Task 7. Approximation χ^2

e-mail: andrzej.kedziorski@fizyka.umk.pl
phone: 56611-3274
office: 485B
http://www.fizyka.umk.pl/~tecumseh/EDU/MNII/

Task 7

Write a program that approximates the function $y(x)$ by means of linear least squares method (generally χ^2) via construction and solving of normal equations. Program takes as an input the 'measurement points' (x_i, y_i, σ_i), where σ_i estimate the errors of y_i, $i = 1, 2, ..., N$. Calculate the variance-covariance matrix.

Linear least squares - general formulation

- Input data (x_i, y_i, σ_i), $i = 1, ..., N$, where σ_i estimates error of y_i; we assume that the error of x_i is negligible
- In general we do not know the actual relation $y(x)$
- Our goal is to fit to the input data the model function (linear in a_k parameters)

$$y(x) \approx \sum_{k=1}^{M} a_k F_k(x),$$

  where $F_k(x)$ are known 'basis' functions, a_k - parameters of our model, $N \geqslant M$
- Values of parameters a_k we derive from the following condition

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - \sum_{k=1}^{M} a_k F_k(x_i)}{\sigma_i} \right)^2 = \min,$$

  which leads to the system of M linear equations (so-called normal equations)

$$\frac{\partial(\chi^2)}{\partial a_\ell} = 0, \quad \ell = 1, ..., M,$$

  where a_k parameters are the solution of normal equations.

Derivation of normal equations

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - \sum_{k=1}^{M} a_k F_k(x_i)}{\sigma_i} \right)^2$$

$$\frac{\partial(\chi^2)}{\partial a_\ell} = 2 \sum_{i=1}^{N} \left(\frac{y_i - \sum_{k=1}^{M} a_k F_k(x_i)}{\sigma_i} \right) \left(\frac{-F_\ell(x_i)}{\sigma_i} \right), \quad \ell = 1, 2, ..., M$$

$$\frac{\partial(\chi^2)}{\partial a_\ell} = 0 \quad \Rightarrow \quad \sum_{k=1}^{M} \left[\sum_{i=1}^{N} \frac{F_\ell(x_i)}{\sigma_i} \frac{F_k(x_i)}{\sigma_i} \right] a_k = \sum_{i=1}^{N} \frac{F_\ell(x_i)}{\sigma_i} \frac{y_i}{\sigma_i}$$

Introducing $A_{ik} = \frac{F_k(x_i)}{\sigma_i}$ and $b_i = \frac{y_i}{\sigma_i}$ we obtain the system of linear equations

$$\sum_{k=1}^{M} \left(\sum_{i=1}^{N} A_{i\ell} A_{ik} \right) a_k = \sum_{i=1}^{N} A_{i\ell} b_i, \quad \ell = 1, 2, ..., M$$

By denoting $\alpha_{\ell k} = \sum_{i=1}^{N} A_{i\ell} A_{ik}$ and $\beta_\ell = \sum_{i=1}^{N} A_{i\ell} b_i$ we obtain the final form of **normal equations**

$$\sum_{k=1}^{M} \alpha_{\ell k} a_k = \beta_\ell, \quad \ell = 1, 2, ..., M$$

Normal equations

System of M linear equations on a_k, $k = 1, ..., M$ parameters

$$\boldsymbol{\alpha}\mathrm{a} = \boldsymbol{\beta},$$

where $\boldsymbol{\alpha} = \mathrm{A}^{\mathrm{Tr}}\mathrm{A}$, $\boldsymbol{\beta} = \mathrm{A}^{\mathrm{Tr}}\mathrm{b}$ and

$$\boldsymbol{A} = \begin{pmatrix} \frac{F_1(x_1)}{\sigma_1} & \frac{F_2(x_1)}{\sigma_1} & \dots & \frac{F_M(x_1)}{\sigma_1} \\ \frac{F_1(x_2)}{\sigma_2} & \frac{F_2(x_2)}{\sigma_2} & \dots & \frac{F_M(x_2)}{\sigma_2} \\ \vdots & \vdots & \dots & \vdots \\ \frac{F_1(x_N)}{\sigma_N} & \frac{F_2(x_N)}{\sigma_N} & \dots & \frac{F_M(x_N)}{\sigma_N} \end{pmatrix} \quad \boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{pmatrix} \quad \boldsymbol{b} = \begin{pmatrix} \frac{y_1}{\sigma_1} \\ \frac{y_2}{\sigma_2} \\ \vdots \\ \frac{y_N}{\sigma_N} \end{pmatrix}$$

Sizes of the matrices and vectors (columns)

$$\boldsymbol{A}_{N\times M}, \ \boldsymbol{a}_{M\times 1}, \ \boldsymbol{b}_N, \ \boldsymbol{\alpha}_{M\times M}, \ \boldsymbol{\beta}_{M\times 1}$$

Overdetermined system of linear equations (we assume $y_i \approx \sum_{k=1}^{M} a_k F_k(x_i)$)

$$\boldsymbol{A}\boldsymbol{a} \approx \boldsymbol{b}$$

leads to another method of determination of the parameters a_k (to be continued...)

Accuracy of the fit, parameters, parameter dependency, condition

- Good estimate when $\chi^2 \sim N - M$, where $N \geqslant M$
- What happens for $N = M$?
- Variance-covariance matrix $\boldsymbol{\alpha}^{-1}$
  - Variance of a_k

$$\sigma^2(a_k) = (\boldsymbol{\alpha}^{-1})_{kk}$$

  - Covariance of a_i and a_k, where $i \neq k$

$$Cov(a_i, a_k) = (\boldsymbol{\alpha}^{-1})_{ik}$$

- Condition number of α

$$cond(\boldsymbol{\alpha}) = cond(\boldsymbol{A}^{Tr}\boldsymbol{A}) = cond(\boldsymbol{A})^2$$

If **A** is ill-conditioned, then $\boldsymbol{\alpha}$ is even worse...

To do

1. Generate *noisy* data for known relation $y(x)$; assume that $F_k(x) = x^{k-1}$, where $k = 1, ..., M$
2. Construct matrices and vectors $\mathbf{A}$, $\mathbf{b}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$
3. Calculate a_k parameters by solving normal equations
4. Assess the numerical accuracy of a_k
   4.1 Calculate the condition number of $\boldsymbol{\alpha}$
   4.2 Perform test calculations of *non-noisy* data
5. Assess the quality of the fit
   5.1 Calculate χ^2
   5.2 Calculate the variance-covariance matrix

Notes on generation of the input data

- Non-noisy data
  1. Generate x_i, where $N \sim 20$
  2. Generate y_i according to $y(x) = 4 + 3x + 2x^2 + x^3$
  3. All $\sigma_i = 1$
- Noisy data
  1. Generate x_i, where $N \sim 20$
  2. Generate $y_i = y(x_i) + \Delta_i$, where $y(x) = 4 + 3x + 2x^2 + x^3$ and Δ_i are relatively small random numbers [1]
  3. All $\sigma_i = |\Delta_i|$

[1] Preferably normally distributed random numbers, see e.g. randn() of Matlab and/or http://fizyka.umk.pl/∼tecumseh/EDU/MNII/randn.txt