

Interpolace funkčních závislostí

$y = f(x_1, x_2, \dots, x_k)$... teoretická závislost (fyzikální zákon)

- V experimentu měníme hodnotu jedné nebo několika veličin x_i a studujeme závislost veličiny y .
 - např. měníme $x_1 \equiv x$, ostatní x_i bereme jako parametry ($\alpha, \beta, \gamma, \dots$):

$$y = f(x | \alpha, \beta, \gamma, \dots)$$

- Chceme posoudit platnost závislosti y na x_i z výsledků experimentu.
 - tj. chceme získat odhady parametrů $\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \dots$
- např. pro N hodnot x_1, x_2, \dots, x_N jsme naměřili N hodnot y_1, y_2, \dots, y_N

Předpokládáme, že známe funkční závislost f a že přesnost nastavení hodnot veličiny x je řádově větší, než přesnost měření závisle proměnné y (která má obecně pro každý bod jinou dispersi).

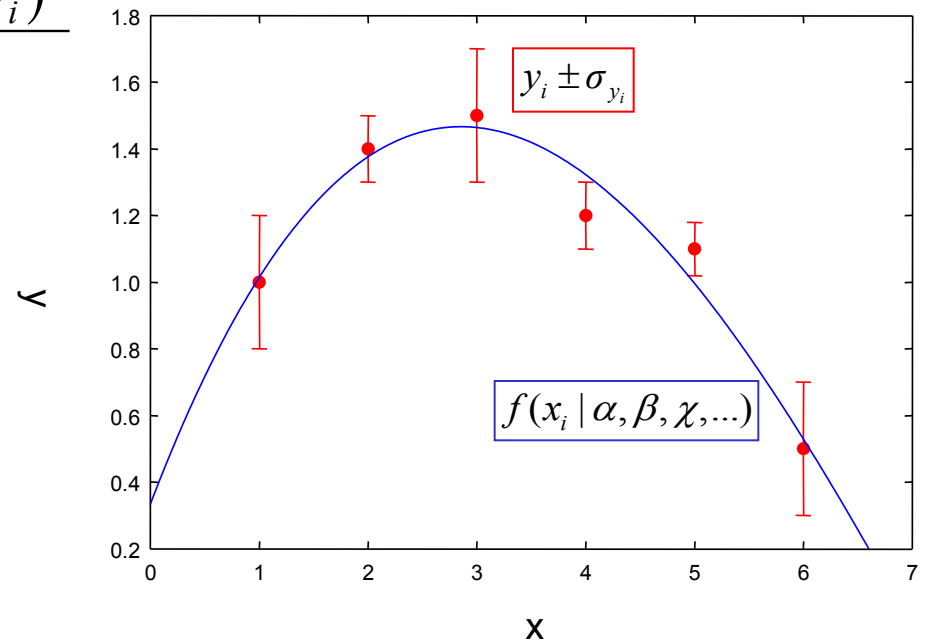
Metoda nejmenších čtverců

- Metoda početní interpolace.
- Používá se pro získání odhadů parametrů $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \dots)$:

1) Zkonstruujeme veličinu

$$\chi^2(\alpha, \beta, \gamma, \dots) = \sum_{i=1}^N \frac{(f(x_i | \alpha, \beta, \gamma, \dots) - y_i)^2}{\sigma_{y_i}^2}$$

2) Hledáme minimum $\chi^2(\alpha, \beta, \gamma, \dots)$.



Metoda nejmenších čtverců – přímka procházející počátkem

- $y = mx$

- $\chi^2(m) = \sum_{i=1}^N \frac{(mx_i - y_i)^2}{\sigma_{y_i}^2}$

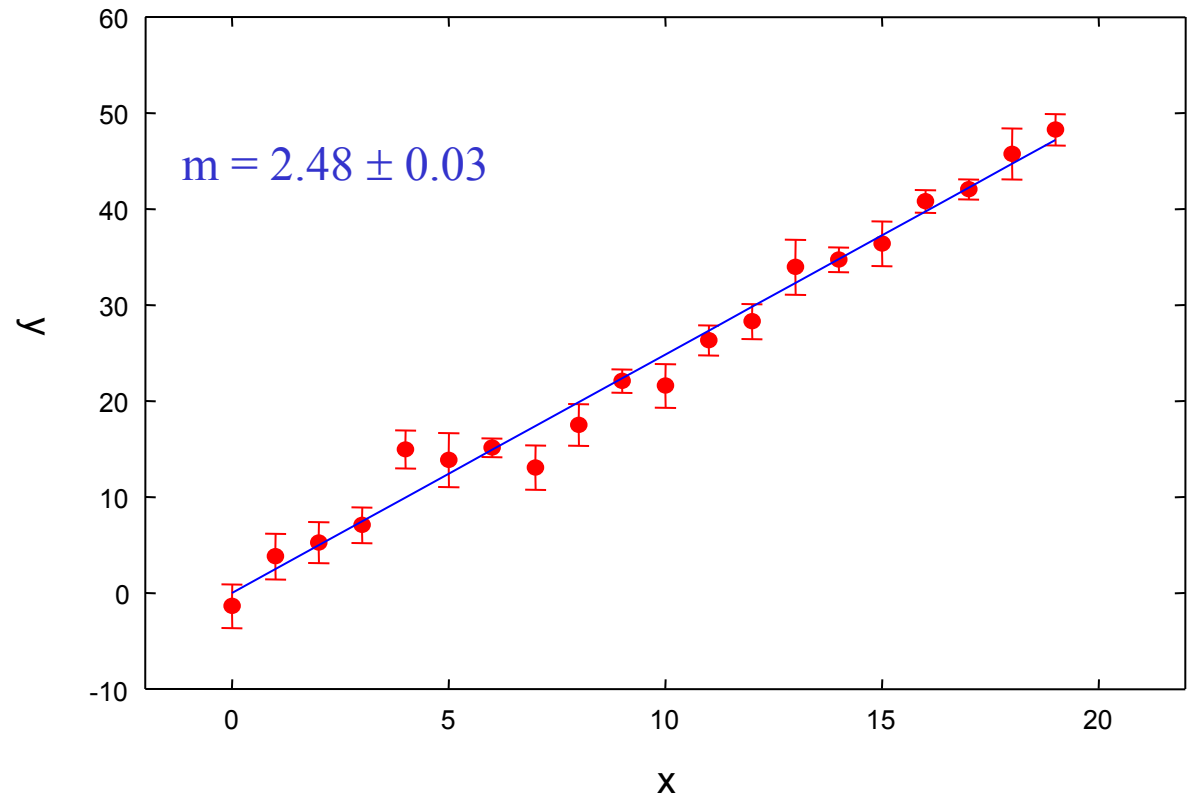
- minimalizace χ^2 :

$$\tilde{m} = \frac{\sum_{i=1}^N \frac{y_i x_i}{\sigma_{y_i}^2}}{\sum_{i=1}^N \frac{x_i^2}{\sigma_{y_i}^2}}$$

- disperze m : $\sigma_{\tilde{m}}^2 = \frac{1}{\sum_{i=1}^N \frac{x_i^2}{\sigma_{y_i}^2}}$

- $m = \tilde{m} \pm \sigma_{\tilde{m}}$

- problém: co když neznáme σ_{y_i}



Metoda nejmenších čtverců – přímka procházející počátkem

- Pokud jsou σ_{y_i} neznámé ale stejné, $\sigma_{y_i} = \sigma_y$

... potom
$$\sigma_{\tilde{m}}^2 = \frac{\sigma_y^2}{\sum_{i=1}^N x_i^2}$$

- Pro neznámou disperzi σ_y pak lze spočítat odhad:
$$\tilde{\sigma}_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{m}x_i)^2$$

ozn.
$$R_1^2 \equiv \sum_{i=1}^n (y_i - \tilde{m}x_i)^2 \quad \dots \text{ minimální suma čtverců odchylek}$$

- nevychýlený odhad:
$$(\tilde{\sigma}_y^*)^2 = \frac{R_1^2}{n-1}$$

- Odhad disperze m je tedy:

$$(\sigma_{\tilde{m}}^*)^2 = \frac{1}{\sum_{i=1}^N x_i^2} \frac{R_1^2}{n-1}$$

Obecná přímka, obecná lineární regrese

• obecná přímka: $y = \beta_0 + \beta_1 x + \varepsilon$

naměřené hodnoty: $[x_i, y_i] \quad i = 1, \dots, n$

• minimalizace χ^2 : $\frac{\partial \chi^2}{\partial \beta_1} = 0 \quad \frac{\partial \chi^2}{\partial \beta_0} = 0$

nejistoty závislé veličiny y_i : $\varepsilon_i \in N(0, \sigma)$

vede na soustavu lineárních rovnic:

$$\beta_0 \sum_{i=1}^n \frac{x_i}{\varepsilon_i^2} + \beta_1 \sum_{i=1}^n \frac{x_i^2}{\varepsilon_i^2} + \sum_{i=1}^n \frac{\varepsilon_i x_i}{\varepsilon_i^2} = \sum_{i=1}^n \frac{x_i y_i}{\varepsilon_i^2}$$

Jak jsou parametry β_0 a β_1 (ne)závislé?
 $\rightarrow \text{Cov}(\beta_0, \beta_1)$

$$\beta_0 \sum_{i=1}^n \frac{1}{\varepsilon_i^2} + \beta_1 \sum_{i=1}^n \frac{x_i}{\varepsilon_i^2} + \sum_{i=1}^n \frac{\varepsilon_i}{\varepsilon_i^2} = \sum_{i=1}^n \frac{y_i}{\varepsilon_i^2}$$

• obecná funkční závislost: $y = y(x, \beta_1, \dots, \beta_m)$

\leftarrow lineární v parametrech β_i tj.

$$\beta_1 \sum_{i=1}^n f_1(x_i) f_1(x_i) + \dots + \beta_m \sum_{i=1}^n f_m(x_i) f_1(x_i) = \sum_{i=1}^n f_1(x_i) y_i$$

$$y = \sum_{k=1}^m \beta_k f_k(x)$$

$$\beta_1 \sum_{i=1}^n f_1(x_i) f_m(x_i) + \dots + \beta_m \sum_{i=1}^n f_m(x_i) f_m(x_i) = \sum_{i=1}^n f_m(x_i) y_i$$

Maticové vyjádření

$$y = \sum_{k=1}^m \beta_k f_k(x)$$

Naměřené hodnoty: $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

$$\mathbf{y} = \mathbf{A}\boldsymbol{\beta}$$

Hledané parametry: $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$

Matice plánu (konstrukční matice, design matrix):

$$\mathbf{A} = \begin{pmatrix} f_1(x_1) & \cdots & f_m(x_1) \\ \vdots & \ddots & \vdots \\ f_1(x_n) & \cdots & f_m(x_n) \end{pmatrix} \quad \leftarrow \text{matice } m \times n, m \leq n$$

$$\frac{\partial}{\partial \boldsymbol{\beta}} \|\mathbf{A}\boldsymbol{\beta} - \mathbf{y}\|^2 = 0 \quad \rightarrow \text{řešení pro parametry: } \boldsymbol{\beta} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y} = \mathbf{H}\mathbf{y}$$

Jak jsou parametry (ne)závislé?
 $\text{Cov}(\beta_0, \beta_1)$

\rightarrow kovarianční matice:

$$U_{ij} = \text{Cov}(\beta_i, \beta_j) \\ V_{ij} = \text{Cov}(y_i, y_j)$$

$$\mathbf{U} = \mathbf{H}\mathbf{V}\mathbf{H}^T$$

Fitování

- Konstrukce křivky (funkce), která co nejlépe odpovídá naměřeným hodnotám.
 - může podléhat dodatečným podmínkám
- Lineární vs. nelineární regrese

metoda největšího spádu

Gaussova-Newtonova metoda

algoritmus Levenberg–Marquardt

simplex

- Interpolace a vyhlazování (spline)
- Regresní analýza a extrapolace
- Softwarové nástroje
 - Excel, Matlab, Origin, ...
 - gnuplot, Python, R, ...

