). The dataset—Boston Housing Data—can be download at <https://bit.ly/2NBaO7e>. The details of the data are as follows:

- I_1 : CRIM—per capita crime rate by town
- I_2 : ZN—Proportion of residential land zoned for lots over 25,000 sq.ft.
- I_3 : INDUS—Proportion of non-retail business acres per town
- I_4 : CHAS—Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
- I_5 : NOX—Nitric Oxides Concentration (parts per 10 million)
- I_6 : RM—Average number of rooms per dwelling
- I_7 : AGE—Proportion of owner-occupied units built prior to 1940
- I_8 : DIS—Weighted distances to five Boston employment centres
- I_9 : RAD—Index of accessibility to radial highways
- I_{10} : TAX—full-value property-tax rate per \$10,000
- I_{11} : PTRATIO—Pupil-teacher ratio by town
- I_{12} : Black— $1000(Bk - 0.63)^2$ where Bk is the proportion of blacks by town
- I_{13} : LSTAT—% lower status of the population
- O_{14} : MEDV—Median value of owner-occupied homes in \$1000's

1. Import regression-datasets-housing.csv file to python

2. Assign values from column 1 to column 13 to X and value from column 14 to y , then convert both variables to numpy.

3. Follow the exercise 3 to 7 in Section 1

4. Normalise this data to zero-mean & unit standard deviation

5. Train a model with normalised data and compare the result with Question 3.

6. Re-write the Gradient Descent function in vector format in order to decrease the number of loops

7. Plot a graph to compare predicted output and actual output

