Regularization

Lab Report #4

|  |  |  |  |
| --- | --- | --- | --- |
| Sidney Carlo P. Lopez (11222522)  sidneylopez18@gmail.com | |  | Jeremy B. Reccion (11210192)  jeremybreccion@yahoo.com |
| Carl Michael O. Sy (11208929)  carl.michael.sy@gmail.com | Wellington S. Villasin (11209534)  wellvill08@gmail.com | | |

*Abstract*— Regularization is commonly used for the selection of models where it would avoid over fitting the given models with utmost values of the parameter and is also considered the process where the establishment of new information in order to solve the problem of over fitting.

Keywords—regularization; linear regression; logistic regression; regression

1. INTRODUCTION

Regularization [1] is mostly used for a process where the models or the new information would avoid getting over fitting. These models are taken in the form of restrictions of the bounds of the space. There are two types of regularization methods which would help in these processes, regularized linear regression [2] and regularized logistic regression [3]. Regular logistic regression is widely used for having a feature function that has produced outstanding generalization performance which that has irrelevant features. Regular linear regression helps avoid over fitting where the models in the relationship within the dependent variable of y and more are denoted with the values of x.

1. OBJECTIVES

* To implement regularized linear regression and regularized logistic regression
* To apply regularized logistic regression to predict which passengers survived the Titanic shipwreck tragedy.

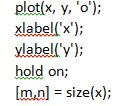
1. EXPERIMENT WITH DATA AND RESULTS
2. Procedure 1

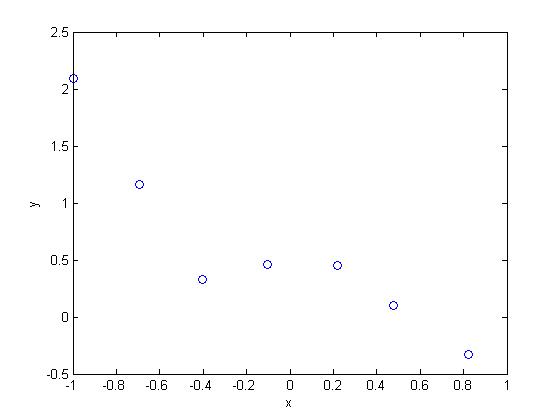
Plot the Data

Load the data files “ml4Linx.dat” and “ml4Liny.dat” into your program. These correspond to the “x” and “y” variables that you will start out with. The input “x” is a single feature, so you can plot y as a function of x on a 2-dimensional graph

C:\Users\sidneycarlolopez\Desktop\1.JPG

Plot





Plot of x and y

1. Procedure 2

Using the normal equation, find values of Theta using the three regularization parameters below:

1. Lambda = 0 (this is the same case as non-regularized linear regression)
2. Lambda =1
3. Lambda=10

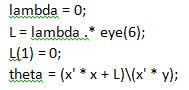
5th order polynomial

C:\Users\sidneycarlolopez\Desktop\1.JPG

Initialize Theta

C:\Users\sidneycarlolopez\Desktop\1.JPG

Set Lambda(0,1,10)



Theta values with the corresponding Lambdas

1. Lambda=0

C:\Users\sidneycarlolopez\Desktop\1.JPG

1. Lambda = 1

C:\Users\sidneycarlolopez\Desktop\1.JPG

1. Lambda =10

C:\Users\sidneycarlolopez\Desktop\1.JPG

1. Procedure 3

When you have found the answers for theta, verify them with the values in the solutions. In addition to listing the values for each elemental theta\_j of the theta vector, we will also provice the L2-norm of theta so you can quickly check if your answer is correct. In Octave, you can calculate the L2-norm of a vector x using the command norm(x).

L2 Norm



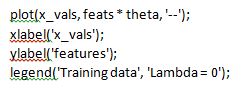
Set x axis

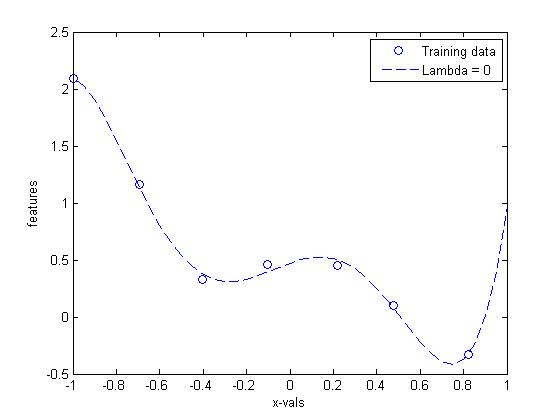
C:\Users\sidneycarlolopez\Desktop\1.JPG

Set regularization line

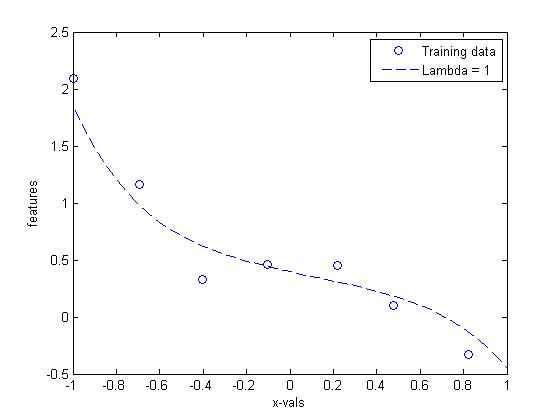
C:\Users\sidneycarlolopez\Desktop\1.JPG

Plot

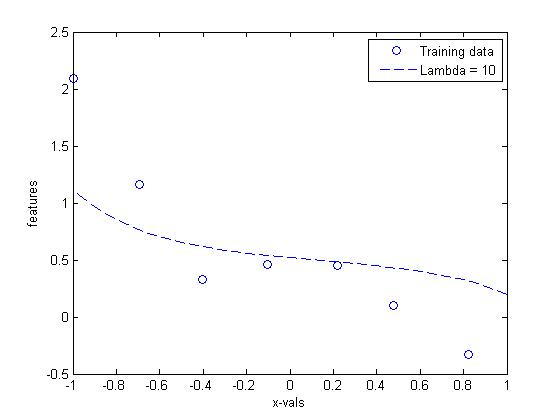




Plot of x\_vals and features at Lambda = 0



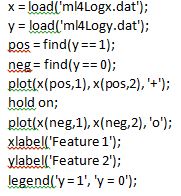
Plot of x\_vals and features at Lambda = 1

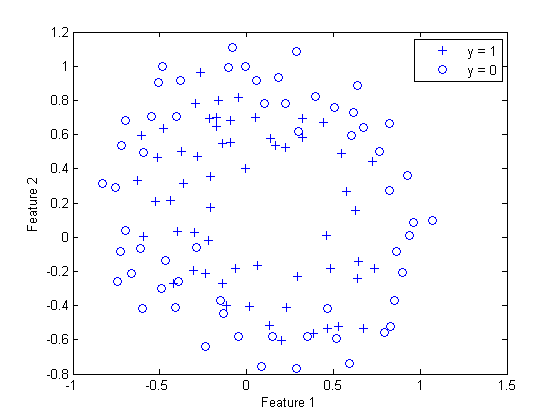


Plot of x\_vals and features at Lambda = 10

1. Procedure 4

After loading the data, plot the points using different markers to distinguish between the two classifications. The commands in Matlab/Octave will be





Plot of Feature 1 and Feature 2

1. Procedure 5

Run Newton’s Method using the three values of lambda below:

1. Lambda = 0 (this is the same case as non-regularized logistic regression)
2. Lambda = 1
3. Lambda = 10

Map feature

C:\Users\sidneycarlolopez\Desktop\1.JPG

Initialize theta

C:\Users\sidneycarlolopez\Desktop\1.JPG

Sigmoid

C:\Users\sidneycarlolopez\Desktop\1.JPG

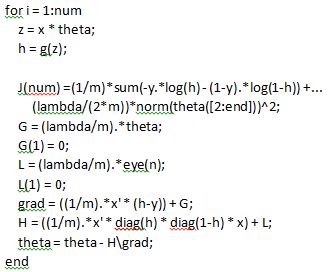
Initialize J and iterations

C:\Users\sidneycarlolopez\Desktop\1.JPG

Set lambda (0,1,10)

C:\Users\sidneycarlolopez\Desktop\1.JPG

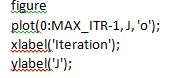
Newton’s method



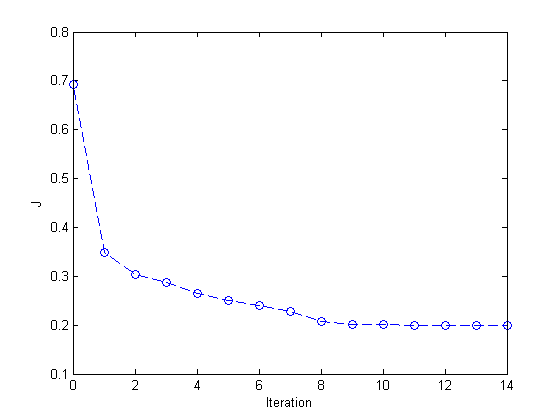
1. Procedure 6

Print out the values of J(theta) during each iteration. J(Theta) should not be decreasing at any point during Newton’s Method. If it is, check that you have defined J(theta) correctly. Also check you definitions of the gradient and Hessian to make sure there are no mistakes in the regularization parts.

Print J for each lambda



Lambda = 0



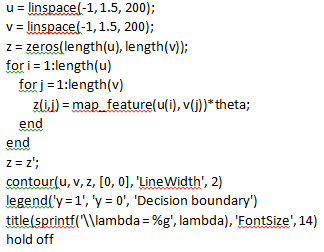
Plot of iteration and J at lambda = 10

1. Procedure 7

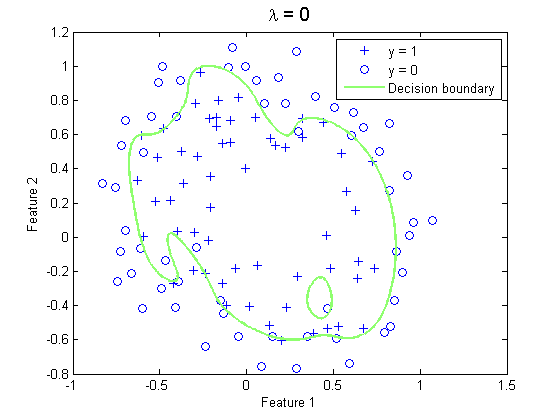
After convergence, use your values of theta to find the decision boundary in the classification problem. The decision boundary is defined as the line.



Decision boundary line

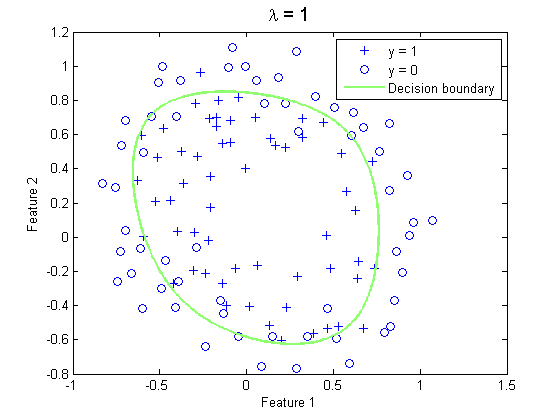


Lambda = 0



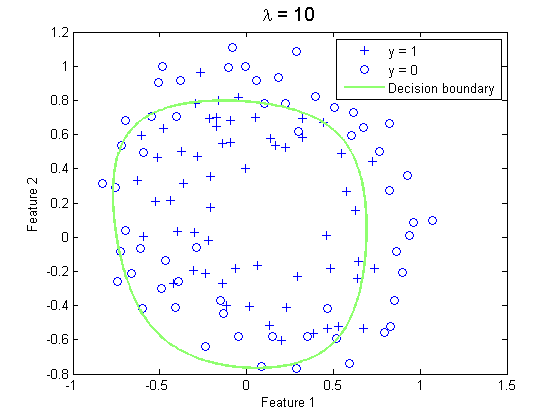
Plot of iteration J at lambda = 0

Lambda = 1



Plot of iteration J at lambda = 1

Lambda 10



Plot of iteration and J at lambda = 10

1. Procedure 8

Finally, because there are 28 elements theta, we will not provide an element-by-element comparison in the solutions. Instead use norm(theta) to calculate the L2- norm of theta, and check it against the norm in the solutions.

L2 Norm

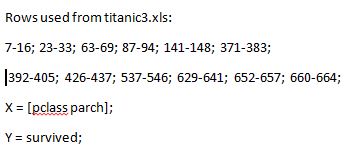
C:\Users\sidneycarlolopez\Desktop\1.JPG

1. Application Objective

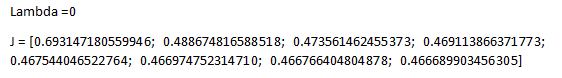
Apply regularized logistic regression to predict which passengers survived the Titanic shipwreck tragedy. Choose two features from the ‘titanic3.xls’ data and utilize it in Regularized Logistic Regression Implementation by repeating procedures 4.4-4.8. Give the necessary plots and analysis for each procedure.

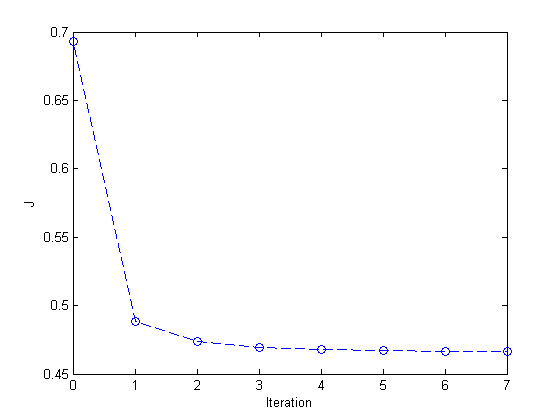
Titanic Data

Rows used from titanic3.xls:



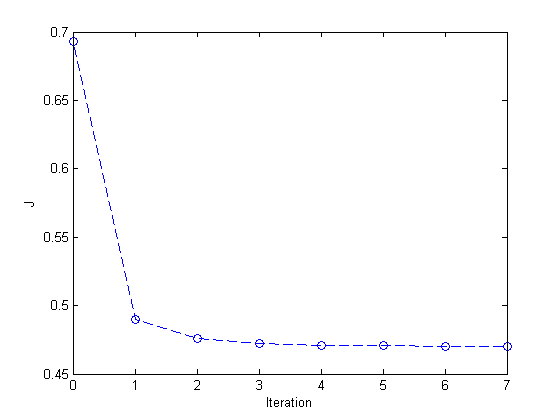
Procedure 6

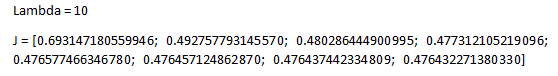


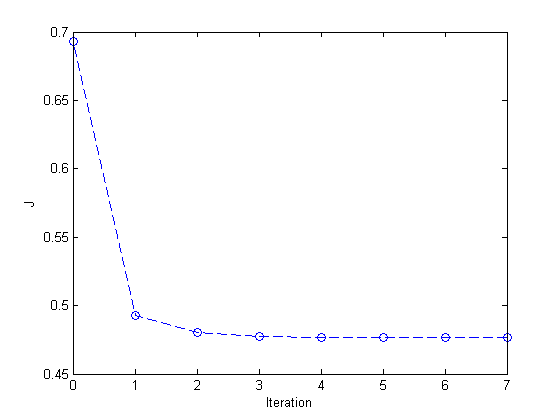


Lambda 1



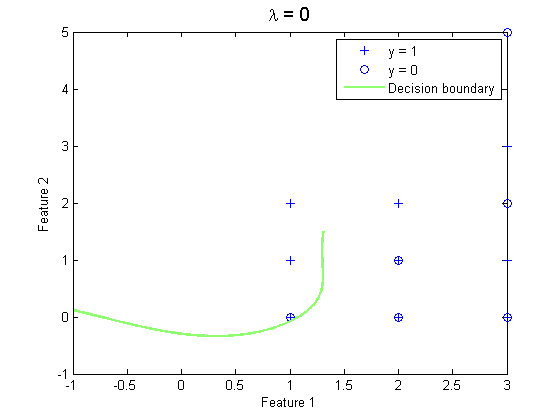




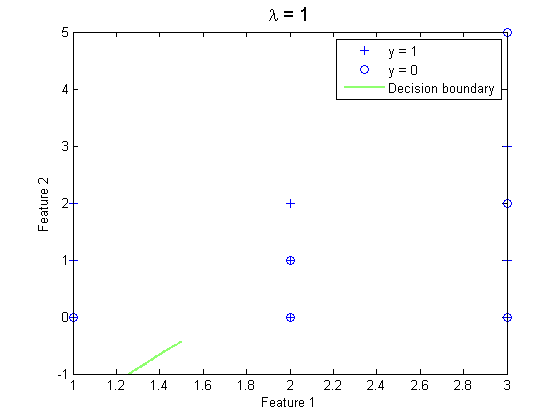


Procedure 7

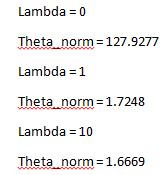
Lambda = 0



Lambda = 1



Procedure 8



1. ANALYSIS AND CONCLUSION

Regularized linear regression is like an approximation, wherein it has increased accuracy readings compared to the simple linear regression and logistic regression. Regularized linear regression guards against under fitting and over fitting as shown in the experiment above, in which the goal of regularization is to minimize the cost function with respect to theta. Lambda is the regularization parameter in which it controls the fitting parameters. When the magnitude of the fitting parameters increases, the penalty on the cost function also increases. The group concluded that when lambda increases, the norm of theta increases, therefore adjusting the lambda value gives more control over the data fitting. In using the Newton’s method, convergence takes about 15 iterations when the value of lambda is 0, and 5 or lower when the value of lambda is 1 or 10. In the experiment, when lambda is equal to 0, there is over fitting and when it is 10, there is under fitting. We concluded that the best lambda for the 2nd experiment is 1 for the decision boundary separates the positives and negatives fairly well.

REFERENCES

[1] 1]P. Bickel, 'Regularization in Statistics', Berkeley, 2015. [Online]. Available: https://www.stat.berkeley.edu/~bickel/Test\_BickelLi.pdf. [Accessed: 29- Sep- 2015].

[2] [2]C. Guestrin, 'Regularization for Logistic Regression', University of Washington, 2013. [Online]. Available: http://courses.cs.washington.edu/courses/cse599c1/13wi/slides/l2-regularization-online-perceptron.pdf. [Accessed: 29- Sep- 2015].

[3] [3]A. Rosenberg, 'Linear Regression with Regularization', The City University of New York, 2009. [Online]. Available: http://eniac.cs.qc.cuny.edu/andrew/gcml/lecture5.pdf. [Accessed: 29- Sep- 2015].