stima dei parametri di una relazione lineare

anticipando quanto vedremo alla fine del corso

vi servirà prima per la parte di elettromagnetismo

rivedremo, con molto più dettaglio

Metodo dei Minimi Quadrati:

dato un campione di n variabili casuali $y_i,\ \ i = 1, n$, indipendenti

con incertezze statistiche σ_i e

valori di aspettazione μ_i che dipendono da parametri non noti,

le stime migliori dei parametri sono i valori che minimizzano la forma quadratica

$$\boldsymbol{X^2 = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{\sigma_i^2}}$$

nel nostro caso, ipotizziamo che tra le grandezze fisiche X e Y ci sia una relazione lineare

$$\boldsymbol{Y = mX + q}$$

e di voler misurare m e q

misure: n coppie di valori misurati (x_i, y_i) → campione $(\boldsymbol{y_1}, \boldsymbol{y_2}, \dots . \boldsymbol{y_n})$

con incertezze statistiche trascurabili su x_i, e σ_i su y_i

$$\mathrm{E}[y_i] = \boldsymbol{\mu_i = mx_i + q}$$

minimizzare $$\boldsymbol{X^2 = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{\sigma_i^2}}$$

$$\begin{cases} \dfrac{\partial X^2}{\partial m} = 0 \\ \dfrac{\partial X^2}{\partial q} = 0 \end{cases}$$

nel nostro caso, ipotizziamo che tra le grandezze fisiche X e Y ci sia una relazione lineare

$$\boldsymbol{Y = mX + q}$$

e di voler misurare m e q

misure: n coppie di valori misurati (x_i, y_i) → campione $(\boldsymbol{y_1, y_2, \dots . y_n})$
con incertezze statistiche trascurabili su x_i, e σ_i su y_i

$$\mathrm{E}[y_i] = \boldsymbol{\mu_i = m x_i + q}$$

minimizzare $$\boldsymbol{X^2 = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{\sigma_i^2}}$$

$$\left\{ \begin{array}{l} \dfrac{\partial X^2}{\partial m} = 0 \\[2ex] \dfrac{\partial X^2}{\partial q} = 0 \end{array} \right. \quad \Rightarrow \quad \begin{array}{l} \hat{m} = \dfrac{1}{D}(S_{00}S_{11} - S_{10}S_{01}) \\[2ex] \hat{q} = \dfrac{1}{D}(S_{01}S_{20} - S_{11}S_{10}) \end{array}$$

$$D = S_{00}S_{20} - S_{10}^2$$

$$S_{jk} = \sum_{i=1}^{n} \frac{x_i^j y_i^k}{\sigma_i^2}$$

con la legge di propagazione della varianza $\sigma_{\hat{m}}^2 = \sum_i \left(\frac{\partial \hat{m}}{\partial y_i}\right)^2 \sigma_i^2$, ...

$$\boldsymbol{\sigma_{\hat{m}}^2 = \frac{S_{00}}{D}} \qquad \boldsymbol{\sigma_{q}^2 = \frac{S_{20}}{D}}$$

e anche $$\boldsymbol{cov(\hat{m}, \hat{q}) = -\frac{S_{10}}{D}}$$

note

a) bisogna sempre controllare i risultati (calcoli sbagliati, relazione ipotizzata non corretta…): grafico con valori misurati, incertezze, e retta con parametri stimati

   se le incertezze non si apprezzano, grafico con differenza tra valori misurati e calcolati in una scala espansa

$$\delta_i = y_i - (\widehat{m} x_i + \hat{q}) \qquad \sigma_{\delta_i} = \sigma_i$$

note

b) se $Y = mX$ formule simili, più semplici, facili da ricavare

c) se l’incertezza su x_i non è trascrurabile ma quella su y_i si, basta scambiare $X\ e\ Y$

d) se entrambe le incertezze sono non trascurabili

al primo ordine
si stimamo una proma volta i parametri come se le incertezze su x_i fossero trascurabili → $\widehat{m}_0,\ \hat{q}_0$

si propaga l’errore $\sigma^2_{y_i,x_1} = \widehat{m}_0^2\, \sigma^2_{x_i}$, $\sigma'^2_i = \sigma_i^2 + \sigma^2_{y_i,x_i}$
si stimano di nuovo i parametri …. fino ad ottenere risultati stabili

note

e) se la relazione tra $\boldsymbol{X}$ e $\boldsymbol{Y}$ non è lineare…

spesso di può ‘’linearizzare’’

ad es $V(t) = V_0 e^{-t/c}$ ‘’campione’’ (t_i, V_i)

però $\ln V = \ln\ V_0 - t/c$

e quindi, introducendo $Y = lnV$ e $X = t$, ci si riconduce al caso precedente

NB: $\sigma_i^2\ = \cdots\ \sigma_{V_i}^2$

stima dei parametri di una relazione lineare

sarà parte della prima esercitazione:

- scrivere un programma (riutilizzabile, meglio se routine) che calcoli stime, loro incertezze e correlazione
- verificare graficamente l'accordo con i dati

riprenderemo il lavoro in esperienze successive