Metoda nejmenších čtverců – lineární model

- sada naměřených hodnot $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ (nezávislé proměnné)

  $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ (závislé proměnné) $y_i \in N(\mu_i, \sigma_i)$

- modelová funkce $\lambda(x|\boldsymbol{\theta})$ (modelujeme závislost $y(x)$)

$$\boxed{\lambda(x_i|\boldsymbol{\theta}) = \sum_{j=1}^{m} a_j(x_i)\theta_j = \sum_{j=1}^{m} A_{ij}\theta_j}$$

$\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$ (parametry modelové závislosti)

- minimalizujeme tzv. „chí kvadrát“ $\chi^2(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{x}) = \sum_{i=1}^{n} \frac{\left(y_i - \sum_{j=1}^{m} A_{ij}\theta_j\right)^2}{\sigma_i^2}$

pokud $\sigma_i = \sigma$

$$\chi^2 = \frac{1}{\sigma^2}(\boldsymbol{y} - \boldsymbol{A\theta})^T(\boldsymbol{y} - \boldsymbol{A\theta})$$

obecně $V_{ij} = \mathrm{cov}(y_i, y_j)$

$$\chi^2 = (\boldsymbol{y} - \boldsymbol{A\theta})^T \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{A\theta})$$

Metoda nejmenších čtverců – lineární model

- minimalizujeme tzv. „chí kvadrát“ $\chi^2(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{\sigma},\boldsymbol{x}) = \sum_{i=1}^{n} \frac{\left(y_i - \sum_{j=1}^{m} A_{ij}\theta_j\right)^2}{\sigma_i^2}$

$$0 = \frac{\partial\chi^2}{\partial\theta_k} = -\frac{2}{\sigma^2}\sum_{i=1}^{n}\left(y_i - \sum_{j=1}^{m} A_{ij}\theta_j\right)A_{ik} \quad\Rightarrow\quad \sum_{i=1}^{n} A_{ik}y_i = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ik}A_{ij}\theta_j \quad\Rightarrow\quad \boldsymbol{A}^T\boldsymbol{y} = \boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\theta}$$

- odhad parametrů $\hat{\theta}$

pro $\sigma_i = \sigma$	$\widehat{\boldsymbol{\theta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{y} \equiv \boldsymbol{B}\boldsymbol{y}$
obecně	$\widehat{\boldsymbol{\theta}} = (\boldsymbol{A}^T\boldsymbol{V}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{V}^{-1}\boldsymbol{y} \equiv \boldsymbol{B}\boldsymbol{y}$

- odhad kovariance parametrů $\hat{\theta}$

$$U_{ij} = \mathrm{cov}(\theta_i,\theta_j) \qquad\qquad \boldsymbol{U} = \boldsymbol{B}\boldsymbol{V}\boldsymbol{B}^T$$

$$U_{ij}^{-1} = \frac{1}{2}\left[\frac{\partial\chi^2}{\partial\theta_i\partial\theta_j}\right]_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$$

Metoda nejmenších čtverců – fit polynomu stupně m

- n naměřených hodnot $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ $\boxed{\sigma_i = \sigma}$
- $m + 1$ parametrů modelové funkce $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$
- obecný polynom stupně m $\lambda(x_i|\boldsymbol{\theta}) = \sum_{j=0}^{m} A_{ij}\theta_j = \sum_{j=0}^{m} \theta_j x_i^j$
- matice $A_{ij} = x_i^j$

$$\boldsymbol{A} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{pmatrix} \quad n \times (m+1)$$

$$\boldsymbol{A}^T\boldsymbol{A} = \begin{pmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \dots & \sum_{i=1}^{n} x_i^m \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \dots & \sum_{i=1}^{n} x_i^{m+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} x_i^m & \sum_{i=1}^{n} x_i^{m+1} & \sum_{i=1}^{n} x_i^{m+2} & \dots & \sum_{i=1}^{n} x_i^{m+m} \end{pmatrix} \quad (m+1) \times (m+1)$$

- odhady parametrů

$$\boxed{\widehat{\boldsymbol{\theta}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{y} \qquad \boldsymbol{U} = \frac{1}{\sigma^2}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T[(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T]^T}$$

Metoda nejmenších čtverců – fit paraboly

- polynom 2. stupně ($m = 3$)

$$\lambda(x_i|\boldsymbol{\theta}) = \sum_{j=0}^{m} A_{ij}\theta_j = \theta_0 + \theta_1 x_i + \theta_2 x_i^2$$

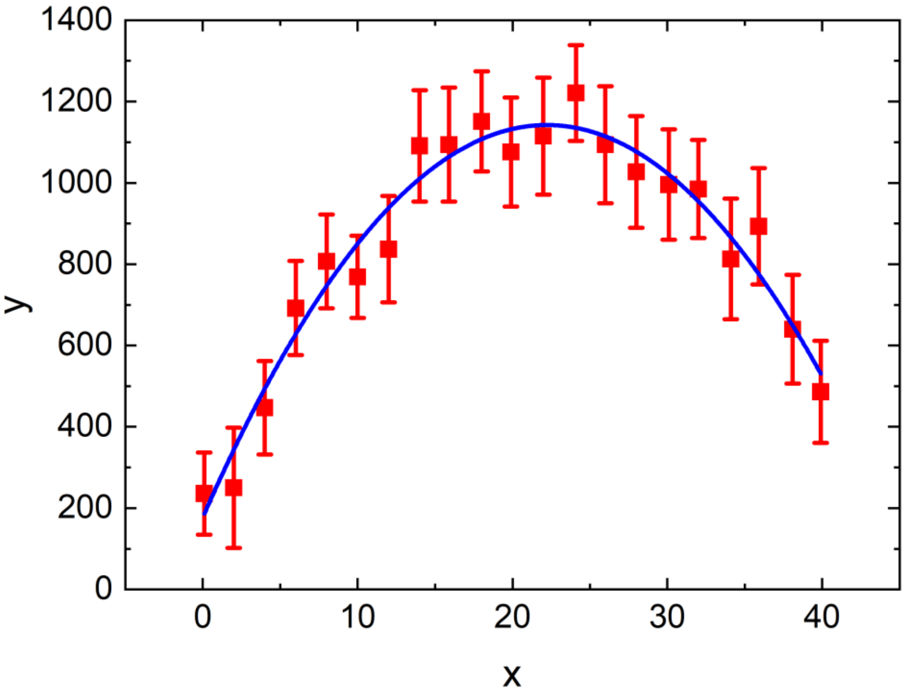
$$\boldsymbol{A} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix}$$

$$\boldsymbol{A}^T\boldsymbol{V}^{-1}\boldsymbol{A} = \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} \\ \sum_{i=1}^n \frac{x_i}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^3}{\sigma_i^2} \\ \sum_{i=1}^n \frac{x_i^2}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^3}{\sigma_i^2} & \sum_{i=1}^n \frac{x_i^4}{\sigma_i^2} \end{pmatrix}$$

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{A}^T\boldsymbol{V}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{V}^{-1}\boldsymbol{y} \equiv \boldsymbol{B}\boldsymbol{y}$$

$$\boldsymbol{U} = \boldsymbol{B}\boldsymbol{V}\boldsymbol{B}^T$$

$$U_{ij} = \mathrm{cov}(\theta_i, \theta_j)$$



$\theta_0 = 180 \pm 70$ $\quad$ $\mathrm{cov}(\theta_1, \theta_2) = -490$

$\theta_1 = 87 \pm 9$ $\quad$ $\mathrm{cov}(\theta_1, \theta_3) = 10.1$

$\theta_2 = -2.0 \pm 0.2$ $\quad$ $\mathrm{cov}(\theta_2, \theta_3) = -1.72$

Metoda nejmenších čtverců – linearizace

`Arrhenius.py`

- modelová funkce $\lambda(x|\boldsymbol{\theta})$ nelineární vzhledem k parametrům $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$.

- nelineární model

$$\lambda(T|\nu_0, Q) = \nu_0 \exp\left(-\frac{Q}{kT}\right)$$

$$\chi^2(\nu_0, Q|\boldsymbol{y}) = \sum_{i=1}^{n} \frac{\left(y_i - \lambda(T_i|\nu_0, Q)\right)^2}{\sigma_i^2}$$

$$\frac{\partial \chi^2}{\partial Q} = 0 \qquad \frac{\partial \chi^2}{\partial \nu_0} = 0$$

- soustava nelineárních rovnic

→ numerické (přibližné řešení)

- lineární model

$$\ln \lambda(T) = -\frac{Q}{kT} + \ln \nu_0 = A\frac{1}{T} + B$$

transformace $\qquad z_i = \ln y_i \qquad \sigma_{z_i} = \frac{\sigma_i}{y_i}$

$$\chi^2(A, B|\boldsymbol{y}) = \sum_{i=1}^{n} \frac{\left(\ln y_i - \ln \lambda\,(T_i|A, B)\right)^2}{\sigma_i^2 / y_i^2}$$

$$\frac{\partial \chi^2}{\partial A} = 0 \qquad \frac{\partial \chi^2}{\partial B} = 0$$

- soustava lineárních rovnic

→ analytické (přesné řešení)

Metoda nejmenších čtverců – linearizace

`Arrhenius.py`

- modelová funkce $\lambda(x|\boldsymbol{\theta})$ nelineární vzhledem k parametrům $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$.

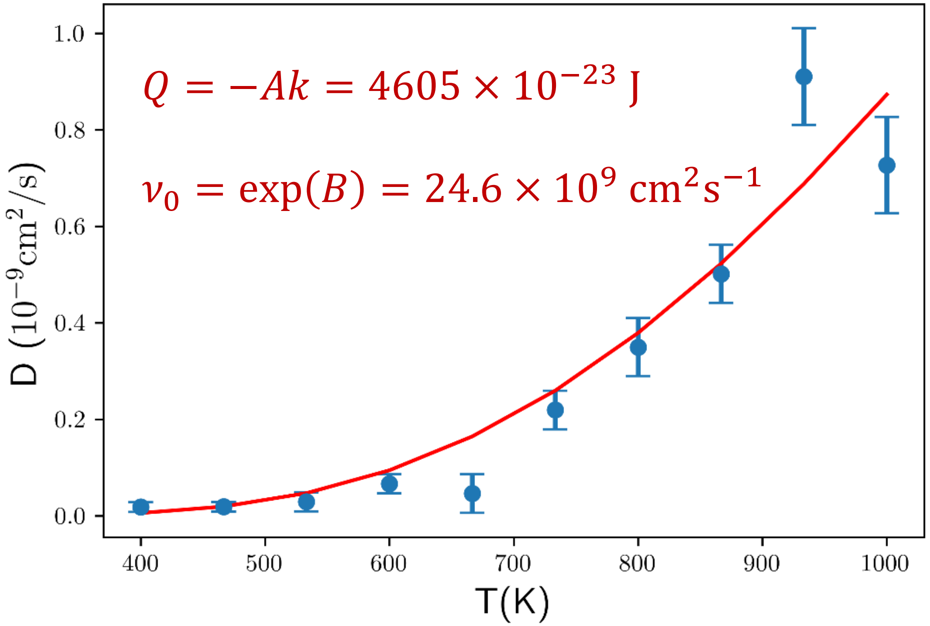
- nelineární model

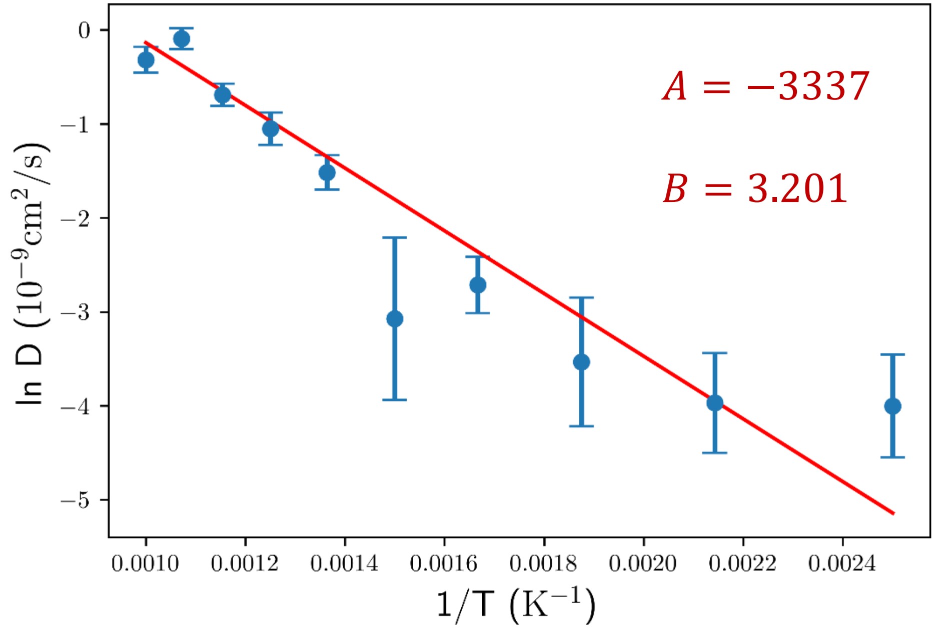
$$\lambda(T|\nu_0, Q) = \nu_0 \exp\left(-\frac{Q}{kT}\right)$$

- lineární model

$$\ln \lambda(T) = -\frac{Q}{kT} + \ln \nu_0 = A\frac{1}{T} + B$$

transformace $z_i = \ln y_i$ $\sigma_{z_i} = \dfrac{\sigma_i}{y_i}$





Python funkce `np.polyfit`

Metoda nejmenších čtverců – linearizace

Arrhenius.py

- modelová funkce $\lambda(x|\boldsymbol{\theta})$ nelineární vzhledem k parametrům $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$.

- nelineární model

$$\lambda(T|\nu_0, Q) = \nu_0 \exp\left(-\frac{Q}{kT}\right) \longrightarrow$$

- lineární model

$$\ln \lambda(T) = -\frac{Q}{kT} + \ln \nu_0 = A\frac{1}{T} + B$$

výsledky fitu:

$$A = (-3300 \pm 200)\ \mathrm{K}$$

$$B = (3.2 \pm 0.3)\ \ln(10^9\ \mathrm{cm^2 s^{-1}})$$

$\longleftarrow$ výsledné parametry:

$$Q = -Ak$$
$$Q = 3300\ \mathrm{K} \cdot 1.38 \times 10^{-23}\ \mathrm{JK^{-1}} = 0.284\ \mathrm{eV}$$

$$\sigma_Q^2 = (-k\sigma_A)^2 \Rightarrow \sigma_Q = k\sigma_A$$
$$\sigma_Q = 276 \times 10^{-23}\ \mathrm{J} = 0.017\ \mathrm{eV} \doteq 0.02\ \mathrm{eV}$$

$$\nu_0 = \exp(B)$$
$$\nu_0 = 24.5 \times 10^9\ \mathrm{cm^2 s^{-1}}$$

$$\sigma_{\nu_0}^2 = [\exp(B)\sigma_B]^2 = (\nu_0\sigma_B)^2 \Rightarrow \sigma_{\nu_0} = \nu_0\sigma_B$$
$$\sigma_{\nu_0} = 7.4 \times 10^9\ \mathrm{cm^2 s^{-1}} \doteq 7 \times 10^9\ \mathrm{cm^2 s^{-1}}$$

zápis výsledku

$$Q = (0.28 \pm 0.02)\ \mathrm{eV}$$
(aktivační energie)

$$\nu_0 = (25 \pm 7) \times 10^9\ \mathrm{cm^2 s^{-1}}$$
(difúzní koeficient pro $T \to \infty$)

Metoda nejmenších čtverců – nelineární fit `sterbina.py`

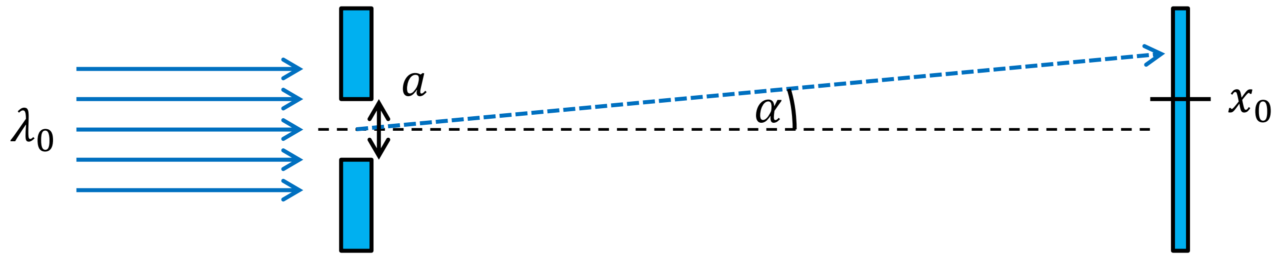
- difrakce na štěrbině

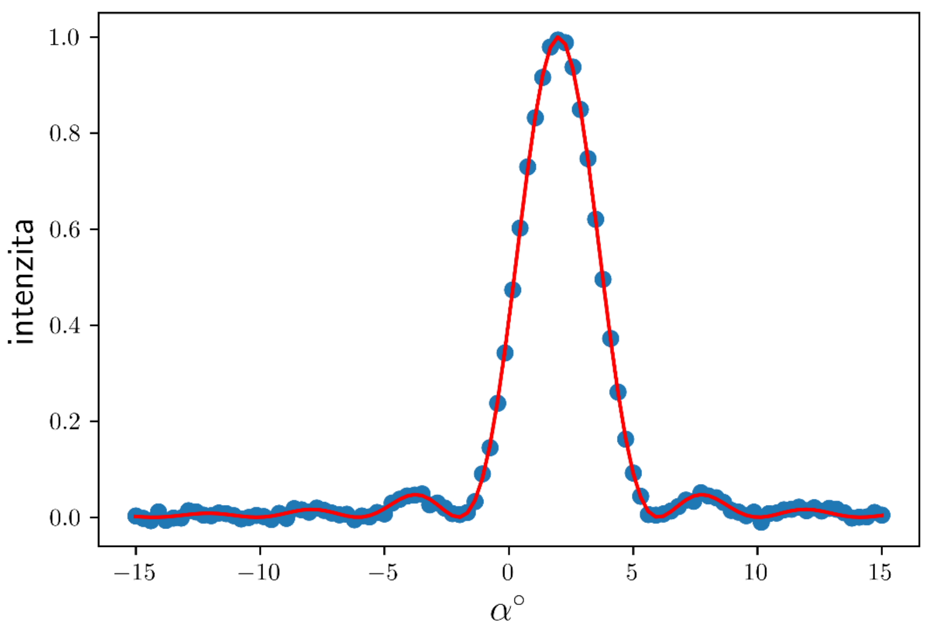
šířka štěrbiny

$$\lambda(\alpha|a,\alpha_0) = \left[\operatorname{sinc}\left(\frac{\pi a}{\lambda_0}\sin(\alpha-\alpha_0)\right)\right]^2$$

$$\operatorname{sinc}(x) = \frac{\sin x}{x}$$

poloha hlavního maxima





$$\chi^2(a,\alpha_0|\boldsymbol{y}) = \sum_{i=1}^{n}\frac{\left(y_i-\lambda(\alpha_i|a,\alpha_0)\right)^2}{\sigma_i^2}$$

$$a = (2.499 \pm 0.004)\ \mu\mathrm{m}$$

$$\alpha_0 = (1.999 \pm 0.003)\ °$$

Python funkce `np.curve_fit`

Metoda nejmenších čtverců – nelineární fit

`dvojsterbina.py`

- difrakce na dvojštěrbině

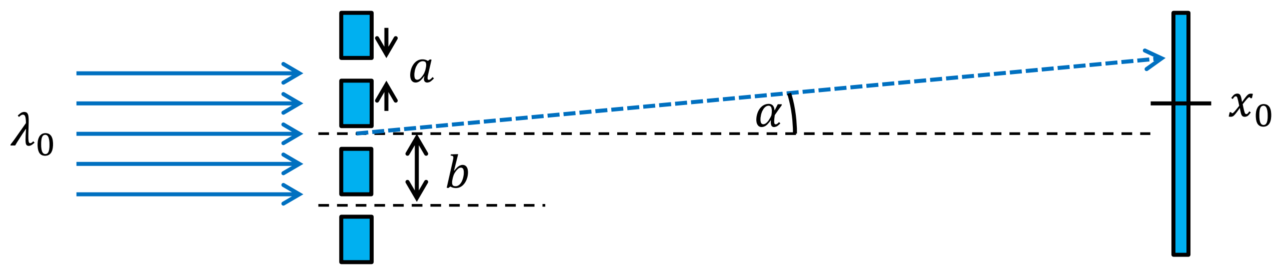
šířka štěrbin

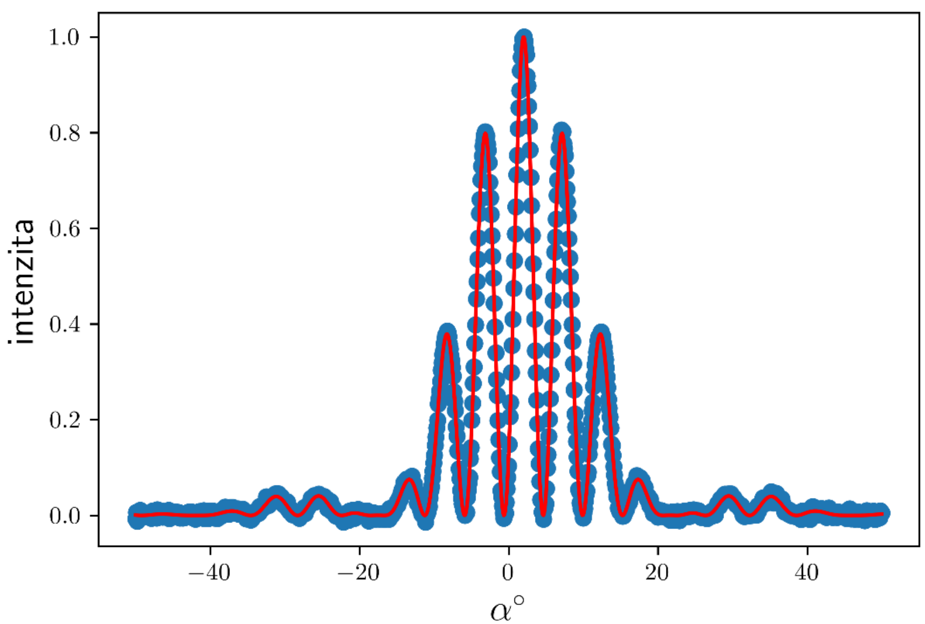
vzdálenost štěrbin

$$\lambda(\alpha|a,b,\alpha_0) = \left[\operatorname{sinc}\left(\frac{\pi a}{\lambda_0}\sin(\alpha-\alpha_0)\right)\right]^2 \left[\cos\left(\frac{\pi b}{\lambda_0}\sin(\alpha-\alpha_0)\right)\right]^2$$

$$\operatorname{sinc}(x) = \frac{\sin x}{x}$$

poloha hlavního maxima





$$\chi^2(a,b,\alpha_0|\boldsymbol{y}) = \sum_{i=1}^{n} \frac{\left(y_i - \lambda(\alpha_i|a,b,\alpha_0)\right)^2}{\sigma_i^2}$$

$$a = (0.4998 \pm 0.0004)\ \mu\text{m}$$

$$b = (5.999 \pm 0.001)\ \mu\text{m}$$

$$\alpha_0 = (1.999 \pm 0.001)\ ^\circ$$

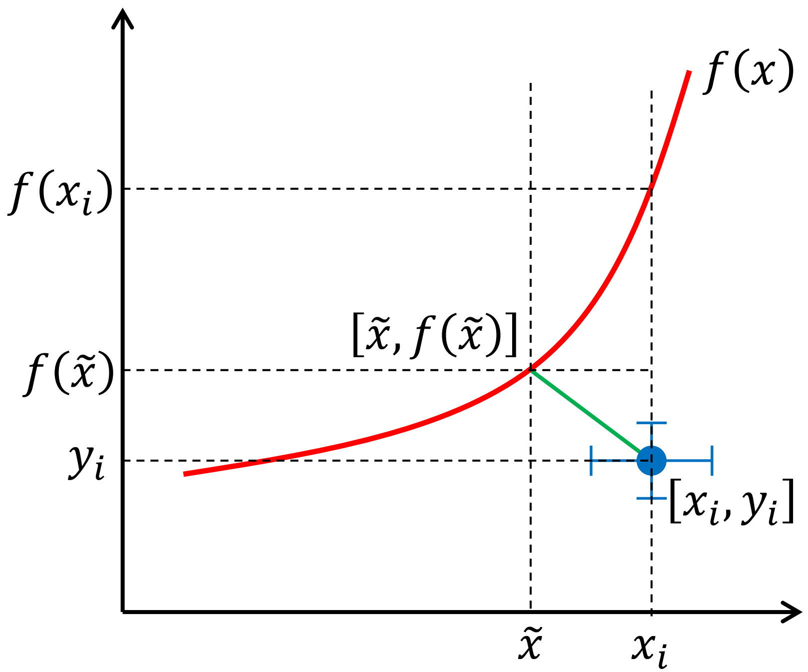
Python funkce `np.curve_fit`

Metoda nejmenších čtverců – chyby obou proměnných

- sada naměřených hodnot

$\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ (náhodné proměnné s chybami σ_{x_i})

$\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ (náhodné proměnné s chybami σ_{y_i})



- vzdálenost bodu $[x_i, y_i]$ od modelové funkce $f(x)$

$$d_i^2 = (x_i - \tilde{x})^2 + \big(y_i - f(\tilde{x})\big)^2$$

- „vážená“ vzdálenost bodu $[x_i, y_i]$ od modelové funkce $f(x)$

$$d_i^2 = \frac{1}{\sigma_{x_i}^2}(x_i - \tilde{x})^2 + \frac{1}{\sigma_{y_i}^2}\big(y_i - f(\tilde{x})\big)^2$$

- Taylorův rozvoj modelové funkce $f(\tilde{x})$ v okolí bodu x_i

$$f(\tilde{x}) = f(x_i) + f'(x_i)(\tilde{x} - x_i)$$

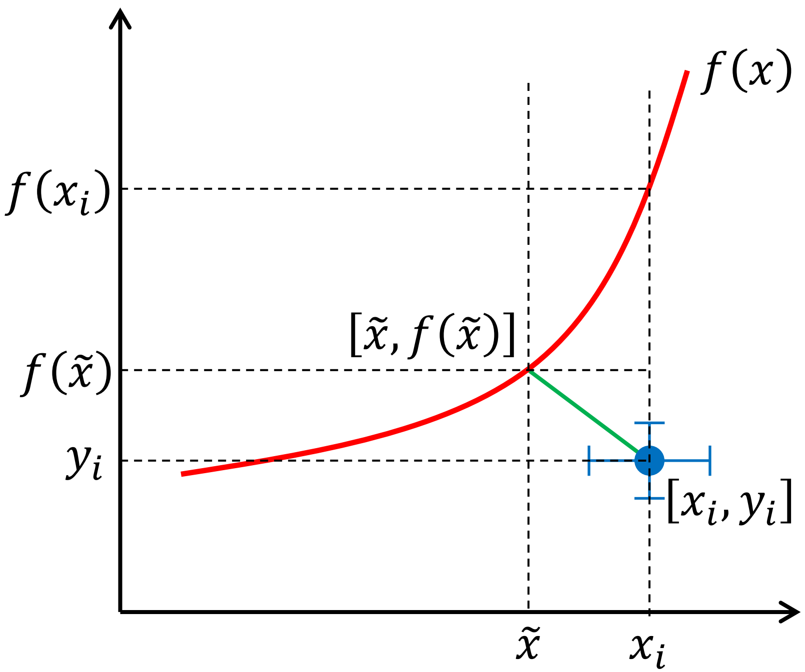
$$d_i^2 = \frac{1}{\sigma_{x_i}^2}(x_i - \tilde{x})^2 + \frac{1}{\sigma_{y_i}^2}[y_i - f(x_i) + f'(x_i)(x_i - \tilde{x})]^2$$

Metoda nejmenších čtverců – chyby obou proměnných

- sada naměřených hodnot

$\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ (náhodné proměnné s chybami σ_{x_i})

$\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ (náhodné proměnné s chybami σ_{y_i})



- minimalizujeme „váženou“ vzdálenost

$$0 = \frac{\partial d_i^2}{\partial \tilde{x}} = -2\frac{1}{\sigma_{x_i}^2}(x_i - \tilde{x}) - 2\frac{f'(x_i)}{\sigma_{y_i}^2}[y_i - f(x_i) + f'(x_i)(x_i - \tilde{x})]$$

$$\Rightarrow (x_i - \tilde{x}) = -\frac{\frac{1}{\sigma_{y_i}^2} f'(x_i)[y_i - f(x_i)]}{\frac{1}{\sigma_{x_i}^2} + \frac{1}{\sigma_{y_i}^2} f'(x_i)^2}$$

- minimální „vážená vzdálenost“

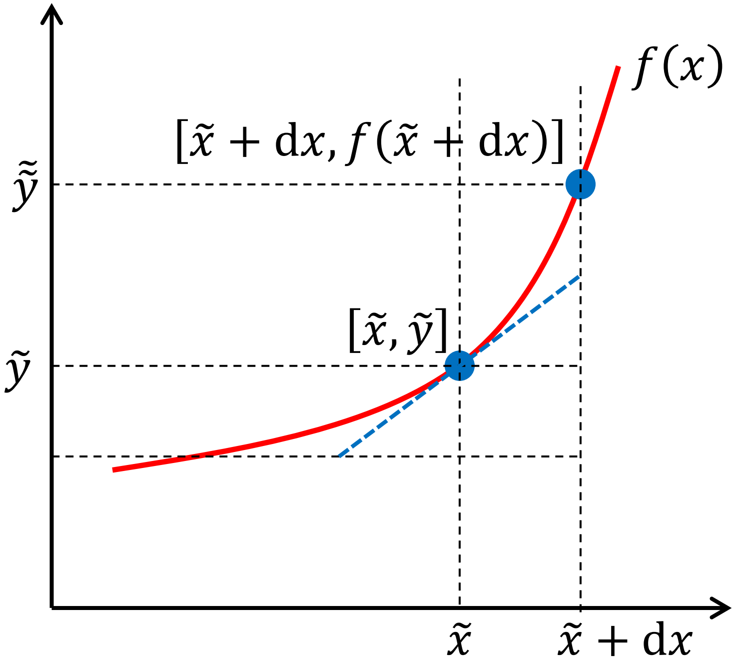
$$d_i^2 = \frac{[y_i - f(x_i)]^2}{\sigma_{y_i}^2 + \sigma_{x_i}^2 f'(x_i)^2}$$

Metoda nejmenších čtverců – chyby obou proměnných

- sada naměřených hodnot

$\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ (náhodné proměnné s chybami σ_{x_i})

$\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ (náhodné proměnné s chybami σ_{y_i})



- minimalizace „chí kvadrátu“

$$\boxed{\chi^2 = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \frac{[y_i - f(x_i)]^2}{\sigma_{y_i}^2 + \sigma_{x_i}^2 f'(x_i)^2}}$$

- Poznámka: Taylorův rozvoj

$$y(x + \mathrm{d}x) = f(x + \mathrm{d}x) \approx f(x) + f'(x)\mathrm{d}x$$

$$\tilde{x} \to \tilde{x} + \mathrm{d}x$$

$$y(\tilde{x}) \to y(\tilde{x} + \mathrm{d}x) + \mathrm{d}y = f(\tilde{x}) + f'(\tilde{x})\mathrm{d}x + \mathrm{d}y$$

celková chyba $\sigma^2 = \sigma_y^2 + f'^2(x)\sigma_x^2$

Metoda nejmenších čtverců – chyby obou proměnných

- sada naměřených hodnot $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ (náhodné proměnné s chybami σ_{x_i})

  $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ (náhodné proměnné s chybami σ_{y_i})

- modelová funkce $f(x|a,b) = ax + b$

- minimalizace „chí kvadrátu“

$$\chi^2 = \sum_{i=1}^{n} \frac{[y_i - f(x_i)]^2}{\sigma_{y_i}^2 + \sigma_{x_i}^2 f'(x_i)^2}$$

$$f(x_i|a,b) = ax_i + b \Rightarrow f'(x_i) = a$$

$$\chi^2 = \sum_{i=1}^{n} \frac{[y_i - f(x_i)]^2}{\sigma_{y_i}^2 + a^2\sigma_{x_i}^2}$$