Zad. 7. Aproksymacja χ^2

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

Zadanie 7

Napisać program dokonujący aproksymacji zależności $y = f(x)$ liniową metodą najmniejszych kwadratów (w ogólności χ^2) poprzez konstrukcję i rozwiązywanie równań normalnych. Program na wejściu pobiera punkty pomiarowe (x_i, y_i, σ_i), gdzie σ_i jest oszacowaniem błędu y_i, $i = 1, 2, ..., N$. Wyznaczyć macierz kowariancji.

Ogólna liniowa metoda najmniejszych kwadratów

- Postać danych (x_i, y_i, σ_i), $i = 1, ..., N$, gdzie σ_i jest oszacowaniem błędu y_i; zakładamy, że błąd x_i jest zaniedbywalnie mały
- Do danych chcemy dopasować funkcję (model liniowy ze względu na parametry)

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

gdzie $F_k(x)$ są ustalonymi funkcjami, a_k - parametry modelu, $N \geqslant M$

- Wartości parametrów a_k wyznaczamy z warunku

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

co prowadzi do układu M równań normalnych (liniowych) na parametry a_k

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

Wyprowadzenia

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

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

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

Wprowadzając oznaczenia $A_{ik} = \frac{F_k(x_i)}{\sigma_i}$ oraz $b_i = \frac{y_i}{\sigma_i}$ uzyskujemy układ równań w postaci

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

Wprowadzając kolejne oznaczenia $\alpha_{\ell k} = \sum_{i=1}^{N} A_{i\ell} A_{ik}$ oraz $\beta_\ell = \sum_{i=1}^{N} A_{i\ell} b_i$

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

Równania normalne

Układ M równań liniowych na parametry a_k, $k = 1, ..., M$

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

gdzie $\boldsymbol{\alpha} = \mathbf{A}^{\mathrm{Tr}}\mathbf{A}$, $\boldsymbol{\beta} = \mathbf{A}^{\mathrm{Tr}}\mathbf{b}$ oraz

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

Widać, że

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

Nadokreślony układ równań liniowych (zakładamy, że $y_i \approx \sum_{k=1}^{M} a_k F_k(x_i)$)

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

prowadzi do innego sposobu rozwiązania problemu...

Dokładność dopasowania, parametrów, zależność parametrów, uwarunkowanie

- Względnie dobre dopasowanie $\chi^2 \sim N - M$, gdzie $N \geqslant M$
- Co się dzieje dla $N = M$?
- Macierz kowariancji $\boldsymbol{\alpha}^{-1}$
  - Wariancja parametrów a_k

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

  - Kowariancja pary różnych parametrów a_i i a_k, gdzie $i \neq k$

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

- Uwarunkowanie macierzy α

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

Jeżeli **A** jest źle uwarunkowana, to $\boldsymbol{\alpha}$ jest jeszcze gorzej uwarunkowana

Do zrobienia

1. Wygenerować zaszumione dane dla znanego modelu[1]; załóżmy, że $F_k(x) = x^{k-1}$, gdzie $k = 1, ..., M$
2. Skonstruwać macierze/wektory $\mathbf{A}$, $\mathbf{b}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$
3. Wyznaczyć paramtery a_k rozwiązując równania normalne
4. Ocenić dokładność numeryczną rozwiązania równań normalnych
   4.1 Wyznaczyć wskaźnik uwarunkowania macierzy $\boldsymbol{\alpha}$
   4.2 Wykonać testowe dopasowania dla niezaszumionych danych
5. Ocenić jakość aproksymacji
   5.1 Wyznaczyć χ^2
   5.2 Wyznaczyć macierz kowariancji

[1] Niech tym znanym modelem jest $y(x) = 4 + 3x + 2x^2 + x^3$.

DODATEK: Jak wygenerować (nie)zaszumione dane

Założono model

$$y = f(x) = 4 + 3x + 2x^2 + x^3,$$

czyli $M = 4$. Ponadto niech $N = 20$.

1. Wybrać N wartości x_i, np. 1, 2, 3, ..., 20
2. Wyznaczyć niezaszumione wartości $y_i = f(x_i)$ w oparciu o założony model
3. Dla niezaszumionych danych przyjąć $\sigma_i = 1$
4. Zaszumione dane $\tilde{y}_i = f(x_i) + \sigma_i$, gdzie $\sigma_i = 0.1\delta_i$, gdzie δ_i są liczbami kwazi-losowymi o rozkładzie normalnym, zob. np. http://fizyka.umk.pl/~tecumseh/EDU/MNII/randn.txt
5. Do wyznaczenia macierzy **A** oraz kolumny **b** wziąć wartości bezwzględne z powyższego szumu, czyli $|\sigma_i|$