

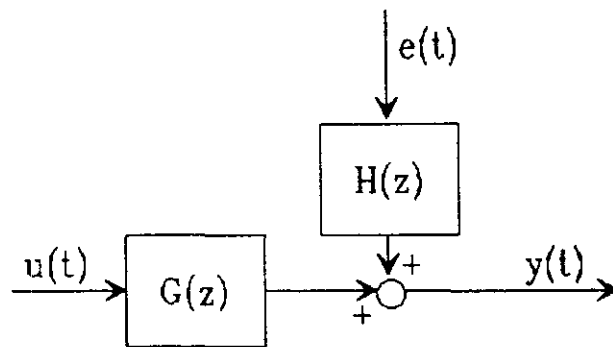
FAMIGLIE DI MODELLI DINAMICI

- **Struttura dei modelli affetti da disturbi stocastici**
- **Predizione ad un passo della variabile di uscita**
- **Identificazione a minimizzazione dell'errore di predizione**
- **Modelli ARX, ARMAX, ARIMAX, FIR, OE**

Le principali famiglie di modelli dinamici considerate nell'identificazione possono essere viste come casi particolari del modello

$$y(t) = G(z)u(t) + H(z)e(t)$$

dove $y(t)$ è l'uscita, $u(t)$ l'ingresso manipolabile e $e(\cdot) \sim WN(0, \sigma^2)$. Ci si limita qui a considerare il caso di sistemi ad un ingresso ed un'uscita, quindi $G(z)$ e $H(z)$ sono funzioni di trasferimento date dal rapporto di polinomi, di cui si suppone noto l'ordine, e i cui parametri sono da stimarsi con opportune procedure di identificazione.



Il segnale $v(t) = H(z)e(t)$ può rappresentare disturbi agenti sul sistema e/o fenomeni non modellizzati da $G(z)u(t)$. Per il teorema della fattorizzazione spettrale si può supporre che $H(z)$ sia il **fattore spettrale canonico**, il suo denominatore ha quindi tutti i poli con modulo minore di uno, mentre le radici del numeratore hanno modulo minore o uguale a uno.

Si ipotizza inoltre che $G(z)$ sia strettamente causale, cioè con denominatore di grado maggiore del numeratore. $H(z)$ è causale, cioè con numeratore e denominatore di ugual grado, e il coefficiente della potenza di ordine massimo di numeratore e denominatore è 1.

Specializzando a casi particolari le strutture di $G(z)$ e $H(z)$ si ottengono diverse **famiglie di modelli**, nel loro complesso chiamati modelli *black-box*, o *a scatola nera*.

PREDIZIONE AD UN PASSO

Poiché i metodi di identificazione più comunemente utilizzati si basano sulla minimizzazione dell'errore di predizione ad un passo, cioè sulla determinazione dei parametri di $G(z)$ e $H(z)$ che minimizzano tale errore, è ora necessario ricavare il predittore del sistema scritto nella sua forma più generale.

Per questo, moltiplichiamo il sistema per $H(z)^{-1}$:

$$H(z)^{-1}y(t) = H(z)^{-1}G(z)u(t) + e(t)$$

sommiamo ora $y(t)$ a entrambi i membri:

$$y(t) = y(t) - H(z)^{-1}y(t) + H(z)^{-1}G(z)u(t) + e(t)$$

cioè

$$y(t) = (1 - H(z)^{-1})y(t) + H(z)^{-1}G(z)u(t) + e(t)$$

Poiché numeratore e denominatore di $H(z)$ sono polinomi monici di ugual grado, risulta

$$1 - H(z)^{-1} = 1 - \frac{z^n + \gamma_1 z^{n-1} + \dots + \gamma_n}{z^n + \delta_1 z^{n-1} + \dots + \delta_n} = \frac{\eta_1 z^{n-1} + \dots + \eta_n}{z^n + \delta_1 z^{n-1} + \dots + \delta_n} = \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots$$

Il segnale $(1 - H(z)^{-1})y(t) = \alpha_1 y(t-1) + \alpha_2 y(t-2) + \dots$ dipende quindi al più da $y(t-1)$, così come il segnale $H(z)^{-1}G(z)u(t)$ dipende al più da $u(t-1)$. Poiché e è un rumore bianco, la miglior predizione $y^p(t/t-1)$ di $y(t)$ a partire dai dati disponibili, cioè da y e u , sino all'istante $t-1$ è data da

$$y^p(t/t-1) = (1 - H(z)^{-1})y(t) + H(z)^{-1}G(z)u(t)$$

IDENTIFICAZIONE A MINIMIZZAZIONE DELL'ERRORE DI
PREDIZIONE

Supponiamo di aver raccolto delle sequenze di ingresso $u(1), u(2), \dots, u(N)$ e uscita $y(1), y(2), \dots, y(N)$ e di voler identificare un modello. Il primo passo è quello di **determinare la struttura del modello**, cioè l'ordine di $G(z)$ e $H(z)$, la struttura ed eventualmente l'uguaglianza dei loro polinomi a numeratore e/o a denominatore. Queste scelte (struttura, ordine) sono complesse e verranno considerate nel seguito.

Il passo successivo è quello di **stimare i parametri** delle funzioni di trasferimento $G(z)$ e $H(z)$, cioè i coefficienti dei rispettivi numeratori e denominatori. Supponiamo che questi parametri siano raccolti in un vettore Θ da stimare. Nel metodo *PEM* (*Prediction Error Minimization*) tale vettore viene determinato in modo da **minimizzare l'errore di predizione ad un passo**. Quindi, poiché il predittore ottimo ad un passo $y^p(t/t-1)$ di $y(t)$ dipende da Θ , tale vettore viene determinato come

$$\Theta = \operatorname{argmin}_{\Theta} \sum_{t=n}^N (y(t) - y^p(t/t-1))^2 = \sum_{t=n}^N \varepsilon(t)^2$$

dove $\varepsilon(t) = y(t) - y^p(t/t-1)$ è l'errore di predizione ad un passo e N è il numero di dati di ingresso e uscita disponibili. Si noti che, poiché per calcolare $y^p(t/t-1)$ è necessario conoscere $y(t-1), \dots, y(t-n), u(t-1), \dots, u(t-n)$ dove n è l'ordine massimo di $G(z)$ e $H(z)$, la sommatoria non parte da 1.

Il problema precedente può essere visto come un **problema ai minimi quadrati** del tutto analogo a quelli considerati in precedenza e potrà essere risolto in forma chiusa se la dipendenza di y^p da Θ è di tipo lineare, o con algoritmi iterativi se è non lineare.

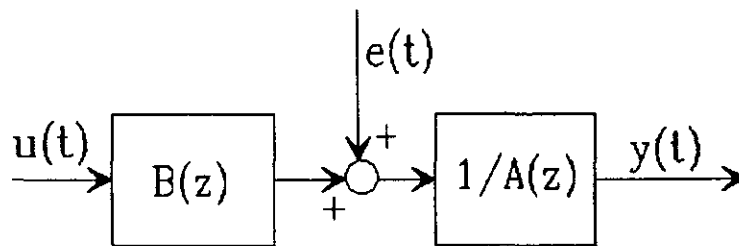
Vengono ora analizzate alcune famiglie di modelli e le relative tecniche di identificazione.

MODELLI A ERRORE D'EQUAZIONE
 Modelli ARX (AutoRegressive eXogenous)

Si ponga

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n}, \quad H(z) = \frac{z^n}{A(z)} = \frac{z^n}{z^n + a_1 z^{n-1} + \dots + a_n}$$

Il corrispondente schema a blocchi è



che mostra come il disturbo passi attraverso il termine $1/A(z)$, cioè agisca sullo stato del sistema.

Il modello, scritto in termini di relazione ingresso/uscita, è

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) + b_1 u(t-1) + b_2 u(t-2) + \dots + b_n u(t-n) + e(t)$$

e il predittore ottimo ad un passo è

$$y^p(t/t-1) = \left(1 - \frac{A(z)}{z^n}\right) y(t) + \frac{B(z)}{z^n} u(t) = y(t) - e(t)$$

Si osservi che definendo

$$\Theta = [a_1 \ a_2 \ \dots \ a_n \ b_1 \ b_2 \ \dots \ b_n]'$$

$$\phi(t) = [-y(t-1) \ -y(t-2) \ \dots \ -y(t-n) \ u(t-1) \ u(t-2) \ \dots \ u(t-n)]'$$

risulta

$$y^p(t/t-1) = \phi(t)\Theta$$

il predittore è lineare nei parametri e il problema di identificazione

$$\Theta = \underset{\Theta}{\operatorname{argmin}} \sum_{t=n}^N (y(t) - y^p(t/t-1))^2 = \underset{\Theta}{\operatorname{argmin}} \sum_{t=n}^N (y(t) - \phi(t)\Theta)^2$$

porta alla stima

$$\Theta = \left[\sum_{t=n}^N \phi(t)\phi(t)' \right]^{-1} \sum_{t=n}^N \phi(t)y(t)$$

dove si è ipotizzato che la matrice $S(N) = \sum_{t=n}^N \phi(t)\phi(t)'$ sia non singolare.

E' anche importante notare che il predittore ad un passo dipende dalle misure passate di y e u , cioè la predizione è effettuata sulla base dell'uscita vera misurata e non dell'uscita predetta agli istanti precedenti. In altri termini **un modello con buone capacità predittive non è necessariamente un buon modello di simulazione.**

Analisi probabilistica

Nell'ipotesi che la struttura del sistema vero

$$y(t) = \phi(t)\Theta^0 + e(t)$$

corrisponda al modello ipotizzato e che effettivamente $e(\cdot) \sim WN(0, \sigma^2)$ si possono utilizzare i risultati dell'analisi probabilistica della stima ai minimi quadrati per concludere che:

1. Lo stimatore è non polarizzato

$$E[\Theta] = \Theta^0$$

2. La stima della varianza del rumore si può determinare con la formula

$$\sigma^2 = \frac{1}{(N - n + 1) - 2n} \sum_{t=n}^N (y(t) - y_{\Theta}^p(t/t-1))^2$$

3. La stima della varianza dei parametri si può valutare come

$$Var[\Theta] = \sigma^2 S(N)^{-1}$$

Modelli FIR (Finite Impulse Response)

Possono essere interpretati come un caso particolare di modelli ARX in cui $A(z)=z^n$, cioè

$$y(t) = B(z)u(t) + e(t) = b_1 u(t-1) + b_2 u(t-2) + \dots + b_n u(t-n) + e(t)$$

e, posto

$$\Theta = [b_1 \ b_2 \ \dots \ b_n]'$$

$$\phi(t) = [u(t-1) \ u(t-2) \ \dots \ u(t-n)]$$

risulta

$$y^p(t/t-1) = \phi(t)\Theta$$

Il problema di identificazione

$$\Theta = \underset{\Theta}{\operatorname{argmin}} \sum_{t=n}^N (y(t) - y^p(t/t-1))^2 = \sum_{t=n}^N (y(t) - \phi(t)\Theta)^2$$

è ancora risolto da

$$\Theta = \left[\sum_{t=n}^N \phi(t)\phi(t)' \right]^{-1} \sum_{t=n}^N \phi(t)y(t)$$

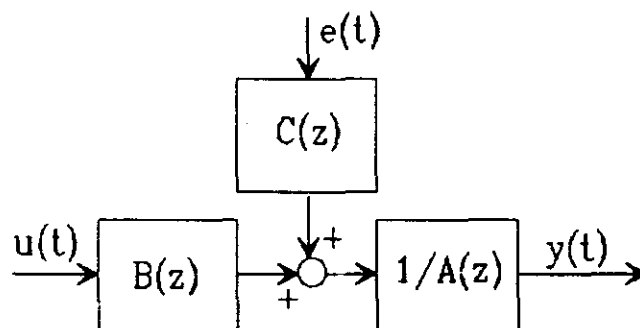
dove si è ipotizzato che la matrice $S(N) = \sum_{t=n}^N \phi(t)\phi(t)'$ sia non singolare.

Modelli ARMAX (AutoRegressive MovingAverage eXogenous)
o CARMA (Controlled AutoRegressive Moving Average)

In molti casi l'ipotesi di bianchezza di e non può essere verificata utilizzando un modello ARX. Per rendere più flessibile il modello si ponga

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n}, \quad H(z) = \frac{C(z)}{A(z)} = \frac{z^n + c_1 z^{n-1} + \dots + c_n}{z^n + a_1 z^{n-1} + \dots + a_n}$$

Il corrispondente schema a blocchi è



e risulta

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) + b_1 u(t-1) + b_2 u(t-2) + \dots + b_n u(t-n) + e(t) + c_1 e(t-1) + c_2 e(t-2) + \dots + c_n e(t-n)$$

e il predittore ottimo ad un passo è

$$y^p(t/t-1) = \frac{C(z) - A(z)}{C(z)} y(t) + \frac{B(z)}{C(z)} u(t)$$

Si osservi che, poiché $A(z)$ e $C(z)$ sono polinomi monici e di ugual grado, il predittore dipende al più da $y(t-1)$.

Il predittore è non lineare nei parametri da stimare, pertanto non si può ricavare una forma chiusa per la stima, ma è necessario minimizzare il funzionale quadratico ricorrendo ad un metodo evolutivo.

Si noti tuttavia che se $C(z)$ fosse noto, si potrebbero calcolare ad ogni istante t i segnali

$$y_c(t) = \frac{z^n}{C(z)} y(t) \quad , \quad u_c(t) = \frac{z^n}{C(z)} u(t)$$

e il predittore si potrebbe scrivere come

$$y^p(t/t-1) = \frac{C(z) - A(z)}{z^n} y_c(t) + \frac{B(z)}{z^n} u_c(t) =$$

lineare nei parametri di $A(z)$ e $B(z)$. In questo caso si tornerebbe al caso dei modelli *ARX*.

Un esempio

Si consideri il caso

$$A(z) = z + a \quad , \quad B(z) = b \quad , \quad C(z) = z + c$$

dalle definizioni precedenti

$$y_c(t) = -cy_c(t-1) + y(t) \quad , \quad u_c(t) = -cu_c(t-1) + u(t)$$

e quindi

$$y^p(t/t-1) = (c-a)y_c(t-1) + bu_c(t-1)$$

lineare nelle incognite a e b (c è noto).

□

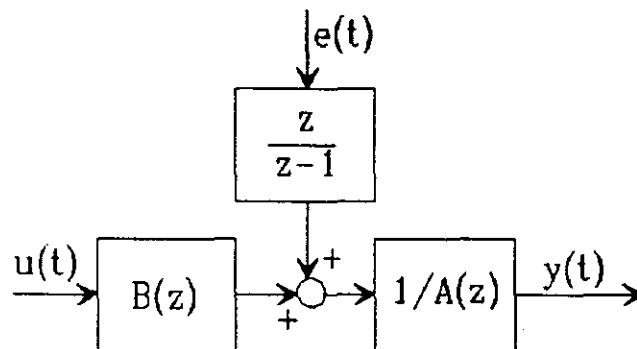
Per questi motivi spesso si ipotizza che $C(z)$ sia noto. La scelta di un polinomio $C(z)$ possibile è analizzata nel seguito. In alternativa sono stati proposti schemi di identificazione che alternano la fase di stima di $C(z)$ a quella di $A(z)$ e $B(z)$.

Modelli ARIMAX (AutoRegressive Integrated MovingAverage eXogenous)
o CARIMA (Controlled AutoRegressive Intergrated Moving Average)

Un problema classico nell'identificazione è costituito dalla presenza di offset, disturbi a bassa frequenza, lente derive dei segnali. In questi casi è opportuno un filtraggio di tipo passa-alto sui segnali per ridurre gli effetti. Una possibilità per introdurre tale filtraggio è quella di considerare modelli in cui il rumore passa attraverso un opportuno sistema dalle caratteristiche note. Ad esempio, qualora si vogliono eliminare polarizzazioni nei segnali di ingresso uscita e limitandosi per semplicità a considerare il caso $C(z)=z^n$, si potrà considerare il modello seguente

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n}, \quad H(z) = \frac{z^n}{A(z) D(z)} = \frac{z^{n+1}}{(z^n + a_1 z^{n-1} + \dots + a_n)(z-1)}$$

Il corrispondente schema a blocchi è



che mostra che l'ipotesi fatta è che sul sistema agisca un rumore bianco integrato, cioè un moto Browniano.

Il modello da identificare può essere scritto nella forma

$$y(t) = \frac{B(z)}{A(z)}u(t) + \frac{z^{n+1}}{A(z)(z-1)}e(t)$$

o anche, definendo

$$y_d(t) = \frac{z-1}{z}y(t) = y(t) - y(t-1) \quad , \quad u_d(t) = \frac{z-1}{z}u(t) = u(t) - u(t-1)$$

come

$$y_d(t) = \frac{B(z)}{A(z)}u_d(t) + \frac{z^n}{A(z)}e(t)$$

che ha la struttura dei modelli *ARX* precedentemente considerati, ma in cui i segnali di ingresso e uscita sono differenziati.

Con considerazioni analoghe, pur di definire correttamente il blocco $\frac{z^v}{D(z)}$, è possibile filtrare i dati in modo da eliminare, ad esempio, componenti periodiche di pulsazione nota.

Questa tecnica può ovviamente essere estesa anche nel caso $C(z) \neq z^n$ o ad altre strutture di modelli.

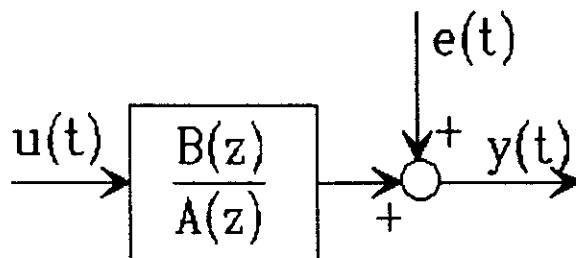
MODELLI OE (Output Error o a errore d'uscita)

Si è visto nel caso di modelli ARX che una buona capacità predittiva non corrisponde necessariamente ad un buon modello di simulazione. Nei modelli a errore d'equazione inoltre il polinomio $A(z)$ compare a denominatore sia di $G(z)$ che di $H(z)$, ipotesi a volte troppo restrittiva.

Allo scopo di rilassare tali ipotesi e di tener in maggior conto delle capacità di simulazione del modello identificato, si consideri la seguente scelta

$$G(z) = \frac{B(z)}{A(z)} = \frac{b_1 z^{n-1} + b_2 z^{n-2} + \dots + b_n}{z^n + a_1 z^{n-1} + \dots + a_n}, \quad H(z) = 1$$

a cui corrisponde lo schema a blocchi



Impiegando la formula consueta il predittore è

$$y^p(t/t-1) = \frac{B(z)}{A(z)} u(t)$$

non lineare nei parametri di $A(z)$ e $B(z)$. Ciò comporta l'uso di un metodo iterativo di minimizzazione. E' interessante osservare dalla relazione precedente che

$$y^p(t/t-1) = -a_1 y^p(t-1/t-2) - a_2 y^p(t-2/t-3) - \dots - a_n y^p(t-n/t-n-1) + b_1 u(t-1) + b_2 u(t-2) + \dots + b_n u(t-n)$$

il predittore non dipende dall'uscita effettiva y , ma dai suoi passati valori. Un **buon modello di predizione** è quindi da considerarsi anche un **buon simulatore del sistema reale**.

CONCLUSIONI

- La tecnica a minimizzazione dell'errore di predizione è usata per la stima di tutti i modelli considerati
- La stima dei modelli ARX, FIR è semplice (il predittore è lineare nei parametri e la soluzione è in forma chiusa). I modelli identificati sono certamente affidabili come predittori ad un passo, meno come simulatori del sistema
- Il predittore nei modelli ARMAX è non lineare nei parametri. Per semplificare il problema di stima spesso si ipotizza noto il polinomio $C(z)$ o si impiega un procedimento iterativo
- Per identificare modelli che siano affidabili come veri e propri simulatori del sistema è opportuno considerare modelli a errore d'uscita. Il predittore è però non lineare nei parametri

PREFILTRAGGIO E IDENTIFICAZIONE IN BANDA

- Con l'uso di opportuni filtri (pesi in frequenza) è possibile focalizzare la stima nell'intervallo di pulsazioni di interesse
- Cosa comporta a questo riguardo scegliere un modello a errore di predizione o a errore di uscita?

E' possibile effettuare un'analisi nel dominio della frequenza dei modelli identificati mediante l'approccio a minimizzazione dell'errore di predizione. Ciò consente di comprendere più approfonditamente le caratteristiche del metodo e di introdurre modifiche atte alla stima di modelli affidabili in una banda di frequenze a priori specificata. A questo riguardo si ricorda che nel controllo è importante avere una buona conoscenza della risposta in frequenza del sistema nell'intorno della pulsazione di taglio, mentre errori in bassa e/o alta frequenza possono essere tollerati purché la loro eventuale presenza sia considerata nella fase di sintesi del regolatore.

Per prima cosa si osservi che il funzionale da minimizzare nei metodi PEM può essere modificato nel modo seguente:

$$\Theta = \underset{\Theta}{\operatorname{argmin}} \sum_{t=n}^N (L(z)(y(t) - y^p(t/t-1)))^2 = \sum_{t=n}^N (L(z)\varepsilon(t))^2$$

dove $L(z)$ rappresenta la funzione di trasferimento di un sistema lineare e asintoticamente stabile, mentre gli altri simboli hanno il consueto significato. Per comprendere il significato di tale modifica, si ricorda che

$$y^p(t/t-1) = (1 - H(z)^{-1})y(t) + H(z)^{-1}G(z)u(t)$$

e quindi

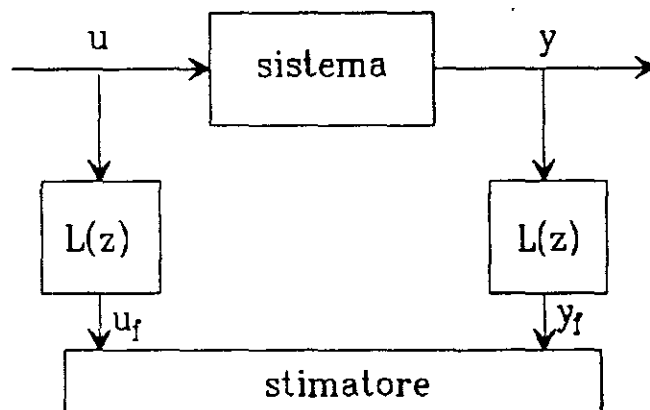
$$\varepsilon(t) = y(t) - y^p(t/t-1) = H(z)^{-1}y(t) - H(z)^{-1}G(z)u(t)$$

pertanto

$$L(z)\varepsilon(t) = L(z)H(z)^{-1}y(t) - L(z)H(z)^{-1}G(z)u(t)$$

il che equivale a modificare il modello del rumore da $H(z)e(t)$ a $H^*(z) = L(z)^{-1}H(z)e(t)$. Si osservi che questo accorgimento è già stato utilizzato nel caso dei modelli ARIMAX dove si è posto $L(z) = (z-1)/z$.

Da un punto di vista pratico, l'introduzione di $L(z)$ nel funzionale da minimizzare non pone alcun problema, in quanto è sufficiente impiegare nella fase di calcolo i segnali $y_f(t)=L(z)y(t)$ e $u_f(t)=L(z)u(t)$, come schematicamente rappresentato nella figura seguente:



Per comprendere intuitivamente il significato dell'uso di $L(z)$, si osservi che lo stimatore opera a partire dai segnali $y_f(t)$ e $u_f(t)$. Se questi segnali, per l'azione filtrante di $L(z)$, non hanno contribuito informativo a certe frequenze, il modello stimato cercherà di render conto dei dati di ingresso e uscita solamente nella banda di frequenze in cui l'ingresso e l'uscita hanno un contenuto armonico significativo. **Il filtro $L(z)$ deve quindi essere scelto di tipo passa-banda in modo da selezionare la banda di frequenze nella quale il modello identificato deve essere una buona rappresentazione del sistema reale.**

Questa osservazione motiva il fatto che nei modelli *ARIMAX* il differenziare i segnali di identificazione fa sì che venga eliminato l'effetto di polarizzazioni e/o disturbi a bassa frequenza.

Nel seguito viene data una giustificazione più formale di quanto detto. Si mostra inoltre come la scelta della struttura del modello (ARX, OE, ...) equivale a privilegiare la stima in particolari intervalli di pulsazioni.

***Analisi in frequenza della stima
a minimizzazione dell'errore di predizione***

Supponiamo che il sistema vero sia descritto da

$$y(t) = G_0(z)u(t) + H_0(z)e_0(t)$$

dove e_0 è un rumore bianco gaussiano con varianza σ^2 , H_0 ha numeratore stabile e con grado uguale a quello del denominatore. Si riscriva ora il sistema nella forma

$$y(t) = G_0(z)u(t) + v(t)$$

in cui il rumore $v(t)$ ha lo spettro

$$\Phi_v(\omega) = \sigma^2 |H_0(e^{j\omega})|^2$$

Sia poi $\Phi_u(\omega)$ lo spettro del segnale di ingresso.

Sotto deboli ipotesi si può affermare che la minimizzazione dell'errore di predizione equivale asintoticamente alla minimizzazione di

$$J = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{|G_0(e^{j\omega}) - G(e^{j\omega})|^2 \Phi_u(\omega) + \Phi_v(\omega)}{|H^*(e^{j\omega})|^2} d\omega$$

rispetto alle incognite $G(z)$ e $H^*(z) = L(z)^{-1}H(z)$. A partire da questa osservazione si possono dedurre le caratteristiche di vari casi particolari di interesse specifico e criteri di scelta per il filtro $L(z)$.

Dimostrazione

Si può dimostrare che risulta (Parseval)

$$\frac{1}{N} \sum_{t=1}^N \varepsilon(t)^2 \xrightarrow{N \rightarrow \infty} E[\varepsilon(t)^2] = \frac{1}{4\pi} \int_{-\pi}^{\pi} \Phi_{\varepsilon}(\omega) d\omega$$

dove $\Phi_{\varepsilon}(\omega)$ è lo spettro dell'errore di predizione ε . Ma poiché

$$\varepsilon(t) = H(z)^{-1} [y(t) - G(z)u(t)] = H(z)^{-1} [(G_0(z) - G(z))u(t) + v(t)]$$

risulta

$$\Phi_{\varepsilon}(\omega) = \frac{|G_0(e^{j\omega}) - G(e^{j\omega})|^2 \Phi_u(\omega) + \Phi_v(\omega)}{|H^*(e^{j\omega})|^2}$$

e il risultato segue.

Modello del rumore fissato

Supponendo che $H^*(z)$ sia fissato a priori (si ricordi quanto detto a proposito dei modelli *ARMAX* relativamente alla determinazione di $C(z)$), gli stimatori *PEM* portano a minimizzare asintoticamente il funzionale

$$J = \frac{1}{4\pi} \int_{-\pi}^{\pi} |G_0(e^{j\omega}) - G(e^{j\omega})|^2 Q(\omega) d\omega \quad , \quad Q(\omega) = \frac{\Phi_u(\omega)}{|H^*(e^{j\omega})|^2}$$

cioè il modulo al quadrato dell'errore sulla funzione di trasferimento tra ingresso e uscita pesato in frequenza da $Q(\omega)$. Questo peso può essere interpretato come una sorta di rapporto segnale/rumore alle diverse frequenze.

In particolare nei *modelli a errore d'uscita* $H(z)=1$ e, se $L(z)=1$, il termine $|G_0(e^{j\omega}) - G(e^{j\omega})|^2$ viene pesato in frequenza soltanto dallo spettro del segnale di ingresso.

Modelli ARX

Ricordando che in questo caso $G(z)=B(z)/A(z)$ e $H(z)=z^n/A(z)$, asintoticamente viene minimizzato

$$J = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left[|G_0(e^{j\omega}) - G(e^{j\omega})|^2 \Phi_u(\omega) + \Phi_v(\omega) \right] |A(e^{j\omega})|^2 d\omega$$

Poiché $A(z)/z^n$ è nella maggior parte dei casi la funzione di trasferimento di un filtro passa-alto, in assenza di rumore ($\Phi_v(\omega)=0$) la minimizzazione porta a attribuire un peso maggiore alle alte frequenze.

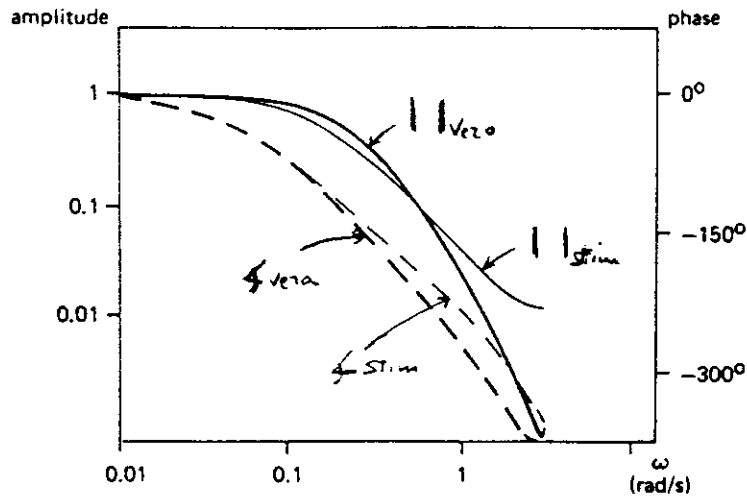
Un esempio

Si è considerato un sistema non affetto da disturbi e descritto da

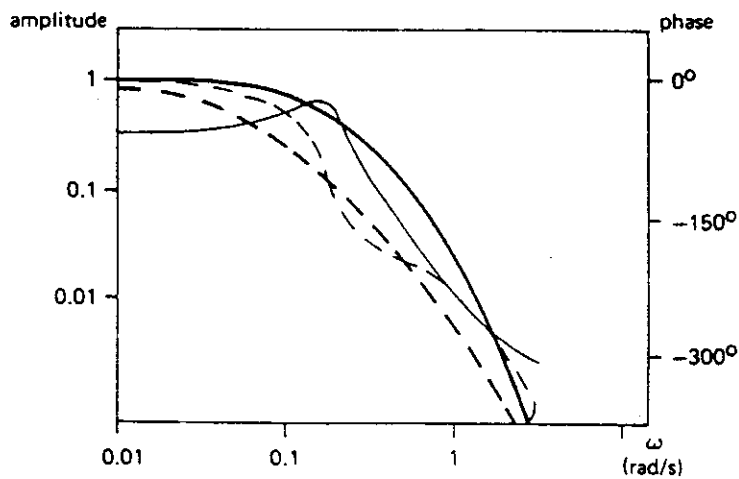
$$G_0(z) = \frac{0.001(10z^3 + 7.4z^2 + 0.924z + 0.1764)}{z(z^4 - 2.14z^3 + 1.553z^2 - 0.4387z + 0.042)}$$

come ingresso si è utilizzato un segnale caratterizzato da $\Phi_u(\omega) \cong 1$ a tutte le frequenze. Inoltre si è sempre impiegata nel seguito una funzione di trasferimento $G(z)$ del II ordine, chiaramente si è in presenza di una sottoparametrizzazione e il modello comunque identificato non potrà mai essere uguale al sistema vero.

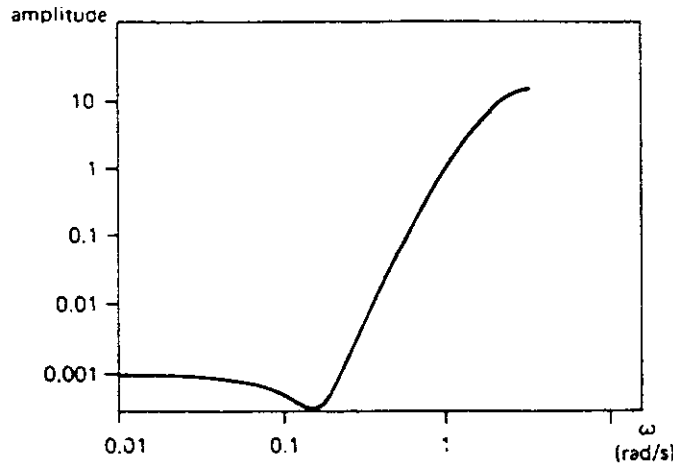
Si è quindi inizialmente usato nella stima un modello ad errore d'uscita ($H(z)=1$) con $L(z)=1$. I diagrammi di Bode del modello stimato sono confrontati con quelli del sistema vero nella figura seguente



Successivamente è stato considerato in fase di stima un modello *ARX* che ha portato ai risultati seguenti:



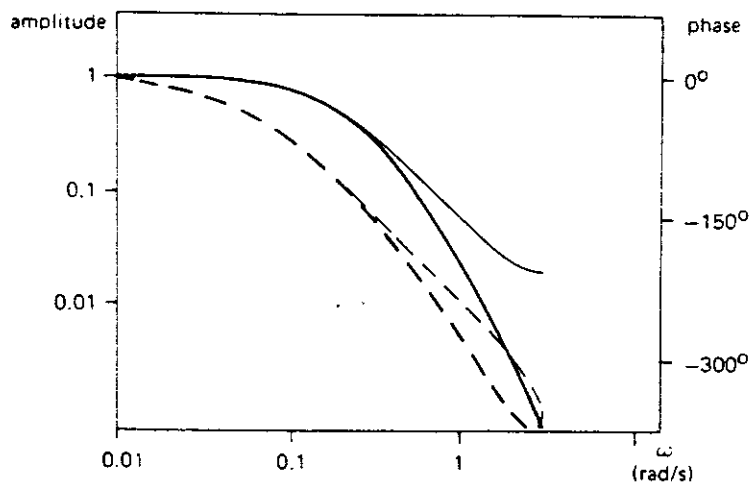
E' evidente che si ha una stima nettamente peggiore, soprattutto a bassa frequenza (dove normalmente è bene che il modello sia preciso). Ciò è dovuto al peso in frequenza del termine $A(z)/z^n$, il cui diagramma di Bode dei moduli e relativo alle stime ottenute è riportato di seguito.



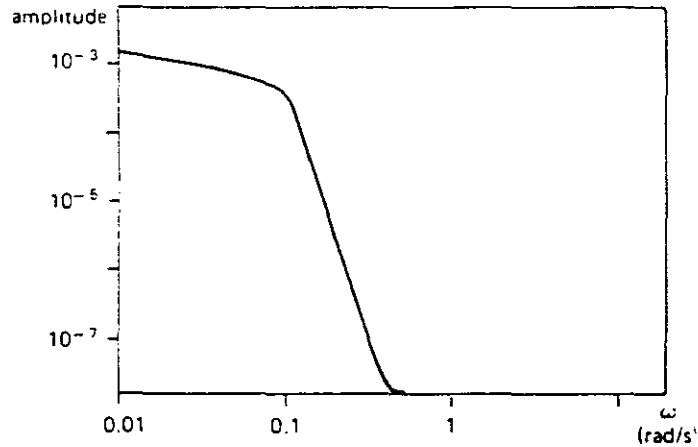
Per ridurre gli effetti precedenti dovuti all'uso di un modello *ARX* ad errore d'equazione (che, ricordiamolo, ha il vantaggio di portare ad una formula esplicita per la stima dei parametri), è possibile ricorrere ad una struttura di tipo *ARMAX* in cui il polinomio $C(z)$ non viene stimato, ma è fisso e pari ad una stima a priori $A^*(z)$ del polinomio $A(z)$. In modo equivalente si può interpretare questo procedimento come l'impiego di un prefiltro $L(z) = z^n / A^*(z)$.

Si noti che in generale una qualche stima a priori della posizione dei poli del sistema è disponibile o facilmente ricavabile con semplici prove d'impianto, ad esempio scalini d'ingresso.

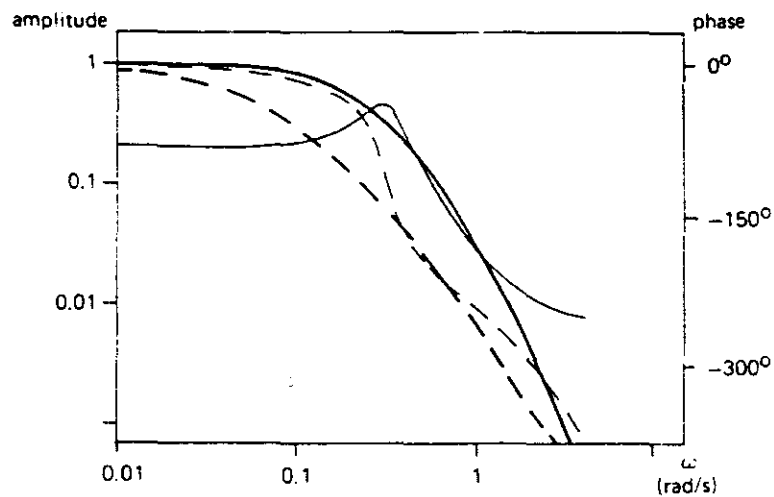
In assenza di una stima a priori di $A(z)$, si può utilizzare un filtro $L(z)$ che forzi la stima nella banda di interesse. Nell'esempio precedente in cui i segnali di ingresso e uscita sono filtrati con $L(z)$ costruito come un filtro di Butterworth del V ordine e pulsazione di taglio 0.1 rad/sec , impiegando un modello *ARX* si ottengono i risultati seguenti



Questi risultati mostrano chiaramente come il modello identificato sia estremamente affidabile sino alla pulsazione di taglio di $L(z)$. Nella figura seguente è inoltre riportato il peso in frequenza corrispondente



A ulteriore conferma di quanto visto e sempre con riferimento all'esempio trattato, se si impiega un modello ad errore d'uscita e un filtro del V ordine di tipo passa-alto con pulsazione di taglio 0.5 rad/sec la stima che si ottiene è affidabile a frequenze medio/alte (si ricorda che senza l'uso di $L(z)$ si ottiene per il modello ad errore d'uscita una buona stima in bassa frequenza).



CONCLUSIONI

- La scelta di una particolare famiglia di modelli implica l'impiego di un diverso peso in frequenza sull'errore di stima. Nella stima di un modello ARX normalmente si privilegia la stima ad alta frequenza (per avere una buona predizione ad un passo). Nell'identificazione dei modelli ad errore d'uscita non si usa alcun peso in frequenza
- Per migliorare la qualità del modello identificato in un particolare intervallo di frequenze è sufficiente prefiltrare i dati di ingresso e uscita con un opportuno filtro $L(z)$. Questo non è computazionalmente oneroso e può cambiare sostanzialmente la qualità del risultato

SCELTA DELL'ESPERIMENTO

- Quando è possibile, si deve scegliere il segnale di ingresso al sistema in modo da ottenere dati per l'identificazione con un adeguato contributo informativo. Il segnale di ingresso deve cioè essere "persistentemente eccitante" di ordine opportuno
- Il segnale di ingresso migliore è il rumore bianco
- In alternativa vengono spesso utilizzate le PRBS, segnali pseudo-casuali di facile generazione

Scelta dell'esperimento

In alcuni casi non è possibile scegliere l'esperimento di identificazione, mentre in altri la scelta di $u(t)$ può essere effettuata liberamente. A questo riguardo si è già visto come lo spettro del segnale di ingresso $\Phi_u(\omega)$ influenzi la stima asintotica ottenuta con i metodi a minimizzazione dell'errore di predizione. Preliminare a ciò vi è tuttavia il problema di **utilizzare un segnale di ingresso che consenta l'identificazione del modello.**

Per approfondire questo aspetto si consideri il problema di identificare un modello *ARX*

$$y(t) = -a_1 y(t-1) - a_2 y(t-2) - \dots - a_n y(t-n) - b_1 u(t-1) + b_2 u(t-2) + \dots + b_n u(t-n) + e(t)$$

per il quale esiste la formula esplicita di soluzione

$$\Theta = S(N)^{-1} \sum_{t=n}^N \phi(t) y(t)$$

dove si è ipotizzato che la matrice $S(N) = \sum_{t=n}^N \phi(t) \phi(t)'$ sia non singolare. In questo caso il problema principale è quindi scegliere $u(t)$ in modo che questa ipotesi sia verificata.

Un esempio

Si consideri il semplice sistema ARX

$$y(t) = -a_1 y(t-1) + b_1 u(t-1) + e(t)$$

per cui si possono definire i vettori

$$\Theta = \begin{bmatrix} -a_1 \\ b_1 \end{bmatrix}, \quad \phi(t) = \begin{bmatrix} y(t-1) \\ u(t-1) \end{bmatrix}$$

e quindi

$$\Phi(t)\Phi(t)' = \begin{bmatrix} y(t-1)^2 & y(t-1)u(t-1) \\ u(t-1)y(t-1) & u(t-1)^2 \end{bmatrix}$$

Ne segue che la matrice

$$R(N) = \frac{1}{N} \begin{bmatrix} \sum_{t=1}^N y(t-1)^2 & \sum_{t=1}^N y(t-1)u(t-1) \\ \sum_{t=1}^N u(t-1)y(t-1) & \sum_{t=1}^N u(t-1)^2 \end{bmatrix}$$

rappresenta una media campionaria su N dati e, se u e y sono processi stazionari, sotto deboli ipotesi e per $N \rightarrow \infty$, converge a

$$\bar{R} = \begin{bmatrix} \bar{R}_{yy}(0) & \bar{R}_{yu}(0) \\ \bar{R}_{uy}(0) & \bar{R}_{uu}(0) \end{bmatrix}$$

dove $\bar{R}_{yy}(\tau)$, $\bar{R}_{uu}(\tau)$ rappresentano le funzioni di correlazione dei processi y , u .

□

Scelta dell'esperimento

Più in generale $R(n) = S(N)/N$ tenderà per $N \rightarrow \infty$ alla matrice

$$\bar{R} = \begin{vmatrix} \bar{R}_{yy} & \bar{R}_{yu} \\ \bar{R}_{uy} & \bar{R}_{uu} \end{vmatrix}$$

dove

$$\bar{R}_{yy} = \begin{vmatrix} \bar{R}_{yy}(0) & \bar{R}_{yy}(1) & \dots & \bar{R}_{yy}(n-1) \\ \bar{R}_{yy}(1) & \bar{R}_{yy}(0) & \dots & \bar{R}_{yy}(n-2) \\ \dots & \dots & \dots & \dots \\ \bar{R}_{yy}(n-1) & \bar{R}_{yy}(n-2) & \dots & \bar{R}_{yy}(0) \end{vmatrix}$$

$$\bar{R}_{uu} = \begin{vmatrix} \bar{R}_{uu}(0) & \bar{R}_{uu}(1) & \dots & \bar{R}_{uu}(n-1) \\ \bar{R}_{uu}(1) & \bar{R}_{uu}(0) & \dots & \bar{R}_{uu}(n-2) \\ \dots & \dots & \dots & \dots \\ \bar{R}_{uu}(n-1) & \bar{R}_{uu}(n-2) & \dots & \bar{R}_{uu}(0) \end{vmatrix}$$

$$\bar{R}_{yu} = \bar{R}_{uy} = \begin{vmatrix} \bar{R}_{yu}(0) & \bar{R}_{yu}(1) & \dots & \bar{R}_{yu}(n-1) \\ \bar{R}_{yu}(1) & \bar{R}_{yu}(0) & \dots & \bar{R}_{yu}(n-2) \\ \dots & \dots & \dots & \dots \\ \bar{R}_{yu}(n-1) & \bar{R}_{yu}(n-2) & \dots & \bar{R}_{yu}(0) \end{vmatrix}$$

La condizione necessaria e sufficiente per l'unicità della stima è dunque l'invertibilità della matrice \bar{R} , almeno per un elevato numero di dati.

A sua volta questa condizione implica che *condizione necessaria* (soltanto!) sia la non singolarità di \bar{R}_{uu} .

Definizione

Il segnale u è persistentemente eccitante di ordine n se \bar{R}_{uu} è invertibile (definita positiva).

Impulso - $u(t)=imp(t)$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N u(t)u(t+\tau) = 0, \quad \forall \tau$$

cioè $\bar{R}_{uu}=0$, l'impulso quindi non è un segnale persistentemente eccitante di alcun ordine. La giustificazione di ciò è dettata dal fatto che la risposta all'impulso di un sistema asintoticamente stabile tende a zero e quindi per $N \rightarrow \infty$ l'uscita coincide con il rumore che la corrompe.

Scalino - $u(t)=sca(t)$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N u(t)u(t+\tau) = 1, \quad \forall \tau$$

cioè $\bar{R}_{uu} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{bmatrix}$, lo scalino è quindi un segnale persistentemente

eccitante di ordine 1 e consente di stimare in maniera adeguata un solo parametro (per esempio il guadagno del sistema).

Rumore bianco - $u(t) \sim WN(0, \sigma^2)$

In questo caso $\bar{R}_{uu} = \sigma^2 I$, non singolare per ogni n . Per l'identificazione evidentemente il rumore bianco è il segnale migliore!

Commenti

- come mostrato nell'analisi in frequenza, può essere conveniente eccitare il sistema con un segnale che sia non solo persistentemente eccitante di ordine opportuno, ma il cui spettro abbia componenti significative nella banda di frequenze di interesse.
- La condizione $\bar{R}_{uu} > 0$ è soltanto necessaria. Non è sufficiente quando il modello da identificare è di ordine superiore al sistema di generazione dei dati. Ad esempio, si supponga che il sistema da identificare sia un *ARX* del primo ordine e che si voglia identificare dai dati un modello del secondo ordine. E' chiaro che vi sono infiniti modelli che possono essere identificati, e precisamente quelli ottenibili moltiplicando il numeratore e il denominatore della vera funzione di trasferimento $G_0(z)$ per lo stesso polinomio del primo ordine. I problemi connessi con la mancata verifica della condizione $\bar{R}_{uu} > 0$ sono chiamati problemi di *identificabilità sperimentale*, mentre questi ultimi sono chiamati problemi di *identificabilità strutturale*.

Proprietà dei segnali persistentemente eccitanti

Proprietà 1

Sia $u(t)$ un segnale scalare persistentemente eccitante di ordine n . Allora la sua densità spettrale $\Phi_u(\omega)$ è diversa da zero in almeno n frequenze.

Proprietà 2

L'uscita di un sistema con funzione di trasferimento $G(z)$ e con ingresso un rumore bianco è persistentemente eccitante di ogni ordine.

Proprietà 3

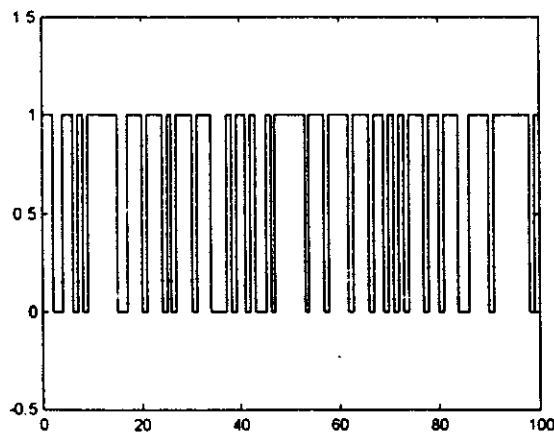
Si consideri il segnale $y(t) = H(z)u(t)$, con $u(t)$ persistentemente eccitante di ordine n e $H(z)$ funzione di trasferimento di un sistema asintoticamente stabile con k zeri sulla circonferenza di raggio unitario centrata nell'origine del piano complesso. E' possibile mostrare che $y(t)$ è persistentemente eccitante di ordine m con $n-k \leq m \leq n$.

Proprietà 4

Un segnale periodico di periodo M può essere al più persistentemente eccitante di ordine M (attenzione alle onde quadre o segnali simili).

Pseudo Random Binary Signals (PRBS)

E' un segnale che commuta tra due livelli prefissati. Pur essendo periodico, se opportunamente generato è in grado di approssimare con sufficiente accuratezza le caratteristiche del rumore bianco (tuttavia per la Proprietà 4 precedente non può essere persistentemente eccitante di ogni ordine). Ha il vantaggio, a differenza del rumore bianco, di avere un'ampiezza controllata (e in ogni caso il rumore bianco non è semplice da generare).



Un (ovvio) commento Nel caso di sistemi lineari, come tutti quelli relativi alle famiglie di modelli considerate, l'ampiezza del segnale di ingresso non è rilevante, mentre se si volessero identificare i parametri di un modello non lineare (come nel seguito discusso), ciò non è più vero.

Generazione di PRBS

Si definisca dapprima la somma modulo due secondo la seguente tabella

a	b	$a \oplus b$
0	0	0
1	0	1
0	1	1
1	1	0

Un segnale $y(t)$ di tipo PRBS che commuti tra i valori a e b può quindi essere generato secondo le seguenti formule

$$x_1(t+1) = a_1 x_1(t) \oplus a_2 x_2(t) \oplus \dots \oplus a_n x_n(t)$$

$$x_2(t+1) = x_1(t)$$

$$x_3(t+1) = x_2(t)$$

...

$$x_n(t+1) = x_{n-1}(t)$$

$$u(t) = x_n(t)$$

$$y(t) = a + (b-a)u(t)$$

dove i coefficienti a_i , $i=1, \dots, n$, possono assumere valori 0 o 1 e ovviamente le condizioni iniziali $x_i(0)$ non possono essere tutte nulle. In funzione dei valori scelti per questi coefficienti si determinano sequenze periodiche di periodo massimo $M=2^n-1$. Per avere il massimo ordine di persistente eccitazione è necessario, fissato n , scegliere in modo opportuno i valori degli a_i . PRBS di periodo $M=2^n-1$ vengono chiamate *di lunghezza massima*. Per la scelta degli a_i che producono PRBS di lunghezza massima esistono condizioni necessarie e sufficienti per le quali si rimanda ai testi specializzati (in particolare: Davies W.D.T. (1970). *System Identification for Self-Adaptive Control*, Wiley Interscience, London).

Analogia tra PRBS e rumore bianco

E' infine possibile mostrare che il segnale $y(t)$, PRBS di lunghezza massima e periodo $M=2^n-1$ con $b=1$ e $a=-1$, è caratterizzato da

$$m_y = \frac{1}{M} \sum_{t=1}^M y(t) = \frac{1}{M} \xrightarrow{M \rightarrow \infty} 0$$

e

$$r_y(\tau) = \frac{1}{M} \sum_{t=1}^M [(y(t+\tau) - m_y)(y(t) - m_y)] = \begin{cases} 1 - \frac{1}{M^2} \xrightarrow{M \rightarrow \infty} 1 & \text{per } \tau = 0 \\ -\frac{1}{M} - \frac{1}{M^2} \xrightarrow{M \rightarrow \infty} 0 & \text{per } \tau \neq 0 \end{cases}$$

per M sufficientemente elevato quindi la covarianza di y è simile a quella del rumore bianco di varianza unitaria. I due segnali sono quindi simili per quanto riguarda le loro proprietà spettrali. Inoltre i segnali PRBS opportunamente modificati, ad esempio filtrandoli o aumentando il periodo minimo di commutazione da un campione a diversi campioni, possono produrre segnali con proprietà spettrali specificate.

Per queste caratteristiche, e per altre qui non discusse, questi sono tra i segnali più utilizzati nell'identificazione.

CONCLUSIONI

- Se i dati di ingresso e uscita non sono adeguati, non vi è speranza di identificare correttamente il modello
- Poiché spesso non è possibile utilizzare un rumore bianco, vale la pena di impiegare le PRBS
- E' necessario fare attenzione all'uso di segnali periodici