Training vs Testing

Attended & did the lab work [Y/N]?

DANTING, Jules A. (10934375)

DE VERA, Luis Paolo M. (11113944)

LIM, Alyssa Raphaella ()

RAMOS, Brian Justin S. (11127120)

VALDEZ, Esteen Rae Salvador ()

1 Introduction and Main Concepts

LINEAR Regression is a simple approach to supervised learning. It is a simple approach for modeling the relationship between dependent variable y and independent variables x.This approach is also used to predict the values of a variable given the values of another variable. Plotting variable/s X and Y will form a straight line.

Cost Function is a function that measures the error in predicting the dependent variable, y. The goal in linear regression is to minimize the cost function. The higher the cost function, the higher is the error.

Gradient Descent is an algorithm that can help in solving the linear regression of the given data. It is an algorithm that minimizes its function using iteration. There will be an initial values and it will iteratively moves toward a set of parameter values that will help in reducing the function. It is the best way to minimize the cost function.

The experiment will not use the gradient descent but it will make use of differential equation. The cost function will be differentiated with respect to a parameter thetaj

2 Methodology

2.1 Implementation

In the `ex1/` directory of the starter code package you will find the file `ex1_linreg.m` which contains the makings of a simple linear regression experiment. This file performs most of the boiler-plate steps for you:

1. The data is loaded from housing.data. An extra 1 feature is added to the dataset so that 1 will act as an intercept term in the linear function.

2. The examples in the dataset are randomly shuffled and the data is then split into a training and testing set. The features that are used as input to the learning algorithm are stored in the variables train.X and test.X. The target value to be predicted is the estimated house price for each example. The prices are stored in train.y and test.y, respectively, for the training and testing examples. You will use the training set to find the best choice of for predicting the house prices and then check its performance on the testing set.

3. The code calls the minFunc optimization package. minFunc will attempt to find the best choice of by minimizing the objective function implemented in `linear_regression.m.` It will be your job to implement `linear_regression.m` to compute the objective function value and the gradient with respect to the parameters.

4. After minFunc completes (i.e., after training is finished), the training and testing error is printed out. Optionally, it will plot a quick visualization of the predicted and actual prices for the examples in the test set.

The `ex1_linreg.m` file calls the `linear_regression.m` file that must be filled in with your code. `The linear_regression.m` file receives the training data X, the training target values (house prices) y, and the current parameters .

Complete the following steps for this exercise:

Fill in the `linear_regression.m` file to compute J() for the linear regression problem as defined earlier. Store the computed value in the variable f.

You may complete both of these steps by looping over the examples in the training set (the columns of the data matrix X) and, for each one, adding its contribution to f and g. We will create a faster version in the next exercise.

2.2 Evaluation

Once you complete the exercise successfully, the resulting plot should look like the plot from the manual

3 Results and Discussion

Based on the procedure the main functionality of the experiment is to oversee linear regression given a set of data from housing.data file. By using the starter code package `ex1_linreg.m`

By affixing our signatures, we, the authors, pledge that: we have completed this course work on our own; we have not used any unauthorized material/assistance/help on this course work; and we have not given directly or indirectly to any other student/unauthorized person/means any access to any part of the specified course work. Coram Deo.

Lab Work Date: Oct. 28, 2015
Submission Date: Oct. 26, 2015

it will be implemented as the main function of the experiment. The task of the group was to implement a linear regression on the given data. The linear regression will be based on the variables of the training data and testing data. The training data will predict the estimated house price while the testing data will check the performance. The variables are then called by the main function `ex1_linreg.m` in order to display the results that were simulated in the `linear_regression.m` function. A plot diagram of House Price vs House Numbers will be displayed having both the actual and predicted price dipicted in the graph to see the differences between the two.

3.1 MATLAB CODE

```
function [f,g] = linear_regression(theta, X,y)
  %
  % Arguments:
  %   theta - A vector containing the parameter values to optimize.
  %   X - The examples stored in a matrix.
  %       X(i,j) is the i'th coordinate of the j'th example.
  %   y - The target value for each example.
y(j) is the target for example j.
  %

  m=size(X,2)
  n=size(X,1)
  %XS=size(X)
  %t=size(theta)
  %yS=size(y)
    f=0;
  g=zeros(size(theta));

  % TODO:  Compute the linear regression objective by looping over the examples in X.
  %        Store the objective function value in 'f'.
  %
  % TODO:  Compute the gradient of the objective with respect to theta by looping over
  %        the examples in X and adding up the gradient for each example.
Store the
  %        computed gradient in 'g'.

    %%% YOUR CODE HERE %%%
    %Linear Regression
    for i = 1:m
    f = f + (theta' * X(:,i) - y(i))^2;
    end
    f = 1/2*f;
    % Step 2: Compute gradient
    for j = 1:n
    for i = 1:m
        g(j) = g(j) + X(j,i)*(theta' * X(:,i) - y(i));
    end
    end
end
```

4 Conclusion

In this experiment the group was able to simulate a linear regression graph using the training data provided. The group used the function minfunc which is also similar to fminunc which is used in nonlinear programming. In this experiment the training data forms a nonlinear graph. The group used the minfunc function inorder to get the estimated price for the house with the help of the training data provided.

5 Answers to Questions

References