Homework 1

Question 1

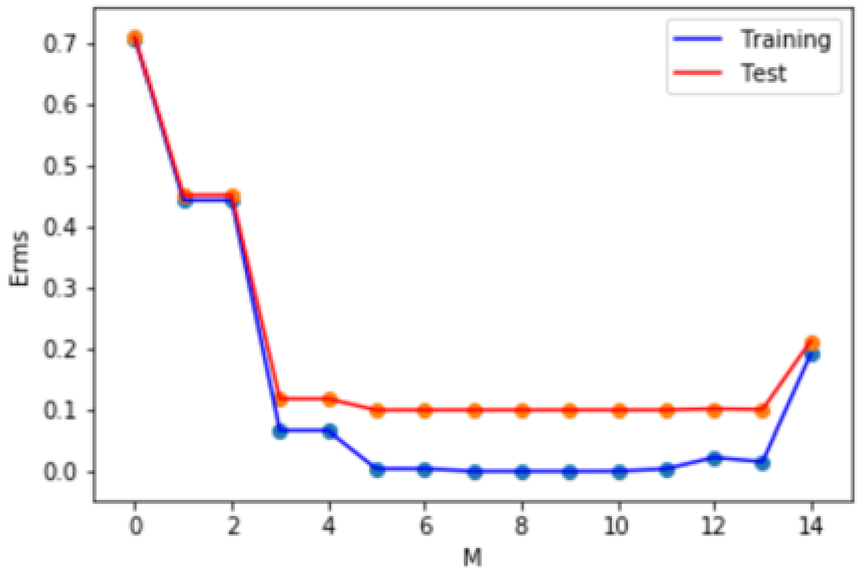
CODE (Python 3.0):

```
#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, E_{RMS} 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 YX^T is defined operation. The size of YX^T is N*M.
4 aX is not defined.
5 a^TX is defined operation. The size of a^TX is 1*M.
6 aX^T is not defined.
7 a^Tb is not defined operation.
8 b^Tb is defined operation. The size of it is 1.
9 bb^T 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 XX^H and $X^H X$ have the same nonzero eigenvalues.

Assume λ is the nonzero eigenvalue and u is the eigenvector of XX^H.
So, $XX^H u = \lambda u$ (1)
Then the vector $\vec{u}$ is defined by $\vec{u} = X^H u$ (2)
Then put (2) into (1): $X^H X\vec{u} = X^H XX^H u = X^H \lambda u = \lambda X^H u = \lambda\vec{u}$
Result can be inferred from the last function. λ is the eigenvalue of $X^H X$ with eigenvector $\vec{u}$.

Question 4: Consider f(x) = $3x^T x + 4y^T x - 1$ where x, y $\in R^d$.

(1) What is $\frac{\partial f}{\partial x}$?

$$\frac{\partial f}{\partial x} = \frac{\partial(3x^Tx+4y^Tx-1)}{\partial x} = 3\frac{\partial x^Tx}{\partial x} + 4\frac{\partial y^Tx}{\partial x} = 3(x^T + x^T) = 6x^T + 4y^T$$

(2) What is $\frac{\partial^2 f}{\partial x^2}$?

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(6x^T + 4y^T) = 6I$$

I is an identity matrix in result.

Question 5: Consider f(x)=$-10x^TQx+4y^Tx+2$ where x, y$\in R^d$,Q$\in R^{d\times d}$ and Q is symmetric.

(1) What is $\frac{\partial f}{\partial x}$?

$$\frac{\partial f}{\partial x} = \frac{\partial(-10x^TQx+4y^Tx+2)}{\partial x} = -10\frac{\partial(x^TQx)}{\partial x} + 4\frac{\partial(y^Tx)}{\partial x}$$ **Because** $\frac{\partial(x^TQx)}{\partial x} = x^T(Q^T + Q)$,

So, $-10\frac{\partial(x^TQx)}{\partial x} + 4\frac{\partial(y^Tx)}{\partial x} = -10x^T(Q^T + Q) + 4y^T = -20x^TQ + 4y^T$

(2) What is $\frac{\partial^2 f}{\partial x^2}$?

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial(-20x^TQ+4y^T)}{\partial x} = \frac{\partial(-20x^TQ)}{\partial x} + \frac{\partial(4y^T)}{\partial x} = -20Q^T = -20Q$$