

della soluzione.

△

4.8 Il filtro di Kalman nel caso di rumore di stato e di osservazione correlati

Nella teoria fin qui sviluppata per il filtro di Kalman, si è sempre ipotizzato che i rumori di stato e di osservazione fossero scorrelati. Nel modello impiegato, in cui per ogni k la variabile aleatoria N_k è standard, ovvero ha covarianza pari all'identità, l'indipendenza dei rumori di stato e di osservazione si traduce nella condizione matriciale $F(k)G^T(k) = 0$ per ogni k .

In questo paragrafo viene mostrato in che modo i risultati fin qui ottenuti possono essere applicati al caso di sistemi in cui i rumori di stato e di osservazione sono correlati, e cioè quando $F(k)G^T(k) \neq 0$, almeno per qualche k . Per semplicità di notazione nel seguito viene

considerato il solo caso di sistemi stazionari. Il sistema in esame sia quindi

$$\begin{aligned} x(k+1) &= Ax(k) + FN_k, & x(0) &= x_0, \\ y(k) &= Cx(k) + GN_k, & FG^T &\neq 0, \end{aligned} \quad (4.8.1)$$

in cui $\{N_k\}$ è una sequenza aleatoria gaussiana bianca e standard. È possibile ricondurre il problema ad uno in cui il rumore di stato e di osservazione sono scorrelati definendo il seguente sistema esteso

$$\begin{aligned} \begin{bmatrix} z_1(k+1) \\ z_2(k+1) \end{bmatrix} &= \begin{bmatrix} A & F \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} \bar{N}_k, \\ y(k) &= [C \quad G] \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix}, & \begin{bmatrix} z_1(0) \\ z_2(0) \end{bmatrix} &= \begin{bmatrix} x_0 \\ N_0 \end{bmatrix}, \end{aligned}$$

in cui la sequenza $\{\bar{N}_k\}$ altro non è che la traslazione di un passo della sequenza $\{N_k\}$, ovvero $\bar{N}_k \triangleq N_{k+1}$. Dalla struttura delle matrici si deduce che $z_2(k) = N_k$ e che $z_1(k) = x(k)$. Il sistema ha ora acquistato la struttura

$$\begin{aligned} z(k+1) &= \bar{A}z(k) + \bar{F}\bar{N}_k, & z(0) &= z_0 \\ y(k) &= \bar{C}z(k) + \bar{G}\bar{N}_k, & \bar{F}\bar{G}^T &= 0 \end{aligned} \quad (4.8.2)$$

$$\text{con } \bar{A} = \begin{bmatrix} A & F \\ 0 & 0 \end{bmatrix}, \quad \bar{F} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad \bar{C} = [C \quad G], \quad \bar{G} = 0, \quad z_0 = \begin{bmatrix} x_0 \\ N_0 \end{bmatrix},$$

in cui il rumore di osservazione è assente, e quindi banalmente indipendente dal rumore di stato. Per questa formulazione è possibile scrivere le usuali equazioni di Riccati per il calcolo del guadagno del filtro di Kalman, tenendo conto che $\bar{G} = 0$

$$\begin{aligned} \bar{P}_p(k+1) &= \bar{A}\bar{P}_p(k)\bar{A}^T + \bar{F}\bar{F}^T, \\ \bar{K}(k+1) &= \bar{P}_p(k+1)\bar{C}^T(\bar{C}\bar{P}_p(k+1)\bar{C}^T)^{-1}, \\ \bar{P}(k+1) &= (I - \bar{K}(k+1)\bar{C})\bar{P}_p(k+1). \end{aligned}$$

Il filtro di Kalman ha la solita struttura

$$\hat{z}(k+1) = \bar{A}\hat{z}(k) + \bar{K}(k+1)(y(k+1) - \bar{C}\bar{A}\hat{z}(k)).$$

Si osservi che ad ogni passo k si ottiene non solo la stima dello stato $x(k)$ ma anche la stima del rumore N_k . Infatti

$$\hat{z}(k) = \begin{bmatrix} \hat{z}_1(k) \\ \hat{z}_2(k) \end{bmatrix} = \begin{bmatrix} \hat{x}(k) \\ \hat{N}_k \end{bmatrix}.$$

È inoltre facile verificare che $\bar{C}\bar{K}(k+1) = I$, e che quindi $\bar{C}\hat{z}(k+1) = y(k+1)$, e cioè sviluppando i blocchi

$$y(k+1) = C\hat{x}(k+1) + G\hat{N}_{k+1}. \quad (4.8.3)$$

Definendo opportunamente la scomposizione di $\bar{K}(k+1)$ in due sottomatrici $K(k+1)$ e $H(k+1)$ è possibile suddividere l'equazione del filtro di Kalman nei due blocchi

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + F(k)\hat{N}_k + K(k+1)(y(k+1) - C(A\hat{x}(k) + F\hat{N}_k)) \\ \hat{N}_{k+1} &= H(k+1)(y(k+1) - CA\hat{x}(k) - CF\hat{N}_k). \end{aligned} \quad (4.8.4)$$

Questa scrittura consente di fornire una interpretazione delle modifiche che il filtro di Kalman subisce quando i rumori di stato e di osservazione sono correlati: la stima del rumore viene utilizzata nel calcolo della predizione ottima dello stato, che per l'appunto diventa

$$\hat{x}(k+1|k) = A\hat{x}(k) + F\hat{N}_k. \quad (4.8.5)$$

Prima di passare alla riscrittura delle equazioni di Riccati, opportunamente suddivise in blocchi, si vuole riportare un'altra implementazione dello stesso filtro appena presentato, che evita il calcolo esplicito della stima del rumore \hat{N}_k . A tale scopo si riscrive la prima equazione del filtro di Kalman utilizzando la predizione ottima, scrivendo per chiarezza $\hat{x}(k|k)$ al posto di $\hat{x}(k)$

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)(y(k) - C\hat{x}(k|k-1)). \quad (4.8.6)$$

Tenendo conto che

$$\hat{N}_k = H(k)(y(k) - C\hat{x}(k|k-1))$$

la predizione ottima si scrive anche

$$\hat{x}(k+1|k) = A\hat{x}(k|k) + FH(k)(y(k) - C\hat{x}(k|k-1)), \quad (4.8.7)$$

e sostituendo $\hat{x}(k|k)$ si ha

$$\begin{aligned} \hat{x}(k+1|k) = A\left(\hat{x}(k|k-1) + K(k)(y(k) - C\hat{x}(k|k-1))\right) + \\ + FH(k)(y(k) - C\hat{x}(k|k-1)). \end{aligned} \quad (4.8.8)$$

Riordinando quest'ultima si ha

$$\hat{x}(k+1|k) = A\hat{x}(k|k-1) + (AK(k) + FH(k))(y(k) - C\hat{x}(k|k-1)). \quad (4.8.9)$$

La predizione ottima $\hat{x}(k+1|k)$ viene impiegata poi per il calcolo della stima ottima

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)(y(k+1) - C\hat{x}(k+1|k)). \quad (4.8.10)$$

Riassumendo, il filtro di Kalman nel caso di rumori di stato e di uscita correlati può essere implementato in due passi successivi: nel primo viene calcolata la predizione ottima all'istante $k+1$ utilizzando l'osservazione all'istante k , mentre nel secondo passo viene calcolata la stima ottima all'istante $k+1$ utilizzando l'osservazione nello stesso istante.

Resta a questo punto solamente da risolvere il problema del calcolo delle matrici di guadagno $K(k)$ e $H(k)$. Per fare ciò occorre suddividere in blocchi le equazioni di Riccati. Si definisca innanzitutto la seguente partizione della matrice di covarianza dell'errore di stima

$$\bar{P}(k) \triangleq \begin{bmatrix} P(k) & R(k) \\ R^T(k) & Q(k) \end{bmatrix}, \quad (4.8.11)$$

in cui $P(k)$ è la covarianza dell'errore di stima dello stato, $Q(k)$ è la covarianza dell'errore di stima del rumore e $R(k)$ è la mutua covarianza tra i due errori di stima. Lo sviluppo del calcolo della covarianza dell'errore di predizione porta a

$$\bar{P}_p(k+1) = \begin{bmatrix} P_p(k+1) & 0 \\ 0 & I \end{bmatrix},$$

$$\text{in cui } P_p(k+1) = AP(k)A^T + FR^T(k)A^T + AR(k)F^T + FQ(k)F^T. \quad (4.8.12)$$

Si ha poi l'equivalenza

$$\bar{C}\bar{P}_p(k+1)\bar{C}^T + \bar{G}\bar{G}^T = CP_p(k+1)C^T + GG^T, \quad (4.8.13)$$

da cui il guadagno $\bar{K}(k+1)$ può essere calcolato

$$\begin{aligned} \bar{K}(k+1) &= \begin{bmatrix} K(k+1) \\ H(k+1) \end{bmatrix} = \\ &= \begin{bmatrix} P_p(k+1)C^T(CP_p(k+1)C^T + GG^T)^{-1} \\ G^T(CP_p(k+1)C^T + GG^T)^{-1} \end{bmatrix}. \end{aligned} \quad (4.8.14)$$

Il calcolo della covarianza dell'errore di stima porta a

$$\bar{P}(k+1) = \begin{bmatrix} (I - K(k+1)C)P_p(k+1) & -K(k+1)G \\ -H(k+1)CP_p(k+1) & I - H(k+1)G \end{bmatrix} \quad (4.8.15)$$

Confrontando quest'ultima con la suddivisione a blocchi della $\bar{P}(k)$ si ha che

$$\begin{aligned} P(k+1) &= (I - K(k+1)C)P_p(k+1) \\ R(k+1) &= -K(k+1)G, \quad Q(k+1) = I - H(k+1)G, \\ &\left(\text{e anche } K(k+1)G = (H(k+1)CP_p(k+1))^T \right). \end{aligned}$$

Poiché le matrici $R(\cdot)$ e $Q(\cdot)$ vengono impiegate esclusivamente nel calcolo della $P_p(\cdot)$, invece di essere memorizzate è possibile sostituirle direttamente nell'equazione ricorsiva, ottenendo

$$P_p(k+1) = AP(k)A + FF^T - FG^TK^T(k)A^T - AK(k)GF^T - FH(k)GF^T. \quad (4.8.16)$$

Definendo la matrice di covarianza dell'errore di predizione dell'uscita

$$S(k+1) = CP_p(k+1)C^T + GG^T$$

le espressioni dei guadagni di Kalman $K(k+1)$ e $H(k+1)$ risultano più semplici

$$\begin{aligned} K(k+1) &= P_p(k+1)C^T S^{-1}(k+1), \\ H(k+1) &= G^T S^{-1}(k+1), \end{aligned}$$

come anche la covarianza della stima del rumore

$$Q(k+1) = I - G^T S^{-1}(k+1)G.$$

A questo punto è possibile riassumere le equazioni del filtro di Kalman nel caso di rumori di stato e di osservazione correlati secondo le due implementazioni presentate.

Filtro di Kalman nel caso $FG^T \neq 0$. (A)

(I) Inizializzazione:

$$\begin{aligned} P_p(0) &= \Psi_{x0}, \\ S(0) &= C\Psi_{x0}C^T + GG^T \end{aligned} \quad \begin{cases} K(0) = P_p(0)C^T S^{-1}(0), \\ H(0) = G^T S^{-1}(0). \end{cases}$$

$$\hat{x}(0) = K(0)y(0),$$

$$\hat{N}_0 = H(0)y(0), \quad k = 0.$$

(II) Equazioni di Riccati:

$$\begin{aligned} P_p(k+1) &= AP(k)A - FG^T K^T(k)A^T - AK(k)GF^T + \\ &\quad + F(I - G^T S^{-1}(k)G)F^T, \end{aligned}$$

$$S(k+1) = CP_p(k+1)C^T + GG^T,$$

$$K(k+1) = P_p(k+1)C^T S^{-1}(k+1),$$

$$H(k+1) = G^T S^{-1}(k+1),$$

$$P(k+1) = (I - K(k+1)C)P_p(k+1),$$

(III) Filtro

$$\begin{aligned}\hat{x}(k+1) &= A\hat{x}(k) + F(k)\hat{N}_k + K(k+1) \cdot (y(k+1) - C(A\hat{x}(k) + F\hat{N}_k)) \\ \hat{N}_{k+1} &= H(k+1)(y(k+1) - CA\hat{x}(k) - CF\hat{N}_k).\end{aligned}\quad (4.8.17)$$

IV $k=k+1$. GOTO II.

Se si è interessati a conoscere la covarianza dell'errore di stima del rumore è sufficiente calcolare $Q(k) = I - G^T S^{-1}(k)G$.

L'implementazione alternativa presentata diventa:

Filtro di Kalman nel caso $FG^T \neq 0$. (B)

(I) Inizializzazione:

$$\begin{aligned}P_p(0) &= \Psi_{x0}, \\ S(0) &= C\Psi_{x0}C^T + GG^T\end{aligned}\quad \begin{cases} K(0) = P_p(0)C^T S^{-1}(0), \\ H(0) = G^T S^{-1}(0). \end{cases}$$

$$\hat{x}(0|-1) = 0, \quad k = -1;$$

(II) Calcolo della stima

$$\hat{x}(k+1|k+1) = x(k+1|k) + K(k+1)(y(k+1) - C\hat{x}(k+1|k)),$$

(III) $k = k + 1$;

(IV) Equazioni di Riccati:

$$\begin{aligned}P_p(k+1) &= AP(k)A - FG^T K^T(k)A^T - AK(k)GF^T + \\ &\quad + F(I - G^T S^{-1}(k)G)F^T, \\ S(k+1) &= CP_p(k+1)C^T + GG^T, \\ K(k+1) &= P_p(k+1)C^T S^{-1}(k+1), \\ H(k+1) &= G^T S^{-1}(k+1), \\ P(k+1) &= (I - K(k+1)C)P_p(k+1),\end{aligned}$$

(V) Calcolo della predizione

$$\hat{x}(k+1|k) = Ax(k|k-1) + (AK(k) + FH(k))(y(k) - C\hat{x}(k|k-1)),$$

(VI) GOTO (II).

Si ricordi che affinché la sequenza di Riccati $\{\bar{P}(k)\}$ possa essere costruita per qualunque k e sia sempre definita positiva deve essere $\bar{P}(0) > 0$ e le matrici $[\bar{A} \mid \bar{F}]$ e $[\bar{C} \mid \bar{G}]$ devono avere rango $n+s$ e q rispettivamente.

Ricordando le definizioni delle matrici "barrate", ciò significa che deve essere

$$\begin{aligned} \bar{P}(0) > 0, \quad \rho\left(\begin{bmatrix} A & F & | & 0 \\ 0 & 0 & | & I \end{bmatrix}\right) &= n + s, \\ \rho\left(\begin{bmatrix} C & G & | & 0 \end{bmatrix}\right) &= q. \end{aligned} \quad (4.8.18)$$

È facile verificare che se sono soddisfatte le condizioni che consentono la costruzione del filtro di Kalman nel caso di rumore di stato e di osservazione scorrelati, e cioè

$$P(0) > 0, \quad \rho([A \ F]) = n, \quad \rho([C \ G]) = q, \quad (4.8.19)$$

allora anche le (4.8.18) sono soddisfatte.

Affinché inoltre esista unico lo stato stazionario per l'equazione di Riccati, e quindi esista il guadagno di regime \bar{K}_∞ tale da stabilizzare la matrice $(I - \bar{K}_\infty \bar{C})\bar{A}$, la coppia (\bar{A}, \bar{F}) deve essere stabilizzabile e la coppia (\bar{A}, \bar{C}) deve essere rilevabile. Ancora una volta, in base alle definizioni date per le quantità "barrate", è facile verificare che queste proprietà sono soddisfatte se e solo se la coppia (A, F) è stabilizzabile e la coppia (A, C) è rilevabile. Ad esempio, (\bar{A}, \bar{F}) è stabilizzabile se e solo se

$$\rho\left(\begin{bmatrix} \lambda I - A & -F & | & 0 \\ 0 & \lambda I & | & I \end{bmatrix}\right) = n + s, \quad \forall |\lambda| \geq 1,$$

e vista la struttura delle matrici nei blocchi, questo è vero se e solo se

$$\rho([\lambda I - A \quad -F]) = n, \quad \forall |\lambda| \geq 1,$$

e cioè se e solo se è stabilizzabile la coppia (A, F) . Analogamente si dimostra l'equivalenza della stabilizzabilità delle coppie (\bar{A}, \bar{C}) e (A, C) .

La matrice $(I - \bar{K}_\infty \bar{C})\bar{A}$ si scrive

$$\begin{bmatrix} (I - K_\infty C)A & (I - K_\infty C)F \\ -H_\infty CA & -H_\infty CF \end{bmatrix},$$

ed è stabile se e solo se la coppia (A, F) è stabilizzabile e la coppia (A, C) è rilevabile.

4.12 I minimi quadrati ricorsivi attraverso il filtro di Kalman

Nel capitolo sulla Teoria della Stima è stato analizzato, tra l'altro, il problema della soluzione di un sistema di equazioni algebriche lineari in cui il vettore dei termini noti in generale non è contenuto nello spazio immagine della matrice dei coefficienti (2.8.14). In questa

sezione viene ripresa l'impostazione del problema secondo cui il vettore delle incognite x non è una variabile aleatoria (2.8.14). Si tratta di un problema di stima di massima verosimiglianza per il vettore di parametri $x \in \mathbb{R}^n$ a partire dal risultato y della variabile aleatoria Y definita da

$$Y = Hx + V,$$

in cui H è la matrice dei coefficienti e V è un rumore additivo gaussiano a media nulla. Se Ψ_V è la covarianza del rumore, la stima di massima verosimiglianza è data da

$$\hat{x} = (H^T \Psi_V^{-1} H)^{-1} H^T \Psi_V^{-1} y. \quad (4.12.1)$$

È stato dimostrato inoltre che questa soluzione è anche quella che minimizza l'indice quadratico

$$(y - Hx)^T \Psi_V^{-1} (y - Hx).$$

Pertanto, qualora si assuma $\Psi_V = I$, questa soluzione del problema viene anche detta ai **minimi quadrati**. Quando Ψ_V è diversa dall'identità si parla di minimi quadrati pesati, in quanto la Ψ_V^{-1} pesa in maniera differente le diverse componenti dell'errore $Y - Hx$.

Questo, così come è formulato, è un problema di tipo **statico**. Di esso si può dare una formulazione **dinamica** supponendo che le componenti del vettore Y dei termini noti siano delle osservazioni che vengano acquisite in istanti di tempo successivi. Sia quindi

$$Y(k) = H(k)x + V(k),$$

$$\text{con } \begin{aligned} Y(k) &= \begin{bmatrix} Y(k-1) \\ y(k) \end{bmatrix}, & H(k) &= \begin{bmatrix} H(k-1) \\ h(k) \end{bmatrix}, \\ V(k) &= \begin{bmatrix} V(k-1) \\ v(k) \end{bmatrix}. \end{aligned} \quad (4.12.2)$$

Definendo per semplicità la matrice peso $W(k) \triangleq \Psi_V^{-1}(k)$, la soluzione ai minimi quadrati in due istanti successivi k e $k+1$ è

data da

$$\begin{aligned}\hat{x}(k) &= (H^T(k)W(k)H(k))^{-1}H^T(k)W(k)Y(k), \\ \hat{x}(k+1) &= (H^T(k+1)W(k+1)H(k+1))^{-1} \\ &\quad \cdot H^T(k+1)W(k+1)Y(k+1).\end{aligned}$$

Si osservi che ad ogni istante deve essere eseguita l'inversione di una nuova matrice $n \times n$, e che al crescere di k aumenta il numero di moltiplicazioni e addizioni da svolgere, poiché aumentano le dimensioni delle matrici in gioco. Inoltre occorre disporre di grandi capacità di memoria per conservare i dati relativi ad $H(k)$ e $Y(k)$.

L'algoritmo dei **minimi quadrati ricorsivi** è una tecnica che, con una ridotta complessità numerica ed una minore necessità di memoria, consente di calcolare la stima ottima $\hat{x}(k+1)$ utilizzando la sola conoscenza della stima all'istante precedente $\hat{x}(k)$, della matrice $h(k+1)$ e del vettore $y(k+1)$.

Lo sviluppo delle equazioni dei minimi quadrati ricorsivi nei testi viene generalmente fatto a partire da una formula che consente il calcolo ricorsivo della matrice

$$S(k) \triangleq (H^T(k)W(k)H(k))^{-1}. \quad (4.12.3)$$

Generalmente, per semplicità, la *crescita* della matrice peso $W(k)$ viene supposta del tipo

$$W(k+1) = \begin{bmatrix} W(k) & 0 \\ 0 & w(k+1) \end{bmatrix}, \quad (4.12.4)$$

dando luogo in questo modo ad una matrice diagonale a blocchi. Si ha così l'equazione ricorsiva

$$\begin{aligned}S(k+1) &= (H^T(k)W(k)H(k) + h^T(k+1)w(k+1)h(k+1))^{-1} = \\ &= (S(k)^{-1} + h^T(k+1)w(k+1)h(k+1))^{-1} = \\ &= (I + S(k)h^T(k+1)w(k+1)h(k+1))^{-1}S(k).\end{aligned} \quad (4.12.5)$$

Occorre a questo punto derivare una relazione ricorsiva tra $\hat{x}(k+1)$ e $\hat{x}(k)$. Una strada è quella di sviluppare l'espressione

$$\hat{x}(k+1) = S(k+1)H^T(k+1)W(k+1)Y(k+1)$$

tenendo conto delle suddivisioni a blocchi precedentemente indicate.

In questa sezione si preferisce ricavare le equazioni dei minimi quadrati ricorsivi utilizzando le formule del filtro di Kalman. A tal fine basta riscrivere il problema in forma di sistema dinamico. Poiché il vettore di parametri x è costante, allora le equazioni di stato saranno semplicemente

$$\begin{aligned} x(k+1) &= x(k), \\ y(k) &= h(k)x(k) + v(k). \end{aligned} \quad (4.12.6)$$

Sia $\psi_V(k) = E\{v(k)v(k)^T\} = w^{-1}(k)$, e sia $G(k) = w^{-1/2}(k)$, allora il sistema possiede effettivamente la forma standard

$$\begin{aligned} x(k+1) &= A(k)x(k) + F(k)N_k, \\ y(k) &= C(k)x(k) + G(k)N_k, \end{aligned}$$

con

$$A(k) = I, \quad F(k) = 0, \quad C(k) = h(k), \quad G(k)G^T(k) = w^{-1}(k),$$

e N_k sequenza aleatoria gaussiana bianca standard. L'equazione del filtro di Kalman diventa allora

$$\hat{x}(k+1) = \hat{x}(k) + K(k+1)(y(k+1) - h(k+1)\hat{x}(k)),$$

in cui $K(k+1)$ si calcola attraverso le equazioni di Riccati nella formulazione (1FK) o (2FK). In entrambe si ha che la covarianza dell'errore di predizione è pari a

$$P_p(k+1) = A(k)P(k)A^T(k) + F(k)F^T(k) = P(k),$$

ovvero coincide con la covarianza dell'errore di stima all'istante precedente. Nella formulazione (2FK) si ha

$$\textcircled{3} \quad P(k+1) = [I + P_p(k+1)C(k+1)]^T \cdot (G(k+1)G^T(k+1))^{-1} C(k+1)]^{-1} P_p(k+1),$$

che applicata al problema in esame diventa

$$P(k+1) = [I + P(k)h(k+1)^T w(k+1)h(k+1)]^{-1} P(k). \quad (4.12.7)$$

Confrontando quest'ultima con l'equazione (4.12.5) si vede che, se $P(0) = S(0)$, allora per ogni k si ha $P(k) = S(k)$. La matrice $K(k+1)$ è data da

$$\textcircled{4} \quad K(k+1) = P(k+1)C^T(k+1)(G(k+1)G^T(k+1))^{-1},$$

ovvero nel nostro caso

$$K(k+1) = P(k+1)h^T(k+1)w(k+1).$$

Si ha in definitiva il seguente algoritmo dei minimi quadrati ricorsivi

$$\begin{aligned} P(k+1) &= [I + P(k)h^T(k+1)w(k+1)h(k+1)]^{-1} P(k), \\ \hat{x}(k+1) &= \hat{x}(k) + P(k+1)h^T(k+1)w(k+1) \cdot (y(k+1) - h(k+1)\hat{x}(k)). \end{aligned} \quad (4.12.8)$$

In questa formulazione per calcolare $P(k+1)$ ad ogni passo occorre invertire una matrice $(n \times n)$. Seguendo invece la formulazione del filtro di Kalman 1K ad ogni passo occorre invertire una matrice $(q \times q)$, dove q è la dimensione del vettore delle osservazioni $y(k)$. L'equazione di Riccati che fornisce il guadagno di Kalman è

$$\textcircled{5} \quad K(k+1) = P_p(k+1)C^T(k+1) \cdot (C(k+1)P_p(k+1)C^T(k+1) + G(k+1)G^T(k+1))^{-1},$$

che particolarizzata al nostro problema diventa

$$K(k+1) = P(k)h^T(k+1) \cdot \left(h(k+1)P(k)h^T(k+1) + w^{-1}(k+1) \right)^{-1},$$

o anche

$$K(k+1) = P(k)h^T(k+1)w(k+1) \cdot \left(h(k+1)P(k)h^T(k+1)w(k+1) + I \right)^{-1},$$

e

$$\begin{aligned} P(k+1) &= [I - K(k+1)C(k+1)]P_p(k+1) = \\ &= [I - K(k+1)h(k+1)]P(k). \end{aligned}$$

Come in tutti gli algoritmi ricorsivi si presenta a questo punto il problema dell'inizializzazione. Si ricordi che in questa formulazione il vettore x non è una variabile aleatoria, e quindi non esistono un valore atteso e una covarianza con cui innescare l'algoritmo. Il modo più rigoroso di procedere consiste nell'inizializzare l'algoritmo a partire da un certo istante $\bar{k} \geq 0$ risolvendo, per un volta, il problema in forma non ricorsiva

$$\begin{aligned} P(\bar{k}) &= (H^T(\bar{k})W(\bar{k})H(\bar{k}))^{-1}, \\ \hat{x}(\bar{k}) &= P(\bar{k})H^T(\bar{k})W(\bar{k})Y(\bar{k}). \end{aligned} \quad (4.12.9)$$

Spesso però, nelle applicazioni, il risultato della stima deve essere fornito fin dall'istante iniziale $k = 0$, per cui non è possibile attendere il primo istante \bar{k} in cui è possibile calcolare $P(\bar{k}) = (H^T(\bar{k})W(\bar{k})H(\bar{k}))^{-1}$. Quello che si fa comunemente è di usare una stima iniziale \bar{x} disponibile grazie a qualche informazione a priori sul problema specifico in esame, e di inizializzare $P(0)$ ad un valore *sufficientemente elevato*. La stima all'istante 0 sarà quindi data da

$$\hat{x}(0) = \bar{x} + P(0)h^T(0)w(0)(y(0) - h(0)\bar{x}). \quad (4.12.10)$$

Si ricordi che il vettore incognito x non è una variabile aleatoria e quindi $P(0)$ non è una matrice di covarianza e \bar{x} non è un valore atteso. Nel funzionamento dell'algoritmo però la $P(k)$ gioca un ruolo analogo a quello della matrice di covarianza dell'errore di stima nel filtro di Kalman. Questo fatto suggerisce quindi di scegliere $P(0)$ abbastanza grande per modellare, seppure in modo qualitativo, l'incertezza nella stima a priori \bar{x} . Si osservi inoltre che nell'espressione (4.12.10) il peso del secondo termine (termine di correzione) risulta essere tanto più elevato rispetto al primo termine (di predizione) quanto più $P(0)$ è elevato. In altre parole più $P(0)$ è elevato e meno l'errore commesso sulla stima a priori \bar{x} influenza l'errore sulle stime successive $\hat{x}(k)$.

Quando la matrice di peso $W(k)$ non è la matrice identità si parla genericamente di minimi quadrati pesati. Nella maggior parte delle situazioni però, se non si hanno informazioni sufficienti per pesare in maniera selettiva le diverse componenti dell'errore $Y(k) - H(k)x$, si adotta per $W(k)$ la matrice identità.

Stima di stato e parametri nei sistemi lineari a tempo discreto

IN questo capitolo si prenderà in esame il problema della stima simultanea dello stato e dei parametri nei sistemi lineari e stazionari a tempo discreto in presenza di rumore di stato e di osservazione. Verranno considerati sistemi lineari in cui le matrici che definiscono le equazioni di transizione e di osservazione dipendono da un vettore di parametri incogniti. Questi ultimi non sono modellati come variabili aleatorie e per la loro stima verrà utilizzato il criterio della massima verosimiglianza. Il vettore di stato invece continuerà ad essere una variabile aleatoria gaussiana in quanto i rumori di stato e di osservazione ancora una volta sono costituiti da sequenze aleatorie gaussiane. Per la stima dello stato si adotterà ancora il criterio della minima varianza dell'errore di stima, ricorrendo quindi all'impiego del filtro di Kalman.

5.1 Impostazione del funzionale di verosimiglianza

Si consideri un sistema lineare e stazionario il cui modello è noto a meno di un vettore di parametri $\vartheta \in \mathbb{R}^r$, secondo la seguente struttura

$$\begin{aligned} x(k+1) &= A(\vartheta)x(k) + B(\vartheta)u(k) + F(\vartheta)N_k, & F(\vartheta)G^T(\vartheta) &= 0, \\ y(k) &= C(\vartheta)x(k) + D(\vartheta)u(k) + G(\vartheta)N_k. \end{aligned} \tag{5.1.1}$$

Le ipotesi a cui soddisfano la sequenza degli ingressi aleatori N_k e lo stato iniziale $x(0)$ sono le stesse del Capitolo 4, necessarie per lo sviluppo del filtro di Kalman. Il problema della stima simultanea dello stato e dei parametri, date le osservazioni $y(k)$ dall'istante 0 fino all'istante T , può essere formulato e risolto seguendo le indicazioni riportate alla fine del Capitolo 2, nel paragrafo che tratta della stima simultanea di variabili alatorie e dei parametri di una distribuzione. Occorre pertanto prima stimare il vettore dei parametri secondo il criterio della massima verosimiglianza e poi lo stato secondo il criterio della minima varianza. Per risolvere il primo problema occorre trovare l'espressione della densità di probabilità dell'aggregato del vettore delle osservazioni, definito come

$$Y_T = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(T) \end{bmatrix}. \quad (5.1.2)$$

Y_T è un vettore gaussiano, poiché gaussiani sono lo stato iniziale e la sequenza dei rumori di stato e di osservazione, ed è possibile per esso calcolare il valor medio e la matrice di covarianza in funzione delle matrici del sistema (5.1.1). A tal fine è sufficiente applicare le definizioni di valor medio e covarianza all'espressione dell'osservazione al generico istante k

$$\begin{aligned} y(k) = & C(\vartheta)A^k(\vartheta)x(0) + \\ & + D(\vartheta)u(k) + \sum_{j=0}^{k-1} C(\vartheta)A^{k-j-1}(\vartheta)B(\vartheta)u(j) + \\ & + G(\vartheta)N_k + \sum_{j=0}^{k-1} C(\vartheta)A^{k-j-1}(\vartheta)F(\vartheta)N_k. \end{aligned} \quad (5.1.3)$$

Il valor medio, indicato con $m_{Y_T}(\vartheta)$, è un vettore a $q(T+1)$ componenti, e la matrice di covarianza, indicata con $\Psi_{Y_T}(\vartheta)$ ha dimensioni $q(T+1) \times q(T+1)$. Il calcolo *simbolico* di queste matrici

non presenta difficoltà, e viene omesso, mentre il calcolo *numerico* si presenta tanto più oneroso quando maggiore è il numero delle osservazioni.

La terna statistica nella quale formulare il problema è

$$(\mathbb{R}^{q(T+1)}, \mathcal{B}, \mathcal{P})$$

dove \mathcal{B} è la σ -algebra di Borel su $\mathbb{R}^{q(T+1)}$ e \mathcal{P} è la famiglia delle misure di probabilità, parametrizzata da ϑ , indotta da Y_T su $(\mathbb{R}^{q(T+1)}, \mathcal{B}^{q(T+1)})$. Date le caratteristiche del vettore aleatorio Y_T ogni misura $P_\vartheta \in \mathcal{P}$ è gaussiana, con media $m_{Y_T}(\vartheta)$ e covarianza $\Psi_{Y_T}(\vartheta)$. Poiché inoltre tale misura è assolutamente continua rispetto alla misura di Lebesgue per ogni ϑ , il funzionale di verosimiglianza coincide con la densità di probabilità della misura P_ϑ calcolata in corrispondenza alle osservazioni Y_T , cioè

$$\begin{aligned} \mathcal{L}_{Y_T}(\vartheta) &= \frac{dP_\vartheta}{d\lambda} \Big|_{Y_T} = p(Y_T; \vartheta) = \\ &= \frac{1}{(2\pi)^{q(T+1)/2} |\Psi_{Y_T}(\vartheta)|^{1/2}} \cdot \\ &\quad \cdot \exp\left(-\frac{1}{2}(Y_T - m_{Y_T}(\vartheta))^T \Psi_{Y_T}^{-1}(\vartheta)(Y_T - m_{Y_T}(\vartheta))\right). \end{aligned} \tag{5.1.4}$$

Il valore del parametro ϑ che rende massimo $\mathcal{L}_{Y_T}(\vartheta)$ costituisce la stima di massima verosimiglianza. Tale stima si ottiene annullando il gradiente della (5.1.4) rispetto a ϑ . In generale si dovrà ricorrere ad algoritmi numerici per la ricerca del massimo, in quanto non è pensabile di poter ricavare una soluzione in forma chiusa.

Poiché la maggior parte degli algoritmi numerici per la ricerca del massimo di una funzione si fonda sull'algoritmo del gradiente, che come è noto costituisce la direzione di massima salita per la funzione, è importante valutare la complessità del calcolo del gradiente del funzionale di verosimiglianza. A tal proposito si può osservare che la dimensione della matrice di covarianza $\Psi_{Y_T}(\vartheta)$ cresce al crescere di T , ed in situazioni di interesse applicativo potrebbe avere dimensioni proibitive, considerando che nell'espressione del funzionale compare

sia la radice quadrata del determinante di tale matrice che la sua inversa. L'elevata complessità, intesa come numero di calcoli elementari, di questo approccio consiglia di cercare strade alternative più adatte ad una implementazione in tempo reale dell'algoritmo di stima.

Ciò che rende non percorribile nella pratica l'approccio sopra descritto è il fatto che la matrice di covarianza è *piena*, e non è dotata di alcuna struttura che ne semplifichi il calcolo del determinante e dell'inversa.

Una drastica semplificazione nella soluzione del problema della stima dei parametri si ottiene ricordando la definizione di sequenza di innovazione dell'uscita e la sua proprietà di avere lo stesso contenuto informativo della sequenza delle osservazioni (Teorema di Mitter).

L'innovazione dell'uscita all'istante k è stata definita nel capitolo 4 per il solo sottosistema stocastico. Si ricordi a tale proposito la scomposizione di un modello del tipo (5.1.1) in un sottosistema deterministico ed in un sottosistema stocastico nel modo seguente

$$\begin{aligned} x_d(k+1) &= A(\vartheta)x_d(k) + B(\vartheta)u(k), & x_d(0) &= E\{x(0)\} \\ y_d(k) &= C(\vartheta)x_d(k) + D(\vartheta)u(k) \end{aligned} \quad (5.1.5)$$

$$\begin{aligned} x_s(k+1) &= A(\vartheta)x_s(k) + F(\vartheta)N_k, & F(\vartheta)G^T(\vartheta) &= 0 \\ y_s(k) &= C(\vartheta)x_s(k) + G(\vartheta)N_k, & x_s(0) &= x(0) - E\{x(0)\}. \end{aligned} \quad (5.1.6)$$

L'innovazione dell'uscita per il sottosistema stocastico è stata definita come

$$\nu_o(k) = E\{y_s(k)|\mathcal{F}_k^Y\} - E\{y_s(k)|\mathcal{F}_{k-1}^Y\}. \quad (5.1.7)$$

e può essere riscritta come

$$\nu_o(k) = y_s(k) - C(\vartheta)A(\vartheta)\hat{x}_s(k-1). \quad (5.1.8)$$

Nel capitolo 4 è stato dimostrato che la sequenza delle innovazioni è a media nulla, è gaussiana ed è bianca. La covarianza dell'innovazione

dell'uscita all'istante k risulta essere

$$\Psi_{\nu_o}(k) = E\{\nu_o(k)\nu_o^T(k)\} = C(\vartheta)P_p(k; \vartheta)C^T(\vartheta) + G(\vartheta)G^T(\vartheta). \quad (5.1.9)$$

in cui $P_p(k; \vartheta)$ è la covarianza dell'errore di predizione dello stato all'istante k , ed è funzione del vettore ϑ .

Poiché $y_s(k) = y(k) - y_d(k)$, per la sequenza di innovazione dell'uscita si ha la seguente espressione

$$\begin{aligned} \nu_o(k) &= y(k) - y_d(k) - C(\vartheta)A(\vartheta)\hat{x}_s(k-1) = \\ &= y(k) - C(\vartheta)x_d(k) - D(\vartheta)u(k) - C(\vartheta)A(\vartheta)\hat{x}_s(k-1) = \\ &= y(k) - C(\vartheta)A(\vartheta)x_d(k-1) - C(\vartheta)B(\vartheta)u(k-1) - \\ &\quad - D(\vartheta)u(k) - C(\vartheta)A(\vartheta)\hat{x}_s(k-1) \end{aligned} \quad (5.1.10)$$

e poiché $\hat{x} = x_d + \hat{x}_s$ si ottiene

$$\nu_o(k) = y(k) - D(\vartheta)u(k) - C(\vartheta)A(\vartheta)\hat{x}(k-1) - C(\vartheta)B(\vartheta)u(k-1). \quad (5.1.11)$$

Da questa ultima espressione risulta evidente che $\nu_o(k)$ è un dato disponibile, come funzione di ϑ , in quanto $y(k)$ è l'uscita misurata, l'ingresso deterministico $u(k)$ è noto e $\hat{x}(k-1)$ è la stima dello stato fornita dal filtro di Kalman.

Si definisca a questo punto l'aggregato del vettore delle innovazioni

$$\mathcal{O}_T = \begin{bmatrix} \nu_o(0) \\ \nu_o(1) \\ \vdots \\ \nu_o(T) \end{bmatrix}; \quad (5.1.12)$$

che, per quanto detto nel capitolo 4, ha lo stesso contenuto informativo del vettore delle osservazioni Y_T . Poiché per $\nu_o(k)$ ha le stesse dimensioni dell'osservazione $y(k)$ si ha che $\mathcal{O}_T \in \mathbb{R}^{q(T+1)}$. Inoltre \mathcal{O}_T è un vettore gaussiano a media nulla, la cui covarianza è

diagonale a blocchi

$$\Psi_{\mathcal{O}_T}(\vartheta) = E\{\mathcal{O}_T^T \mathcal{O}_T\} = \begin{bmatrix} \Psi_{\nu_o}(0) & 0 & \cdots & 0 \\ 0 & \Psi_{\nu_o}(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \Psi_{\nu_o}(T) \end{bmatrix} \quad (5.1.13)$$

La terna statistica in esame è

$$(\mathbb{R}^{q(T+1)}, \mathcal{B}, \mathcal{P})$$

dove \mathcal{B} è la σ -algebra di Borel su $\mathbb{R}^{q(T+1)}$ e \mathcal{P} è la famiglia delle misure di probabilità, parametrizzata da ϑ , indotta da \mathcal{O}_T su $(\mathbb{R}^{q(T+1)}, \mathcal{B}^{q(T+1)})$. Anche in questo caso ogni misura $P_\vartheta \in \mathcal{P}$ è gaussiana, e quindi assolutamente continua rispetto alla misura di Lebesgue. Il funzionale di verosimiglianza è pertanto la densità di probabilità della misura \mathcal{P}_ϑ calcolata in corrispondenza ad \mathcal{O}_T

$$\mathcal{L}_{\mathcal{O}_T}(\vartheta) = \frac{d\mathcal{P}_\vartheta}{d\lambda} \Big|_{\mathcal{O}_T} = \frac{1}{(2\pi)^{q(T+1)/2} |\Psi_{\mathcal{O}_T}(\vartheta)|^{1/2}} \cdot \exp\left(-\frac{1}{2} \mathcal{O}_T^T \Psi_{\mathcal{O}_T}(\vartheta)^{-1} \mathcal{O}_T\right) \quad (5.1.14)$$

e con le opportune sostituzioni si ottiene

$$\mathcal{L}_{\mathcal{O}_T}(\vartheta) = \frac{1}{(2\pi)^{q(T+1)/2} \prod_{k=0}^T |\Psi_{\nu_o}(k; \vartheta)|^{1/2}} \cdot \exp\left\{-\frac{1}{2} \sum_{k=1}^T \nu_o^T(k) \Psi_{\nu_o}^{-1} \nu_o(k)\right\}. \quad (5.1.15)$$

La stima di massima verosimiglianza $\hat{\vartheta}$ è il vettore che $\mathcal{L}_{\mathcal{O}_T}(\vartheta)$ che rende massima la (5.1.15). D'altra parte massimizzare $\mathcal{L}_{\mathcal{O}_T}(\vartheta)$ equivale a minimizzare il suo logaritmo cambiato di segno, eliminando tutte le costanti moltiplicative. Si può pertanto definire

il seguente indice di costo da minimizzare per ottenere la stima di massima verosimiglianza

$$\begin{aligned}
 J(\vartheta) = & \frac{1}{2} \sum_{k=0}^T \log |\Psi_{\nu_o}(k)| + \\
 & + \frac{1}{2} \sum_{k=0}^T [y(k) - Du(k) - CBu(k-1) - CA\hat{x}(k-1)]^T \Psi_{\nu_o}^{-1}(k) \cdot \\
 & \cdot [y(k) - Du(k) - CBu(k-1) - CA\hat{x}(k-1)],
 \end{aligned} \tag{5.1.16}$$

(la dipendenza da ϑ è stata omessa per ragioni di spazio).

Algoritmi numerici per la minimizzazione di funzioni di questo tipo possono essere implementati in maniera efficiente. Si noti che le matrici nell'espressione (5.1.16), di cui occorre calcolare i determinanti e le inverse, sono tutte di dimensioni $q \times q$. Il calcolo risulta particolarmente facilitato quando l'uscita è scalare.

Per semplificare ulteriormente il calcolo del funzionale di verosimiglianza è possibile operare alcune approssimazioni. Ad esempio è possibile considerare le matrici di covarianza delle innovazioni *a regime*

$$\Psi_{\nu_o}(\infty; \vartheta) = C(\vartheta)P_p(\infty; \vartheta)C^T(\vartheta) + G(\vartheta)G^T(\vartheta); \tag{5.1.17}$$

in modo da avere il funzionale semplificato

$$\begin{aligned}
 \bar{J}(\vartheta) = & \frac{(T+1)}{2} \log |\Psi_{\nu_o}(\infty)| + \\
 & + \frac{1}{2} \sum_{k=0}^T [y(k) - Du(k) - CBu(k-1) - CA\hat{x}(k-1)]^T \Psi_{\nu_o}^{-1}(\infty) \cdot \\
 & \cdot [y(k) - Du(k) - CBu(k-1) - CA\hat{x}(k-1)].
 \end{aligned} \tag{5.1.18}$$

Questa semplificazione trova corretta applicazione quando si voglia impiegare per la stima le osservazioni non a partire da 0 ma da un istante k_0 sufficientemente elevato da poter considerare a regime le equazioni di Riccati.

Una volta calcolata la stima di massima verosimiglianza $\hat{\vartheta}$ del vettore di parametri, la stima ottima dello stato si ottiene utilizzando il filtro di Kalman, utilizzando le matrici del sistema (5.1.1) valutate in $\hat{\vartheta}$.

Anche se la formulazione del problema utilizzando le innovazioni dell'uscita risulta più semplice rispetto alla formulazione nello spazio delle osservazioni, in generale gli algoritmi numerici per la ricerca del massimo della funzione $J(\vartheta)$ possono essere molto complessi e non adatti ad una implementazione in linea.

Algoritmi di stima più efficienti possono essere ottenuti adottando delle semplificazioni. Ad esempio, si osservi che al crescere di T aumenta la dimensione del vettore delle innovazioni \mathcal{O}_T , e pertanto aumenta il numero di calcoli che occorre eseguire per ricavare la stima ottima di ϑ . Anche la memoria richiesta all'elaboratore aumenta con l'avanzare del tempo. Pertanto nella letteratura scientifica si è cercato di introdurre opportune semplificazioni al fine di ottenere una implementazione ricorsiva della stima dei parametri, ovvero che consenta di esprimere la stima ottima all'istante $T+1$ in funzione della stima all'istante T , dell'osservazione all'istante T e, eventualmente, di un numero limitato di osservazioni dell'uscita precedenti l'istante T .

5.2 Stima di stato e parametri con il filtro di Kalman Estesio

Un algoritmo non rigoroso, ma che in molti casi fornisce buoni risultati, per la stima simultanea di stato e parametri è il filtro di Kalman Estesio applicato ad un'opportuna estensione del sistema lineare in esame. Si consideri nuovamente il sistema lineare (5.1.1) in cui le sole matrici A, B, C e D dipendono da un vettore di parametri incogniti ϑ . Il vettore di parametri, pur essendo costante, può essere annoverato tra le variabili di stato del sistema, definendo opportunamente la funzione di transizione $\vartheta(k+1) = \vartheta(k)$. È

possibile così creare uno stato esteso

$$x_e(k) = \begin{bmatrix} x(k) \\ \vartheta(k) \end{bmatrix}$$

ed un sistema esteso

$$\begin{aligned} x_e(k+1) &= f_e(x_e(k), u(k), k) + F_e(k)N_k, & F_e(j)G^T(k) &= 0, \\ y(k) &= h_e(x_e(k), u(k), k) + G(k)N_k, & \forall j, k, \end{aligned} \quad (5.2.1)$$

in cui

$$\begin{aligned} f_e(x_e(k), u(k), k) &= \begin{bmatrix} A(\vartheta(k))x(k) + B(\vartheta(k))u(k) \\ \vartheta(k) \end{bmatrix}, \\ h_e(x_e(k), u(k), k) &= C(\vartheta(k))x(k) + D(\vartheta(k))u(k), \\ F_e(k) &= \begin{bmatrix} F(k) \\ 0 \end{bmatrix}. \end{aligned}$$

Il sistema esteso, non lineare, ha esattamente lo stesso comportamento del sistema lineare originario ammesso che $\vartheta(0)$ sia correttamente inizializzato al valore vero del vettore di parametri.

L'applicazione del filtro di Kalman Esteso al sistema esteso (5.2.1) fornisce ad ogni istante k una stima dell'intero vettore di stato esteso $x_e(k)$, e quindi fornisce simultaneamente la stima dello stato $\hat{x}(k|k)$ ed la stima del vettore di parametri $\hat{\vartheta}(k|k)$.

Osservazione 5.2.1 – Utilizzando lo stesso accorgimento il filtro di Kalman Esteso può essere applicato anche per la stima simultanea di stato e parametri in sistemi non lineari del tipo

$$\begin{aligned} x(k+1) &= f(x(k), u(k), \vartheta, k) + F(k)N_k, & F(j)G^T(k) &= 0, \\ y(k) &= h(x(k), u(k), \vartheta, k) + G(k)N_k, & \forall j, k, \end{aligned} \quad (5.2.2)$$

Con la stessa definizione di stato esteso $x_e(k) = [x^T(k), \vartheta^T(k)]^T$ si ha nuovamente un sistema non lineare con la struttura (5.2.1), ma

con la seguente definizione per le funzioni $f_e(\cdot)$, $h_e(\cdot)$ e per la matrice $F_e(k)$

$$f_e(x_e(k), u(k), k) = \begin{bmatrix} f(x(k), u(k), \vartheta(k), k) \\ \vartheta(k) \end{bmatrix}, \quad F_e(k) = \begin{bmatrix} F(k) \\ 0 \end{bmatrix},$$

$$h_e(x_e(k), u(k), k) = h(x(k), u(k), \vartheta(k), k).$$

△

Esempio 5.2.2 – Dato un sistema lineare e stazionario del secondo ordine strettamente proprio si vuole stimare la risposta impulsiva $w(k)$ a partire dalla conoscenza dell'ingresso $u(k)$ e dell'uscita osservata $y(k)$.

Date le ipotesi la risposta impulsiva può senz'altro essere scritta come

$$w(k) = \begin{cases} 0 & \text{per } k = 0, \\ R_1 a_1^k + R_2 a_2^k & \text{per } k \geq 1. \end{cases} \quad (5.2.3)$$

Una realizzazione per il sistema in esame è data dalla forma diagonale

$$A = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}, \quad B = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}, \quad (5.2.4)$$

$$C = [1 \quad 1]$$

Assumendo rumori indipendenti per ciascuna componente dello stato e sull'uscita, si ha

$$F = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{bmatrix}, \quad G = [0 \quad 0 \quad \sigma_3], \quad (5.2.5)$$

in cui le deviazioni standard sono note.

Definendo il vettore di parametri ϑ come

$$\vartheta = [a_1 \quad a_2 \quad R_1 \quad R_2]^T, \quad (5.2.6)$$

e lo stato esteso come

$$x_e = [x_1 \quad x_2 \quad \vartheta_1 \quad \vartheta_2 \quad \vartheta_3 \quad \vartheta_4], \quad (5.2.7)$$

si ha

$$f_e(x_e, u) = \begin{bmatrix} \vartheta_1 x_1 + \vartheta_3 u \\ \vartheta_2 x_2 + \vartheta_4 u \\ \vartheta_1 \\ \vartheta_2 \\ \vartheta_3 \\ \vartheta_4 \end{bmatrix}, \quad F = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (5.2.8)$$

$$h_e(x_e) = x_1 + x_2, \quad G = [0 \quad 0 \quad \sigma_3].$$

Le matrici

$$\bar{A}(x_e) \triangleq \frac{\partial f_e}{\partial x_e}, \quad e \quad \bar{C}(x_e) \triangleq \frac{\partial h_e}{\partial x_e}, \quad (5.2.9)$$

sono date da

$$\bar{A} = \begin{bmatrix} \vartheta_1 & 0 & x_1 & 0 & u & 0 \\ 0 & \vartheta_2 & 0 & x_2 & 0 & u \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

e

$$\bar{C} = [1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0]. \quad (5.2.10)$$

Si supponga ora di avere ulteriori informazioni sul sistema. Ad esempio da prove sperimentali si è ricavato che a regime la risposta al gradino unitario vale $y(\infty) = 2$. La trasformata z della risposta al gradino vale

$$Y(z) = \left(\frac{R_1}{z - a_1} + \frac{R_2}{z - a_2} \right) \frac{z}{z - 1} \quad (5.2.11)$$

e la condizione sulla risposta a regime si traduce in

$$y(\infty) = \lim_{z \rightarrow 1} \left(\frac{R_1}{z - a_1} + \frac{R_2}{z - a_2} \right) z, \quad (5.2.12)$$

e quindi

$$\frac{R_1}{z - a_1} + \frac{R_2}{z - a_2} = 2 \quad \Rightarrow \quad R_2 = (1 - a_2) \left(2 - \frac{R_1}{1 - a_1} \right). \quad (5.2.13)$$

Definendo il vettore di parametri $\vartheta = [a_1 \ a_2 \ R_1]^T$ si ha

$$R_2(\vartheta) = (1 - \vartheta_2) \left(2 - \frac{\vartheta_3}{1 - \vartheta_1} \right). \quad (5.2.14)$$

Lo stato esteso ora ha dimensione minore, e la funzione di transizione f_e è pari a

$$f_e(x_e, u) = \begin{bmatrix} \vartheta_1 x_1 + \vartheta_3 u \\ \vartheta_2 x_2 + R_2(\vartheta) u \\ \vartheta_1 \\ \vartheta_2 \\ \vartheta_3 \end{bmatrix}.$$

La matrice $\bar{A}(x_e)$ è ora pari a

$$\bar{A} = \begin{bmatrix} \vartheta_1 & 0 & x_1 & 0 & u \\ 0 & \vartheta_2 & \frac{\partial R_2}{\partial \vartheta_1} u & x_2 + \frac{\partial R_2}{\partial \vartheta_2} u & \frac{\partial R_2}{\partial \vartheta_3} u \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5.2.15)$$