Thuong Duc HOANG

Lec 3

χ^2 fitting to a straight line.

- We consider a set of N data points $(x_i; y_i)$.
- We assume that the uncertainty (deviation) σ_i associated with measurement y_i.
- We want to fit a model: $y = b + ax$
- The chi-square (χ^2) function measures how well model agrees with data:

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{y_i - a x_i - b}{\sigma_i} \right)^2 \qquad (1)$$

- If the errors are normal distribution, the χ^2 function will give the maximum likelihood parameter estimation of a, b.
- by minimizing (1) with respect to a, b to determine a, b.

$$\frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{y_i - a x_i - b}{\sigma_i^2}$$

$$\boxed{(u^\alpha)' = \alpha u^{\alpha - 1}}$$

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} x_i \frac{y_i - a x_i - b}{\sigma_i^2}$$

- Minimum χ^2

$$\Rightarrow \begin{cases} \dfrac{\partial \chi^2}{\partial b} = 0 \\ \partial \chi^2 / \partial a = 0 \end{cases} \Rightarrow \begin{aligned} &\sum_{i=1}^{N} \frac{y_i - a x_i - b}{\sigma_i^2} = 0 \\ &\sum_{i=1}^{N} x_i \frac{y_i - a x_i - b}{\sigma_i^2} = 0 \end{aligned} \qquad (2)$$

* We define some summation:

$$S = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \quad ; \quad S_x = \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \quad ; \quad S_y = \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$S_{xx} = \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \quad ; \quad S_{xy} = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}$$

$$\frac{\partial \chi^2}{\partial a} = \sum_{i=1}^{N} x_i \frac{y_i - a x_i - b}{\sigma_i^2} \quad ; \quad \frac{\partial \chi^2}{\partial b} = \sum_{i=1}^{N} \frac{y_i - a x_i - b}{\sigma_i^2}$$

Minimum $\chi^2 \Rightarrow \begin{cases} \dfrac{\partial \chi^2}{\partial a} = 0 \\ \dfrac{\partial \chi^2}{\partial b} = 0 \end{cases}$

$$\Rightarrow \begin{cases} a S_{xx} + b S_x = S_{xy} \\ a S_x + b S = S_y \end{cases} \quad ; \quad D = \begin{vmatrix} S_{xx} & S_x \\ S_x & S \end{vmatrix} = S S_{xx} - (S_x)^2 = \Delta$$

$$D_a = \begin{vmatrix} S_{xy} & S_x \\ S_y & S \end{vmatrix} = S S_{xy} - S_x S_y \quad ; \quad D_b = \begin{vmatrix} S_{xx} & S_{xy} \\ S_x & S_y \end{vmatrix} = S_y S_{xx} - S_x S_{xy}$$

$$\left.\begin{cases} a = \dfrac{D_a}{\Delta} = \dfrac{S S_{xy} - S_x S_y}{S S_{xx} - (S_x)^2} \\ b = \dfrac{D_b}{\Delta} = \dfrac{S_y S_{xx} - S_x S_{xy}}{S S_{xx} - (S_x)^2} \end{cases}\right\}$$

The solution for the best model fitting parameter for a, b

* We must estimate the uncertainties:

$$\sigma_a^2 = \frac{S}{\Delta} \quad ; \quad \sigma_b^2 = \frac{S_{xx}}{\Delta}.$$

mple.) fit $y = ax + b$ for the data; estimate parameter and χ^2

$x = 1, 2, 3, 4, 5$

$y = 2.2, 2.9, 4.3, 5.2, 6.3$

$\delta = 0.2, 0.4, 0.3, 0.1, 0.35$

$y_i = ax_i + b$; $\partial\chi$

$$\chi^2(a,b) = \sum_i^N \left(\frac{y_i - ax_i - b}{\delta_i} \right)^2$$

$$\frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^N x_i \frac{(y_i - ax_i - b)}{\delta_i^2}$$

$$\frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^N \frac{y_i - ax_i - b}{\delta_i^2}$$

$S = \sum_1^N \frac{1}{\delta_i^2}$; $S_x = \sum_i^N \frac{x_i}{\delta_i^2}$

$S_y = \sum_{i=1}^N \frac{y_i}{\delta_i^2}$; $S_{xy} = \sum_{i=1}^N \frac{x_i y_i}{\delta_i^2}$

$S_{xx} = \sum_{i=1}^N \frac{x_i^2}{\delta_i^2}$

$$S = \sum_{i=1}^N \frac{1}{\delta_i^2} = \frac{1}{0.2^2} + \frac{1}{0.4^2} + \frac{1}{0.3^2} + \frac{1}{0.1^2} + \frac{1}{0.35^2} = 150.5$$

$$S_x = \frac{1}{0.2^2} + \frac{2}{0.4^2} + \frac{3}{0.3^2} + \frac{4}{0.1^2} + \frac{5}{0.35^2} = 511.65$$

$$S_y = \frac{2.2}{0.2^2} + \frac{2.9}{0.4^2} + \frac{4.3}{0.3^2} + \frac{5.2}{0.1^2} + \frac{6.3}{0.35^2} = 692.33$$

$$S_{xx} = \frac{1^2}{0.2^2} + \frac{2^2}{0.4^2} + \frac{3^2}{0.3^2} + \frac{4^2}{0.1^2} + \frac{5^2}{0.35^2} = 1954.08$$

$$S_{xy} = \frac{2.2 \times 1}{0.2^2} + \frac{2 \times 2.9}{0.4^2} + \frac{3 \times 4.3}{0.3^2} + \frac{4.5 \times 2}{0.1^2} + \frac{5 \times 6.3}{0.35^2} = 2571.726$$

$$a = \frac{S\,S_{xy} - S_x S_y}{SS_{xx} - (S_x)^2} = \frac{150 * 2571.726 - 511.65 * 692.33}{150.5 * 1954.08 - (511.65)^2}$$

$$= 1.016$$

$$b = \frac{S_y S_{xx} - S_x S_{xy}}{SS_{xx} - (S_x)^2} = 1.145.$$

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_i - a x_i - b}{\sigma_i} \right)^2 = 0.698.$$

* We must estimate the uncertainties in the estimate of a, b

$$\sigma_b^2 = \frac{S_{xx}}{\Delta} \quad ; \quad \sigma_a^2 = \frac{S}{\Delta}$$

---

* in detail: If the data are independent, there are uncertainty to the parameters. The propagation of errors

$$\sigma_f^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial f}{\partial y_i}\right)^2$$

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx}(S_y)' - S_x (S_{xy})'}{\Delta} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta}$$

$$\frac{\partial b}{\partial y_i} = \frac{S S_{xy} - S_x S_y}{\Delta} = \frac{S x_i - S_x}{\sigma_i^2 \Delta}$$

---

* The coefficient of correlation btw uncertainty in a and uncertainty in b :

$$r_{ab} = \frac{-S_x}{\sqrt{S S_{xx}}}$$