

De La Salle University

Gokongwei College of Engineering

Electronics and Communications Engineering Department

**LBYCP29 Experiment 2**

Submitted By:

**Jan Carlo D. Rabacca**

**Suzette Dela Cruz**

**Marc Chiu**

Submitted To:

**Engr. Melvin Cabatuan**

Date:

**September 16, 2015**

Multivariate

Jan Carlo D. Rabacca

11106344

Suzette Dela Cruz

11118113

Marc Chiu

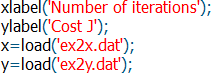
11114525

# Introduction

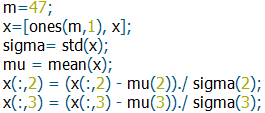
The multivariate linear regression models use a type of analysis that can be used in predicting a value of a certain response by basing it on a set of training data. The predicted value is a vector of the correlation of the training data. It is also used in estimating the linear association of the data and responses. The training data values may be continuous or direct.

# . Procedures

## The first step is to load data to analyze, in this experiment we have,



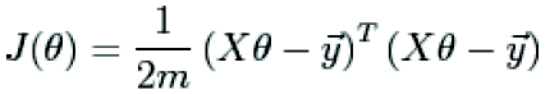
1. Next, set the training examples equal to the length of the loaded data which is 47 samples then concatenate a single column with values of 1 with the x-data. Then, scale both types of inputs by their respective standard deviation and set their means to zero.



1. After that, initialize theta and cost function values to zero and define the learning rate.



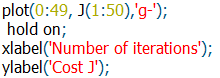
1. Lastly, apply the Cost Function formula. Do not for get to include the gradient descent formula used from the last experiment.



The cost function is excuted in this manner in octave,



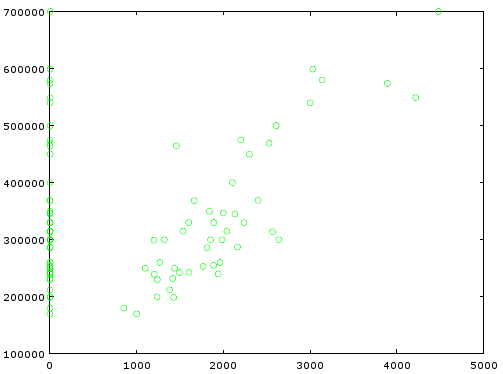
1. Create a 2D plot showing the relationship of the Cost J function with respect to the number of iterations.



# Data and Results

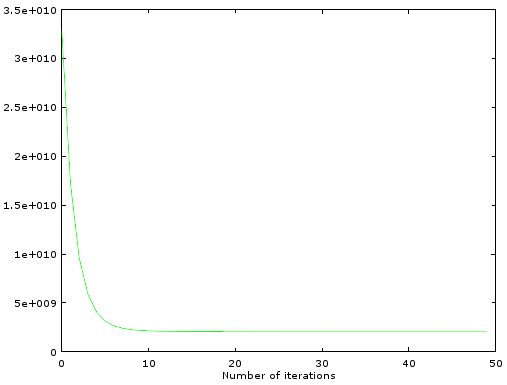
### 1st Procedure

## Plot:

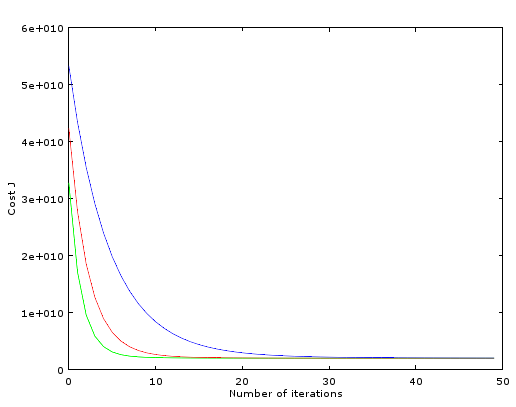


### 2nd Procedure

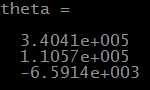
## Plot:



## Plot:

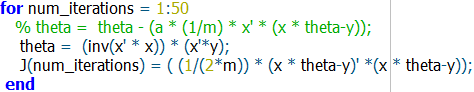


The characteristic curve becomes better in terms of its learning rate as the value is higher, the green colored curve represents the best learning curve because it is faster to have lesser deviation from the reference data meaning it is faster to have the least mistakes. The best learning rate is valued at 0.3.





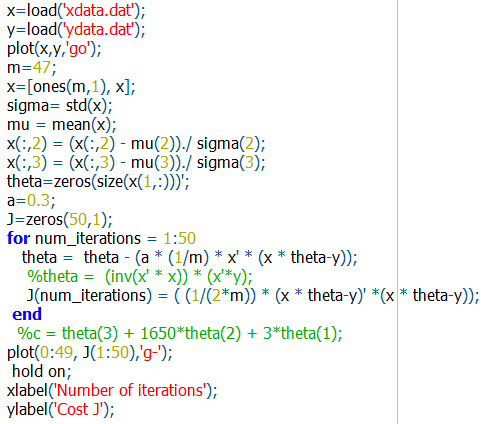
### 3rd Procedure





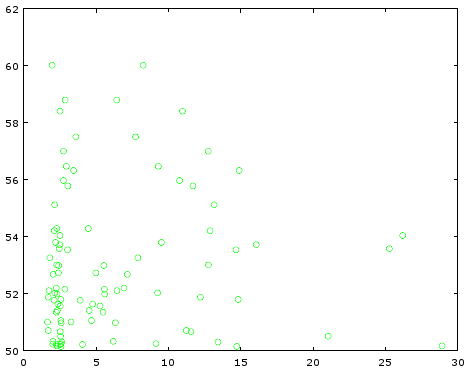
### Problem

1. Same with procedure with the experiment however with different data values.



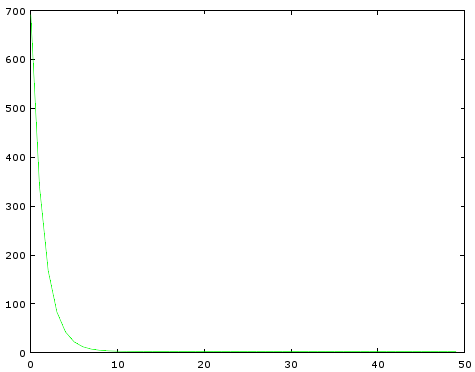
### 1st Procedure

Plot:

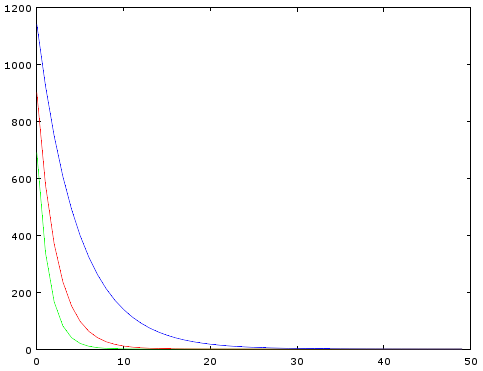


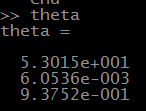
1. 2nd Procedure

Plot:



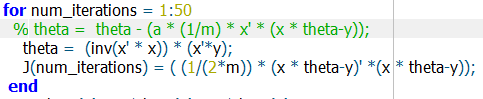
Plot:







1. 3rd Procedure





# Analysis and conclusion

In this experiment, a set of training data with the values of the size of the house, price of house and number of bedrooms in the house were analyzed. These variables are represented by their respective theta value (e.g. theta (1), theta (2), etc.). The important part of this experiment is to predict a certain data based on the training data provided and to apply a closed form solution instead of a gradient descent approach which would theoretically yields the same values.

In conclusion, the application of a closed form solution yields an equal value of data with the gradient descent approach however sometimes with larger numbers, the values tend to deviate a little bit.

# REFERENCE

[1] B. Hidalgo and M. Goodman, 'Multivariate or Multivariable Regression?', Am J Public Health, vol. 103, no. 1, pp. 39-40, 2013.

[2] (n.a). [Online]. Available: http://www.public.iastate.edu/~maitra/stat501/lectures/MultivariateRegression.pdf/. [Accessed: 15- Sep- 2015].

[3] Stattrek.com, 'Linear Regression', 2015. [Online]. Available: http://stattrek.com/regression/linear-regression.aspx. [Accessed: 08- Sep- 2015].

[4]G. implementation, 'Gradient Descent Matlab implementation', Stackoverflow.com, 2015. [Online]. Available: http://stackoverflow.com/questions/21799435/gradient-descent-matlab-implementation. [Accessed: 08- Sep- 2015].

[5]O. Niburski, 'Hello.Oskar', Raksonibs.github.io, 2015. [Online]. Available: http://raksonibs.github.io/blog/22/. [Accessed: 08- Sep- 2015].