Terminologia:
Sistemi LTITC: Sistemi Lineari Tempo Invarianti Tempo Continuo
Sistemi LTITD: Sistemi Lineari Tempo Invarianti Tempo Discreto

Dx = derivata di x rispetto al tempo

5.8 Verifiche dirette di Raggiungibilità

La condizione di completa raggiungibilità è verificabile per ispezione diretta delle matrici del sistema in alcuni casi.

Se la matrice dinamica $\mathbf{A}$ di un sistema SISO è diagonale, $\mathbf{A} = \mathsf{diag}\,(\lambda_1, \ldots, \lambda_n)$, la matrice di raggiungibilità è

$$\mathbf{R} = \left[\begin{array}{c|c|c|c|c} \mathbf{B}_1 & \lambda_1\mathbf{B}_1 & \lambda_1^2\mathbf{B}_1 & \cdots & \lambda_1^{n-1}\mathbf{B}_1 \\ \mathbf{B}_2 & \lambda_2\mathbf{B}_2 & \lambda_2^2\mathbf{B}_2 & \cdots & \lambda_2^{n-1}\mathbf{B}_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{B}_n & \lambda_n\mathbf{B}_n & \lambda_n^2\mathbf{B}_n & \cdots & \lambda_n^{n-1}\mathbf{B}_n \end{array}\right]$$

Perchè questa sia di pieno rango, è necessario che tutte le componenti di $\mathbf{B}$ siano diverse da zero, e che tutti gli autovalori siano distinti (si avrebbero altrimenti righe nulle o righe proporzionali). Queste condizioni sono anche sufficienti: infatti si può riscrivere

$$\mathbf{R} = \mathsf{diag}\,(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n) \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \lambda_n & \lambda_n^2 & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

ed il determinante della matrice a destra (che è una matrice di Vandermonde) non si annulla per λ_i distinti.

Nel caso più generale di sistema MIMO con forma di Jordan nota, la verifica di raggiungibilità può essere fatta ricorrendo al *Lemma P.B.H.* (Popov, Belevitch, Hautus):

Teorema Il sistema LTI con matrici $(\mathbf{A}, \mathbf{B})$ è raggiungibile se e solo se la matrice

$$\mathbf{P}(\lambda) = \left[\begin{array}{c|c} \lambda\mathbf{I} - \mathbf{A} & \mathbf{B} \end{array}\right] \tag{21}$$

ha rango pieno per ogni $\lambda \in \mathbb{C}$.

Dimostrazione. Basterà considerare $\lambda = \lambda_i \in s(\mathbf{A})$, $\forall i$. Supponiamo che per qualche λ_i, $\mathsf{rank}\,(P(\lambda_i)) < n$: allora esisterà un vettore $\mathbf{q} \neq 0$ tale che $\mathbf{q}^T\mathbf{P}(\lambda_i) = 0$, ovvero tale che al contempo

$$\mathbf{q}^T\mathbf{B} = 0 \;\text{ e } \mathbf{q}^T\mathbf{A} = \lambda_i\mathbf{q}^T \tag{22}$$

Postmoltiplicando la seconda per $\mathbf{B}$, si ha

$$\mathbf{q}^T\mathbf{A}\mathbf{B} = \lambda_i\mathbf{q}^T\mathbf{B} = 0$$

e postmoltiplicando per $\mathbf{AB}$ si ha

$$\mathbf{q}^T\mathbf{A}^2\mathbf{B} = \lambda_i\mathbf{q}^T\mathbf{A}\mathbf{B} = 0$$

Iterando questa procedura, si trova che

$$\mathbf{q}^T\left[\begin{array}{c|c|c|c} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{array}\right] = 0$$

per cui $\mathbf{R}$ non ha rango n, e il sistema non è raggiungibile.

Di converso, supponiamo che il sistema non sia raggiungibile; senza perdere in generalità possiamo assumere che $\mathbf{A}, \mathbf{B}$ siano in forma standard di raggiungibilità, e dobbiamo far vedere che, se $\mathbf{A}_N$ ha dimensione $n - r > 0$, allora

esiste un $\bar{\mathbf{q}}$ che soddisfa 22.
Per questo $\bar{\mathbf{q}}$ sarà

$$\bar{\mathbf{q}}^T\mathbf{B} = \bar{\mathbf{q}}^T \left[\begin{array}{c} \mathbf{B}_R \\ \hline \mathbf{0} \end{array}\right] = 0 \Rightarrow \bar{\mathbf{q}}^T = \left[\begin{array}{c|c} \mathbf{0} & \bar{\mathbf{q}}_R^T \end{array}\right]$$

e inoltre basterà scegliere

$$\bar{\mathbf{q}}_R^T\mathbf{A}_N = \lambda_i \bar{\mathbf{q}}_R^T$$

Applichiamo il lemma PBH al caso di una coppia $(\mathbf{A}, \mathbf{B})$ con $\mathbf{A}$ in forma di Jordan, con p miniblocchi:

$$\mathbf{A} = \left[\begin{array}{c|c|c} \begin{array}{cccc} \lambda_1 & 1 & \cdots & 0 \\ 0 & \lambda_1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \lambda_1 \end{array} & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \ddots & \mathbf{0} \\ \hline \mathbf{0} & \cdots & \begin{array}{cccc} \lambda_p & 1 & \cdots & 0 \\ 0 & \lambda_p & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \lambda_p \end{array} \end{array}\right] ; \quad \mathbf{B} = \left[\begin{array}{c} \mathbf{B}_{11} \\ \mathbf{B}_{12} \\ \vdots \\ \mathbf{B}_{1,m_1} \\ \hline \vdots \\ \hline \mathbf{B}_{p,1} \\ \mathbf{B}_{p,2} \\ \vdots \\ \mathbf{B}_{p,m_p} \end{array}\right]$$

Risulta che, per essere raggiungibile, le ultime righe $\mathbf{B}_{i,m_i}$ per ogni miniblocco corrispondente ad autovalori coincidenti, devono essere linearmente indipendenti. Infatti la matrice $\lambda_i\mathbf{I}-\mathbf{A}$ ha tante righe nulle quante sono i miniblocchi associati a λ_i, cioè la molteplicità geometrica μ_i di λ_i , ed il rango di $P(\lambda_i)$ non diminuisce se e solo se le corrispondenti righe estratte da $\mathbf{B}$ hanno rango μ_i.

In particolare, per un sistema SISO, è necessario che la molteplicità geometrica di tutti gli autovalori sia pari a uno, e che $\mathbf{B}$ abbia almeno tanti elementi diversi da zero quanti gli autovalori distinti di $\mathbf{A}$

Se in un sistema SISO vi sono più catene di autovalori generalizzati corrispondenti ad uno stesso autovalore, solo i modi di (al più) la più lunga di queste catene potranno apparire nella risposta forzata del sistema. In altri termini, nella f.d.t. il polo corrispondente apparirà con molteplicità pari (al più) alla dimensione del miniblocco di ordine più elevato.

Esempio: Per il sistema

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix}; \mathbf{C} = \begin{bmatrix} C_1 \\ C_2 \\ C_3 \end{bmatrix}^T ; \mathbf{D} = 0$$

si ha

$$G(s) = \frac{\mathbf{C}\begin{bmatrix} (\lambda+1)^2 & (\lambda+1) & 0 \\ 0 & (\lambda+1)^2 & 0 \\ 0 & 0 & (\lambda+1)^2 \end{bmatrix}\mathbf{B}}{(\lambda+1)^3} = \frac{C_1B_2 + (C_1B_1 + C_2B_2 + C_3B_3)(\lambda+1)}{(\lambda+1)^2}$$

Un sistema con μ_i miniblocchi associati ad un unico autovalore λ_i può essere raggiungibile solo se ha almeno μ_i ingressi.

5.9 Forma canonica di controllo

Per n sistema SISO con matrici $(\mathbf{A}_c, \mathbf{B}_c)$ nella particolare forma

$$\mathbf{A}_c = \left[\begin{array}{c|cccc} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \hline -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{array}\right]; \quad \mathbf{B}_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

(la forma di $\mathbf{A}_c$ si dice *compagna orizzontale inferiore*), la matrice di raggiungibilità vale

$$\mathbf{R}_c = \left[\begin{array}{c|c|c|c|c} 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & -a_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \star \\ 0 & 0 & 1 & \star & \star \\ 0 & 1 & -a_{n-1} & \star & \star \\ 1 & -a_{n-1} & -a_{n-2} + a_{n-1}^2 & \star & \star \end{array}\right]$$

(dove $\star$ indica elementi il cui calcolo esplicito è tralasciato), che è invertibile. Il sistema è quindi raggiungibile.

Dato un qualsiasi altro sistema SISO $(\mathbf{A}, \mathbf{B})$ con matrice di raggiungibilità $\mathbf{R}$:
se $(\mathbf{A}, \mathbf{B})$ è algebricamente equivalente a $(\mathbf{A}_c, \mathbf{B}_c)$ ($\mathbf{A} = \mathbf{T}\mathbf{A}_c\mathbf{T}^{-1}$, $\mathbf{B} = \mathbf{T}\mathbf{B}_c$), allora è raggiungibile: infatti $\mathbf{R} = \mathbf{T}\mathbf{R}_c$ ha rango pieno;
Se $(\mathbf{A}, \mathbf{B})$ è raggiungibile, $\mathbf{R}$ ha rango pieno, quindi esiste una $\mathbf{T}$ invertibile,

data da $\mathbf{T} = \mathbf{R}\mathbf{R}_c^{-1}$, che trasforma le coordinate in modo tale da avere $\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{A}_c$, $\mathbf{T}^{-1}\mathbf{B} = \mathbf{B}_c$.

La forma $(\mathbf{A}_c, \mathbf{B}_c)$ in cui possono essere messi tutti e soli i sistemi completamente raggiungibili si chiama *canonica di controllo*.

La matrice $\mathbf{R}_c^{-1}$ ha una forma particolare di facile memorizzazione (verificare per esercizio):

$$\mathbf{R}_c^{-1} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & 1 \\ a_2 & a_3 & \cdots & 1 & 0 \\ a_3 & a_4 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Il polinomio caratteristico di una matrice $\mathbf{A}_{c,n}$ in forma compagna orizzontale inferiore di dimensione n, si può calcolare ricorsivamente sviluppando il determinante secondo gli elementi della prima colonna:

$$\begin{aligned}
\det(s\mathbf{I} - \mathbf{A}_{c,n}) &= \det\left(\left[\begin{array}{c|cccc} s & -1 & 0 & \cdots & 0 \\ 0 & s & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ \hline a_0 & a_1 & a_2 & \cdots & s + a_{n-1} \end{array}\right]\right) = \\
&= (-1)^{n-1} a_0 (-1)^{n-1} + s\det(s\mathbf{I} - \mathbf{A}_{c,n-1}) = \\
&= a_0 + \det\left(\left[\begin{array}{c|cccc} s & -1 & 0 & \cdots & 0 \\ 0 & s & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ \hline a_1 & a_2 & a_3 & \cdots & s + a_{n-1} \end{array}\right]\right)
\end{aligned}$$

da cui

$$\det(s\mathbf{I} - \mathbf{A}_{c,n}) = a_0 + a_1 s + \cdots + a_{n-1}s^{n-1} + s^n = \pi(s)$$

I coefficienti dell'ultima riga della forma compagna orizzontale inferiore sono i coefficienti del polinomio caratteristico della matrioce stessa, ordinati secondo le potenze crescenti di s da sinistra a destra e cambiati di segno.

Si osservi anche esplicitamente che il polinomio caratteristico di una matrice non cambia per trasformazioni di similitudine:

$$\det(s\mathbf{I} - \mathbf{A}) = \det(s\mathbf{T}^{-1}\mathbf{T} - \mathbf{T}^{-1}\mathbf{A}\mathbf{T}) = \det(\mathbf{T}^{-1})\det(s\mathbf{I} - \mathbf{A}0\det(\mathbf{T})$$

Sulla base di questa e delle osservazioni precedenti, è facile portare un sistema $(\mathbf{A}, \mathbf{B})$ (LTI raggiungibile SISO) nella forma canonica di controllo. Si procede infatti come segue:

1. Si calcola il polinomio caratteristico di $\mathbf{A}$, $\det(s\mathbf{I}-\mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0$;
2. Si costruisce la matrice di raggiungibilità $\mathbf{R}$ e se ne verifica il rango;
3. Si costruiscono direttamente $\mathbf{A}_c, \mathbf{B}_c$;
4. Si calcola $\mathbf{T} = \mathbf{R}\mathbf{R}_c^{-1}$;
5. Si trova $\mathbf{C}_c = \mathbf{C}\mathbf{T}$

Ricordiamo che, quando il sistema fosse dato in termini della sua equazione normale, come spesso avviene derivando un modello fisico da equazioni differenziali o incrementali di ordine superiore, del tipo

$$D^n y(t) = \sum_{i=0}^{n-1} -a_i D^i y(t) + b_0 u(t)$$

il metodo più diretto per riscrivere il sistema nello spazio di stato consiste nel porre $\mathbf{x}_1 = y$, $\mathbf{x}_2 = \dot{y}$, etc., ottenendo così l'equazione di stato nella forma canonica di controllo.

Nel caso più generale in cui appaiano nella forma normale anche le derivate (differenze) di ordine più elevato dell'ingresso, cioè

$$D^n y(t) = \sum_{i=0}^{n-1} -a_i D^i y(t) + \sum_{i=0}^{m} b_j D^j u(t)$$

si può scrivere ancora il sistema in forma canonica di controllo utilizzando le proprietà dei sistemi lineari di sovrapposizione degli effetti degli ingressi e di risposta alla derivata di un segnale come derivata della risposta.

Infatti, detta $z(t)$ la soluzione della equazione "ausiliaria"

$$\begin{array}{rcl} D^n z(t) & = & \sum_{i=0}^{n-1} -a_i D^i z(t) + u(t), \\ & & D^i z(0) = z_0^{(i)}, \quad i = 1, \ldots, n-1 \end{array}$$

si ha che la soluzione corrispondente a $v(t) = \sum_{i=0}^{m} b_j D^j u(t)$ con le stesse condizioni iniziali, è data da

$$y(t) = \sum_{i=0}^{m} b_j D^j z(t).$$

Per dimostrarlo, basterà far vedere che nel caso $v(t) = Du(t)$, la funzione $y(t) = Dz(t)$ $(y(0) = Dz(0))$ risolve l'equazione

$$\begin{array}{rcl} D^n y(t) & = & \sum_{i=0}^{n-1} -a_i D^i y(t) + Du(t), \\ & & D^{i-1} y(0) = z_0^{(i)}, \quad i = 1, \ldots, n-1, \end{array}$$

il resto seguendo per reiterazione e dal principio di sovrapposizione degli effetti. Che questo sia vero lo si osserva facilmente sostituendo $y = Dz$ nella equazione, integrandone entrambe i membri tra 0 e t, ed eliminando le condizioni iniziali.

Si può quindi ottenere una forma di stato (canonica di controllo) che nel caso di sistemi strettamente propri ($m < n$) vale

$$D\mathbf{z} = \left[\begin{array}{c|ccccccc} 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \\ \hline -a_0 & -a_1 & -a_2 & \cdots & -a_m & -a_{m+1} & \cdots & -a_{n-1} \end{array}\right]\mathbf{z} + \left[\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{array}\right]\mathbf{u}$$

$$\mathbf{y} = \left[\begin{array}{cccccccc} b_0 & b_1 & b_2 & \cdots & b_m & 0 & \cdots & 0 \end{array}\right]\mathbf{z} \stackrel{def}{=} \mathbf{C}_c\mathbf{z}$$

La forma canonica di controllo appare pertanto fisicamente motivata comeuna forma di stato con n stati in cui viene naturalmente scritta una equazione differenziale (o alle differenze) di ordine n.

Nel caso di sistemi propri non strettamente, la forma della matrice delle uscite diviene la seguente:

$$\mathbf{y} = \left[\begin{array}{ccccc} b_o - b_n a_0 & b_1 - b_n a_1 & \mathbf{b}_2 - b_n a_2 & \dots & b_{n-1} - b_n a_{n-1} \end{array}\right]\mathbf{x} + [b_n]\,u$$

Dato un sistema LTI SISO in forma normale, è dunque sempre possibile scrivere un sistema in forma di stato con matrici $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ ad esso equivalente con $(\mathbf{A}, \mathbf{B})$ in forma canonica di controllo (quindi raggiungibile). "Equivalente" si intende nel senso che, dato un ingresso e delle condizioni iniziali per il primo, esistono degli stati iniziali tali per cui l'uscita del secondo sistema corrispondente allo stesso ingresso è la stessa.

Esempio: Data l'equazione $\ddot{y} + y = 2u + \dot{u}$, un sistema in forma di stato che la realizza (in forma canonica di controllo) è

$$\begin{aligned} \mathbf{A} &= \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right], \ \mathbf{B} = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]; \\ \mathbf{C} &= \left[\begin{array}{cc} 2 & 1 \end{array}\right], \ \mathbf{D} = 0. \end{aligned}$$

Lo stesso rapporto ingresso-uscita è realizzato anche ad esempio ponendo $\mathbf{x}_1 = y$, $\mathbf{x}_2 = \dot{y}$ e $\mathbf{x}_3 = y + \dot{y}$, quindi dal sistema

$$\begin{aligned} \mathbf{A}' &= \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & 1 & 0 \end{bmatrix}, \quad \mathbf{B}' = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}; \\ \mathbf{C}' &= \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}, \quad \mathbf{D}' = 0 \end{aligned}$$

In questo secondo caso però la coppia $(\mathbf{A}', \mathbf{B}')$ non è raggiungibile.

Le funzioni di trasferimento per i due sistemi sono naturalmente le stesse (avendo rapporto solo col rapporto ingresso/uscita, quindi con la equazione in forma normale): nel secondo caso, si osserva una cancellazione polo/zero.

La funzione di trasferimento per un sistema SISO strettamente proprio in forma canonica di controllo vale $G(s) = \mathbf{C}_c(s\mathbf{I} - \mathbf{A}_c)^{-1}\mathbf{B}_c$. Ricordando che

$$\det(s\mathbf{I} - \mathbf{A}) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0$$

e che

$$\begin{aligned} &\mathbf{C}_c \,\mathsf{adj}\,(s\mathbf{I} - \mathbf{A})\mathbf{B}_c = \\ &= \begin{bmatrix} b_0 & \cdots & b_m & 0 & \cdots & 0 \end{bmatrix} \left[\begin{array}{ccc|c} \star & \cdots & \star & 1 \\ \star & \cdots & \star & s \\ \vdots & \ddots & \vdots & \vdots \\ \star & \cdots & \star & s^{n-1} \end{array}\right] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \end{aligned}$$

si ha

$$G(s) = \frac{b_m s^m + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}$$

In altre parole, nella forma canonica di controllo si trovano i coefficienti del polinomio caratteristico nell'ultima riga della matrice dinamica $\mathbf{A}_c$, e i coefficienti del polinomio degli zeri nella matrice delle uscite $\mathbf{C}_c$.

6 Osservabilità

Si è in passato già sottolineato il fatto che, nel modello in forma di stato di un sistema, l'uscita (di misura) ha il significato di specificare quelle grandezze di cui, a differenza dello stato, si ha a disposizione una misura ad ogni istante del processo. Essendo la conoscenza degli stati necessaria per effettuare la loro retroazione, è perciò fondamentale determinare se, dalla conoscenza delle uscite (e degli ingressi, che sono ovviamente a nostra disposizione), è possibile conoscere (o più in generale stimare) lo stato attuale del sistema.

La capacità di stimare lo stato di un sistema è d'altronde importante anche in senso diretto in molte applicazioni, dove è importante risalire dalla osservazione di fenomeni misurabili alla situazione "interna" del sistema, di per sé non accessibile.

Se i dati di ingresso/uscita sono misurati in un intervallo $[0, t]$, il problema di ricavare informazione sullo stato $\mathbf{x}(0)$ al tempo 0 si dice di *osservazione dello stato*; viceversa, il problema di ricavare lo stato $\mathbf{x}(t)$ al tempo t si dice di *ricostruzione dello stato*. Si vedrà che i due problemi sono equivalenti per sistemi LTITC, ma non per sistemi LTITD.

Nelle questioni di osservazione/ricostruzione, ci troveremo spesso di fronte a problemi che possono essere ricondotti a sistemi di equazioni lineari sovradeterminati (con meno incognite che equazioni), che in pratica saranno spesso inconsistenti. Per questi problemi si porranno quindi questioni di approssimazione ottima ad una soluzione impossibile.

Consideriamo un sistema tempo invariante in forma di stato

$$\begin{array}{rcl} D\mathbf{x}(t) & = & \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \\ \mathbf{y}(t) & = & \mathbf{h}(\mathbf{x}(t), \mathbf{u}(t)) \end{array}$$

con $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^p$ e $\mathbf{u} \in \mathbb{R}^m$. Sia al solito $\mathbf{x}(\bar{\mathbf{x}}, \mathbf{u}(\cdot), t)$ il valore della soluzione corrispondente a $\mathbf{x}(0) = \bar{\mathbf{x}}$ e controllo $\mathbf{u}(\tau)$, $\tau \in [0, t]$, al tempo t; si indichi poi con $\mathbf{y}(\bar{\mathbf{x}}, \mathbf{u}(\cdot), t) = \mathbf{h}\left(\mathbf{x}(\bar{\mathbf{x}}, \mathbf{u}(\cdot), t), \mathbf{u}(t)\right)$.

Ci chiediamo se, conoscendo

- il modello del sistema
- l'ingresso $\mathbf{u}(\cdot)$ sull'intervallo $[0, t]$;
- l'uscita $\mathbf{y}(\cdot)$ sull'intervallo $[0, t]$,

è possibile osservare $\mathbf{x}(0) = \bar{\mathbf{x}}$ (ricostruire $\mathbf{x}(t)$).

È ovvio (almeno in linea teorica) che, se è possibile osservare $\bar{\mathbf{x}}$, la conoscenza dell'ingresso, del modello e quindi della sua soluzione (che si suppone unica nel futuro per il determinismo del modello) implica che si possa determinare univocamente $\mathbf{x}(t) = \mathbf{x}(\bar{\mathbf{x}}, \mathbf{u}, t)$. Il viceversa non è detto: per i sistemi TD, in particolare, abbiamo già visto che le traiettorie di un sistema non sono univocamente determinate nel passato (si pensi alla possibile non invertibilità di $\mathbf{A}^t$ per sistemi LTITD).

Poniamo la questione della osservabilità dello stato in termini diversi. Due stati iniziali $\bar{\mathbf{x}}$ e $\bar{\mathbf{x}}_1$ si dicono *indistinguibili* (nel futuro) nell'intervallo $[0, t]$ (si scrive $\bar{\mathbf{x}} I_t \bar{\mathbf{x}}_1$) se, qualsiasi ingresso $\mathbf{u}(\cdot) \in U$ venga applicato al sistema, le uscite corrispondenti alle evoluzioni relative sono uguali, cioè

$$\bar{\mathbf{x}} I_t \bar{\mathbf{x}}_1 \Leftrightarrow \forall \mathbf{u} \in U,\ \mathbf{y}(\bar{\mathbf{x}}, \mathbf{u}, \tau) = \mathbf{y}(\bar{\mathbf{x}}_1, \mathbf{u}, \tau)\, \forall \tau \in [0, t]$$

Gli stati sono detti *indistinguibili* tout–court se sono indistinguibili per ogni t (si scrive $\bar{\mathbf{x}} I \bar{\mathbf{x}}_1$).

Si definisce insieme dei punti *indistinguibili* nell'intervallo $[0, t]$ da $\bar{\mathbf{x}}$ l'insieme

$$\mathcal{I}_t(\bar{\mathbf{x}}) \stackrel{def}{=} \{\bar{\mathbf{x}}_1 \in \mathbb{R}^n | \mathbf{y}(\bar{\mathbf{x}}, \mathbf{u}, \tau) = \mathbf{y}(\bar{\mathbf{x}}_1, \mathbf{u}, \tau)\, \forall \tau \in [0, t], \forall \mathbf{u} \in U\}\,.$$

Ovviamente, se questo insieme contiene altri punti diversi da $\bar{\mathbf{x}}$ stesso, non sarà possibile determinare univocamente lo stato dalle uscite.

6.1 Insieme indistinguibile per un sistema LTITC

Consideriamo l'insieme di indistinguibilità per il sistema

$$\begin{array}{rcl} \dot{\mathbf{x}} & = & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \\ \mathbf{y} & = & \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{array} \tag{23}$$

con $\mathbf{x}(0) = \bar{\mathbf{x}}$, e la corrispondente soluzione per le uscite:

$$\mathbf{y}(\bar{\mathbf{x}}, \mathbf{u}, \tau) = \mathbf{C}e^{\mathbf{A}\tau}\bar{\mathbf{x}} + \int_0^\tau \mathbf{C}e^{\mathbf{A}(\tau-\sigma)}\mathbf{B}\mathbf{u}(\sigma)d\sigma + \mathbf{D}\mathbf{u}(\tau)$$

Confrontando le uscite per due stati iniziali diversi, si ha

$$\bar{\mathbf{x}}_1 I_t \bar{\mathbf{x}} \Leftrightarrow \mathbf{C}e^{\mathbf{A}\tau}\bar{\mathbf{x}}_1 = \mathbf{C}e^{\mathbf{A}\tau}\bar{\mathbf{x}},\ \forall \tau \in [0, t] \Leftrightarrow \mathbf{C}e^{\mathbf{A}\tau}(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}) = 0,\ \forall \tau \in [0, t]$$

per cui lo studio della indistinguibilità di $\bar{\mathbf{x}}_1$ da $\bar{\mathbf{x}}$ può essere riformulato come lo studio della indistinguibilità dello stato $\tilde{\mathbf{x}} \stackrel{def}{=} \bar{\mathbf{x}}_1 - \bar{\mathbf{x}}$ dall'origine. Inoltre è evidente che in questo problema, gli ingressi non giocano alcun ruolo (questo è ero solo perchè il sistema è lineare!)

Uno stato $\tilde{\mathbf{x}}$ indistinguibile dall'origine nell'intervallo $[0, t]$ viene detto *non osservabile* in $[0, t]$. Non si specifica l'intervallo se questo vale per ogni intervallo. Se per un sistema l'insieme dei punti non osservabili contiene solo l'origine, si dice che *il sistema* è (completamente) osservabile.

Dunque, la distinguibilità di uno stato $\bar{\mathbf{x}}_1$ da $\bar{\mathbf{x}}$ su $[0,t]$ dipende dalla eguaglianza dei due corrispondenti vettori di uscita per ogni istante dell'intervallo. Supponiamo ancora una volta che gli ingressi applicati al sistema (e quindi anche le sue soluzioni) siano analitici. L'uguaglianza di due funzioni analitiche in ogni punto di un intervallo implica ed è implicata dalla uguaglianza delle funzioni e di tutte le loro derivate nel punto iniziale dell'intervallo. Considerando quindi ancora l'espressione della k–esima derivata dello stato

$$\mathbf{x}^{(k)} = \mathbf{A}^k\mathbf{x} + \sum_{i=1}^{k} \mathbf{A}^{k-i}\mathbf{B}\mathbf{u}^{(i-1)}$$

, si ha

$$\begin{array}{rcl} \mathbf{y}(\bar{\mathbf{x}},\mathbf{u},0) - \mathbf{y}(\bar{\mathbf{x}}_1,\mathbf{u},0) & = & \mathbf{C}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_1) \\ \dot{\mathbf{y}}(\bar{\mathbf{x}},\mathbf{u},0) - \dot{\mathbf{y}}(\bar{\mathbf{x}}_1,\mathbf{u},0) & = & \mathbf{C}\mathbf{A}(\bar{\mathbf{x}} - \bar{\mathbf{x}}_1) \\ \vdots & = & \vdots \\ \mathbf{y}^{(k)}(\bar{\mathbf{x}},\mathbf{u},0) - \mathbf{y}^{(k)}(\bar{\mathbf{x}}_1,\mathbf{u},0) & = & \mathbf{C}\mathbf{A}^k(\bar{\mathbf{x}} - \bar{\mathbf{x}}_1) \end{array}$$

Queste equazioni, che definiscono gli stati $\bar{\mathbf{x}}_1$ indistinguibili da $\bar{\mathbf{x}}$, possono essere riscritte in forma matriciale:

$$\begin{bmatrix} \tilde{\mathbf{y}}(0) \\ \tilde{\mathbf{y}}^{(1)}(0) \\ \tilde{\mathbf{y}}^{(2)}(0) \\ \vdots \\ \tilde{\mathbf{y}}^{(k)}(0) \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \vdots \\ \mathbf{C}\mathbf{A}^k \\ \vdots \end{bmatrix} \tilde{\mathbf{x}}$$

dove $\tilde{\mathbf{y}}(t) = \mathbf{y}(\bar{\mathbf{x}},\mathbf{u},t) - \mathbf{y}(\bar{\mathbf{x}}_1,\mathbf{u},t)$.

Dunque $\tilde{\mathbf{x}}$ è non osservabile (ovvero $\bar{\mathbf{x}}_1$ è indistinguibile da $\bar{\mathbf{x}}$) se appartiene allo spazio nullo della matrice composta con infinite righe sopra riportata. In seguito al teorema di Cayley–Hamilton, sappiamo comunque che ogni gruppo di righe $\mathbf{C}\mathbf{A}^r$ con $r \geq n$ è combinazione lineare delle prime n righe della matrice stessa. Possiamo dunque dire che l'insieme dei punti non osservabili nell'intervallo $[0,t]$ è in effetti un sottospazio, detto *sottospazio di inosservabilità*, dato da

$$\bar{\mathcal{O}} = \ker(\mathbf{O}) \stackrel{def}{=} \ker\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix}\right)$$

La matrice $\mathbf{O}$ viene detta *matrice di osservabilità* del sistema.

Il sottospazio di inosservabilità nell'intervallo $[0, t]$ non dipende in effetti dalla durata dell'intervallo di osservazione t: se uno stato è indistinguibile dall'origine in un tempo t, tale rimarrà per ogni durata della osservazione. Viceversa, se uno stato è osservabile con osservazioni della uscita di una certa durata, lo sarà anche con osservazioni di durata arbitrariamente breve. Quindi, per sistemi in TC, si ometterà di specificare la durata del tempo di osservazione. Un sistema è completamente osservabile se e solo se $\ker(\mathbf{O}) = \{0\}$.

L'insieme dei punti indistinguibili dal generico punto $\bar{\mathbf{x}}$ è dunque dato da

$$\mathcal{I}_t(\bar{\mathbf{x}}) = \left\{\mathbf{x} = \bar{\mathbf{x}} + \mathbf{r}, \forall \mathbf{r} \in \bar{\mathcal{O}}\right\}$$

ed è quindi un iperpiano, parallelo al sottospazio di inosservabilità, passante per $\bar{\mathbf{x}}$ ed anch'esso indipendente da t. Nessun punto, eccetto $\bar{\mathbf{x}}$, è indistinguibile da $\bar{\mathbf{x}}$ nel caso che il sistema sia osservabile.

6.4 Sistemi LTI: sottosistemi osservabili e non

Il sottospazio di inosservabilità, sia nel caso LTITC , ha anch'esso una caratterizzazione geometrica: esso è il più grande sottospazio $\mathbf{A}$–invariante contenuto in $\ker(\mathbf{C})$.

È ovvio infatti che

$$\ker \left[\begin{array}{c} \mathbf{C} \\ \hline \mathbf{CA} \\ \hline \vdots \\ \hline \mathbf{CA}^{(n-1)} \end{array} \right]$$

è contenuto in $\ker(\mathbf{C})$, e che è $\mathbf{A}$–invariante. Inoltre, se $\mathcal{S}$ è contenuto in $\ker(\mathbf{C})$ ed è $\mathbf{A}$–invariante, esso è contenuto anche in $\ker(\mathbf{CA}^k)$, $\forall k$, e quindi $\mathcal{S} \subseteq \bar{\mathcal{O}}$.

Sia $\mathbf{T}_{\bar{\mathcal{O}}} \in \mathbb{R}^{n \times \bar{o}}$ una matrice di base per il sottospazio di inosservabilità $\bar{\mathcal{O}} = \ker(\mathbf{O})$ del sistema LTI

$$\begin{array}{rcl} D\mathbf{x} & = & \mathbf{Ax} + \mathbf{Bu}, \\ \mathbf{y} & = & \mathbf{Cx} + \mathbf{Du} \end{array}$$

e $\mathbf{T}_O \in \mathbb{R}^{n\times(n-\bar{o})}$ una matrice di base complementare. Nelle nuove coordinate descritte da

$$\mathbf{x} = \mathbf{T}\mathbf{z} \stackrel{def}{=} \left[\begin{array}{c|c} \mathbf{T}_O & \mathbf{T}_{\bar{O}} \end{array}\right] \left[\begin{array}{c} \mathbf{z}_O \\ \hline \mathbf{z}_{\bar{O}} \end{array}\right]$$

il sistema diviene

$$\begin{array}{rcl} D\mathbf{z} & = & \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{z} + \mathbf{T}^{-1}\mathbf{B} \\ \mathbf{y} & = & \mathbf{C}\mathbf{T}\mathbf{z} + \mathbf{D}\mathbf{u} \end{array}$$

ovvero

$$\begin{array}{rcl} \left[\begin{array}{c} D\mathbf{z}_O \\ \hline D\mathbf{z}_{\bar{O}} \end{array}\right] & = & \left[\begin{array}{c|c} \mathbf{A}_O & \mathbf{0} \\ \hline \mathbf{A}_{O\bar{O}} & \mathbf{A}_{\bar{O}} \end{array}\right] \left[\begin{array}{c} \mathbf{z}_O \\ \hline \mathbf{z}_{\bar{O}} \end{array}\right] + \left[\begin{array}{c} \mathbf{B}_O \\ \hline \mathbf{B}_{\bar{O}} \end{array}\right] \mathbf{u} \\ \mathbf{Y} & = & \left[\begin{array}{c|c} \mathbf{C}_O & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \mathbf{z}_O \\ \hline \mathbf{z}_{\bar{O}} \end{array}\right] + \mathbf{D}\mathbf{u} \end{array}$$

dove $\mathbf{z}_O \in \mathbb{R}^{n-\bar{o}}$ e $\mathbf{z}_{\bar{O}} \in \mathbb{R}^{\bar{o}}$.

In queste coordinate, il sistema è dunque riscritto nella forma

$$\begin{array}{rcl} D\mathbf{z}_O & = & \mathbf{A}_O\mathbf{z}_O + \mathbf{B}_O\mathbf{u} \\ D\mathbf{z}_{\bar{O}} & = & \mathbf{A}_{O\bar{O}}\mathbf{z}_O + \mathbf{A}_{\bar{O}}\mathbf{z}_{\bar{O}} + \mathbf{B}_{\bar{O}}\mathbf{u} \\ \mathbf{y} & = & \mathbf{C}_O\mathbf{z}_O + \mathbf{D}\mathbf{u} \end{array} \tag{25}$$

cioè scomposto in due sottosistemi, dei quali il secondo è completamente inosservabile: infatti, lo stato $\mathbf{z}_{\bar{O}}$ non influenza l'uscita ne direttamente, nè attraverso lo stato $\mathbf{z}_O$ (il contributo di $\mathbf{z}_O$ alla evoluzione del secondo sottosistema si può guardare come un ulteriore ingresso, che mostreremo ora essere noto).

Per quanto riguarda il sottosistema con stato $\mathbf{z}_O$ e matrici $(\mathbf{A}_O, \mathbf{B}_O, \mathbf{C}_O, \mathbf{D})$, esso è completamente osservabile: infatti, la matrice di osservabilità dell'intero sistema nelle nuove coordinate è

$$\mathbf{O}' = \left[\begin{array}{c|c} \mathbf{C}_O & \mathbf{0} \\ \hline \mathbf{C}_O\mathbf{A}_O & \mathbf{0} \\ \hline \vdots & \vdots \\ \hline \mathbf{C}_O\mathbf{A}_O^{n-1} & \mathbf{0} \end{array}\right] = \mathbf{O}\mathbf{T}$$

quindi, avendo $\mathbf{O}'$ rango $n - \bar{o}$ come $\mathbf{O}$, le sue prime $n - \bar{o}$ colonne, le sole non nulle, sono indipendenti.

La forma 25 viene detta *forma standard di osservabilità* del sistema. Il sottosistema $(\mathbf{A}_O, \mathbf{B}_O, \mathbf{C}_O, \mathbf{D})$ è detto *sottosistema osservabile*; il sottosistema $(\mathbf{A}_{\bar{O}}, \mathbf{B}_{\bar{O}}, \mathbf{C}_{\bar{O}} = \mathbf{0}, \mathbf{D})$ è detto *sottosistema non osservabile.*

La scelta delle matrici di base $\mathbf{T}_{\bar{O}}$ e complementare $\mathbf{T}_O$ è arbitraria: qualsiasi altra scelta porterebbe ad una forma analoga, con blocchi diagonali diversi ma simili (algebricamente equivalenti) a quelli ottenuti in altra base.

Gli autovalori di $\mathbf{A}_{\bar{O}}$ e quelli di $\mathbf{A}_O$ sono quindi invarianti in numero e in posizione con i cambiamenti di coordinate, e sono quindi proprietà strutturali del sistema. I primi vengono detti *autovalori interni al sottospazio di inosservabilità*, i secondi *esterni*.

La verifica di osservabilità può essere fatta anche ricorrendo al *Lemma P.B.H.* (Popov, Belevitch, Hautus):

Teorema Il sistema LTI con matrici $(\mathbf{A}, \mathbf{C})$ è osservabile se e solo se la matrice

$$\mathbf{P}_o(\lambda) = \left[\begin{array}{c} \lambda \mathbf{I} - \mathbf{A} \\ \hline \mathbf{C} \end{array} \right] \tag{26}$$

ha rango pieno per ogni $\lambda \in \mathbb{C}$.

Dimostrazione. La dimostrazione verte sulla possibilità di trovare un vettore $\mathbf{q}$ tale che $\mathbf{P}_o(\lambda)\mathbf{q} = 0$. Si riconduce direttamente a quella del lemma PBH per la raggiungibilità considerando l'equazione $\mathbf{q}^T \mathbf{P}_o^T(\lambda) = 0$.

Applichiamo il lemma PBH al caso di una coppia $(\mathbf{A}, \mathbf{C})$ con $\mathbf{A}$ in forma di Jordan, con p miniblocchi:

$$
\mathbf{A} = \left[\begin{array}{cccc|c|cccc}
\lambda_1 & 1 & \cdots & 0 & & & & & \\
0 & \lambda_1 & \cdots & 0 & \cdots & & \mathbf{0} & & \\
\cdots & \cdots & \ddots & \cdots & & & & & \\
0 & 0 & \cdots & \lambda_1 & & & & & \\
\hline
 & \mathbf{0} & & & \ddots & & \mathbf{0} & & \\
\hline
 & & & & & \lambda_p & 1 & \cdots & 0 \\
 & \mathbf{0} & & & \cdots & 0 & \lambda_p & \cdots & 0 \\
 & & & & & \cdots & \cdots & \ddots & \cdots \\
 & & & & & 0 & 0 & \cdots & \lambda_p
\end{array}\right];
$$

$$
\mathbf{C} = \left[\begin{array}{cccc|c|cccc} \mathbf{C}_{11} & \mathbf{C}_{21} & \cdots & \mathbf{C}_{m_1,1} & \cdots & \mathbf{C}_{p,1} & \mathbf{C}_{2,p} & \cdots & \mathbf{C}_{m_p,p} \end{array}\right]
$$

Risulta che, per essere osservabile, le prime colonne per ogni miniblocco corrispondente ad autovalori coincidenti, devono essere linearmente indipendenti. In particolare, per un sistema SISO, è necessario che la molteplicità geometrica di tutti gli autovalori sia pari a uno, e che $\mathbf{C}$ abbia almeno tanti elementi diversi da zero quanti gli autovalori distinti di $\mathbf{A}$. Un sistema con μ_i miniblocchi associati ad un unico autovalore λ_i può essere osservabile solo se ha almeno μ_i uscite indipendenti.

6.5 Forma canonica di osservazione

Per un sistema SISO con matrici dinamica e di uscita nella particolare forma

$$
\begin{aligned}
\mathbf{A}_o &= \left[\begin{array}{cccc|c}
0 & 0 & \cdots & 0 & -a_0 \\
\hline
1 & 0 & \cdots & 0 & -a_1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & -a_{n-1}
\end{array}\right]; \\
\mathbf{C}_o &= \left[\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right]
\end{aligned}
$$

(la forma di $\mathbf{A}_o$ si dice *compagna verticale destra*), la matrice di osservabilità ha la stessa forma della matrice di raggiungibilità della forma canonica di controllo.