Regularization

|  |  |  |
| --- | --- | --- |
| Jan Carlo D. Rabacca  11106344 | Suzette Dela Cruz  11118113 | Marc Janssen Chiu  11114525 |

# Introduction

Regularization is method that reduces overfitting by adding complexity penalty to the loss function, adding more variables may cause overfitting, causing the machine to poorly predict outcomes [1].

# Procedures

1. First, load the training data and create a 2D representation plot (refer to Fig.1).



Fig. 1

1. Set the hypotheses by using high order polynomials. (refer to Fig.2).



Fig. 2

1. Initialize the required variables (refer to Fig.3).

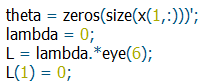
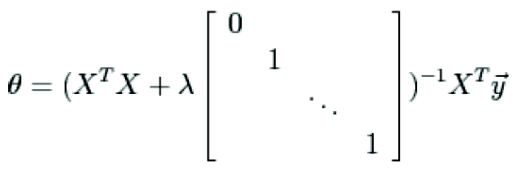


Fig. 3

1. Find the best parameters of the model by using the normal equation (refer to Eqn.1).



(Eqn. 1)

The implementation should look like this: (refer to Fig.4)



Fig. 4

1. The second set of data is for regularized logistic regression, first load the training data (refer to Fig.5).



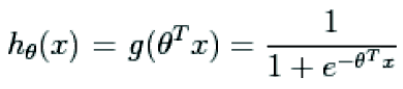
Fig. 5

1. After loading the data, separate two classifications by assigning “+” and “o” markers (refer to Fig.6).



Fig. 6

1. Fit a regularized regression model to the training data using the logistic hypotheses function (refer to Eqn.2).



(Eqn. 2)

The implementation should look like this: (refer to Fig.7)



Fig. 7

Remember to initialize the function wherein: (refer to Fig. 8)



Fig. 8

1. Assign the x in the sigmoid function to be monomials up to the six power (refer to Fig. 9).

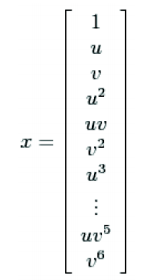


Fig. 9

1. For convenience, use the map\_feature.m (refer to Fig. 10)



Fig. 10

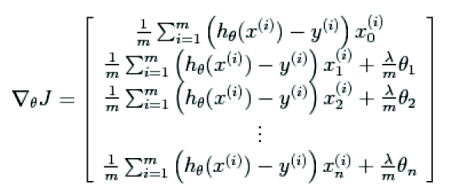
1. Implement the cost j function and implement newton’s method update rule where the hessian and gradient functions are utilized (refer from Eqn. 3 to Eqn. 6).



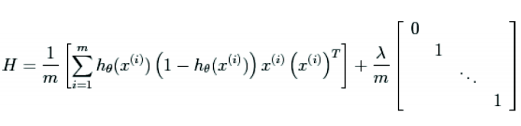
Eqn. 3 Cost J function



Eqn. 4 Newton’s update rule



Eqn. 5 Gradient



Eqn. 6 Hessian

The implementation of these equations should look like: (refer to Fig. 11)

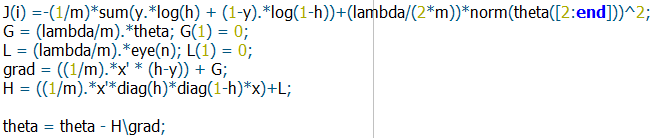
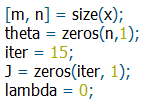


Fig. 11

1. Do not forget to initialize the important values to be used on the formulas (refer to Fig. 12).

  
Fig. 12

# Data and Results

## Procedure 4.1

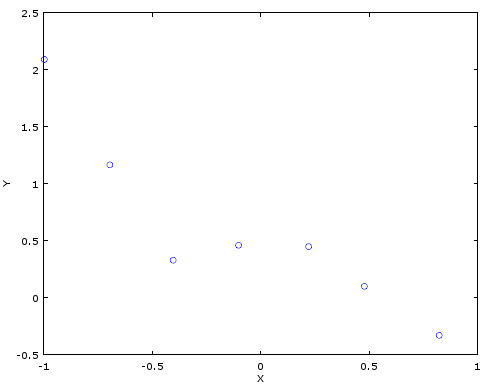


Fig. 13 X-data vs. Y-data Regularized Linear Regression

## Procedure 4.2

|  |  |  |  |
| --- | --- | --- | --- |
|  | λ = 0 | λ = 1 | λ = 10 |
| norm ϴ | 8.1687 | 0.80977 | 0.59307 |
|  | 0.47253 | 0.39760 | 0.520471 |
|  | 0.68135 | -0.42067 | -0.182507 |
|  | -1.38013 | 0.12959 | 0.060643 |
|  | -5.97769 | -0.39747 | -0.148177 |
|  | 2.44173 | 0.17526 | 0.074330 |
|  | 4.73711 | -0.33939 | -0.127957 |

Table 1. Theta values obtained using regularization parameters (lambda)

## Procedure 4.3

## lambda = 0

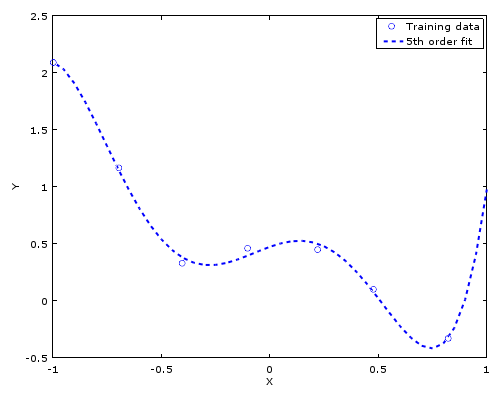


Fig. 14

* *lambda = 1*

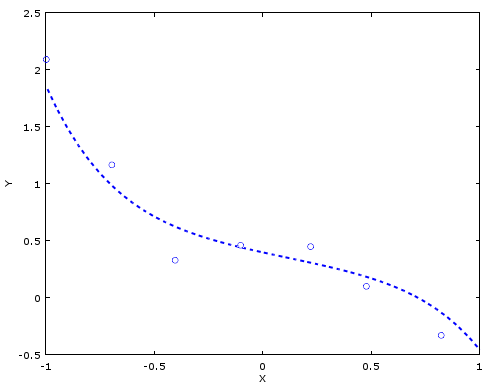


Fig. 15

* *lambda = 10*

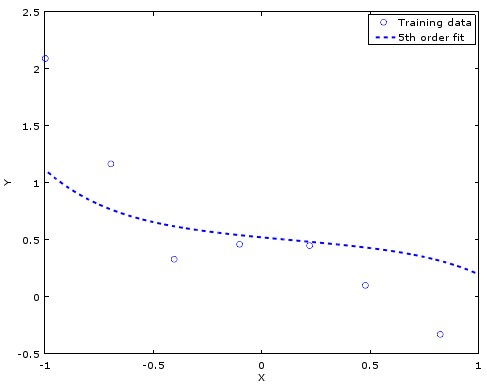


Fig. 16

## Procedure 4.4

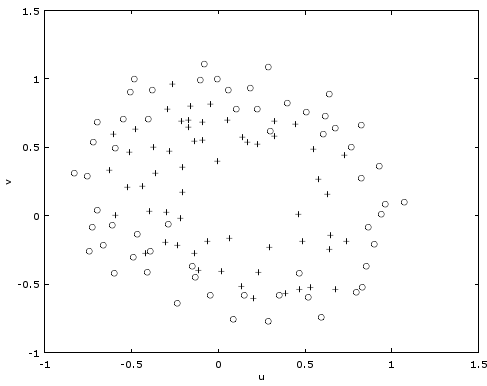


Fig.17 Regularized

## Procedure 4.5

|  |  |  |  |
| --- | --- | --- | --- |
|  | λ = 0 | λ = 1 | λ = 10 |
| norm ϴ | 7172.7 | 4.2400 | 0.03842 |

Table 2. Theta values obtained using regularization parameters (lambda)

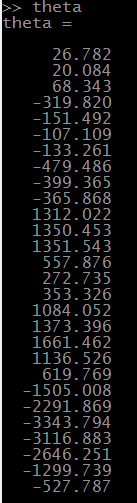
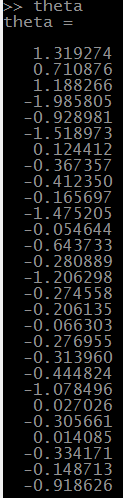
 

Fig. 18 λ = 0 Fig. 19 λ = 1

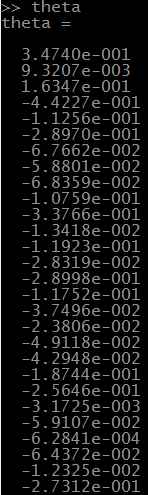


Fig 20 λ = 10

## Procedure 4.6

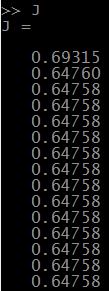


Fig. 21 J values

## Procedure 4.7

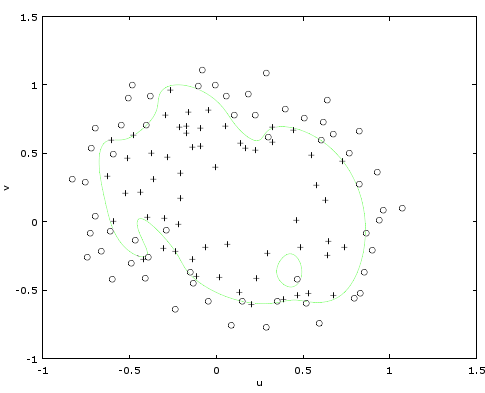


Fig. 22 Decision Boundary lambda = 0

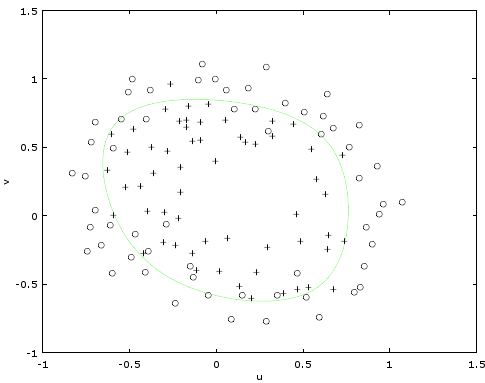


Fig. 23 Decision Boundary lambda = 1

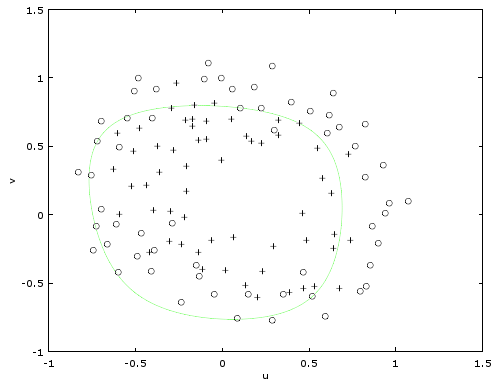


Fig. 24 Decision Boundary lambda = 10

## Titanic data

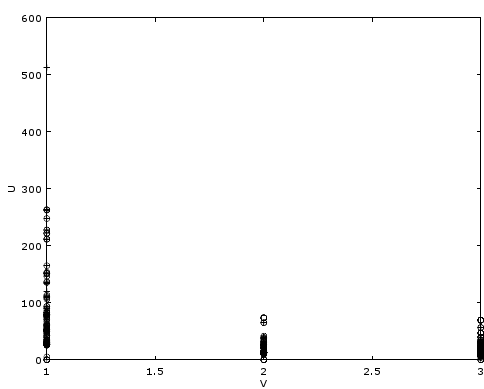


Fig. 25 Data plot using class and fare variables

|  |  |  |  |
| --- | --- | --- | --- |
|  | λ = 0 | λ = 1 | λ = 10 |
| norm ϴ | 4.69e-008 | 4.6876e-008 | 4.6876e-008 |

Table 2. Theta values obtained using regularization parameters (lambda)

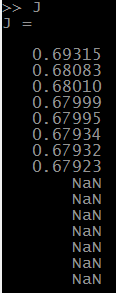


Fig. 26 J values

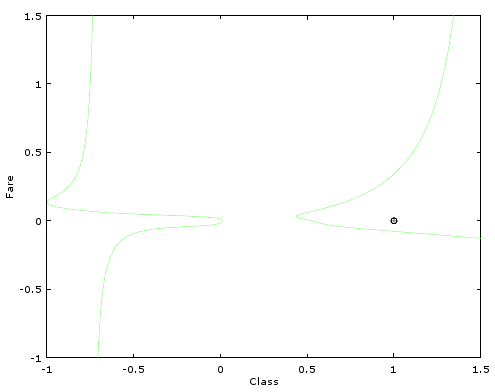


Fig. 27 Decision Boundary lambda = 0

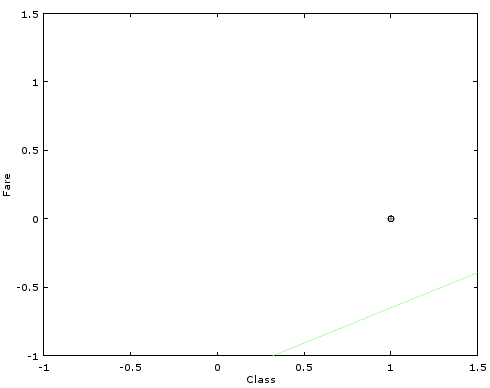


Fig. 28 Decision Boundary lambda = 1

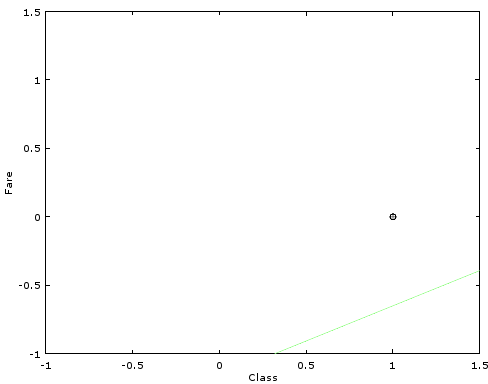


Fig. 28 Decision Boundary lambda = 10

# Analysis and conclusion

When more data is added, more often than not, the hypotheses becomes overfitted or it tried too much to estimate the precise values causing the machine to fail to generalize the data making a poor estimate or prediction of the outcome based on the training data.

Regularization prevents overfit by adding penalty or introducing additional information, this important to maintain the learning state of the machine, to make the machine able to generalize and yield acceptable predictions.

# REFERENCES

1. 2015. [Online]. Available: http://www.eecs.berkeley.edu/~russell/classes/cs194/f11/lectures/CS194%20Fall%202011%20Lecture%2004.pdf. [Accessed: 29- Sep- 2015].
2. Quora.com, 'What is regularization in machine learning? - Quora', 2015. [Online]. Available: https://www.quora.com/What-is-regularization-in-machine-learning. [Accessed: 29- Sep- 2015].