01AYS - FONDAMENTI DI AUTOMATICA

Stima parametrica classica

proff. Marina Indri e Michele Taragna
Dip. di Automatica e Informatica
Politecnico di Torino

Anno Accademico 2003/2004

Versione 1.2

Identificazione classica (I)

Ipotesi #1: il sistema $\mathcal{S}$ è descritto da un modello parametrico $\mathcal{M}$ che dipende da un vettore di parametri "veri" non noti $p^o \in \Re^n$:

$$\mathcal{S} = \mathcal{M}\left(p^o\right)$$

Ipotesi #2: su $\mathcal{S}$ si ha solo un numero limitato N di informazioni sperimentali, corrotte da un rumore di misura non noto $e^N \in \Re^N$, costituito da una sequenza aleatoria con momenti del I e del II ordine noti:

$$\underbrace{E\left(e^N\right)}_{\text{valor medio}} = 0 \qquad \underbrace{E\left(e^N \cdot \left(e^N\right)^T\right)}_{\text{matrice di covarianza}} = \Sigma_e \text{ nota}$$

Identificazione classica (II)

Problema di stima parametrica:

data la relazione:

$$\underbrace{z^N}_{\substack{\text{dati}\\ \text{noti}}} = \underbrace{F\left(p^o\right)}_{\substack{\text{funzione}\\ \text{nota}}} + \underbrace{e^N}_{\substack{\text{rumore}\\ \text{non noto}}}$$

trovare una stima $\widehat{p}$ di p^o mediante un algoritmo di stima (stimatore) ϕ applicato ai dati z^N:

$$p^o \cong \widehat{p} = \phi\left(z^N\right)$$

Esempio (I)

<u>Stima della resistenza di un resistore reale</u>

Si effettuano N misure di tensione-corrente supponendo che:

- la caratteristica statica del resistore sia lineare e quindi il modello del dispositivo sia dato dalla legge di Ohm:

$$v_R = R \cdot i_R$$

- le misure siano corrotte da un rumore casuale $e^N = [e_1, \ldots, e_N]^T$

Si ottiene così il sistema di equazioni:

$$\begin{aligned} v_{R,1} &= R \cdot i_{R,1} + e_1 \\ v_{R,2} &= R \cdot i_{R,2} + e_2 \\ &\vdots \\ v_{R,N} &= R \cdot i_{R,N} + e_N \end{aligned}$$

Esempio (II)

In forma matriciale:

$$\underbrace{\begin{bmatrix} v_{R,1} \\ v_{R,2} \\ \vdots \\ v_{R,N} \end{bmatrix}}_{z^N} = \underbrace{\begin{bmatrix} i_{R,1} \\ i_{R,2} \\ \vdots \\ i_{R,N} \end{bmatrix}}_{L} \cdot \underbrace{[R\,]}_{p^o} + \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}}_{e^N}$$

è nella forma standard:

$$\underbrace{z^N}_{\substack{\text{dati}\\ \text{noti}}} = \underbrace{F\left(p^o\right)}_{\substack{\text{funzione}\\ \text{nota}}} + \underbrace{e^N}_{\substack{\text{rumore}\\ \text{non noto}}}$$

con:

$F\left(p^o\right) = L \cdot p^o =$ funzione lineare del parametro non noto p^o

<u>Obiettivo</u>: trovare una stima $\widehat{R}$ di R mediante un algoritmo di stima (stimatore) ϕ applicato ai dati z^N:

$$\widehat{R} = \phi\left(z^N\right) \cong R$$

Proprietà delle stime statistiche (I)

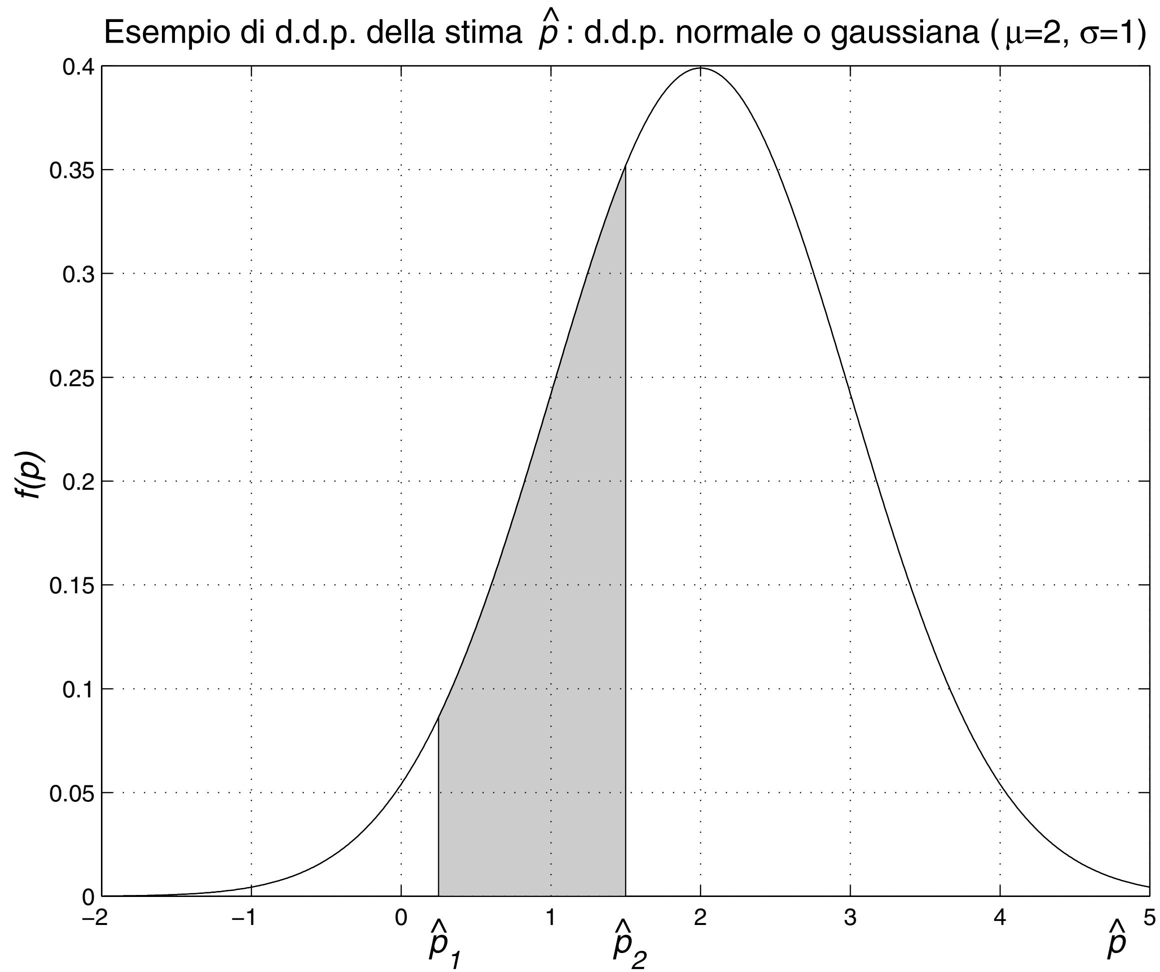
$$e^N \text{ è una variabile aleatoria}$$
$$\Downarrow$$
$$z^N = F(p^o) + e^N \text{ è una variabile aleatoria}$$
$$\Downarrow$$
$$\widehat{p} = \phi\left(z^N\right) \text{ è una variabile aleatoria}$$

Problema: come si misura la "bontà" della stima $\widehat{p}$?

Per confrontare due differenti algoritmi di stima che portano a due diverse stime $\widehat{p}$ e $\tilde{p}$, paragono le densità di probabilità delle rispettive stime sulla base delle seguenti caratteristiche:

- correttezza;
- efficienza;
- consistenza

Proprietà delle stime statistiche (II)

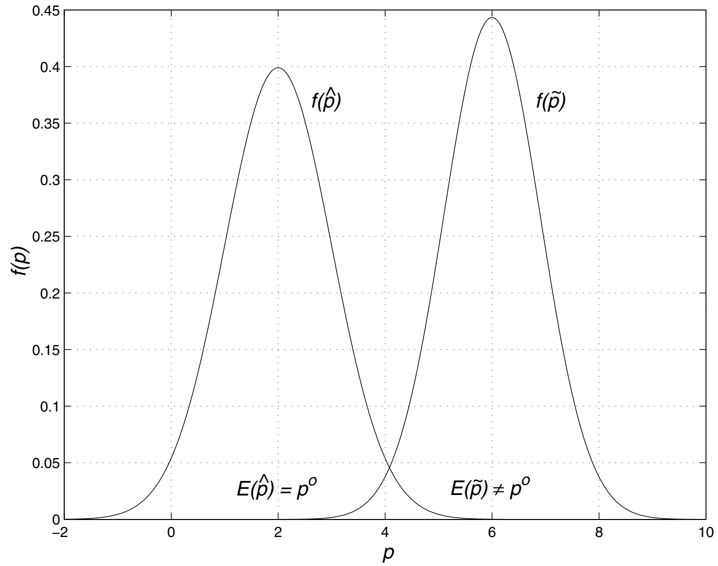


Area = Probabilità$(\widehat{p}_1 \leq \widehat{p} \leq \widehat{p}_2)$

Proprietà delle stime statistiche (III)

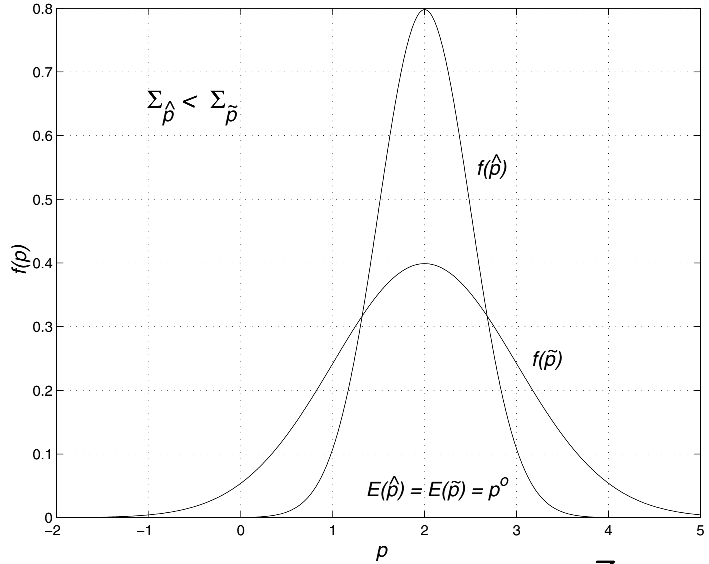
- **Correttezza** (o **non polarizzazione**)
  confronto i momenti del I ordine e scelgo $\widehat{p}$ in modo che:

$$\underbrace{E(\widehat{p})}_{\substack{\text{valor medio}\\ \text{della stima}}} \quad = \quad \underbrace{p^o}_{\substack{\text{valore vero}\\ \text{dei parametri}}}$$



- **Efficienza**
  se entrambe le stime sono corrette, confronto i momenti del II ordine e scelgo $\widehat{p}$ in modo che:

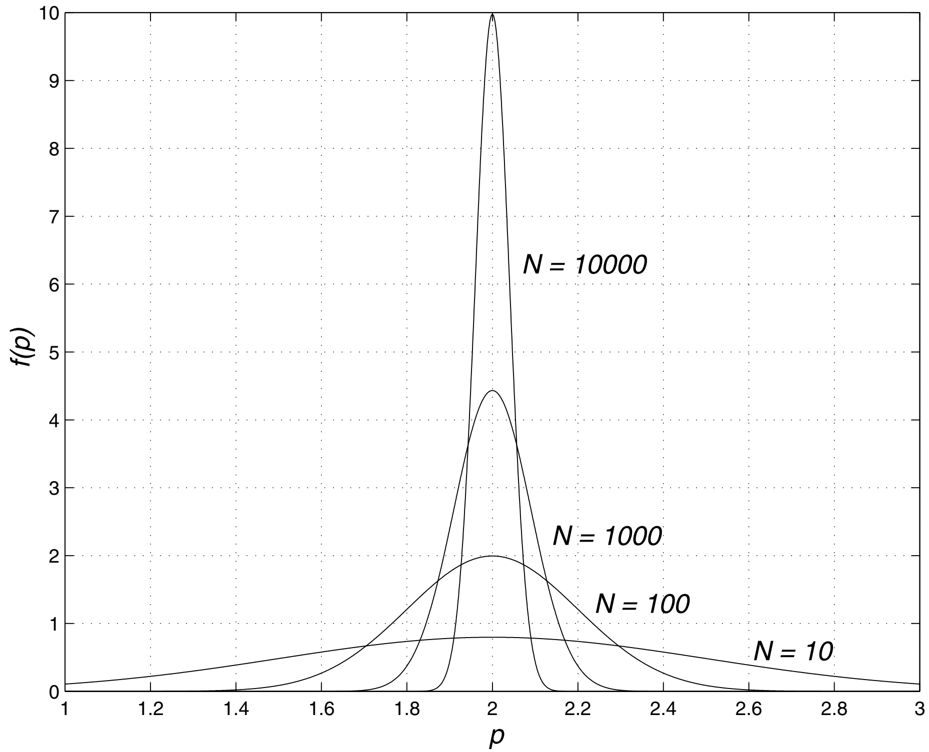
$$\sum\nolimits_{\widehat{p}} \leq \sum\nolimits_{\tilde{p}}$$



Proprietà delle stime statistiche (IV)

- **Consistenza**
  sia per stime efficienti
  sia per stime non efficienti,
  guardo se vale la seguente
  proprietà asintotica:

$$\lim_{N \to \infty} \Sigma_{\widehat{p}} = 0$$



Stima di massima verosimiglianza (I)

Si riescono a trovare stimatori corretti, efficienti, consistenti?

Definisco **funzione di verosimiglianza** la d.d.p. condizionata

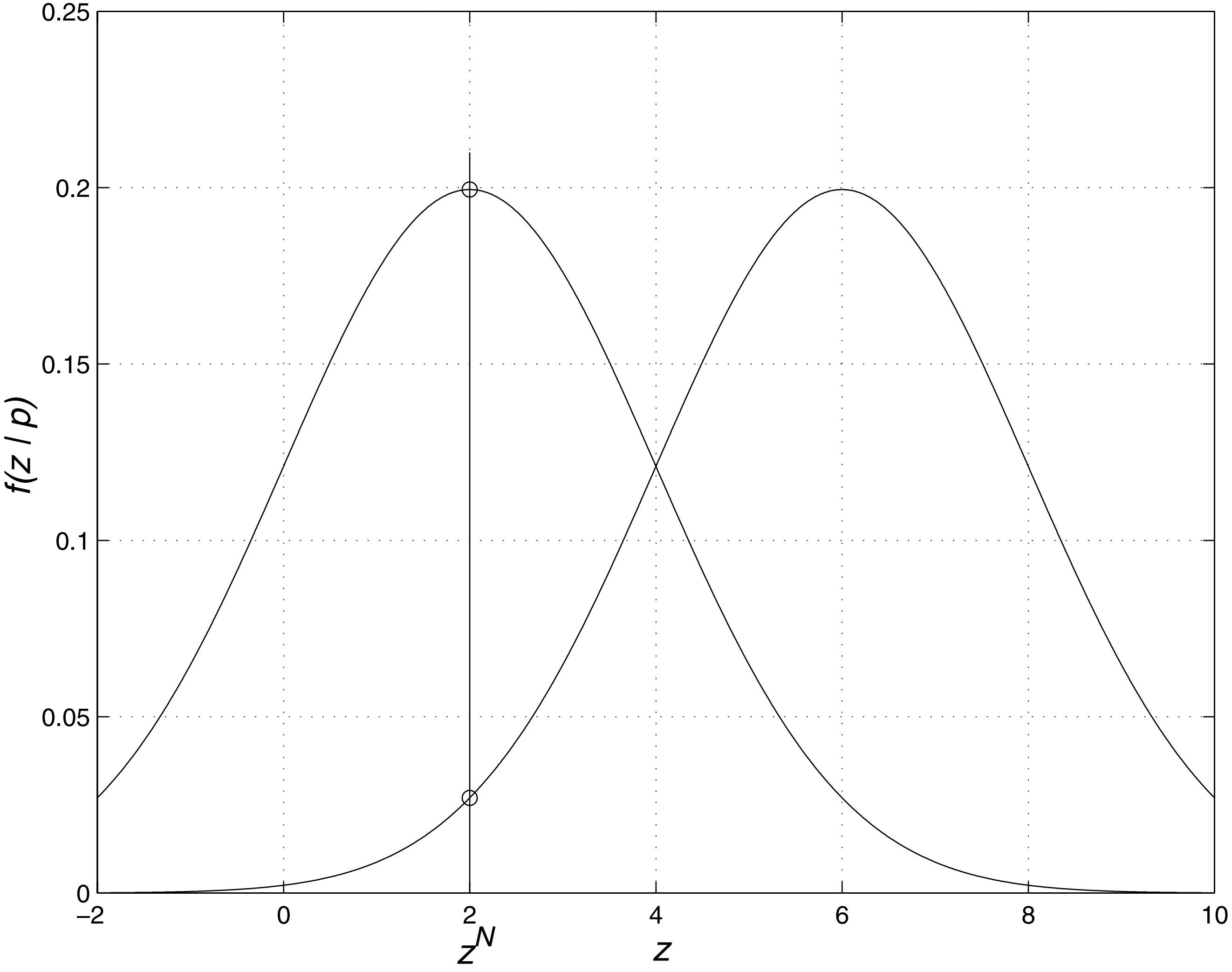
$f(z|\,p) =$ d.d.p. dei dati z se fossero generati dal sistema $\mathcal{M}(p)$

e definisco **stima di massima verosimiglianza:**

$$\widehat{p}^{MV} = \arg\max_{p \in \Re^n} f\left(z = z^N \middle|\, p\right)$$

la stima corrispondente al valore dei parametri per cui i dati misurati z^N sono i più probabili

Stima di massima verosimiglianza (II)



Proprietà della stima di massima verosimiglianza

$\widehat{p}^{\,MV}$ è una stima:

- asintoticamente corretta: $E\left(\widehat{p}^{\,MV}\right) \xrightarrow[N \to \infty]{} p^o$
- asintoticamente efficiente: $\sum_{\widehat{p}^{MV}} \leq \sum_{\tilde{p}} \quad \forall \tilde{p}$ se $N \to \infty$
- consistente: $\lim\limits_{N\to\infty} \sum_{\widehat{p}^{MV}} = 0$
- asintoticamente gaussiana (per $N \to \infty$)

Calcolo della stima di massima verosimiglianza (I)

Nel caso di errori gaussiani:

$$f\left(e^N\right) = \underbrace{\mathcal{N}\left(0, \sum_e\right)}_{\substack{\text{distribuzione} \\ \text{normale o gaussiana}}} = \frac{1}{\sqrt{(2\pi)^N \cdot \det\left(\sum_e\right)}} \exp\left(-\frac{1}{2}\left(e^N\right)^T \cdot \sum_e^{-1} \cdot e^N\right)$$

Caso generale: $F(p)$ = generica funzione <u>non lineare</u> dei parametri

$$z^N = F(p) + e^N$$

$$\Downarrow \quad {\scriptstyle e^N = z^N - F(p)}$$

$$f\left(z^N \middle| p\right) = \mathcal{N}\left(F(p), \sum_e\right) = \frac{1}{\sqrt{(2\pi)^N \cdot \det\left(\sum_e\right)}} \exp\left(-\frac{1}{2}\left[z^N - F(p)\right]^T \cdot \sum_e^{-1} \cdot \left[z^N - F(p)\right]\right)$$

Calcolo della stima di massima verosimiglianza (II)

$$f\left(z^N \middle| p\right) \text{ è una funzione esponenziale in } p$$

$$\Downarrow$$

$$\widehat{p}^{MV} = \arg\max_{p\in\Re^n} f\left(z^N \middle| p\right) = \arg\min_{p\in\Re^n} \underbrace{\left\{\left[z^N - F(p)\right]^T \cdot \Sigma_e^{-1} \cdot \left[z^N - F(p)\right]\right\}}_{R(p)}$$

Problema: bisogna trovare il minimo globale di $R(p)$ al variare di p, che però può avere minimi locali se $F(p)$ è una generica funzione non lineare dei parametri; gli usuali algoritmi di ottimizzazione non lineare non garantiscono di arrivare al minimo globale

Calcolo della stima di massima verosimiglianza (III)

Caso particolare: $F(p) =$ funzione lineare dei parametri $= Lp$

$$z^N = Lp + e^N$$

$$\Downarrow$$

$R(p)$ è quadratica in p : $R(p) = \left[z^N - Lp\right]^T \cdot \sum_e^{-1} \cdot \left[z^N - Lp\right]$

$$\Downarrow$$

esiste un solo minimo di $R(p)$

$$\Downarrow$$

$\widehat{p}^{MV} = \left(L^T \sum_e^{-1} L\right)^{-1} L^T \sum_e^{-1} z^N =$ **stima di Gauss-Markov** $= \widehat{p}^{GM} =$
$=$ *stima dei minimi quadrati pesati* con la covarianza degli errori

Nel caso $\sum_e = \sigma_e^2 I_N$ (errori indipendenti identicamente distribuiti):

$\widehat{p}^{MV} = \widehat{p}^{GM} = \left(L^T L\right)^{-1} L^T z^N =$ **stima dei minimi quadrati**

Proprietà della stima di Gauss-Markov

$\widehat{p}^{GM}$ è una stima:

- corretta: $E\left(\widehat{p}^{GM}\right) = p^o$

- efficiente: $\sum_{\widehat{p}^{GM}} \leq \sum_{\widetilde{p}} \quad \forall \widetilde{p}$

- consistente: $\lim_{N\to\infty} \sum_{\widehat{p}^{GM}} = 0$

- gaussiana