Test1 **Patel Femina**

Machine Learning **N03094115**

```
% converting wine.csv to wine.mat
Wine=csvread('wine.csv');
save('wine.mat','Wine')
```

Question 1

```
% Taking the first 1280 data rows and calling them the training set
trainingset=Wine(2:1281,:);
```

Question 2

```
% Taking the remaining rows and calling them the test set
testset=Wine(1282:1600,:);
```

Question 3

```
% Model M1 using training set
x=trainingset(:,[1 7]);
% columns 1 and 7 are for acidity and sulphite level of wine respectively
y=trainingset(:,9);
% column 9 is for quality of wine
[J_h,t_h]=testGradientDescend(x,y,1.0,0.5,0.5,0.01,2000)
% Given values for Θ0, Θ1, Θ2, α and max Iterations are 1.0, 0.5, 0.5, 0.01, 2000 respectively
```

% value of J_h goes from 0.62720 to 0.35760 and the corresponding Θ values are 1, 0.5, 0.5 to %2.96316, 0.19935, 1.40583. Lower the value of cost function, better the model.

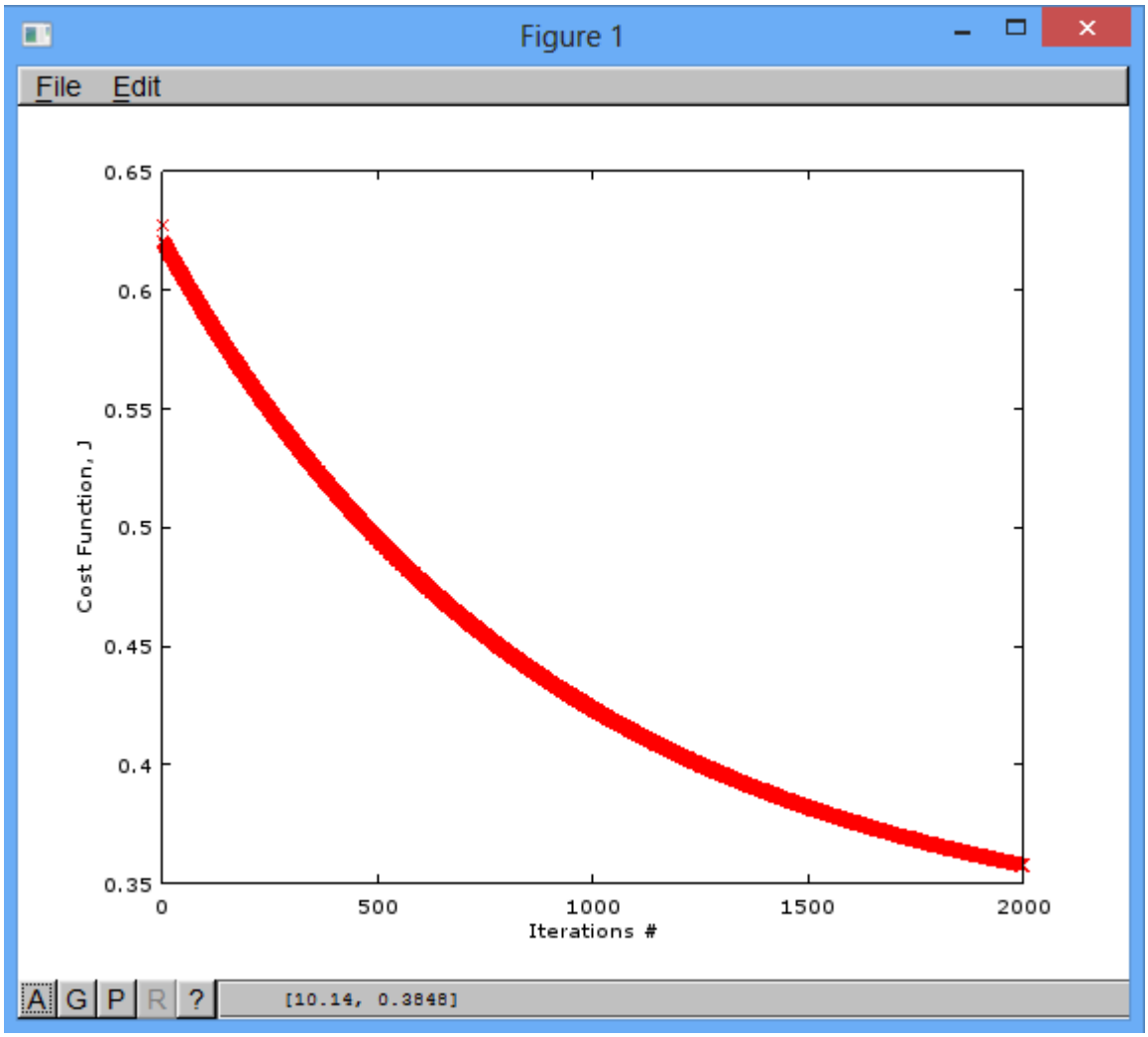
% So θ_0=2.96316, θ_1=0.19935 and θ_3=1.40583

% Graph of the cost function as the iterations proceed for Model M1

```
plot(J_h,'rx');

xlabel('Iterations #');

ylabel('Cost Function, J');
```



Hypothesis function for model M1, $h_\Theta(x) = \Theta_0 * x_0 + \Theta_1 * x_1 + \Theta_2 * x_2$

By default we take $x_0=1$

So equation is $\Theta_0 + \Theta_1 * x_1 + \Theta_2 * x_2$

Placing the values of Θ_0, Θ_1 and Θ_2 from above we get

$h_\Theta(x) = 2.96316 + 0.19935 * x_1 + 1.40583 * x_2$

Question 4

% Model M2 using training set

x1=trainingset(:,[3 6]);

% columns 3 and 6 are for sugar and pH level of wine respectively

y1=trainingset(:,9);

% column 9 is for quality of wine

[J_h,t_h]=testGradientDescend(x1,y1,1.0,0.5,0.5,0.01,2000)

% Given values for Θ_0, Θ_1, Θ_2, α and max Iterations are 1.0, 0.5, 0.5, 0.01, 2000 respectively

% value of J_h goes from 2.04119 to 0.35864 and the corresponding Θ values are 1, 0.5, 0.5 to %1.450857, 0.038766, 1.243855. Lower the value of cost function, better the model.

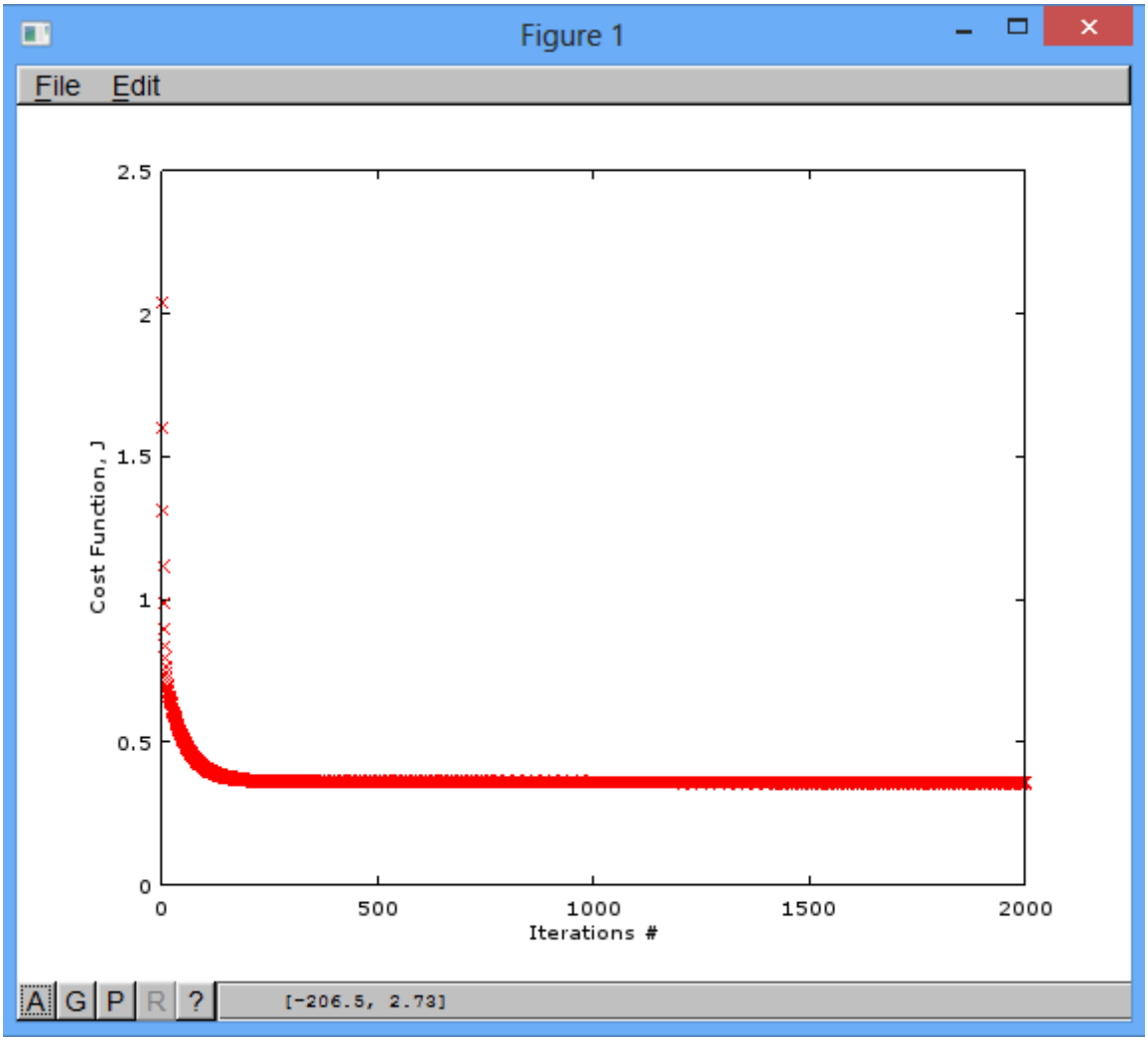
% So **θ_0=1.450857, θ_1=0.038766 and θ_3=1.243855**

% Graph of the cost function as the iterations proceed for Model M2

plot(J_h,'rx');

xlabel('Iterations #');

ylabel('Cost Function, J');



Hypothesis function for model M2, $h_\Theta(x) = \Theta_0 * x_0 + \Theta_1 * x_1 + \Theta_2 * x_2$

By default we take $x_0=1$

So equation is $\Theta_0 + \Theta_1 * x_1 + \Theta_2 * x_2$

Placing the values of Θ_0, Θ_1 and Θ_2 from above we get

$h_\Theta(x) = 1.450857 + 0.038766 * x_1 + 1.243855 * x_2$

Question 5

% Evaluating cost function using Θ values of M1 to predict the quality of wine in the test data set

```
x3=testset(:,[1 7]);
y3=testset(:,9);
costFunForMultivariable(x3, y3, 2.96316,0.19935,1.40583)
```

ans = 0.32174

% Evaluating cost function using Θ values of M2 to predict the quality of wine in the test data set

```
x4=testset(:,[3 6]);
y4=testset(:,9);
costFunForMultivariable(x4, y4, 1.450857,0.038766,1.243855)
```

ans = 0.33412

As the value of cost function of Model M1 is lower than that of Model M2, **M1 is a better model**.