# Programming Exercise 1: Linear Regression

## 1. Linear regression with one variable

In this exercise, you will implement linear regression with one variable to predict profits for a food truck. 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 to select which city to expand to next.

The file ex1data.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.

Complete the following requirements:

### a) Plotting the data

Before starting on any task, it is often useful to understand the data by visualizing it. For this dataset, you can use a scatter plot to visualize the data, since it has only two properties to plot (profit and population). Figure 1 is the example result of this step.

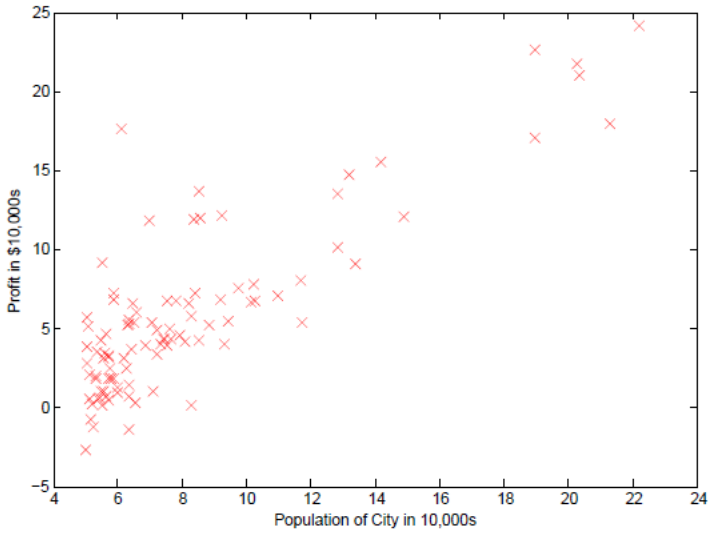


Figure 1: Scatter plot of training data

### b) Implementing gradient descent

In this part, you will fit the linear regression parameters $\theta$ to our dataset using gradient descent.

After you are finished (i.e. $\theta$ is found), you should use your final parameters to plot the linear fit. The result should look something like Figure 2.

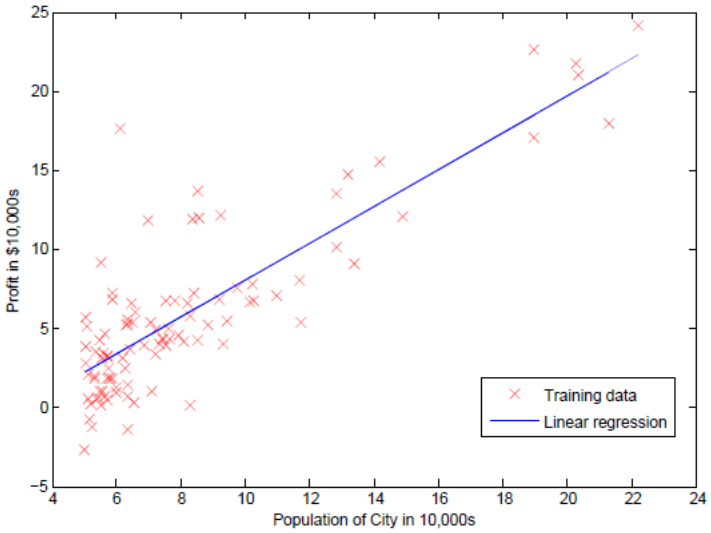


Figure 2: Training data with linear regression fit

## 2. Linear regression with multiple variables

In this part, you will implement linear regression with multiple variables to predict the prices of house. Suppose you are selling your house, and you want to know what a good market price would be. One way to do this is to first collect information on recent houses sold and make a model of housing prices.

The file ex1data2.txt contains a training set of housing prices in Porland, Oregon. The first column is the size of the house (in square feet), the second column is the number of bedrooms, and the third column is the price of the house.

Complete the following requirements:

### a) Feature normalization

By looking at the data, note that house sizes are about 1,000 times the number of bedrooms. When features differ by orders of magnitude, performing feature scaling can make gradient descent converge much quicker.
Your task here is to implement Z-Score Standardization algorithm to scale features.

Formula of Z-Score Standardization is in Equation (1).

$$X^* = \frac{X - mean(X)}{StandardDeviation(X)} \quad (1)$$

### b) Gradient descent

Previously, you implemented gradient descent on a univariate regression problem. The only different now is that there is one more feature in the matrix X. The hypothesis function and the batch gradient descent update rule remain unchanged.