

Learning the Model

In this exercise, you will implement a linear regression model to predict how many miles per gallon (MPG) a particular car can drive. Each car will be described by seven attributes provided in the data file `CarData.mat`. Our goal is to learn the weights w associated with every attribute such that our predicted MPG value would be as close to the true value as possible. Formally, we want to learn the weights w by minimizing the following learning objective:

$$E(w) = \frac{1}{2} \sum_{i=1}^n [y^{(i)} - w^T x^{(i)}]^2$$

where $y^{(i)}$ denotes a real valued ground truth MPG value for a data point i , and $x^{(i)}$ depicts a 7-dimensional feature vector that describes 7 car attributes associated with a data point i . Now assume that all features are stored into a $n \times 7$ dimensional matrix X , where each row i represents a data point i and every column in the matrix represents one of the seven features. Furthermore assume that a $n \times 1$ dimensional vector y stores every ground truth MPG values. Then, to obtain the weights w that minimize the error function $E(w)$ we need to solve the following system of linear equations:

$$(X^T X)w = X^T y$$

Your job is to solve this system and find the weights w . Your function will take a feature matrix X and a label vector y as its inputs, and return a vector storing the 7-dimensional weights w . Please note that you are not allowed to use the matlab function `inv()` since it introduces numerical errors and is very slow in practice. Instead, you should use the Matlab's backslash `"\"` operator for solving this system of equations.

In case you want to work on this problem outside of the EdX environment, we provide the data file here: [CarData.mat](http://courses.edx.org/asset-v1:PennX+ROBO2x+2T2017+type@asset+block@CarData.mat) (<http://courses.edx.org/asset-v1:PennX+ROBO2x+2T2017+type@asset+block@CarData.mat>).

Note that the parts 3a, 3b, 3c of this assignment are related and build on top of each other. Therefore, if you are working outside of the EdX environment, you need to complete each part before proceeding to another one. If you are working inside the EdX environment, it is highly recommended but not necessary to complete one part before starting another. Note that even if your code in the earlier parts is incorrect, you can still get the full credit for the later parts if they are implemented correctly.

Your Function

 Save  Reset  MATLAB Documentation (<https://www.mathworks.com/help/>)

```
1 function w = LinearRegression(X,y)
2     % learn a linear regression model
3     %
4     % Input:
5     % - X: n x 7 dimensional feature matrix where every row depicts a particular observation and every column d
6     % - y: n x 1 dimensional ground truth MPG labels
7     % Output:
8     % - w: 7 x 1 dimensional weights of a learned linear regression model
9
10    A = X' * X;
11    b = X' * y;
12    w= A\b;
13
14 end
15
```

Code to call your function

 Reset

```
1
2
```

```
load('CarData.mat');  
w=LinearRegression(trainsetX,trainsetY);
```

▶ Run Function



Previous Assessment: All Tests Passed

Submit



✔ Are the Learned Weights Correct?

✔ Does the Solution Exclude a Built in 'inv()' Function?