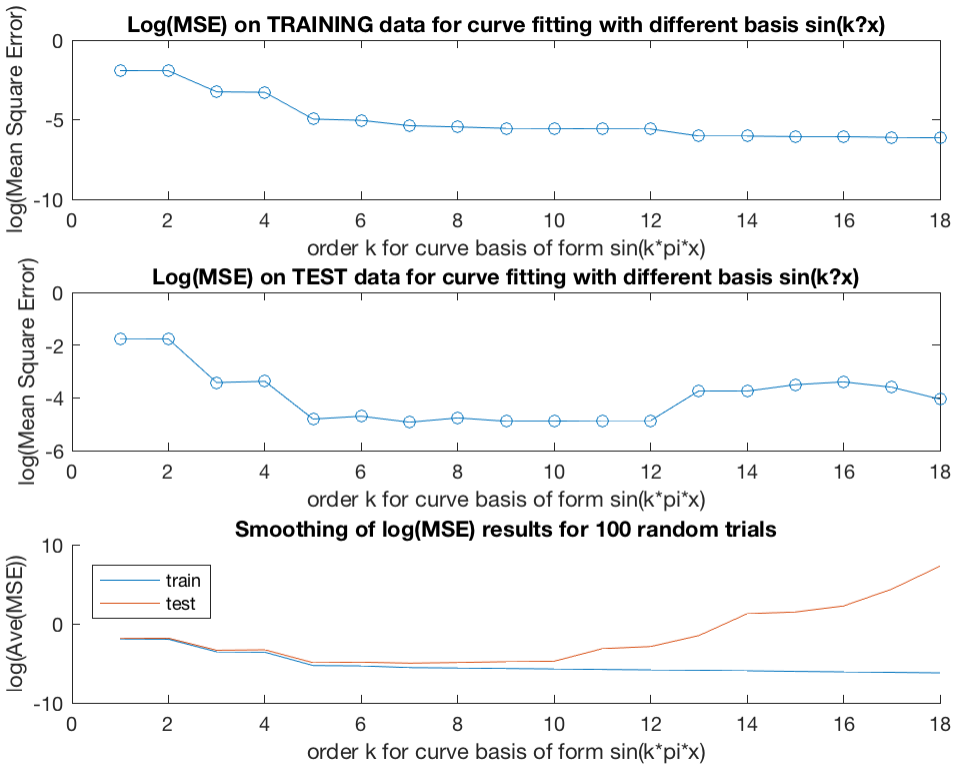
```
%GI07 Homework 1 Question 2 Part 3
%Klaudia Ludwisiak

clear all;
close all;

%%==Repeat==Q2.2=parts=C=and=D=======

sd = 0.07;

knew = [1:18]';
% generate new test and train sets to be used with an extended amount
 of
% polynomial basis (k rangin from 1 to 18)
% x coordinates:
xtrain = rand(30,1);
xtest = rand(1000,1);

% use previously created function to obtain noisy sin function
 approximation:
Gtrain = Gofx(xtrain,0,sd);
Gtest = Gofx(xtest,0,sd);

% Initialize vector for new polynomial basis:
% Initialize vector to store log(MSE) results for test set:
logMSEtrain = zeros(size(knew));
logMSEtest = zeros(size(knew));

% calculate log(MSE(k)) :
for i = 1:size(knew,1)
    % set as temporary variables as they change in size with every
 loop
    Ftrain = MySIN2(xtrain,i); %polynomial coefficients
    Ftest = MySIN2(xtest,i);
    ww = Ftrain\Gtrain; %weight factor
    % calculate logMSE for Test and Tarin data (parts C and D)
    logMSEtrain(i,1) = log(sum((Gtrain - Ftrain*ww).^2)/
size(Gtrain,1));
    logMSEtest(i,1) = log(sum((Gtest - Ftest*ww).^2)/size(Gtest,1));
end

% repeat plot from part C:
figure;
subplot(3,1,1);
plot(knew,logMSEtrain,'-o')
title('Log(MSE) on TRAINING data for curve fitting with different
 basis sin(k?x)');
ylabel('log(Mean Square Error)')
xlabel('order k for curve basis of form sin(k*pi*x)');

% repeat plot from part D:
subplot(3,1,2);
```

```
plot(knew', logMSEtest,'-o');
xlabel('order k for curve basis of form sin(k*pi*x)');
ylabel('log(Mean Square Error)');
title('Log(MSE) on TEST data for curve fitting with different basis
 sin(k?x)');

%%==Repeat==Q2.2==Part=E=====log(MSE)==Smoothing========

% Setup:

% number of runs:
run = [1:100];

% Initialize matrix to store MSE results for train and test set:
MSEtrain = zeros(length(run),size(knew,1));
MSEtest = zeros(length(run),size(knew,1));

for t=1:length(run)
% generate new test and train for every loop
xtrain = rand(30,1);
xtest = rand(1000,1);

% use previously created function to obtain noisy sin function
 approximation:
Gtrain = Gofx(xtrain,0,sd);
Gtest = Gofx(xtest,0,sd);

    % calculate MSE:
    for i = 1:size(knew,1)
        % set as temporary variables as they change in size with every
 loop
        Ftrain = MySIN2(xtrain,i); %polynomial coefficients
        Ftest = MySIN2(xtest,i);
        ww = Ftrain\Gtrain; %weight factor
        % calculate MSE for Test and Tarin data
        MSEtrain(t,i) = sum((Gtrain - Ftrain*ww).^2)/size(Gtrain,1);
        MSEtest(t,i) = sum((Gtest - Ftest*ww).^2)/size(Gtest,1);
    end
end

%now take the log of average MSE results from each run for all k
 values:
logAveMSEtrain = log(mean(MSEtrain));
logAveMSEtest = log(mean(MSEtest));

subplot(3,1,3);
hold on;
plot(knew',logAveMSEtrain);
plot(knew',logAveMSEtest);
xlabel('order k for curve basis of form sin(k*pi*x)');
ylabel('log(Ave(MSE))');
legend('train','test','Location','northwest');
title('Smoothing of log(MSE) results for 100 random trials');
hold off;
```



Published with MATLAB® R2016b