呂浩宇

Introduction to Data Analysis

2023. Mar. 03

B. Fitting a straight line using χ^2 fitting (實作見課程網頁 python code)

(power-law 的 fitting 亦可以同樣的方法處理，只是要注意在 linear scale 下的 Gaussian noise 在 log scale 並非 Gaussian)

model: $y = ax + b$

$$\chi^2 = \chi^2(a,b) = \sum_{i=0}^{N-1}\left(\frac{y_i - ax_i - b}{\sigma_i}\right)^2$$

視 χ^2 為 a 與 b 的函數

※ This method is terribly sensitive to outliers that are having small σ_i but are largely offset from the model

can be alleviated with Robust Estimation

為專有名詞

(i) deriving best-fit parameters.

finding the minimum of χ^2 at:

$$\begin{cases} \dfrac{\partial\chi^2}{\partial a} = -2\displaystyle\sum_{i=0}^{N-1}\frac{x_i(y_i - ax_i - b)}{\sigma_i^2} = 0 \\ \dfrac{\partial\chi^2}{\partial b} = -2\displaystyle\sum_{i=0}^{N-1}\frac{y_i - ax_i - b}{\sigma_i^2} = 0 \end{cases}$$

Defining

$$\begin{cases} S \equiv \displaystyle\sum_{i=0}^{N-1}\frac{1}{\sigma_i^2}, \quad S_x \equiv \sum_{i=0}^{N-1}\frac{x_i}{\sigma_i^2}, \quad S_y \equiv \sum_{i=0}^{N-1}\frac{y_i}{\sigma_i^2} \\ S_{xx} \equiv \displaystyle\sum_{i=0}^{N-1}\frac{x_i^2}{\sigma_i^2}, \quad S_{xy} \equiv \sum_{i=0}^{N-1}\frac{x_i y_i}{\sigma_i^2} \end{cases}$$

$$\Rightarrow \begin{cases} bS + aS_x = S_y \\ bS_x + aS_{xx} = S_{xy} \end{cases}$$

$$\Rightarrow \boxed{\begin{aligned} b &= \frac{S_{xx}S_y - S_x S_{xy}}{\Delta} \\ a &= \frac{S S_{xy} - S_x S_y}{\Delta} \end{aligned}} \qquad (\because \Delta \equiv S S_{xx} - (S_x)^2)$$

(ii) uncertainty of parameters

Concept: viewing a and b as functions of y_i's

Therefore, perturbing individual y_i measurements will also perturbe a and b.

The uncertainties in a and b are thus related to the uncertainties in y_i's.

for any function $f(y_0, y_1, y_2, \cdots, y_{N-1})$, considering that we make infinitely many random realization of y_i's around the actually measured y values based on the noise statistics, the variance of f is given by

$$\sigma_f^2 \equiv Var(f(y)) = E\left[\left(f(y) - f(y_{measured})\right)^2\right]$$

一階泰勒展開

$$\sim E\left[\left(\sum_{i=0}^{N-1}\left.\frac{\partial f}{\partial y_i}\right|_{y_{measured}}\delta y_i\right)^2\right]$$

assuming that the measurements are independent

$$E[\delta y_i \delta y_j] = \begin{cases}\sigma_i^2 & \text{for } j = i \\ 0 & \text{for } j \neq i\end{cases}$$

$$\sim \sum_{i=0}^{N-1}\sigma_i^2\left(\frac{\partial f}{\partial y_i}\right)^2$$

To know the variance of a, and b, we first evaluate

$$\begin{cases}\dfrac{\partial a}{\partial y_i} = \dfrac{S x_i - S_x}{\sigma_i^2\Delta} \\ \dfrac{\partial b}{\partial y_i} = \dfrac{S_{xx} - S_x x_i}{\sigma_i^2\Delta}\end{cases}$$

$$\Rightarrow \begin{cases}\sigma_a^2 = S/\Delta \\ \sigma_b^2 = S_{xx}/\Delta\end{cases}$$

(iii) goodness-of-fitting: discussed later

呂喆宇

Introduction to Data Analysis 2023. Mar. 03

C. General linear problem using χ^2 fitting

Problems that depend linearly on the parameters a_k, $k=0, \cdots M$

For example,

1. fitting polynomials : $y(x) = a_0 + a_1 + a_2 x^2 + \cdots + a_{M-1} x^{M-1}$
2. expanding with basis functions: $y(x) = \sum_{k=0}^{M-1} a_k \underbrace{X_k(x)}_{\text{members of the basis functions}}$
   (e.g. fourier analysis)

$$\chi^2 = \sum_{i=0}^{N-1} \left[\frac{y_i - \sum_{k=0}^{M-1} a_k X_k(x_i)}{\sigma_i} \right]^2$$

, which has a minimum at

對 a_k 偏微分，找微分等於 0 處

$$0 = \sum_{i=0}^{N-1} \frac{1}{\sigma_i^2} \left[y_i - \sum_{j=0}^{M-1} a_j X_j(x_i) \right] X_k(x_i) \quad , \quad k = 0, \cdots, M-1 \qquad \text{——— Eq. (1)}$$

defining $\alpha_{kj} = \sum_{i=0}^{N-1} \frac{X_j(x_i) X_k(x_i)}{\sigma_i^2}$, $\beta_k = \sum_{i=0}^{N-1} \frac{y_i X_k(x_i)}{\sigma_i^2}$

Then Eq. (1) can be expressed as a matrix equation.

$$\overleftrightarrow{\alpha} \cdot \vec{a} = \vec{\beta}$$

(i) determining best-fit parameters

defining $\underbrace{\overleftrightarrow{C}}_{\text{covariant matrix}} \equiv \overleftrightarrow{\alpha}^{-1}$

此處自己寫程式時可以 quote 求反矩陣之演算法 → 如 numpy.linalg.inv carr

(the inverse matrix of $\overleftrightarrow{\alpha}$, i.e., $\overleftrightarrow{C} \cdot \overleftrightarrow{\alpha} = \mathbb{I}$)

若反矩陣不存在，如 $\overleftrightarrow{\alpha}$ 為 singul情形時，此方法會 crash.
發生在解為 ambiguous 情形，例多組不同的 a_k 對數據的描述一樣好.
此時應用 "Singular Value Decomposit. 求 χ^2 極值.

$$\Rightarrow a_j = \sum_{k=0}^{M-1} \overleftrightarrow{\alpha}_{jk}^{-1} \beta_k = \sum_{k=0}^{M-1} C_{jk} \left[\sum_{i=0}^{N-1} \frac{y_i X_k(x_i)}{\sigma_i^2} \right]$$

$$\Rightarrow \frac{\partial a_j}{\partial y_i} = \sum_{k=0}^{M-1} C_{jk} X_k(x_i) / \sigma_i^2$$

similar to the simpler case of fitting a line, we can know the varience of a_k after knowing this term :

視 a_k 為 y_i 的函數

$$\Rightarrow \sigma^2(a_j) = \sum_{i=0}^{N-1} \sigma_i^2 \left(\frac{\partial a_j}{\partial y_i} \right)^2$$

$$= \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} C_{jk} C_{jl} \underbrace{\left[\sum_{i=0}^{N-1} \frac{X_k(x_i) X_l(x_i)}{\sigma_i^2} \right]}_{\overleftrightarrow{\alpha} \text{ by definition}}$$

verify by yourself

$$= C_{jj}$$

The variance of a parameter is given by the diagonal term of the covariant matrix