# La funzione tau_estrap.m come esempio di un metodo di BOOTSTRAP

Il nome bootstrap a che fare con l'allacciarsi gli scarponi (o gli stivali); il significato é paradossale: é possibile sollevarsi da terra allacciandosi gli stivali?

Nel caso in esame: abbiamo una serie di dati $E_0(a)$, ovvero determinazioni di $E_0$ per diversi valori di $a$. Ci proponiamo di determinare i parametri di una LEGGE (descrizione dei dati)

K = grado del polinomio $\left\{ \underbrace{\hat{E}_0(a) = \sum_{m=0}^{K} c_m a^{K-m}} \right.$ (NB Noi in realtà estrapoliamo in $a^2$!)

Evidentemente ci interessa $\underbrace{c_K = \lim_{a\to 0} E_0(a) = E_0}$, ma vogliamo tener conto degli errori $dE_0(a)$ (c'è un errore associato alla determinazioni di E ad ogni valore di a).

Conoscete la strada canonica; minimizzazione del funzionale

$$\chi^2 \propto \sum_{i=1}^{N} \left( \frac{E_0(a_i) - \hat{E}_0(a_i)}{dE_0(a_i)} \right)^2$$

che è cosa analiticamente risolubile per $\hat{E}_0(a) = c_0 a + c_1$; solubile anche per quanto riguarda la stima degli errori, così da derminare $E_0 \pm dE_0$ (dove ancora $E_0 = \hat{E}_0(a=0)$)

Possiamo però procedere in altro modo. Numericamente, se abbiamo dei DATI $\{d_i\} \to$ vettore d

ERRORI associati $\{dd_i\} \to$ vettore dd

ottenuti per valori di a $\{a_i\} \to$ vettore a, Matlab ci consente di valutare i coefficienti di un polinomio

$\sum_{m=0}^{K} P_m\, a^{K-m}$, dove al solito a noi interessa la INTERCETTA

$P_K$ (estrapolazione per $a \to 0$ dei dati). Si fa (banalmente)

$\{P_m\} \to$ vettore P $\quad \boxed{P = \text{polyfit}(a, d, K)} \quad (*)$

<u>PROBLEMA</u> e gli errori dd? e l'errore su $P_K$?

Ci sono funzioni di Matlab che ci aiuterebbero, ma noi "facciamo da soli" e usiamo un metodo di BOOTSTRAP.

L'idea è di <u>RIPETERE N VOLTE il fit (*)</u> $\to$ otteniamo allora un VETTORE di RISULTATI per $P_K$ $\{P_K^{(j)} \mid j = 1 \dots N\} = PP$ che chiamo (un nome come un altro...) PP. Il nostro risultato sarà $\boxed{P_K = \text{mean}(PP) \quad \text{con ERRORE} \quad dP_K = \text{std}(PP)}$

Come si prepara il campione "esteso" di risultati

$$PP = \{P_K^{(j)} \mid j = 1 \dots N\} \; ?$$

E' molto semplice. Consideriamo tutti i nostri DATI $d$, con gli ERRORI associati $dd$. $\forall i$ abbiamo il DATO $d_i$ e l'ERRORE $dd_i$ (da pensarsi associati al valore $a_i$ della variabile indip.)

<u>NOTA</u> che <u>ogni valore estratto da una GAUSSIANA CENTRATA in $d_i$ e di SIGMA $\sigma = dd_i$ è un RISULTATO PLAUSIBILE!</u>

Per questo gli <u>N FIT</u> sono calcolati (attenzione agli indici; $i$ INDICE in $\{d_i\}$, $j$ INDICE $1, \dots, N$)

Questo è fatto ← "vettorialmente"

(1a) generando $\forall i$ $\tilde{d}_i^{(j)} = d_i + dd_i \cdot \mathrm{randn}()$; $\overset{N}{\underset{j=1}{\forall}}$

(1b) questo mi restituisce un vettore $\tilde{d}^{(j)}$;

(2) faccio poi il FIT j-mo $\tilde{P}_j = \mathrm{polyfit}(a, \tilde{d}^{(j)}, K)$

dove per ogni $\overset{N}{\underset{j=1}{\forall}}$ mi interessa $\tilde{P}_j(K) = P_K^{(j)}$

Notazione vettoriale di Matlab, per non confonderci coi pedici...

nella notazione di PP (pagina precedente)

<u>NB</u> Intuitivamente: il FIT fatto sui VALORI CENTRALI $d$ non ha SIGNIFICATO ASSOLUTO... Ogni fit fatto per ogni set di dati $\tilde{d}^{(j)}$ è fatto per valori plausibili. Mi interessa la MEDIA di questi risultati (che ovviamente mi restituisce lo stesso risultato di $\mathrm{polyfit}(a, d, K)$) e la STANDARD DEVIATION di questo campione. Quest'ultima mi dà una STIMA SENSATA dell'ERRORE sul parametro fittato...