University of Helsinki / Department of Mathematics and Statistics
SCIENTIFIC COMPUTING
Exercise 6, 13.10.2014

N.B. The files mentioned in the exercises (if any) are available on the course homepage

1. The Senior Researcher studies the intelligence quotient (IQ) of the Ubuntu tribe in the Third World. The results of an IQ test are listed here:

IQ -range	# of samples	Normalized samples
61-70	89	
71-80	106	
81-90	84	
91 -100	94	
101-110	35	
111-120	23	
121-130	11	
131-140	1	
141-150	2	

Compute the mean μ and the variance σ^2 of the sample. Fill in the third column, for each row the normalized sample is # of samples divided by the total number of samples.

2. The daily temperature data is observed and the results appear in the table below. Create a file with these thirteen (x,y) pairs of this temperature data.

x	0.0	2.0	4.0	6.0	8.0	10.0	12.0	14.0	16.0	18.0	20.0	22.0	24.0
y	6.3	4.0	6.6	10.9	14.6	19.1	24.3	25.7	22.9	19.5	15.9	10.3	5.4

Let $m = \min\{y\}, M = \max\{y\}$ and set $a = 0.5*(M+m), b = 0.5*(M-m)$. Try to find some reasonable integer value for the parameter c in the interval $[0, 24]$ so that the curve $y = a + b * \sin(2 * \pi * (x - c)/24))$ becomes as close to the data as possible. Carry out the following steps:

(a) Read the data $(x_j, y_j), j = 1, ..., 13$, from the file [or copy these values in a vector] and compute the maximum and minimum temperatures M and m. Then compute a and b.

FILE: ~/MME/demo11/d06/d06.tex — 21. elokuuta 2014 (klo 11.40).

(b) For each $c = 0 : 24$ compute

$$A(c) = \sum_{j=1}^{13} (y_j - (a + b * \sin(2 * \pi * (x_j - c)/24)))^2,$$

and choose the value of c that yields the minimal $A(c)$.

(c) With these values of the parameters a, b, c plot the curve $y = a + b * \sin(2 * \pi * (x - c)/24))$ and the data points in the same picture.

3. To fit a circle (1) $(x - c_1)^2 + (y - c_2)^2 = r^2$ to n sample pairs of coordinates $(x_k, y_k), k = 1, ..., n$ we must determine the center (c_1, c_2) and the radius r. Now (1) $\Leftrightarrow$ (2) $2xc_1 + 2yc_2 + (r^2 - c_1^2 - c_2^2) = x^2 + y^2$. If we set $c_3 = r^2 - c_1^2 - c_2^2$, then the equation takes the form

$$2xc_1 + 2yc_2 + c_3 = x^2 + y^2.$$

Substituting each data point we get

$$\begin{bmatrix} 2x_1 & 2y_1 & 1 \\ & \vdots & \\ 2x_n & 2y_n & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} x_1^2 + y_1^2 \\ \vdots \\ x_n^2 + y_n^2 \end{bmatrix}$$

This system can be solved in the usual way for `c = matrix\rhs`. Then $r = \sqrt{c_3 + c_1^2 + c_2^2}$. Apply this algorithm for the points generated by

```
s=0.5+0.5*rand(10,1);
theta=2*pi*rand(10,1);
clear x
clear y
x=3*s.*cos(theta) ;
y=3*s.*sin(theta);
```

Plot the data and the circle.

4. The number of participants of the weekly problem sessions of a mathematics course during the first six weeks were 21, 24, 17, 21, 14 and 17, respectively. Fit a model of the type

$$y = \lambda_1 \exp(-\lambda_2 x)$$

to this data and predict the number of participants in the 12th problem session.

Hint: It may (or may not) be a good idea to make a linear transform $y' = y/25$, $x' = x/12$ for the fitting, and then use the program parfit.m/Lectures/Section 2 and finally to transform back to the original variables.

5. Familiarize yourself with the program `getpts.m` and use it to plot a closed polygon in the plane. Compute also its area with `polyarea`.

6. Consider the tabulated values `x=0:0.2:3.2; y=d071f(c,d,x)` of the function $d071f(x) = \sum_{j=1}^{m} c_j \sin(d_j * x)$ with `c=[1 2 3 2 1]`, `d=[3 2 1 2 2]` . The data is interpolated to the points `x=0.0:0.05:3.2` by using two different methods; (a) `interp1` , (b) `spline`. Find the maximum error of each method by comparing the interpolation to the values of the function at these points.