University of Massachusetts Lowell ECE
EECE 1070 Introduction to Engineering for ECE
Curve Fitting and Data Analysis using Matlab

Objectives: To learn how to do linear and polynomial curve fitting. To learn Some basic data analysis techniques in Matlab; To learn to use graphical visualization techniques to understand system behavior.

Part 1 Curvefitting: In the table below is the are the winning time, year, and name for the 100-meter dash.

Year Winner and Country Time (secs)
1928 Betty Robinson, USA 12.2
1932 Stella Walsh, POL 11.9
1936 Helen Stephens, USA 11.5
1948 Fanny Blankers-Koen, NED 11.9
1952 Marjorie Jackson, AUS 11.5
1956 Betty Cuthbert, AUS 11.5
1964 Wyomia Tyus, USA 11.4
1968 Wyomia Tyus, USA 11.08
1972 Renate Stecher, E. Ger 11.07
1976 Annegret Richter, W. Ger 11.08
1980 Lyudmila Kondratyeva, USSR 11.06
1984 Evelyn Ashford, USA 10.97
1988 Florence Griffith Joyner, USA 10.54
1992 Gail Devers, USA 10.82
1996 Gail Devers, USA 10.94
2000 Marion Jones, USA 10.75

(a) Using Matlab, create two arrays one for the year and one for the times of the best finisher. Note that there is a steady decrease, albeit irregular decrease in the finishing time over the years 1928 to 2000. Plot year (x-axis) versus finishing time (y-axis). Include a title “Women’s 100-meter time versus year”, x-axis title (“year”) and y’axis title “finishing time (sec)”
(b) Using the polyfit command, find a best first order least squares fit to the data by a line: Hint: Fit1=polyfit(year,finish,1). In the equation 1 is the order (for example 2,3,4 etc) Note that Fit1 is an array of the coefficients $y=a_1x+a_0$.
(c) Add your best fit line to the plot created in part a.
(d) RMS Error (root mean squared) is calculated by taking the difference between the data and the fit, squaring it and taking the average or mean. Then take the square root (root of the mean squared).
(e) In 1960, the 100-meter dash was won by Wilma Rudolf of the USA in 11.0 seconds. How accurately does your line reflect the truth? Plot this separately on the same plot. Calculate

the error between your prediction and Wilma Rudolf's time. This is known as interpolation.

(f) Extend the best-fit line to include the year 2004. What do you predict as the finish time for that year? In 2004 the 100-meter dash was won by Yulia Nesterenko of Belarus in 10.93 seconds. How does your prediction compare to the actual time? Extending into the future is called extrapolation. Calculate the error and compare to the RMS error you calculated in part D. Bye the way extrapolation too far into the future can sometimes get you into trouble.

Part 2: Arbitrary Curve Fitting.

In this section in am going to give you some data that has additional noise in it. I will tell you that the data represents a generalized polynomial that could have degree (3,2 or 1). You have to figure it out.

a) Download the file "ECE1070.mat" from the website and load it using the command Load('ECE1070.mat'). Use polyfit and find the RMS error for degree 3,2,1. The best fit is the one that has the minimum RMS error. Compare your experimental fit and the data.

Part 3: Thinking Randomly.

In this section we are going to learn about the Histogram feature of Matlab to help you look probabilistically at data

a) Create an array x of uniform random numbers by using the command x=rand(100000,1)
b) Plot a histogram of x using the command hist(x,100). This will create a histogram with 100 bins and with arbitrary scaling. Make a comment about the probability of being in each of the bins. Calculate the mean value of x and the standard deviation.
c) Create an array y using the command y=randn(100000,1).
d) Plot a histogram of y using the command hist(y,100). Make observations on the difference between the distribution of values for x and for y.
e) Calculate the mean and standard deviation of y.
f) Write a program to figure out the probability of a point in the y array is >1.0. There are a number of ways to do it.

What to turn in? For each section turn in answers to the questions, your code for each section, and any plots. Remember neatness and clarity of the presentation counts.