

CAPITOLO 3

Algoritmi di identificazione.

3.1. Introduzione.

In questo capitolo ci proponiamo di presentare e di giustificare teoricamente un algoritmo che consente di identificare un sistema Hammerstein-Wiener a partire dalle misurazioni degli ingressi e delle uscite del sistema. L'obiettivo sarà quello di identificare il sistema "a ritroso" a partire dalla conoscenza del segnale di uscita. Partendo da un sistema NLN in tempo continuo, considereremo il suo equivalente in tempo discreto; vedremo come sia possibile, sfruttando una frequenza di campionamento multipla rispetto a quella dell'ingresso, annullare la dipendenza delle equazioni costitutive del sistema dall'ingresso stesso, in modo da avere la possibilità di identificare il sistema a partire dalla sola uscita.

La trattazione che segue è dovuta essenzialmente a E.W. Bai, e per ogni approfondimento si rimanda a [2].

Infine presenteremo una variante di questo algoritmo, che usa la tecnica delle funzioni modulanti per identificare la parte Hammerstein del sistema.

3.2. Considerazioni preliminari.

Consideriamo un sistema non lineare di tipo Hammerstein-Wiener in tempo continuo. Immaginiamo di fornire ingressi campionati mediante uno *zero order hold* e di campionare le uscite del sistema; in questo caso si ottiene il sistema di Figura 3.1:

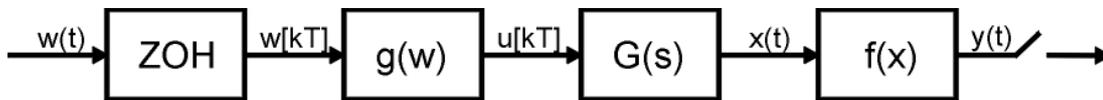


Figura 3.1. Sistema Hammerstein-Wiener campionato

Supponiamo che la rappresentazione nello spazio degli stati della parte lineare $G(s)$ sia la seguente:

$$\begin{cases} \dot{\eta}(t) = A \cdot \eta(t) + b \cdot u(t), & u, x \in \mathbb{R}, \quad A \in \mathbb{R}^{n \times n} \\ x(t) = c \cdot \eta(t) \end{cases} \quad (3.1)$$

Per un dato periodo di campionamento T , la corrispondente equazione nello spazio degli stati in tempo discreto è:

$$\begin{aligned} \Phi &\triangleq e^{AT} \\ \Gamma &\triangleq \int_0^T e^{A(T-t)} \cdot b \cdot dt \end{aligned} \quad (3.2)$$

$$\begin{cases} \eta[(k+1)T] = \Phi \cdot \eta[kT] + \Gamma \cdot u[kT] \\ x[kT] = c \cdot \eta[kT] \end{cases}$$

e la funzione di trasferimento associata è:

$$G(z) = c \cdot (zI - \Phi)^{-1} \cdot \Gamma = \frac{\beta(z)}{\alpha(z)} = \frac{\beta_1 \cdot z^{-1} + \beta_2 \cdot z^{-2} + \dots + \beta_n \cdot z^{-n}}{1 - \alpha_1 \cdot z^{-1} - \alpha_2 \cdot z^{-2} - \dots - \alpha_n \cdot z^{-n}} \quad (3.3)$$

dove gli α_i ed i β_j sono opportuni coefficienti.

Per quanto riguarda la non linearità di ingresso, supporremo che sia statica ma non faremo alcuna assunzione riguardo alla sua struttura.

$$u = g(w) \quad (3.4)$$

Considereremo una non linearità di uscita statica:

$$y = f(x). \quad (3.5)$$

E' molto importante che la funzione $f(x)$ sia invertibile, dato che cercheremo di ricavare il valore del segnale $x(t)$ a partire da quello di $y(t)$. Supporremo inoltre che la $f(x)$ passi per l'origine e che la sua inversa $f^{-1}(y)$ ammetta una rappresentazione polinomiale:

$$x = f^{-1}(y) \cong \sum_{i=1}^m r_i \cdot y^i \quad (3.6)$$

Nella procedura di identificazione stimeremo prima l'inversa della non linearità di uscita, ottenendo così una stima dell'uscita della parte lineare a partire dall'uscita del sistema. Cercheremo poi di stimare i coefficienti polinomiali della funzione diretta che meglio approssimano i dati ottenuti, nel senso dei minimi quadrati. Ovviamente occorre ammettere che anche la $f(x)$ abbia una rappresentazione polinomiale:

$$y = f(x) = \sum_{i=1}^q a_i \cdot x^i \quad (3.7)$$

Si noti che se $f(x)$ è continua, oltre che invertibile, anche $f^{-1}(y)$ è continua; inoltre, se il sistema è stabile e l'ingresso è limitato, anche x e y sono limitati, e quindi l'inversa della non linearità di uscita può essere ben approssimata da un polinomio di grado sufficientemente elevato; questo giustifica teoricamente le assunzioni fatte in precedenza. D'altro canto usare una approssimazione polinomiale di grado molto alto conduce ad un aumento della sensibilità al rumore ed all'incertezza di modello, nonché a problemi numerici; questo limita in una certa misura l'approccio qui seguito.

L'idea di partenza è quella di identificare il sistema a partire dalle sole misurazioni di w e di y , campionando l'uscita ad una frequenza superiore rispetto all'ingresso; questo significa che, dato un certo periodo di campionamento per l'ingresso T , l'uscita sarà campionata con un periodo più breve h :

$$h = \frac{T}{l}, \quad l \geq 1 \quad (3.8)$$

dove l è un intero positivo detto *rapporto di sovracampionamento*.

Un'ipotesi che è necessario fare è che il sistema sia minimo, ossia che la $G(z)$ non abbia cancellazioni polo/zero, il che renderebbe la parametrizzazione non univoca e comprometterebbe l'identificabilità del sistema. In questo senso è importante il seguente lemma:

Lemma 3.1: si consideri il sistema tempo continuo di (3.1) ed il relativo sistema campionato con periodo di campionamento $h=T/l$, con l intero positivo. Il sistema è minimo per ogni l intero positivo se e solo se il sistema è minimo per il periodo di campionamento T .

(per la dimostrazione si veda [3]). Pertanto dobbiamo assumere che il sistema sia minimo quando il periodo di campionamento è T .

Sempre a proposito dell'univocità della parametrizzazione del sistema, dobbiamo ricordare che, per i sistemi Hammerstein-Wiener così come per tutti i sistemi a blocchi, questa non è affatto garantita senza opportune assunzioni. Si supponga, ad esempio, di avere un sistema Hammerstein-Wiener rappresentabile mediante tre blocchi $u=g(w)$, $G(z)$ e $y=f(x)$. Data qualunque terna di numeri a , b , c , il cui prodotto sia 1, il sistema costituito dai blocchi $u=a g(w)$, $b G(z)$, $y=c f(x)$ sarebbe del tutto identico al sistema sopra riportato dal punto di vista della relazione tra ingresso e uscita. Per ottenere una parametrizzazione unica è necessario fare delle assunzioni che permettano di normalizzare due dei tre blocchi. In particolare noi normalizzeremo la parte lineare e la non linearità di uscita con le seguenti posizioni:

$$\begin{cases} \beta_1 = \mathbf{1} \\ r_1 = \mathbf{1} \end{cases} \quad (3.9)$$

.

3.3. Identificazione della non linearità di uscita.

Come si è detto le uniche grandezze del sistema che sono misurabili sono il suo ingresso e la sua uscita. Nel processo di identificazione che segue il primo passo è cercare di ricavare l'inversa della non linearità di uscita a partire dalle sole misure dell'uscita y ; dopodiché si identifica il resto del sistema "a ritroso": prima la parte lineare, poi la non linearità di ingresso.

Sia T sia il periodo di campionamento per l'ingresso ed n l'ordine della $G(z)$, imponiamo che il periodo di campionamento per l'uscita sia:

$$h = \frac{T}{n+1} \quad (3.10)$$

Sebbene sia possibile usare qualunque periodo di campionamento tale che:

$$h = \frac{T}{\bar{n}+1}, \quad \forall \bar{n} \geq n; \quad (3.11)$$

tuttavia, per semplicità, useremo la posizione (3.10).

La funzione di trasferimento per il sistema campionato con periodo h sia:

$$\tilde{G}(z) = \frac{\tilde{\beta}(z)}{\tilde{\alpha}(z)} = \frac{\tilde{\beta}_1 \cdot z^{-1} + \tilde{\beta}_2 \cdot z^{-2} + \dots + \tilde{\beta}_n \cdot z^{-n}}{1 - \tilde{\alpha}_1 \cdot z^{-1} - \tilde{\alpha}_2 \cdot z^{-2} - \dots - \tilde{\alpha}_n \cdot z^{-n}} \quad (3.12)$$

Questa equazione, portata nel dominio del tempo, diventa:

$$x[kh] = \sum_{i=1}^n \tilde{\alpha}_i \cdot x[kh - ih] + \sum_{i=1}^n \tilde{\beta}_i \cdot u[kh - ih] \quad (3.13)$$

Sostituendo alla $x[kh]$ la sua approssimazione polinomiale (3.6) si ottiene:

$$\sum_{j=1}^m r_j \cdot y^j[kh] = \sum_{i=1}^n \tilde{\alpha}_i \cdot \sum_{j=1}^m r_j \cdot y^j[kh - ih] + \sum_{i=1}^n \tilde{\beta}_i \cdot u[kh - ih] + v_1[kh] \quad (3.14)$$

dove si è indicato con v_1 un termine che tiene conto dell'errore complessivo dovuto ad approssimazioni, rumore, ecc.

Ricordando che si è deciso di porre $r_1=1$, si può riscrivere l'equazione (3.14) in forma matriciale, ottenendo la seguente:

$$\phi_1[kh] \triangleq [-y^2[kh], \dots, -y^m[kh], y[(k-1)h], \dots, y[(k-n)h], \\ y^2[(k-1)h], \dots, y^2[(k-n)h], \dots, y^m[(k-1)h], \dots, y^m[(k-n)h]]$$

$$\theta_1 \triangleq \begin{bmatrix} r_2 \\ \vdots \\ r_m \\ \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_n \\ \tilde{\alpha}_1 \cdot r_2 \\ \vdots \\ \tilde{\alpha}_n \cdot r_2 \\ \vdots \\ \tilde{\alpha}_1 \cdot r_m \\ \vdots \\ \tilde{\alpha}_n \cdot r_m \end{bmatrix} \quad (3.15)$$

$$y[kh] = \phi_1[kh] \cdot \theta_1 + \sum_{i=1}^n \tilde{\beta}_i \cdot u[kh - ih] + v_i[kh]$$

Nell'equazione (3.15) il termine ϕ_1 contiene solo elementi che dipendono dalle $y[kh]$ o dalle loro potenze, quindi si può ricavare direttamente dalla misurazione delle

uscite del sistema; il termine θ_1 contiene i parametri che caratterizzano l'inversa della non linearità di uscita ed il denominatore della funzione di trasferimento per $T=h$, che vogliamo trovare, ed i loro prodotti. Tutto il resto, invece, non è noto.

Consideriamo adesso due equazioni consecutive, ottenute per $k=l(n+1)$ e per $k=l(n+1)-1$.

$$y[l(n+1)h] = \phi_1[l(n+1)h] \cdot \theta_1 + \sum_{i=1}^n \tilde{\beta}_i \cdot u[l(n+1)h - ih] + v_1[l(n+1)h] \quad (3.16)$$

$$y[l(n+1)h - h] = \phi_1[l(n+1)h - h] \cdot \theta_1 + \sum_{i=1}^n \tilde{\beta}_i \cdot u[l(n+1)h - ih - h] + v_1[l(n+1)h - h] \quad (3.17)$$

Avendo deciso di utilizzare un periodo di campionamento per l'uscita h , tale che il periodo di campionamento per l'ingresso sia $T = (n+1)h$, l'ingresso del sistema rimane costante per tutti gli intervalli di tempo $[kT, kT+nh]$, $k = 0, 1, 2, \dots$. In particolare si ha che:

$$w[(l-1)T] = w[(l-1)T + h] = \dots = w[(l-1)T + nh], \quad l = 1, 2, \dots \quad (3.18)$$

Visto che la non linearità in ingresso $g(w)$ è statica, si può anche dire che:

$$u[(l-1)T] = u[(l-1)T + h] = \dots = u[(l-1)T + nh], \quad l = 1, 2, \dots \quad (3.19)$$

Considerando ancora che $T = (n+1)h$ si ottiene:

$$u[(l-1)T] = u[l(n+1)h - h - nh] = \dots = u[l(n+1)h - h] = u[(l-1)T + nh] \quad (3.20)$$

Se si definiscono le seguenti:

$$\begin{aligned}
\Delta y[l] &\triangleq y[lT] - y[lT - h] \\
\Delta v[l] &\triangleq v_1[lT] - v_1[lT - h] \\
\Delta \phi_1[l] &\triangleq \phi_1[lT] - \phi_1[lT - h]
\end{aligned}
\tag{3.21}$$

si può dire che:

$$\Delta y[l] = \Delta \phi_1[l] \cdot \theta_1 + \Delta v[l] \tag{3.22}$$

Infatti, nel termine $\Delta y[l]$, la sommatoria contenente termini relativi al segnale u per il tempo lT e quella per il tempo $lT-h$ sono uguali e si annullano tra loro. In effetti noi consideriamo da un lato il valore della y nel momento in cui l'ingresso viene ricampionato, dall'altro nell'istante di campionamento dell'uscita immediatamente precedente. Tenendo conto del fatto che, nella suddetta sommatoria, compaiono valori della u negli n periodi di campionamento dell'uscita precedenti all'istante per il quale si vuole calcolare la y , e considerando che la w (e di conseguenza la u) resta costante per $n+1$ periodi di campionamento dell'uscita, si capisce perché i termini dipendenti dalla u scompaiano. La Figura 3.2 chiarisce meglio questo ragionamento.

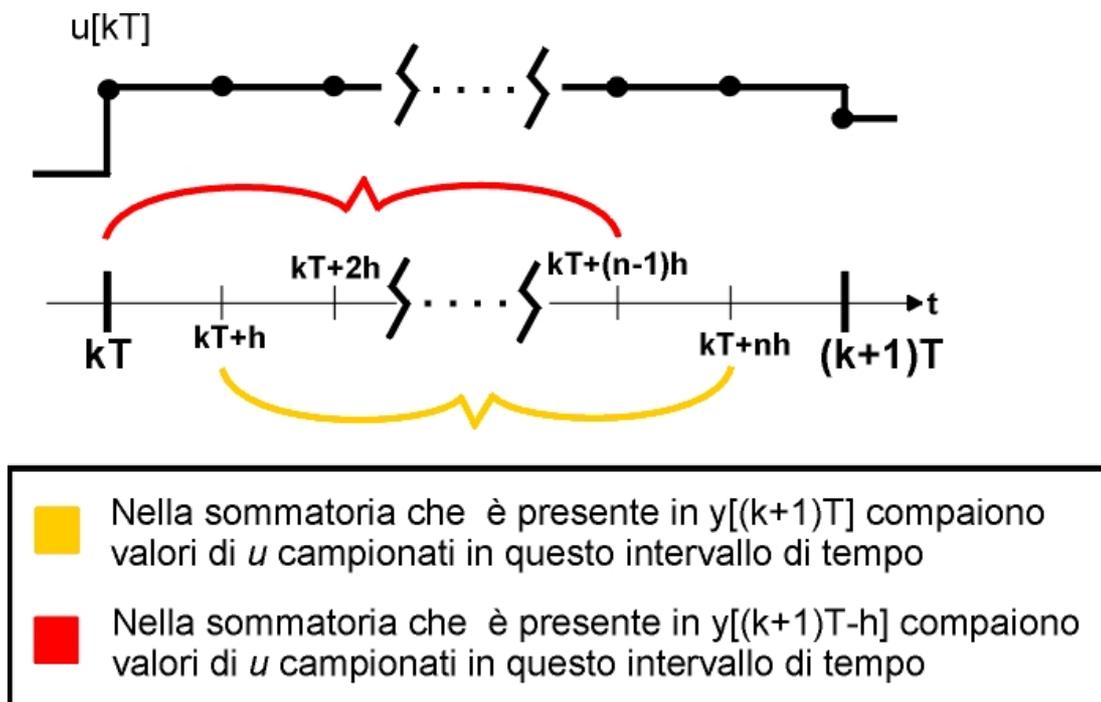


Figura 3.2

L'equazione (3.22) è algebrica lineare nella θ_1 , ovvero nei parametri r_i e $\tilde{\alpha}_i$, e dipende solo dall'uscita y , che è misurabile. Replicando l'equazione (3.22) per diversi l , si ottiene un sistema di equazioni lineari, riportato nella (3.23), in cui sono stati chiamati $\Delta\phi_1$ e Δy , rispettivamente, la matrice del sistema ed il vettore dei termini noti

$$\Delta y = \Delta\phi_1 \cdot \theta_1 \quad (3.23)$$

In particolare, se si misura il segnale di uscita per un periodo sufficientemente lungo, si ottiene un sistema sovradeterminato, il che è necessario per cercare di ridurre l'effetto del rumore. Il vettore θ_1 può dunque essere calcolato con qualunque metodo per la risoluzione dei sistemi lineari; nel seguito useremo sempre il metodo dei

minimi quadrati, che è il metodo standard per la soluzione di sistemi lineari sovradeterminati in ambiente MATLAB.

Una volta ottenuto θ_1 si hanno a disposizione le stime degli $\tilde{\alpha}_i$ e degli r_i . In particolare $\hat{r}_1 = \mathbf{1}$, $\hat{r}_2, \dots, \hat{r}_m$ sono le prime $m-1$ componenti di θ_1 e $\hat{\alpha}_1, \dots, \hat{\alpha}_n$ sono le successive n componenti di θ_1 . A questo punto conosciamo una stima del denominatore di $\tilde{G}(z)$ ed una stima dell'inversa della non linearità sull'uscita:

$$\hat{f}^{-1}(y) = \sum_{i=1}^m r_i \cdot y^i[kh]. \quad (3.24)$$

La $f(x)$ può essere facilmente stimata a partire dalla sua inversa. Infatti dalle misure note della $y[kh]$ si ricavano le stime di $x[kh]$, $\hat{x}[kh]$. Se $f(x)$ è polinomiale, e quindi:

$$y = f(x) = \sum_{i=1}^q a_i \cdot x^i, \quad (3.25)$$

le stime dei coefficienti a_i possono essere trovate cercando il polinomio che meglio approssima le coppie $(\hat{x}[kh], y[kh])$ nel senso dei minimi quadrati, ovvero:

$$\hat{a} = \arg \left\{ \min \sum_k \left(y[kh] - \sum_{i=1}^q \hat{a}_i \cdot \hat{x}^i[kh] \right)^2 \right\} \quad (3.26)$$

Si può notare che, in realtà, il vettore θ_1 non fornisce solo i parametri del sistema sopraindicati, ma anche i loro prodotti. Questa sovrabbondanza di informazioni può essere usata a nostro vantaggio, per avere una stima più robusta dei parametri (soprattutto nel caso in cui l'uscita misurata sia affetta da rumore).

Supponiamo che la stima del vettore θ_1 trovata sia:

$$\hat{\theta}_1^T = [\bar{r}_2 \quad \cdots \quad \bar{r}_n \quad \bar{\alpha}_1 \quad \cdots \quad \bar{\alpha}_n \quad \bar{\alpha}_1 r_2 \quad \cdots \quad \bar{\alpha}_1 r_2 \quad \cdots \quad \bar{\alpha}_1 r_2 \quad \cdots \quad \bar{\alpha}_1 r_2] \quad (3.27)$$

Possiamo stimare i parametri del sistema tenendo conto di tutti i dati a disposizione come segue:

$$\begin{aligned} [\hat{r}_i, \hat{\alpha}_k] &= \arg \left\{ \min \left[\sum_k (\hat{\alpha}_k - \bar{\alpha}_k)^2 + \sum_{i,k} (\hat{\alpha}_k \hat{r}_i - \bar{\alpha}_k r_i)^2 \right] \right\} \\ &= \arg \left\{ \min \left[\left\| \begin{bmatrix} \bar{\alpha}_1 & \bar{\alpha}_1 r_2 & \cdots & \bar{\alpha}_1 r_m \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\alpha}_n & \bar{\alpha}_n r_2 & \cdots & \bar{\alpha}_n r_m \end{bmatrix} - \begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_n \end{bmatrix} \cdot [\hat{r}_1 \quad \hat{r}_2 \quad \cdots \quad \hat{r}_m] \right\|_F \right] \right\} \end{aligned} \quad (3.28)$$

dove F rappresenta la norma di Frobenius. ($\|A\|_F = \sum_i \sum_j |a_{ij}|^2$).

Definiamo:

$$\Theta_1 \triangleq \begin{bmatrix} \bar{\alpha}_1 & \bar{\alpha}_1 r_2 & \cdots & \bar{\alpha}_1 r_m \\ \vdots & \vdots & \ddots & \vdots \\ \bar{\alpha}_n & \bar{\alpha}_n r_2 & \cdots & \bar{\alpha}_n r_m \end{bmatrix}. \quad (3.29)$$

Consideriamo inoltre la scomposizione a valori singolari di Θ_1 :

$$\Theta_1 = \sum_{i=1}^{\min(n,m)} \xi_i \cdot \sigma_i \cdot \eta_i \quad (3.30)$$

dove i σ_i rappresentano i valori singolari della matrice, gli ξ_i e gli η_i i relativi vettori ortonormali di dimensione rispettivamente n ed m .

Si può dimostrare (vedere [1]) che una soluzione del problema di minimizzazione (3.28) è data da:

$$\begin{bmatrix} \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_n \end{bmatrix} = \sigma_1 \cdot s_\eta \cdot \xi_1, \quad \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_m \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \hat{r}_2 \\ \vdots \\ \hat{r}_m \end{bmatrix} = \frac{\mathbf{1}}{s_\eta} \cdot \eta_1, \quad (3.31)$$

dove s_η è la prima componente di η_1 .

3.4. Identificazione della parte lineare.

Nel paragrafo precedente si sono ottenuti i coefficienti che caratterizzano $\tilde{\alpha}(z)$, denominatore della funzione di trasferimento, relativa al periodo di campionamento h , $\tilde{G}(z)$. Per identificare la parte lineare del sistema occorre trovare i coefficienti di numeratore e denominatore della $G(z)$ per il periodo di campionamento T .

Possiamo fattorizzare il polinomio $\tilde{\alpha}(z)$ in modo da evidenziare i poli della $\tilde{G}(z)$ stimata:

$$\hat{\alpha}(z) = \mathbf{1} - \hat{\alpha}_1 z^{-1} - \dots - \hat{\alpha}_n z^{-n} = (\mathbf{1} - \tilde{s}_1 z^{-1}) \cdot \dots \cdot (\mathbf{1} - \tilde{s}_n z^{-n}). \quad (3.32)$$

s è un polo del sistema in tempo continuo se e solo se e^{sT} è un polo della $G(z)$ e se e solo se e^{sh} è un polo della $\tilde{G}(z)$. Questo implica che se \tilde{s}_i è un polo della $\tilde{G}(z)$, allora \tilde{s}_i^{n+1} è un polo della $G(z)$. A questo punto è chiaro come si può stimare il denominatore della $G(z)$; a partire dalla (3.32), applicando ai poli la trasformazione appena mostrata, si ottiene:

$$\hat{\alpha}(z) = (\mathbf{1} - \tilde{s}_1^{n+1} z^{-1}) \cdot \dots \cdot (\mathbf{1} - \tilde{s}_n^{n+1} z^{-n}) = \mathbf{1} - \hat{\alpha}_1 z^{-1} - \dots - \hat{\alpha}_n z^{-n}. \quad (3.33)$$

Adesso resta da trovare il numeratore della $G(z)$.

Consideriamo le due sequenze:

$$\begin{aligned} \{x[kT]\} &\Leftrightarrow X(z) = \sum_{k=0}^{\infty} x[kT]z^{-k} = G(z) \cdot U(z) \\ \{x[kT + T/2]\} &\Leftrightarrow \bar{X}(z) = \sum_{k=0}^{\infty} x[kT + T/2]z^{-k} \triangleq \bar{G}(z) \cdot U(z) \end{aligned} \quad (3.34)$$

dove $U(z)$ è la Z -trasformata della sequenza $\{u[kT]\}$ con periodo di campionamento T , mentre $G(z)$ e $\bar{G}(z)$ sono le funzioni di trasferimento che determinano, a partire da $U(z)$, rispettivamente $X(z)$ e $\bar{X}(z)$.

Ricordando l'equazione di stato in tempo continuo (3.1), si ha che:

$$\begin{aligned} \eta[(k+1)T + T/2] &= e^{AT} \cdot \eta[kT + T/2] + \int_{T/2}^T e^{A(T-t)} \cdot b \cdot dt \cdot u[kT] \\ &+ \int_0^{T/2} e^{A(T-t)} \cdot b \cdot dt \cdot u[(k+1)T] \\ &\triangleq \Phi \cdot \eta[kT + T/2] + \Gamma_1 \cdot u[kT] + \Gamma_2 \cdot u[(k+1)T] \end{aligned} \quad (3.35)$$

$$x[kT + T/2] = c \cdot \eta[kT + T/2]$$

e la funzione di trasferimento associata è:

$$\bar{G}(z) = \frac{\bar{\beta}(z)}{\bar{\alpha}(z)} = c \cdot (zI - \Phi)^{-1} \cdot (\Gamma_1 + \Gamma_2 \cdot z) = \frac{\bar{\beta}_0 + \bar{\beta}_1 z^{-1} + \dots + \bar{\beta}_n z^{-n}}{1 - \bar{\alpha}_1 z^{-1} - \bar{\alpha}_2 z^{-2} - \dots - \bar{\alpha}_n z^{-n}}. \quad (3.36)$$

Dal confronto con la (3.3) è evidente che questa funzione di trasferimento ha lo stesso denominatore della $G(z)$, mentre il numeratore è diverso; si noti inoltre che mentre $G(z)$ è strettamente propria, $\bar{G}(z)$ è propria ma non strettamente.

Come detto sopra, per le sequenze $\{x[kT]\}$ e $\{x[kT+T/2]\}$ valgono le seguenti relazioni:

$$\begin{cases} X(z) = G(z) \cdot U(z) \\ \bar{X}(z) = \bar{G}(z) \cdot U(z) \end{cases} \quad (3.37)$$

dalle quali, tenendo presente che $G(z)$ e $\bar{G}(z)$ hanno lo stesso denominatore, è chiaro che:

$$\begin{aligned}\bar{G}(z) \cdot X(z) - G(z) \cdot \bar{X}(z) &= \bar{G}(z) \cdot G(z) \cdot U(z) - G(z) \cdot \bar{G}(z) \cdot U(z) = \mathbf{0} \\ \bar{\beta}(z) \cdot X(z) - \beta(z) \cdot \bar{X}(z) &= \bar{\beta}(z) \cdot \frac{\beta(z)}{\alpha(z)} \cdot U(z) - \beta(z) \cdot \frac{\bar{\beta}(z)}{\alpha(z)} \cdot U(z) = \mathbf{0}\end{aligned}\quad (3.38)$$

cioè:

$$(\bar{\beta}_0 + \bar{\beta}_1 z^{-1} + \dots + \bar{\beta}_n z^{-n}) \cdot X(z) - (\beta_1 z^{-1} + \dots + \beta_n z^{-n}) \cdot \bar{X}(z) = \mathbf{0}, \quad (3.39)$$

che riportata nel dominio del tempo diventa:

$$\begin{aligned}\beta_1 \cdot x[kT + T/2] \\ = [-x[kT - T + T/2], \dots, -x[kT - (n-1)T + T/2], \\ x[kT + T], \dots, x[kT - (n-1)T]] \cdot \begin{bmatrix} \beta_2 \\ \vdots \\ \beta_n \\ \bar{\beta}_0 \\ \vdots \\ \bar{\beta}_n \end{bmatrix} \hat{=} \bar{\phi}_2[k] \cdot \theta_2\end{aligned}\quad (3.40)$$

Ponendo, come detto in precedenza, $\beta_1=1$ si ottiene:

$$x[kT + T/2] = \bar{\phi}_2[k] \cdot \theta_2 \quad (3.41)$$

Abbiamo nuovamente trovato un'equazione lineare nei parametri, ma stavolta è necessario conoscere le sequenze $\{x[kT]\}$ e $\{x[kT+T/2]\}$, che non sono misurabili. Tuttavia possiamo valutarle, sfruttando la stima dell'inversa della non linearità di

uscita trovata nel paragrafo precedente e data da (3.24), a partire dalle sequenze, queste sì misurabili, $\{y[kT]\}$ e $\{y[kT+T/2]\}$. Inglobando l'errore dovuto alla sostituzione delle sequenze con le loro stime nel termine v_2 e chiamando ϕ_2 la matrice $\bar{\phi}_2$ in cui al posto delle x si sono usate le loro stime, si ottiene:

$$\hat{x}[kT + T/2] = \phi_2[k] \cdot \theta_2 + v_2[k]. \quad (3.42)$$

Per ottenere θ_2 si prende un certo numero di equazioni analoghe alla (3.42), per diversi k , e si risolve il sistema di equazioni lineari così ottenuto. Il vettore trovato è costituito dalle stime dei coefficienti dei numeratori della $G(z)$ e della $\bar{G}(z)$:

$$\hat{\theta}_2 = \begin{bmatrix} \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_n \\ \hat{\beta}_0 \\ \vdots \\ \hat{\beta}_n \end{bmatrix}; \quad (3.43)$$

Da questo, considerando che $\beta_1=1$, si ottengono le stime delle due funzioni di trasferimento:

$$\hat{G}(z) = \frac{\hat{\beta}_1 z^{-1} + \hat{\beta}_2 z^{-2} + \dots + \hat{\beta}_n z^{-n}}{1 - \hat{\alpha}_1 z^{-1} - \hat{\alpha}_2 z^{-2} - \dots - \hat{\alpha}_n z^{-n}}$$

$$\hat{\bar{G}}(z) = \frac{\hat{\beta}_0 + \hat{\beta}_1 z^{-1} + \hat{\beta}_2 z^{-2} + \dots + \hat{\beta}_n z^{-n}}{1 - \hat{\alpha}_1 z^{-1} - \hat{\alpha}_2 z^{-2} - \dots - \hat{\alpha}_n z^{-n}}. \quad (3.44)$$

$$\hat{\alpha}_i = \hat{\alpha}_i, i = 1, \dots, n$$

3.5. Identificazione della non linearità di ingresso.

In questo lavoro non si è imposto che la non linearità di ingresso abbia una struttura particolare, per cui, per identificarla, dovremo determinare le coppie (w, u) al variare della w . Poiché, mentre la $w[kT]$ è nota, la $u[kT]$ non lo è, lo scopo che ci prefiggiamo in questo paragrafo è stimare la $u[kT]$.

Dal momento che l'ingresso viene campionato con periodo T , $u[kT] = u[kT+T/2]$.

Se almeno una della due tra $G(z)$ e $\bar{G}(z)$ è a fase minima, allora si può ricavare $U(z)$ da:

$$\begin{aligned} U(z) &= G^{-1}(z) \cdot X(z) \\ \text{oppure} & \quad \cdot \\ U(z) &= \bar{G}^{-1}(z) \cdot \bar{X}(z) \end{aligned} \quad (3.45)$$

L'equivalente nel dominio del tempo di queste equazioni è:

$$\begin{aligned} u[kT] &= -\beta_2 \cdot u[kT - T] - \dots - \beta_n \cdot u[kT - (n-1)T] \\ &+ x[kT + T] - \alpha_1 \cdot x[kT] - \dots - \alpha_n \cdot x[kT - (n-1)T] \end{aligned} \quad (3.46)$$

oppure

$$\begin{aligned} u[kT] &= \frac{1}{\beta_0} \cdot (-\bar{\beta}_1 \cdot u[kT - T] - \dots - \bar{\beta}_n \cdot u[kT - nT]) \\ &+ x[kT + T/2] - \bar{\alpha}_1 \cdot x[kT + T/2 + T] - \dots - \bar{\alpha}_n \cdot x[kT + T/2 - nT] \end{aligned} \quad (3.47)$$

In precedenza abbiamo visto come calcolare le stime di $\beta(z)$, $\bar{\beta}(z)$, $x[kT]$ e $x[kT+T/2]$, perciò siamo in grado di calcolare una stima di $u[kT]$, $\hat{u}[kT]$.

Se né $G(z)$ né $\bar{G}(z)$ fossero a fase minima il procedimento appena descritto non sarebbe applicabile, perché le inverse delle due funzioni di trasferimento non

sarebbero stabili. Tuttavia, se $G(z)$ e $\bar{G}(z)$ non hanno nessuno zero a comune, è possibile ovviare a questa difficoltà. Infatti, per l'identità di Bezout, esistono due funzioni di trasferimento stabili, $F(z)$ e $\bar{F}(z)$, tali che:

$$F(z) \cdot \bar{G}(z) + \bar{F}(z) \cdot G(z) = \mathbf{1} \quad (3.48)$$

e quindi:

$$F(z) \cdot \bar{X}(z) + \bar{F}(z) \cdot X(z) = [F(z) \cdot \bar{G}(z) + \bar{F}(z) \cdot G(z)] \cdot U(z) = U(z) \quad (3.49)$$

per cui è possibile ricavare le $u[kT]$ sulla base di queste equazioni, trovando la $F(z)$ e la $\bar{F}(z)$, a partire dalle sequenze $x[kT]$ e $x[kT+T/2]$.

Una volta ottenuta la stima delle uscite dalla non linearità di ingresso, è possibile ricostruirne l'andamento riportando in un grafico le coppie $(w[kT], u[kT])$. Se si ipotizza che la non linearità di ingresso abbia una particolare struttura (polinomiale o altro), è comunque possibile ricavarne i parametri caratteristici a partire dalle coppie di punti trovate.

3.6. Algoritmo di identificazione.

I risultati ottenuti nei paragrafi precedenti consentono di tracciare lo schema di un algoritmo per identificare un sistema Hammerstein-Wiener a partire dalle misurazioni degli ingressi e delle uscite del sistema:

1. Considerando il sistema illustrato in Figura 3.1 e dato un periodo di campionamento T , si raccolgono le misure degli ingressi agli istanti kT e quelle delle uscite agli istanti kT , kh e $kT+T/2$, con $h = T/(n+1)$.

2. Si costruisce il vettore Δy e la matrice $\Delta\phi_1$ descritti in (3.16), (3.17), (3.21) e si stima il vettore θ_1 risolvendo il sistema lineare di (3.23). A partire da θ_1 si calcola

l'inversa della non linearità di uscita $x = f^{-1}(y) \cong \sum_{i=1}^m r_i \cdot y^i$. Si ricava, poi, la non

linearità di uscita $y = f(x) = \sum_{i=1}^q a_i \cdot x^i$ che meglio si adatta alla sua inversa, appena

calcolata. Si stimano infine, a partire dalla approssimazione di $f^{-1}(y)$ e dai campioni di y , i campioni di x agli istanti kT e $kT+T/2$.

3. A partire da θ_1 si ricava una stima del denominatore della funzione di trasferimento $G(z)$, $\alpha(z)$, utilizzando le (3.31-33). Si costruisce poi la matrice ϕ_2 (vedi (3.40)) e si stima il vettore θ_2 ; da questo, oltre che dalla stima di $\alpha(z)$, si ricavano le funzioni di trasferimento $G(z)$ e $\bar{G}(z)$.

4. Se almeno una della due tra $G(z)$ e $\bar{G}(z)$ è a fase minima, la si utilizza per stimare usando la (3.45) o alternativamente la (3.47). Diversamente, se $G(z)$ e $\bar{G}(z)$ non hanno nessuno zero a comune, si calcolano $F(z)$ e $\bar{F}(z)$ in accordo alla (3.48) e si stimano gli $\hat{u}[kT]$ a partire da queste.

5. Si stima la non linearità di ingresso a partire dalle coppie $(w[kT], \hat{u}[kT])$ ottenute al passo 4.

3.7. Considerazioni generali.

3.7.1. Non linearità di ingresso.

Nella trattazione fatta in precedenza non abbiamo mai fatto assunzioni sulla struttura della non linearità di ingresso. Questo consente di poter identificare, con l'algoritmo descritto nel paragrafo 3.5, sistemi che abbiano una non linearità di ingresso di qualunque genere (purché statica), e comunque senza dover disporre di particolari informazioni sulla sua struttura. D'altronde la $g(w)$ viene identificata stimando genericamente le coppie (w, u) . Queste forniscono informazioni sulla $g(w)$ sufficienti per poterla poi parametrizzare come si preferisce.

E' anche vero che, se si avessero delle informazioni sulla struttura della $g(w)$, nulla vieterebbe di utilizzarle per sviluppare un algoritmo diverso. Data una parametrizzazione della non linearità di ingresso lineare nei coefficienti, si può ad esempio immaginare di ricavare tutti i parametri del sistema risolvendo un sistema lineare derivato dall'equazione (3.13), similmente a quanto fatto in precedenza per ricavare la non linearità di uscita. Ovviamente questo approccio comporterebbe il problema della dimensione, a cui si potrebbe ovviare pensando ad un metodo iterativo in cui si ricavi di volta in volta una parte dei coefficienti, lasciando fissati gli altri.

Possiamo fare un'ultima considerazione riguardo la stima della non linearità di ingresso. Infatti nel paragrafo 3.5 si è detto che è possibile ricavare la $g(w)$ solo se la almeno una della due tra $G(z)$ e $\bar{G}(z)$ è a fase minima, oppure se queste due funzioni di trasferimento non hanno zeri in comune. In realtà, se anche nessuna di queste condizioni fosse verificata, si potrebbe riuscire ad identificare la non linearità di ingresso ripetendo la procedura di identificazione con un tempo di campionamento diverso da T . Infatti, cambiando il tempo di campionamento, cambiano anche le funzioni di trasferimento tempo discreto $G(z)$ e $\bar{G}(z)$ ed è senz'altro possibile che per la nuova coppia di f.d.t. almeno una delle condizioni sopraesposte sia vera. Altrimenti si può tentare con un periodo di campionamento ancora diverso, e così

via. Chiaramente questo è un metodo del tutto empirico, senza nessuna garanzia di funzionamento, perciò le condizioni sopradette restano le uniche che permettano di avere una certezza sul buon esito dell'algoritmo.

3.7.2. Scelta del periodo di campionamento.

Come spiegato nel paragrafo 3.3, il periodo di campionamento per l'uscita deve essere superiore al periodo di campionamento relativo al segnale di ingresso. In precedenza si è fissato, per semplicità, in $h = T/(n+1)$, dove T è il periodo di campionamento per l'ingresso ed n è l'ordine della $G(z)$. In questo modo si riusciva ad annullare la dipendenza del sistema lineare ottenuto in (3.15) dal segnale di ingresso. Tuttavia qualunque $h = T/l (n+1)$ con l intero positivo può essere scelto ottenendo, in teoria, lo stesso risultato; infatti il segnale di ingresso resterebbe comunque costante per almeno $n + 1$ periodi di campionamento del segnale in uscita. E' peraltro chiaro che nella pratica il *rapporto di sovracampionamento* l non può essere troppo elevato; infatti per l molto grande, ovvero per h molto piccolo, il segnale di uscita cambierebbe troppo poco da un periodo di campionamento all'altro, causando così malcondizionamenti numerici nell'algoritmo di identificazione. Per lo stesso motivo, a parità di l , non è possibile scegliere un periodo di campionamento troppo basso per il segnale di ingresso.

3.7.3. Non linearità di uscita.

Nella trattazione fatta al paragrafo 3.3 si è sempre assunto che il grado m della funzione inversa della non linearità di uscita fosse noto. In generale questo non è vero. In teoria tanto più grande è m , tanto migliore è l'approssimazione di $f^{-1}(y)$, quindi si potrebbe credere che è opportuno scegliere un m molto elevato. In realtà, al crescere di m , aumenta la sensibilità al rumore ed agli errori di modello, oltre che i problemi numerici (ricordiamo che nella matrice $\Delta\phi_1$ compaiono le potenze fino alla m -esima di y). Pertanto con l'aumentare di m da un lato migliora l'approssimazione della $f^{-1}(y)$, dall'altro peggiorano le prestazioni dell'algoritmo, soprattutto in una situazione rumorosa. Poiché il miglioramento nella stima di $f^{-1}(y)$ all'aumentare di m

è trascurabile per m al di sopra di una certa soglia, è chiaro che è insensato utilizzare valori di m particolarmente elevati. In ogni caso è necessario trovare l' m che fornisca il miglior compromesso tra bontà nella stima di $f^{-1}(y)$ e buone prestazioni dell'algoritmo. Per trovare il giusto valore di m si può procedere così: partendo da un valore basso di m , ad esempio 1, si ripete la procedura di identificazione, aumentando m , e si calcola ogni volta l'errore quadratico medio sull'uscita; questo tenderà inizialmente a decrescere per poi stabilizzarsi su un certo valore, o addirittura aumentare. Il valore di m da scegliere sarà quello più basso per il quale la differenza tra l'errore relativo ad m e quello relativo ad $m-1$ sia sufficientemente piccola; in altre parole converrà utilizzare il valore di m oltre il quale l'errore sull'uscita non ha sostanziali miglioramenti.

Riguardo alla non linearità di uscita si può fare una seconda considerazione. Sino ad ora abbiamo sempre assunto che la $f(x)$ sia polinomiale ed abbiamo sempre parametrizzato la sua inversa usando una base polinomiale. In realtà questa ipotesi non deve essere necessariamente verificata: infatti i ragionamenti seguiti in questo capitolo sono ripetibili considerando qualunque parametrizzazione dell'inversa della non linearità di uscita che sia lineare nei coefficienti, cioè:

$$f^{-1}(y) = \sum_{i=1}^m r_i \cdot h_i(y) \quad (3.50)$$

dove $h_i(y)$ è un qualunque insieme di funzioni che possa costituire una base per la $f^{-1}(y)$. Sostituendo la (3.50), al posto della parametrizzazione polinomiale espressa nella (3.6), nella (3.13) si ottengono delle equazioni lineari algebriche del tutto analoghe a quella di (3.22), modificando opportunamente $\Delta\phi[l]$, in cui compariranno le funzioni di base $h_i(y)$ scelte anziché le potenze della y , e sostituendo al termine $\Delta y[l]$ un termine che tiene conto delle $h_i(y)$. In sintesi la (3.15) diventa:

$$\begin{aligned}
\phi_1[kh] \triangleq & [-h_2(y[kh]), \dots, -h_m(y[kh]), h_1(y[(k-1)h]), \dots, h_1(y[(k-n)h]), \\
& h_2(y[(k-1)h]), \dots, h_2(y[(k-n)h]), \dots, h_m(y[(k-1)h]), \dots, \\
& h_m(y[(k-n)h])] \\
\theta_1 \triangleq & \begin{bmatrix} r_2 \\ \vdots \\ r_m \\ \tilde{\alpha}_1 \\ \vdots \\ \tilde{\alpha}_n \\ \tilde{\alpha}_1 \cdot r_2 \\ \vdots \\ \tilde{\alpha}_n \cdot r_2 \\ \vdots \\ \tilde{\alpha}_1 \cdot r_m \\ \vdots \\ \tilde{\alpha}_n \cdot r_m \end{bmatrix} \tag{3.50}
\end{aligned}$$

$$h_1(y[kh]) = \phi_1[kh] \cdot \theta_1 + \sum_{i=1}^n \tilde{\beta}_i \cdot u[kh - ih] + v_i[kh]$$

3.7.4. Scelta del segnale di ingresso.

Il segnale di ingresso che viene fornito al sistema durante una procedura di identificazione non può essere qualsiasi. In generale è opportuno che il sistema sia sollecitato su tutta la sua banda caratteristica e con un insieme sufficientemente ampio di valori dell'ingresso, in modo che si manifestino tutte le caratteristiche del sistema, sia da un punto di vista dinamico che da un punto di vista statico.

Come viene evidenziato in [2], con l'algoritmo proposto in questo capitolo è necessario fornire al sistema un ingresso tale che le matrici $\Delta\phi_1$ e ϕ_2 siano persistentemente eccitate (PE). Questo avviene a patto che le linee spettrali del segnale $u[kT]$ non siano concentrate in meno di $2n$ punti (un'analogia condizione può essere posta per l'ingresso $w[kT]$). Questa condizione è necessaria per la robustezza dell'algoritmo al rumore e all'incertezza di modello e garantisce l'identificabilità.

Nel seguito si considereranno sempre segnali di ingresso costituiti da sequenze di gradini di ampiezza casuale uniformemente distribuita all'interno di un opportuno intervallo. A questo proposito dobbiamo notare che il segnale di ingresso non dovrebbe assumere valori eccessivamente elevati; infatti, in questo caso, è possibile che le non linearità presenti nel sistema lo amplifichino producendo dei valori di uscita particolarmente alti, che potrebbero causare problemi numerici. E' invece chiaro che l'intervallo di valori su cui spazia l'ingresso non può essere troppo piccolo, pena una insufficiente eccitazione del sistema.

3.8. Uso della tecnica delle funzioni modulanti.

L'algoritmo presentato nei paragrafi precedenti identifica i parametri di un sistema Hammerstein-Wiener a partire dalle misure dell'uscita. In particolare il primo passo dell'algoritmo di identificazione è stimare la funzione inversa della non linearità di uscita. Una volta ottenuta questa, è chiaro che ci si può ricondurre ad un sistema di tipo Hammerstein semplice. Si consideri la Figura 3.3.

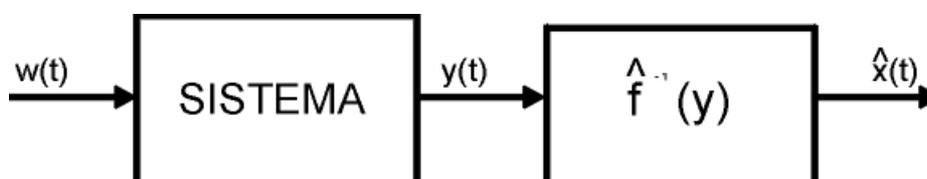


Figura 3.3. Ricostruzione del segnale $x(t)$

Si vede che, qualunque segnale di ingresso venga fornito al sistema, “filtrando” la sua uscita per mezzo della stima di $f^{-1}(y)$ si ottiene una stima del segnale $x(t)$ di uscita dal blocco lineare. $\hat{x}(t)$ non riproduce esattamente $x(t)$, sia perché l'inversa della funzione di uscita è stata stimata con un inevitabile errore, sia perché l'uscita del sistema sarà affetta da rumore. Tuttavia le coppie $\{w(t), \hat{x}(t)\}$ costituiscono una approssimazione dell'ingresso e dell'uscita della parte Hammerstein del sistema Hammerstein-Wiener, cioè della non linearità di ingresso seguita dal blocco lineare. Una volta stimata la $f^{-1}(y)$, e quindi isolata la parte Hammerstein del sistema, questa può essere identificata con uno dei molti metodi sviluppati a questo scopo, senza bisogno di procedere con l'algoritmo sinora descritto.

Una buona possibilità per l'identificazione della parte Hammerstein è offerta dalla tecnica delle funzioni modulanti, illustrata nel Capitolo 2, caratterizzata da una intrinseca robustezza al rumore e dal fatto di operare direttamente in tempo continuo.

Pertanto possiamo proporre un algoritmo alternativo per l'identificazione di sistemi Hammerstein-Wiener, descritto di seguito:

1. Considerando il sistema illustrato in Figura 3.1 e dato un periodo di campionamento T , si raccolgono le misure degli ingressi agli istanti kT e quelle delle uscite agli istanti kT , kh e $kT+T/2$, con $h = T/(n+1)$.

2. Si costruisce il vettore Δy e la matrice $\Delta \phi_1$ come descritto in (3.16), (3.17), (3.21) e si stima il vettore θ_1 risolvendo il sistema lineare di (3.23). A partire da θ_1 si ricava l'inversa della non linearità di uscita $x = f^{-1}(y) \cong \sum_{i=1}^m r_i \cdot y^i$. Si ricava poi la non

linearità di uscita $y = f(x) = \sum_{i=1}^q a_i \cdot x^i$ che meglio si adatta alla sua inversa appena calcolata.

3. Si applica al sistema un ingresso adatto all'identificazione con la tecnica delle funzioni modulanti (vedi paragrafo 2.4.1) e se ne misura l'uscita.

4. A partire dall'uscita ottenuta al punto 3, si stima l'uscita del blocco lineare attraverso l'inversa della non linearità di uscita ricavata al punto 2.

5. Si applica la tecnica delle funzioni modulanti per identificare il sistema Hammerstein avente come ingresso quello usato al punto 3 e come uscita il segnale ricavato al punto 4 (vedi paragrafo 2.6).

L'algoritmo appena illustrato è, in sostanza, un'estensione della tecnica delle funzioni modulanti a sistemi non lineari a blocchi di tipo Hammerstein-Wiener.