

Il metodo delle osservazioni indirette

Teoria della stima ai minimi quadrati

Il criterio di massima verosimiglianza

Sia data una grandezza η e si abbiano n osservazioni indipendenti l_i ($i=1, \dots, n$) di essa.

Siano v_i ($i=1, \dots, n$) gli errori di osservazione:

$$v_1 = l_1 - \eta$$

$$v_2 = l_2 - \eta$$

...

$$v_i = l_i - \eta$$

...

$$v_n = l_n - \eta$$

Problema: determinare il valore più probabile di η sulla base delle n osservazioni indipendenti l_i ($i=1, \dots, n$).

Così formulato, il problema equivale a trovare i più probabili valori per gli scarti v_i ($i=1, \dots, n$). I quali sono n variabili casuali, con distribuzione normale, la cui densità di probabilità è

$$f(v_i) = \frac{1}{\sigma_i \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{v_i}{\sigma_i}\right)^2}$$

La densità di probabilità congiunta è data dal prodotto delle probabilità delle singole componenti (teorema delle probabilità composte, per n variabili fra loro indipendenti):

$$\begin{aligned} f(v_1, v_2, \dots, v_n) &= \\ &= \frac{1}{\sigma_1 \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{v_1}{\sigma_1}\right)^2} \cdot \frac{1}{\sigma_2 \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left(\frac{v_2}{\sigma_2}\right)^2} \cdot \dots \cdot \frac{1}{\sigma_n \sqrt{2\pi}} \cdot \\ &= \frac{1}{(\sqrt{2\pi})^n} \cdot \left(\frac{1}{\sigma_1} \cdot \frac{1}{\sigma_2} \cdot \dots \cdot \frac{1}{\sigma_n} \right) \cdot e^{-\frac{1}{2} \sum_{i=1}^n \left(\frac{v_i}{\sigma_i}\right)^2} \end{aligned}$$

La densità di probabilità $f(v_1, v_2, \dots, v_n)$ è massima per valori minimi dell'esponente

$$\sum_{i=1}^n \left(\frac{v_i}{\sigma_i} \right)^2 = \min$$

Riscriviamo la condizione di minimo introducendo un fattore di scala σ_0 :

$$\sum_{i=1}^n \frac{\sigma_0^2 v_i^2}{\sigma_i^2} = \min$$

e chiamando P_i il peso dell'osservazione:

$$P_i = \frac{\sigma_0^2}{\sigma_i^2}$$

In forma vettoriale:

$$v^t P v = \min$$

Un'osservazione con $\sigma_i^2 = \sigma_0^2$ ha peso unitario.

Pertanto σ_0^2 viene detta VARIANZA DELL'UNITA' DI PESO.

Il principio dei minimi quadrati

I sistemi sovradeterminati (numero di equazioni superiore al numero di variabili) non hanno una soluzione unica. Per essi si cerca la soluzione migliore, cioè quella di minima varianza. Si supponga di poter scrivere per ognuna delle n osservazioni (misure di distanza, di angoli o dislivello) una relazione del tipo $f_i(x, \alpha) = 0$, dove x è il vettore delle $m < n$ incognite e α è il vettore delle r quantità misurate. A causa degli errori di misura (di tipo accidentale) presenti nelle quantità misurate α , compare un residuo v (dovuto agli errori di misura) nella relazione tra osservazioni e incognite, che diviene $f_i(x, \alpha) = v_i$.

Si presentano due problemi:

- 1) spesso le equazioni si presentano in forma non lineare: occorre linearizzarle.
- 2) non tutte le osservazioni contano allo stesso modo: bisogna pesarle.

1) Linearizzazione delle equazioni di osservazione.

Se supponiamo che le quantità misurate non siano eccessivamente disperse (solo errori di tipo accidentale e di piccola entità), anche le x_k non differiranno eccessivamente fra di loro. Sarà pertanto possibile trovare dei valori approssimati x_0 delle incognite x e sviluppare, a partire da essi, la relazione $f_i(x, \alpha) = v_i$ in termini di serie di Taylor arrestata al primo ordine.

$$X = X_0 + \delta X$$

$$f_i(x_0, \alpha) + \left[\frac{\partial f_i}{\partial x_1} \right]_0 \delta x_1 + \left[\frac{\partial f_i}{\partial x_2} \right]_0 \delta x_2 + \dots + \left[\frac{\partial f_i}{\partial x_m} \right]_0 \delta x_m = v_i$$

dove

$f_i(x_0, \alpha) = l_i$ è una quantità nota e le

δx_k sono le nuove incognite di cui i coefficienti

$\left[\frac{\partial f_i}{\partial x_k} \right]_0$ costituiscono una matrice A.

Dalla linearizzazione si ottiene un sistema lineare di n equazioni in m incognite δx (e n incognite v) di forma:

$$A \delta x + l = v$$

2) Analisi ponderale delle equazioni.

Il problema si può ricondurre alla determinazione della varianza dell'osservazione, a sua volta riconducibile - nell'ambito delle misure - alla varianza dello strumento di misura (bisognerebbe eseguire l'analisi della distribuzione di ogni singola osservazione).

Il peso da assegnare a ciascuna equazione sarà inversamente proporzionale alla varianza dell'osservazione. La matrice dei pesi sarà una matrice diagonale del tipo:

$$P = \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & P_n \end{pmatrix}$$

La matrice di varianza-covarianza delle osservabili, è uguale - a meno di un fattore di scala σ_0^2 - alla matrice inversa dei pesi P^{-1} .

$$C_{LL} = \sigma_0^2 \cdot P^{-1} = \sigma_0^2 \cdot Q$$

Il sistema sottodeterminato insieme alla condizione di minimo pesato della somma degli scarti quadratici costituisce il **sistema normale**:

$$A \delta x + l = v$$

$$\sum P v^2 = v^t P v = \min$$

Imporre la condizione di minimo equivale a differenziare la forma quadratica $v^t P v$ in questo modo:

$$d(v^t P v) = 0$$

$$d v^t P v + v^t P d v = 0$$

$$2 d v^t P v = 0$$

$$d v^t P v = 0$$

I passaggi precedenti si spiegano facilmente considerando che la forma quadratica $v^t P v$ è uno scalare, infatti:

$$(v_1 \ v_2 \ \dots \ v_n) \cdot \begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & P_n \end{pmatrix} \cdot \begin{pmatrix} dv_1 \\ dv_2 \\ \dots \\ dv_n \end{pmatrix}$$

$$(1, n) \cdot (n, n) \cdot (n, 1) = (1, 1)$$

Il trasposto di una matrice (1,1) è la matrice stessa:

$$v^t P d v = (v^t P d v)^t = d v^t P v$$

essendo $P^t=P$.

Trasponendo e differenziando l'espressione $A \delta x + l = v$ si ottiene:

$$v^t = \delta X^t A^t + l^t$$

$$d v^t = d(\delta X^t A^t)$$

Sostituendo nella condizione di minimo, essa diviene:

$$d v^t P v = 0$$

$$d(\delta X^t A^t) P v = 0$$

cioè:

$$v = A \delta X + l$$

$$d(\delta X^t A^t) P (A \delta X + l) = 0$$

Dovendo essere la precedente equazione soddisfatta per ogni differenziale $d(\delta X^t)$ si ha:

$$A^t P (A \delta X + l) = 0$$

$$A^t P A \delta X + A^t P l = 0$$

Alla fine:

$$\delta X = \begin{pmatrix} \delta X_1 \\ \delta X_2 \\ \dots \\ \delta X_m \end{pmatrix} = -(A^t P A)^{-1} A^t P l$$

Dimensionalmente:

$$A \rightarrow (n, m)$$

$$A^t \rightarrow (m, n)$$

$$P \rightarrow (n, n)$$

$$l \rightarrow (n, 1)$$

$$A^t P A \rightarrow (m, m)$$

$$(A^t P A)^{-1} \rightarrow (m, m)$$

$$A^t P l \rightarrow (m, 1)$$

$$(A^t P A)^{-1} A^t P l \rightarrow (m, 1)$$

La stima delle varianze delle incognite δx è data dai termini diagonali h_{ii} della matrice normale inversa

$$(A^t P A)^{-1}$$

moltiplicati per un termine

$$\hat{\sigma}_0^2$$

che è la stima della varianza dell'unità di peso.

La forma esplicita è la seguente

$$C_{ii} = \hat{\sigma}_0^2 \cdot h_{ii}$$

dove

$$\hat{\sigma}_0^2$$

è dato dalla sommatoria degli scarti pesati al quadrato divisa per la ridondanza (n-m) del sistema:

$$\hat{\sigma}_0^2 = \frac{\sum P v_i^2}{n - m} = \frac{v^t P v}{n - m}$$

Le covarianze C_{ik} tra le incognite determinate sono date dalla varianza di peso unitario

$$\hat{\sigma}_0^2$$

moltiplicata per i termini extra-diagonali della matrice normale inversa h_{ik} .

La forma esplicita è la seguente

$$C_{ik} = \hat{\sigma}_0^2 \cdot h_{ik}$$

In forma matriciale:

$$C_{xx} = \hat{\sigma}_0^2 \cdot H$$

Vediamo perché vale la relazione precedente.

Applichiamo alle osservabili la legge di propagazione della varianza, che in generale - per due variabili casuali U e V legate dalla matrice T, tale che $U=TV$ - vale

$$C_{UU} = TC_{VV} T^t$$

Nel nostro caso:

$$C_{XX} = TC_{LL} T^t$$

con

$$T = -(A^t P A)^{-1} A^t P$$

pertanto

$$\begin{aligned} C_{XX} &= TC_{LL} T^t = ((A^t P A)^{-1} A^t P) C_{LL} ((A^t P A)^{-1} A^t P)^t \\ &= ((A^t P A)^{-1} A^t P) (\hat{\sigma}_0^2 Q) ((A^t P A)^{-1} A^t P)^t \\ &= \hat{\sigma}_0^2 (A^t P A)^{-1} A^t P Q P^t A ((A^t P A)^{-1})^t \\ &= \hat{\sigma}_0^2 (A^t P A)^{-1} A^t P A (A^t P A)^{-1} \\ &= \hat{\sigma}_0^2 (A^t P A)^{-1} \\ &= \hat{\sigma}_0^2 \cdot H \end{aligned}$$

Analogamente, la matrice di varianza-covarianza delle osservabili vale:

$$C_{LL} = A C_{XX} A^t$$