Assignment 3

1. **Interpolation** – Use Matlab functions `polyfit`, `polyval`, and `spline` to compare the polynomial and cubic spline interpolations for different types of functions: smooth, like $\sin x$ or xe^x (on any interval), smooth but with large derivatives, like Runge's function $1/(1+x^2)$ on [-5,5] or $\cos x/(1+\sin^2 x)$, functions with discontinuous derivatives, like $1/(1+e^{|x|})$ or $\sin|x|$. For the polynomial interpolation use both the equidistant and Chebyshev's knots, for splines – equidistant knots. In each case determine experimentally how the maximal interpolation error err_{max} depends on the number of knots n, for, say, $5 \leq n \leq 80$. Present your results as the `loglog` graphs of $err_{max} = f(1/n)$. Use Matlab's function `subplot` to show all three graphs (two for the polynomial and one for spline interpolation of a function) on the same page. In your written report

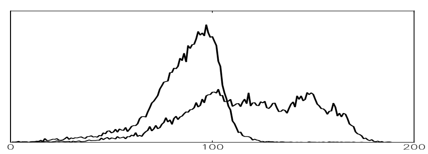
   explain briefly how do the results obtained for the polynomial interpolation of $\sin x$ and xe^x correspond to the theory;
   compare polynomial interpolation with Chebyshev's and equidistant knots;
   compare the spline and polynomial interpolations;
   estimate the convergence rate of spline interpolation using the results obtained.

   Confirm your conclusions using the graphs you have got.

2. **Analysis of FACS histograms** – We ask you to analyze the fluorescence histograms obtained using the Fluorescence-Activated Cell Sorter (FACS) at the Bone Marrow Transplantation Department of Hadassah hospital.

   The binary file `data.mat` contains five vectors of the same length, $f1$, $f2$, $f3$, $f4$, and fm (to get them, download the file and use Matlab's command `>> load data`). Here $f1, ..., f4$ are histograms characterizing the distribution of fluorescence levels in four different populations of cells stained by a fluorochrome (see figure). The vector fm contains a similar histogram for a mixture of these cell populations. Your goal is to estimate the concentrations $c1$, $c2$, $c3$, $c4$ of each of the populations in the mixture by using the least squares method. No program needs to be submitted. In your report describe the model you used (note that $c1+c2+c3+c4=1$) and how you calculated the unknown concentrations. Present the concentration values and a graph showing the histogram of the mixture and its fit by the mixture of histograms.

   Figure: Histograms $f2$ and fm. To build these histograms, the possible range of fluorescence levels was divided into 200 intervals. FACS measured fluorescence levels of about twenty thousands of cells from each population and calculated, say, $f2(i)$ as the number of cells from the 2-nd population in the i-th interval divided by the total number of analyzed cells from that population.



3. **Nonlinear least square models** – The matrix [x,y] in an ascii file `xy2.dat` contains measurements of y for different values of x. Which of the two nonlinear models provides for a better least squares fit to these data?

$$(a) \quad y \approx \tan(a\exp(-t^2)+b), \qquad (b) \quad y \approx a\exp(b/(t+0.5)).$$

   To answer these questions you should approximate the data by each of the models (use Matlab) and compare the values of $\sum(\Delta y_i)^2$. Use data transformations to simplify the fitting. Present graphs showing the data and the best fit curve for each model.