**Saurabh Weather Prediction**

**City-Folklore**

**Attempt 1** - Predict the relationship between humidity and temperature

x = csvread('weatherHistory.csv');

temp = x(2:end,4); % get the Temperature vector

humidity = x(2:end,6);

function plotData(x,y)

plot(x,y,'rx','MarkerSize',8); % Plot the data on y axis with red color X ‘rx’

xlabel('Temperature in Degree Celsius'); % Set the x-axis label

ylabel('Humidity'); % Set the y-axis label

end

m = length(temp);

X = [ones(m,1) temp]; % creating the X matrix with the input variable (temperature in this case) and will predict the theta based on the y vector which is humidity

y = humidity;

% we know the Normal Equation **Theta = (X’X)-1X’y**

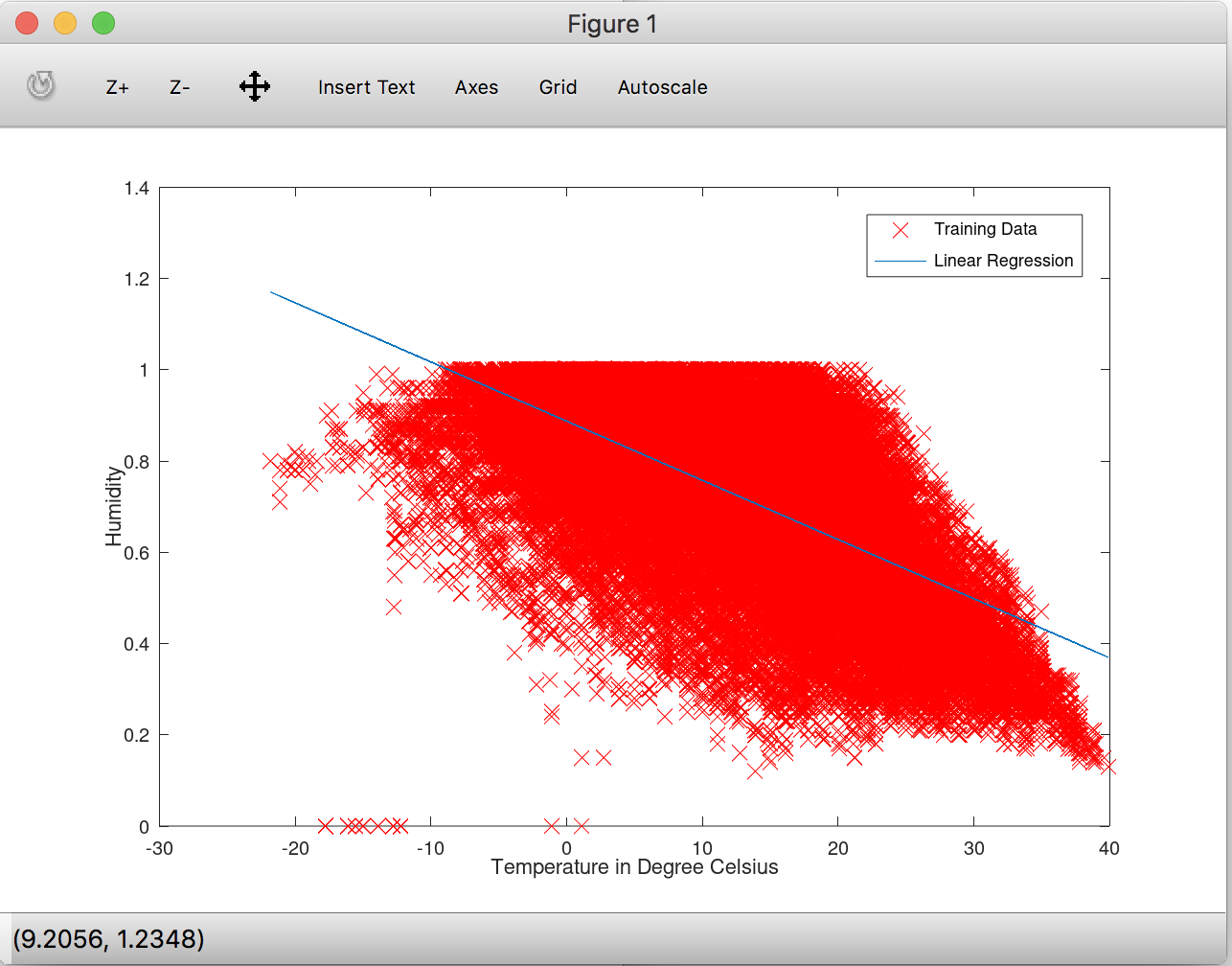
theta = (pinv(X'\*X))\*X'\*y

plotData(temp, humidity);

hold on;

plot(X(:,2), X\*theta, '-');

legend('Training Data', 'Linear Regression');



**Observations:**

Currently Temperature is on X and humidity is on Y axis and we have calculated the theta values as

Theta0 = 0.887274

Theta1 = -0.012970

If you swap the axis the theta will be calculated as

Theta0 = 34.597

Theta1 = -30.95

And this is the poor model for prediction

**How to verify the efficiency of model created?**

test = csvread('weatherTest.csv');

testx = test(2:end,4); %test input

testy = test(2:end,6); %actual prediction

testm = length(testx); % length of test input

TestX = [ones(testm,1) testx]; %Input matrix for testing

TestY = TestX \* theta; % theta is calculated from the above method

sumActual = sum(testy,1);

sumPrediction = sum(TestY,1);

A = testy-TestY;

MSError = A/testm;

MSum=sum(A,1);

MSError=MSum/testm;

Using the above model – Efficiency comes as 97% because the error loss is MSError = 0.035823

Efficiency = 96.4177%

**Attempt 2** - Predict the relationship between humidity and apparent temperature

x = csvread('weatherHistory.csv');

appTemp = x(2:end, 5);

humidity = x(2:end, 6);

function plotData(x,y)

plot(x,y,'rx','MarkerSize',8); % Plot the data on y axis with red color X ‘rx’

xlabel('Temperature in Degree Celsius'); % Set the x-axis label

ylabel('Humidity'); % Set the y-axis label

end

m = length(appTemp);

X=[ones(m,1) appTemp];

y = humidity;

theta = (pinv(X'\*X))\*X'\*y;

theta =

0.852349

-0.011039

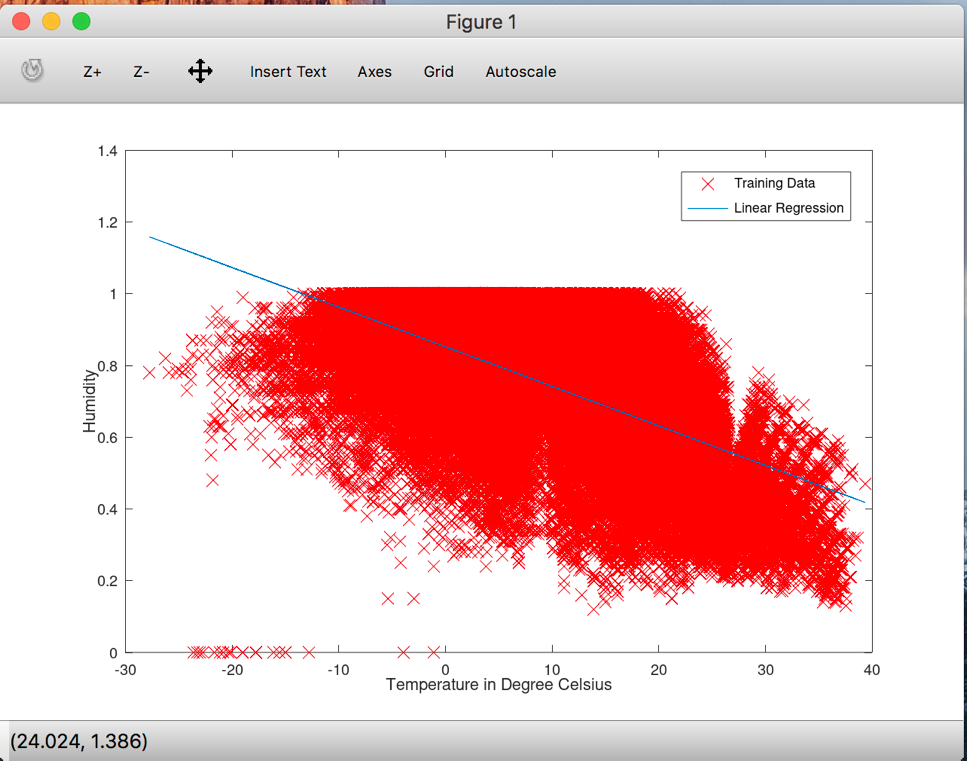
plotData(appTemp, humidity);

hold on;

plot(X(:,2),X\*theta,'-');

legend('Training Data', 'Linear Regression');

Efficiency in this case is = 96.4417%



**Attempt 3** - Model to predict the humidity given the temperature

We have already used the Linear Regression and found that the Attempt 2 is a better prediction model

**Attempt 4** - Predict the relationship between humidity and *temperature and apparent temperature*

a = csvread('weatherHistory.csv');

b = a(2:end,4);

c = a(2:end,5);

d = a(2:end,6);

plot3(b,c,d) % plotting a 3d plot

m = length(b);

function plotData(x,y,z)

plot3(x,y,z,'MarkerSize',8);

xlabel('Temperature in Degree Celsius');

ylabel('Apparent Temperature');

zlabel('Humidity');

end

plotData(b,c,d); % trial of plotting data

X=[ones(m,1) b c];

y = d;

theta = (pinv(X'\*X))\*X'\*y;

plotData(b,c,d);

hold on;

plot3(X(:,2),X(:,3),X\*theta,'x');

**Checking Efficiency**

test = csvread('weatherTest.csv');

testx = test(2:end,4);

testx1 = test(2:end,5);

testy = test(2:end,6);

testm = length(testx);

testm

TestX = [ones(testm, 1) testx testx1];

TestY = TestX\*theta;

sumActual = sum(testy,1);

sumPred = sum(TestY,1)

A= testy-TestY;

MSError = A/testm

MSum = sum(A,1);

MSError = MSum/testm;

MSError

Effeciency = 96.3853%

