# Homework 1

## Question 1

CODE (Python 3.0):

#import needed libraries

import numpy as np

import matplotlib

import matplotlib.pyplot as plt

get\_ipython().magic('matplotlib inline')

import math

import cmath

test\_g = np.random.normal(0,.15,100)

train\_g = np.random.normal(0,.15,100)

def ERMS(m):

    M = m # model order

    #Generate data of training

    #In this script. I creat 100 number of data in sine curve function.

    #Gaussian noise genrate

    g1 = train\_g #(location, scale, Number of data)

    x = np.arange(0,1,0.01) #Range of x from 0 to 1

    t = np.sin(2\*math.pi\*x)       # training t

    train\_t = np.sin(2\*math.pi\*x)+g1    # training t with gaussian

    g2 = test\_g

    X = np.array([x\*\*m for m in range(M+1)]).T # X of training data

    Z = X.T@X

    w = np.linalg.inv(Z)@X.T@t # weight

    # test set

    test\_data = np.arange(0.05,1.05,0.01)

    test\_t = np.sin(2\*math.pi\*test\_data)+g2 # test t

    X\_test\_data = np.array([test\_data\*\*m for m in range(w.size)]).T # X of test data

    test\_ERMS = abs(cmath.sqrt((np.linalg.norm((w.T@X\_test\_data.T-test\_t.T))\*\*2)/100))

    train\_predict\_curve = X@w  #training curve

    train\_ERMS = abs(cmath.sqrt((np.linalg.norm((train\_predict\_curve-train\_t))\*\*2)/100))

    return test\_ERMS,train\_ERMS

m = int(input('m='))

Erms = np.array([ERMS(m) for m in range(m)])

x = np.arange(0,m,1)

p1 = plt.plot(x,Erms[:,1],'b')

p2 = plt.plot(x,Erms[:,0],'r')

plt.scatter(x,Erms[:,1])

plt.scatter(x,Erms[:,0])

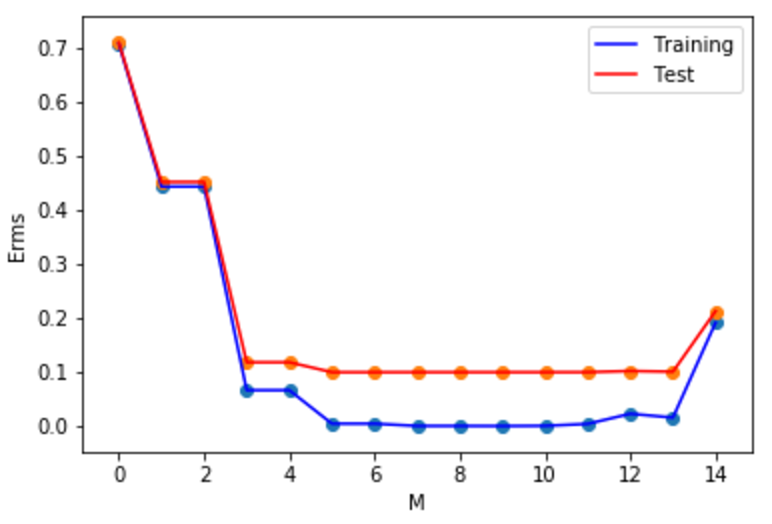
plt.xlabel('M')

plt.ylabel('Erms')

plt.rcParams["axes.titlesize"] = 16

plt.legend((p1[0],p2[0]),('Training', 'Test'), fontsize=10)

Result:



In my plot, the value of model order M should be 13 to avoid over-training. Because at this point, ERMS goes up which means predicted line is going to highly deviated from observational value.

## Question 2: For each of the following problems, state whether or not the operation is defined (i.e., valid and can be computed) and, if it is defined, what is the size of the resulting answer. For all of the following problems let X be a M×N matrix, Y be a N×N matrix, a be a M × 1 vector, b be a N × 1 vector and s be a scalar.

1. XY is defined operation. The size of XY is M\*X.
2. YX is not defined.
3. YXT is defined operation. The size of YXT is N\*M.
4. aX is not defined.
5. aTX is defined operation. The size of aTX is 1\*M.
6. aXT is not defined.
7. aTb is not defined operation.
8. bTb is defined operation. The size of it is 1.
9. bbT is defined operation. The size of it is N\*N.
10. sX+Y is not defined.

## Question 3: If X is a rank r matrix, show that the two square matrices XXH and XH X have the same nonzero eigenvalues.

Assume is the nonzero eigenvalue and u is the eigenvector of .

So, (1)

Then the vector is defined by (2)

Then put (2) into (1): =

Result can be inferred from the last function. is the eigenvalue of with eigenvector .

## Question 4: Consider f(x) = 3xT x + 4yT x − 1 where x, y ∈ Rd.

1. **What is ?**

1. **What is ?**

=6**I I is an identity matrix in result.**

Question 5: **Consider f(x)=−10xTQx+4yTx+2 where x, y∈Rd,Q∈Rd×d and Q is symmetric.**

1. **What is ?**

**Because,**

**So,**

1. **What is ?**