Lezione 1

Complementi di teoria dei sistemi

Richiami sui sistemi dinamici

Come è noto, la scrittura dei modelli matematici in termini di "sistemi dinamici" consente di studiarne le proprietà con strumenti di analisi molto generali, che prescindono dalla natura specifica del sistema. Nel seguito riprenderemo rapidamente gli elementi essenziali alla base della nozione di sistema dinamico, già noti dal primo corso di Automatica.

Ricordiamo che un sistema dinamico si interfaccia con il "resto del mondo" per mezzo di una serie di variabili, che definiremo di ingresso, ed altre che definiremo di uscita. Definiamo di ingresso le variabili con cui dall'esterno si influenza il comportamento del sistema, di uscita quelle che caratterizzano il comportamento del sistema e sulle quali soffermiamo il nostro interesse (tipicamente perché costituiscono l'obiettivo del controllo).

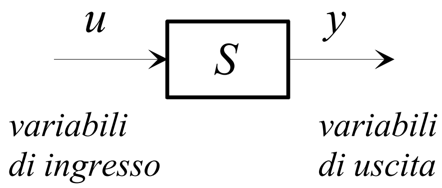

Fig. 1 : Ingressi e uscite di un sistema

E' anche noto che per conoscere il valore delle uscite ad un dato istante t è necessario che siano note un certo numero di condizioni iniziali assegnate ad un istante iniziale t_0 e l'andamento degli ingressi dall'istante iniziale a quello attuale. Il numero minimo di condizioni iniziali che occorre assegnare per determinare tutte le uscite del sistema, noti gli andamenti degli ingressi a partire dall'istante iniziale, prende il nome di *ordine* del sistema e si indica con la lettera n.

Lo **stato** del sistema ad un dato istante riassume tutta la storia passata del sistema fino a quell'istante ed è quindi quanto occorre conoscere per calcolare le uscite da quell'istante in poi, noti gli ingressi. Per quanto affermato sopra, lo stato si può esprimere per mezzo di n variabili, indicate con i simboli $x_1, x_2, \ldots , x_n$, che prendono il nome di **variabili di stato**.

La formalizzazione matematica del sistema dinamico passa allora per la scrittura delle equazioni differenziali (*equazioni di stato*) di cui le variabili di stato sono le soluzioni, noti gli ingressi esterni, e del legame (espresso da trasformazioni algebriche) tra le variabili di uscita e quelle di stato e di ingresso (*trasformazioni di uscita*).

Sia m il numero delle variabili di ingresso e p il numero di variabili di uscita. Introdotti i tre vettori:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix},$$

e le due funzioni vettoriali:

$$
\boldsymbol{f}(\boldsymbol{x},\boldsymbol{u},t)=\begin{bmatrix} f_1(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \\ f_2(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \\ \vdots \\ f_n(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \end{bmatrix},
$$

$$
\boldsymbol{g}(\boldsymbol{x},\boldsymbol{u},t)=\begin{bmatrix} g_1(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \\ g_2(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \\ \vdots \\ g_p(x_1,x_2,...,x_n,u_1,u_2,...,u_m,t) \end{bmatrix}
$$

la formulazione vettoriale del sistema dinamico è la seguente:

$$
\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \\ \boldsymbol{y}(t) &= \boldsymbol{g}(\boldsymbol{x}(t),\boldsymbol{u}(t),t) \end{aligned} .
$$

Si osservi che la dipendenza esplicita dal tempo delle due funzioni vettoriali rende il sistema genericamente *tempo variante*. In realtà nel presente corso ci limiteremo allo studio di sistemi in cui questa dipendenza esplicita non esista (sistemi *tempo invarianti* o *stazionari*).

Definiremo poi come sistemi *SISO* (Single Input Single Output) i sistemi per cui m=p=1, e genericamente *MIMO* (Multiple Input Multiple Output) gli altri. Si dirà *strettamente proprio* un sistema in cui la funzione $\boldsymbol{g}$ non dipende dall'ingresso $\boldsymbol{u}$, genericamente proprio un sistema in cui ciò non accade. Infine si definisce *lineare* un sistema in cui tutte le equazioni di stato e tutte le trasformazioni di uscita siano funzioni lineari delle variabili di stato e delle variabili di ingresso, *non lineari* tutti gli altri.

Ricordiamo che, assegnata una condizione iniziale all'istante t_0:

$$
\boldsymbol{x}(t_0)=\boldsymbol{x}_0 ,
$$

e una funzione di ingresso a partire da t_0:

$$
\boldsymbol{u}(t)=\bar{\boldsymbol{u}}(t), \quad t \geq t_0 ,
$$

diciamo **movimento dello stato** $\bar{\boldsymbol{x}}(t)$ la soluzione delle equazioni di stato corredate dalla condizione iniziale assegnata:

$$
\begin{aligned} \dot{\bar{\boldsymbol{x}}}(t) &= \boldsymbol{f}(\bar{\boldsymbol{x}}(t),\bar{\boldsymbol{u}}(t),t) \\ \bar{\boldsymbol{x}}(t_0) &= \boldsymbol{x}_0 \end{aligned} ,
$$

e movimento dell'uscita la conseguente uscita, ricavabile dalla trasformazione d'uscita:

$$
\bar{\boldsymbol{y}}(t)=\boldsymbol{g}(\bar{\boldsymbol{x}}(t),\bar{\boldsymbol{u}}(t),t).
$$

Ricordiamo infine che i particolari movimenti costanti nel tempo, associati a ingressi costanti, prendono il nome di **equilibri**. Per i sistemi tempo-invarianti, la ricerca di eventuali stati di equilibrio associati all'ingresso $\bar{\boldsymbol{u}}$ si conduce annullando le derivate nelle equazioni di stato e ricercando le eventuali soluzioni dell'equazione vettoriale implicita in $\bar{\boldsymbol{x}}$:

$$
\boldsymbol{f}(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}})=0 .
$$

Sistemi LTI: movimento

Si consideri un sistema LTI di equazioni:

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t)\\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t)\end{aligned}.$$

Siamo interessati a scrivere l'espressione del movimento dello stato e del movimento dell'uscita conseguenti ad una condizione iniziale $\boldsymbol{x}_0$ assegnata all'istante t_0 e ad un ingresso definito a partire dall'istante t_0.

Per questo, occorre preliminarmente definire la nozione di **esponenziale di matrice**. Dati una matrice quadrata $\boldsymbol{A}$ ed uno scalare τ definiamo esponenziale di matrice la matrice che si ottiene dal seguente sviluppo in serie:

$$e^{\boldsymbol{A}\tau} = \sum_{k=0}^{\infty} \frac{(\boldsymbol{A}\tau)^k}{k!} = \boldsymbol{I} + \boldsymbol{A}\tau + \frac{\boldsymbol{A}^2\tau^2}{2!} + \frac{\boldsymbol{A}^3\tau^3}{3!} + \cdots$$

La nozione di esponenziale di matrice generalizza quindi la nozione di esponenziale scalare. Con questa condivide anche la regola di derivazione, dal momento che, definita come derivata rispetto a τ dell'esponenziale la matrice ottenuta derivando tutti gli elementi della matrice esponenziale rispetto a τ, risulta:

$$\frac{d}{d\tau}e^{\boldsymbol{A}\tau} = \boldsymbol{0} + \boldsymbol{A} + 2\frac{\boldsymbol{A}^2\tau}{2!} + 3\frac{\boldsymbol{A}^3\tau^2}{3!} + \cdots = \boldsymbol{A}\left[\boldsymbol{I} + \boldsymbol{A}\tau + \frac{\boldsymbol{A}^2\tau^2}{2!} + \frac{\boldsymbol{A}^3\tau^3}{3!} + \cdots\right] = \boldsymbol{A}e^{\boldsymbol{A}\tau}.$$

Il calcolo della matrice esponenziale risulta poi particolarmente agevole quando la matrice $\boldsymbol{A}$ è diagonalizzabile, ossia quando è simile ad una matrice diagonale:

$$\hat{\boldsymbol{A}} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1} = diag\{\lambda_i, i = 1,\ldots,n\},$$

dove λ_i sono gli autovalori di $\boldsymbol{A}$. Risulta infatti:

$$\begin{aligned}e^{\boldsymbol{A}\tau} &= e^{\boldsymbol{T}^{-1}\hat{\boldsymbol{A}}\boldsymbol{T}\tau} = \boldsymbol{I} + \boldsymbol{T}^{-1}\hat{\boldsymbol{A}}\boldsymbol{T}\tau + \boldsymbol{T}^{-1}\hat{\boldsymbol{A}}\boldsymbol{T}\boldsymbol{T}^{-1}\hat{\boldsymbol{A}}\boldsymbol{T}\frac{\tau^2}{2!} + \cdots = \boldsymbol{T}^{-1}\left[\boldsymbol{I} + \hat{\boldsymbol{A}}\tau + \hat{\boldsymbol{A}}^2\frac{\tau^2}{2!} + \cdots\right]\boldsymbol{T} = \\ &= \boldsymbol{T}^{-1}e^{\hat{\boldsymbol{A}}\tau}\boldsymbol{T} = \boldsymbol{T}^{-1}diag\left\{e^{\lambda_i\tau}\right\}\boldsymbol{T}\end{aligned}$$

Fatta questa premessa, siamo in grado di scrivere le espressioni dei movimenti di stato e uscita, come segue:

$$\begin{aligned}\boldsymbol{x}(t) &= e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0 + \int_{t_0}^{t} e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau \qquad \text{(formula di Lagrange)}\\ \boldsymbol{y}(t) &= \boldsymbol{C}e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0 + \boldsymbol{C}\int_{t_0}^{t} e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau + \boldsymbol{D}\boldsymbol{u}(t)\end{aligned}.$$

Per la verifica della correttezza dell'espressione del movimento dello stato (quello dell'uscita si ottiene per ovvia sostituzione nella trasformazione di uscita), è utile ricordare un risultato noto dall'analisi matematica (derivazione sotto segno di integrale). Data una funzione $\varphi(\tau,t)$ (scalare o matriciale) di due argomenti scalari, risulta:

$$\frac{d}{dt}\int_{t_0}^{t}\varphi(\tau,t)d\tau=\int_{t_0}^{t}\frac{d}{dt}\varphi(\tau,t)d\tau+\varphi(t,t).$$

Assumendo come funzione φ l'espressione $e^{\boldsymbol{A}(t-\tau)}\boldsymbol{Bu}(\tau)$, e ricordando la regola di derivazione dell'esponenziale di matrice, si ottiene:

$$\begin{aligned}
\dot{\boldsymbol{x}}(t)&=\frac{d}{dt}\left[e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0\right]+\int_{t_0}^{t}\frac{d}{dt}\left[e^{\boldsymbol{A}(t-\tau)}\boldsymbol{Bu}(\tau)\right]d\tau+\left[e^{\boldsymbol{A}(t-\tau)}\boldsymbol{Bu}(\tau)\right]_{\tau=t}=\\
&=\boldsymbol{A}e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{A}e^{\boldsymbol{A}(t-\tau)}\boldsymbol{Bu}(\tau)d\tau+\boldsymbol{Bu}(t)=\boldsymbol{A}\left[e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0+\int_{t_0}^{t}e^{\boldsymbol{A}(t-\tau)}\boldsymbol{Bu}(\tau)d\tau\right]+\boldsymbol{Bu}(t)=\\
&=\boldsymbol{Ax}(t)+\boldsymbol{Bu}(t)
\end{aligned}$$

Inoltre:

$$\boldsymbol{x}(t_0)=\boldsymbol{x}_0.$$

Si osservi che, tanto il movimento dello stato, quanto quello dell'uscita, sono esprimibili come la somma di un termine che dipende (linearmente) solo dallo stato iniziale ed un termine che dipende (linearmente) solo dall'ingresso. Il primo termine prende il nome di moto libero, il secondo di moto forzato.

Sistemi LTI: cambiamento di variabili di stato

I sistemi dinamici lineari tempo invarianti (LTI) possono essere sinteticamente rappresentati dalla quaterna di matrici $\boldsymbol{A}(n\times n)$, $\boldsymbol{B}(n\times m)$, $\boldsymbol{C}(p\times n)$ e $\boldsymbol{D}(p\times m)$ che definiscono le equazioni di stato e le trasformazioni di uscita:

$$\begin{aligned}\dot{\boldsymbol{x}}(t)&=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)\\ \boldsymbol{y}(t)&=\boldsymbol{C}\boldsymbol{x}(t)+\boldsymbol{D}\boldsymbol{u}(t)\end{aligned}.$$

E' noto che, una volta definite le variabili di ingresso e uscita, con cui il sistema si interfaccia con l'esterno, la scelta delle variabili di stato è arbitraria. In un sistema LTI, il passaggio da una rappresentazione in variabili di stato ad una rappresentazione alternativa si ottiene sempre moltiplicando il vettore di stato originario per una matrice $(n\times n)$ non singolare:

$$\hat{\boldsymbol{x}}(t)=\boldsymbol{T}\boldsymbol{x}(t),\quad \det(\boldsymbol{T})\neq 0\,.$$

Potremo quindi esprimere anche il vecchio vettore di stato in termini del nuovo attraverso l'inversa della matrice $\boldsymbol{T}$:

$$\boldsymbol{x}(t)=\boldsymbol{T}^{-1}\hat{\boldsymbol{x}}(t)\,.$$

Sostituendo questa espressione nelle equazioni che definiscono il sistema dinamico in $\boldsymbol{x}$, e premoltiplicando entrambi i membri dell'equazione vettoriale di stato per la matrice $\boldsymbol{T}$, si ottiene:

$$\begin{aligned}\dot{\hat{\boldsymbol{x}}}(t)&=\hat{\boldsymbol{A}}\hat{\boldsymbol{x}}(t)+\hat{\boldsymbol{B}}\boldsymbol{u}(t)\\ \boldsymbol{y}(t)&=\hat{\boldsymbol{C}}\hat{\boldsymbol{x}}(t)+\hat{\boldsymbol{D}}\boldsymbol{u}(t)\end{aligned},$$

dove:

$$\hat{\boldsymbol{A}}=\boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1},\quad \hat{\boldsymbol{B}}=\boldsymbol{T}\boldsymbol{B},\quad \hat{\boldsymbol{C}}=\boldsymbol{C}\boldsymbol{T}^{-1},\quad \hat{\boldsymbol{D}}=\boldsymbol{D}\,.$$

Si osservi che la relazione che sussiste tra le matrici $\boldsymbol{A}$ e $\hat{\boldsymbol{A}}$ è di similitudine.

Esempio

Si consideri il circuito elettrico RLC di figura:

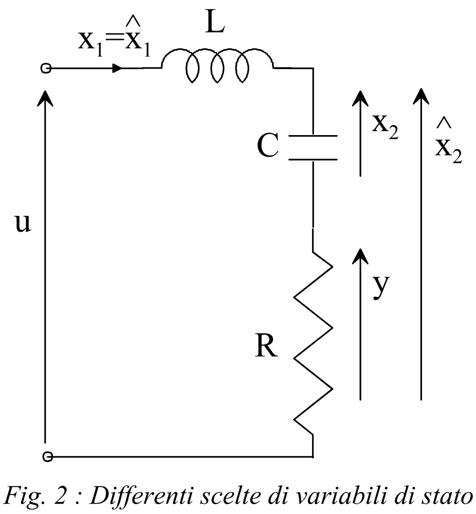

Fig. 2 : Differenti scelte di variabili di stato

Come prima scelta di variabili di stato si prenda la corrente nell'induttore (x_1) e la caduta di tensione ai capi del condensatore (x_2). Assunta come uscita del sistema la caduta di tensione ai capi della resistenza, le equazioni del sistema dinamico si ricavano facilmente:

$$\begin{cases} \dot{x}_1 = -\dfrac{R}{L}x_1 - \dfrac{1}{L}x_2 + \dfrac{1}{L}u \\ \dot{x}_2 = \dfrac{1}{C}x_1 \\ y = Rx_1 \end{cases} \quad \Rightarrow \quad \boldsymbol{A} = \begin{bmatrix} -\dfrac{R}{L} & -\dfrac{1}{L} \\ \dfrac{1}{C} & 0 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \dfrac{1}{L} \\ 0 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} R & 0 \end{bmatrix}, \quad \boldsymbol{D} = 0$$

Prendendo ora come variabili di stato ancora la corrente nell'induttore ($\hat{x}_1$) e la caduta di tensione ($\hat{x}_2$) sulla serie tra resistenza e condensatore, si ottiene:

$$\begin{cases} \dot{\hat{x}}_1 = -\dfrac{1}{L}\hat{x}_2 + \dfrac{1}{L}u \\ \dot{\hat{x}}_2 = \dfrac{1}{C}\hat{x}_1 - \dfrac{R}{L}\hat{x}_2 + \dfrac{R}{L}u \\ y = Rx_1 \end{cases} \quad \Rightarrow \quad \hat{\boldsymbol{A}} = \begin{bmatrix} 0 & -\dfrac{1}{L} \\ \dfrac{1}{C} & -\dfrac{R}{L} \end{bmatrix}, \quad \hat{\boldsymbol{B}} = \begin{bmatrix} \dfrac{1}{L} \\ \dfrac{R}{L} \end{bmatrix}, \quad \hat{\boldsymbol{C}} = \begin{bmatrix} R & 0 \end{bmatrix}, \quad \hat{\boldsymbol{D}} = 0.$$

Il legame tra i due insiemi di variabili di stato è evidentemente:

$$\begin{cases} \hat{x}_1 = x_1 \\ \hat{x}_2 = Rx_1 + x_2 \end{cases},$$

ovvero:

$$\hat{\boldsymbol{x}} = \boldsymbol{Tx}, \quad \boldsymbol{T} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix}.$$

Poiché det($\boldsymbol{T}$) = 1, il cambiamento di variabili di stato è ben posto, e si possono verificare le formule di trasformazione:

$$\hat{\boldsymbol{A}} = \boldsymbol{TAT}^{-1} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix} \begin{bmatrix} -\dfrac{R}{L} & -\dfrac{1}{L} \\ \dfrac{1}{C} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -R & 1 \end{bmatrix} = \begin{bmatrix} 0 & -\dfrac{1}{L} \\ \dfrac{1}{C} & -\dfrac{R}{L} \end{bmatrix},$$

$$\hat{\boldsymbol{B}} = \boldsymbol{TB} = \begin{bmatrix} 1 & 0 \\ R & 1 \end{bmatrix} \begin{bmatrix} \dfrac{1}{L} \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{L} \\ \dfrac{R}{L} \end{bmatrix},$$

$$\hat{\boldsymbol{C}} = \boldsymbol{C}\boldsymbol{T}^{-1} = \begin{bmatrix} R & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -R & 1 \end{bmatrix} = \begin{bmatrix} R & 0 \end{bmatrix},$$

$$\hat{\boldsymbol{D}} = \boldsymbol{D} = 0.$$

Concludiamo ricordando che una proprietà di un sistema LTI si dice **strutturale** se è valida qualunque sia la scelta effettuata sulle variabili di stato. In altre parole, descritto il sistema con un certo vettore $\boldsymbol{x}$ di variabili di stato, ed effettuato il cambiamento di variabili di stato $\hat{\boldsymbol{x}}(t) = \boldsymbol{T}\boldsymbol{x}(t)$, $\det(\boldsymbol{T}) \neq 0$, la proprietà verificata con il primo insieme di variabili di stato rimane valida anche per il nuovo insieme, qualsiasi sia la matrice di trasformazione $\boldsymbol{T}$.

Stabilità

Nel primo corso di Automatica si è fatto riferimento ad una nozione di stabilità valida per sistemi lineari tempo invarianti, che si basa sulla proprietà del sistema di reagire a perturbazioni che intervengono al suo ingresso, ritornando alla condizione preesistente alla perturbazione, o quantomeno non allontanandosene indefinitamente.

Discuteremo ora una nozione più generale di stabilità, che prevede che la perturbazione sulla base della quale si definisce la stabilità intervenga sullo **stato iniziale**. Questa teoria della stabilità (che va sotto il nome di stabilità alla **Lyapunov**, dal nome del matematico che l'ha elaborata) prende le mosse dalla definizione di stabilità di un movimento per un sistema dinamico genericamente non lineare e consente di pervenire all'analisi di stabilità dei sistemi LTI, i cui risultati sono peraltro identici a quelli ricavati con la definizione di stabilità semplificata basati sulla perturbazione impulsiva all'ingresso del sistema.

Prenderemo in considerazione solo sistemi dinamici **tempo invarianti**. Si farà inoltre uso della nozione di **norma** di un vettore che, per i nostri scopi, si potrà sempre definire come la norma Euclidea:

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}^T \boldsymbol{v}} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2} \; .$$

Si consideri dunque un sistema dinamico **tempo invariante**, genericamente non lineare, che per $t \geq 0$ sia soggetto all'ingresso $\overline{\boldsymbol{u}}(t)$.

Definito con $\boldsymbol{x}_{0n}$ lo stato iniziale (per t=0) nominale, sia $\boldsymbol{x}_n(t)$ il movimento che origina da $\boldsymbol{x}_{0n}$ con ingresso $\overline{\boldsymbol{u}}(t)$: definiamo $\boldsymbol{x}_n(t)$ **movimento nominale**.

Sia ora $\boldsymbol{x}_{0p}$ un diverso stato iniziale e $\boldsymbol{x}_p(t)$ il movimento che origina da $\boldsymbol{x}_{0p}$ con *lo stesso ingresso* $\overline{\boldsymbol{u}}(t)$: definiamo $\boldsymbol{x}_p(t)$ **movimento perturbato**.

Ciò posto, il movimento $\boldsymbol{x}_n(t)$ si dice **stabile** se, $\forall \varepsilon>0$, esiste in corrispondenza un valore $\delta_\varepsilon>0$, tale che $\forall$ $\boldsymbol{x}_{0p}$ che soddisfi la condizione:

$$\left\| \boldsymbol{x}_{0p} - \boldsymbol{x}_{0n} \right\| \leq \delta_\varepsilon ,$$

risulti:

$$\left\| \boldsymbol{x}_p(t) - \boldsymbol{x}_n(t) \right\| \leq \varepsilon, \quad \forall t \geq 0.$$

La proprietà di stabilità del movimento formalizza quindi l'idea che perturbando lo stato iniziale del sistema (non l'ingresso) il movimento perturbato non si allontani indefinitamente dal movimento nominale o, più precisamente, che fissato arbitrariamente il massimo scostamento ammissibile tra movimento perturbato e nominale, sia possibile che lo scostamento risulti inferiore a questa soglia, pur di scegliere lo stato iniziale perturbato sufficientemente vicino allo stato iniziale nominale. Nel caso di sistema del primo ordine (nel quale quindi lo stato è rappresentato da una variabile scalare), si può dare un'espressiva rappresentazione grafica di un movimento stabile:

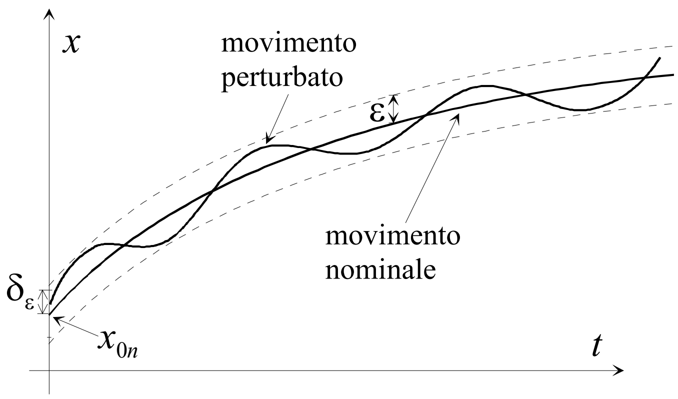

Fig. 3 : Movimento stabile

Un movimento si dirà **instabile** se non è stabile, ossia se non soddisfa la condizione di stabilità precedentemente enunciata. Fissato quindi il massimo scostamento ammissibile tramite il parametro ε, se il movimento è instabile non sarà possibile individuare un corrispondente intorno dello stato iniziale nominale, definito dal parametro δ_ε, comunque piccolo lo si scelga, in modo che per tutti gli stati iniziali perturbati il movimento perturbato non si scosti da quello nominale più del massimo scostamento ammissibile.

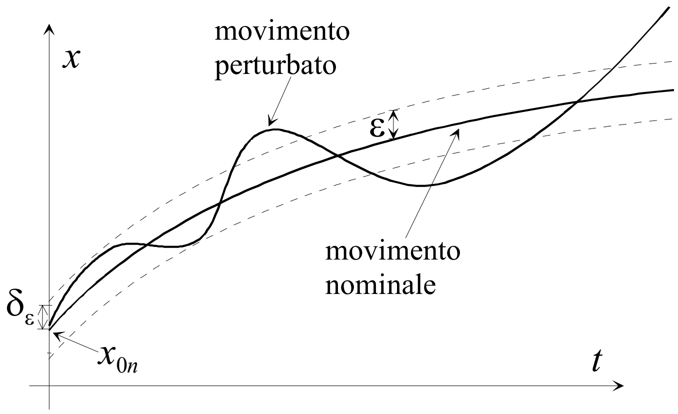

Fig. 4 : Movimento instabile

Infine un movimento $\boldsymbol{x}_n(t)$ si dice **asintoticamente stabile** se, oltre ad essere stabile, soddisfa la condizione:

$$\lim_{t\to\infty}\left\|\boldsymbol{x}_p(t)-\boldsymbol{x}_n(t)\right\|=0,\quad \forall \boldsymbol{x}_{0p} : \left\|\boldsymbol{x}_{0p}-\boldsymbol{x}_{0n}\right\|<\delta_\varepsilon .$$

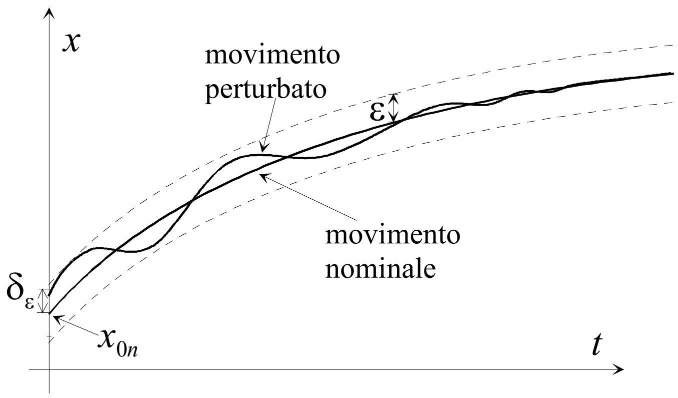

Fig. 5 : Movimento asintoticamente stabile

La proprietà di asintotica stabilità richiede quindi che i movimenti perturbati, oltre a non allontanarsi indefinitamente dal movimento nominale, convergano a questo per $t \rightarrow \infty$.

La definizione di stabilità del movimento si può naturalmente applicare anche a quei particolari movimenti costanti nel tempo che sono gli **equilibri**. Si parlerà quindi di stato di equilibrio stabile, instabile, asintoticamente stabile. Uno stesso sistema dinamico può ammettere in corrispondenza dello stesso ingresso costante diversi stati di equilibrio, con differenti proprietà di stabilità. Si pensi a titolo di esempio ad un pendolo nel piano verticale, soggetto a coppia esterna nulla. Come è noto, il sistema ammette due stati equilibrio fisicamente distinguibili, nel primo dei quali il pendolo è fermo rivolto verso il basso, nel secondo dei quali il pendolo è fermo e rivolto verso l'alto. Dalle definizioni di stabilità precedentemente date, si deduce facilmente che il primo stato di equilibrio è stabile (asintoticamente in presenza di dissipazione), il secondo (pendolo rivolto verso l'alto) è instabile.

Per gli stati di equilibrio asintoticamente stabili si definisce anche la nozione di **regione di attrazione**, intesa come regione nello spazio di stato $\Re^n$ in cui può essere scelto lo stato iniziale perturbato perché il movimento perturbato converga allo stato di equilibrio. Se la regione di attrazione coincide con $\Re^n$, ossia se comunque si scelga lo stato iniziale il movimento perturbato converge allo stato di equilibrio, lo stato di equilibrio si dice **globalmente asintoticamente stabile**.

Esempio

Si consideri un sistema dinamico del primo ordine soggetto ad un ingresso costante $\bar{u}$, e si supponga che il legame tra x e $\dot{x}$ imposto dalla relazione:

$$\dot{x} = f(x, \bar{u})$$

sia rappresentato dalla curva in figura:

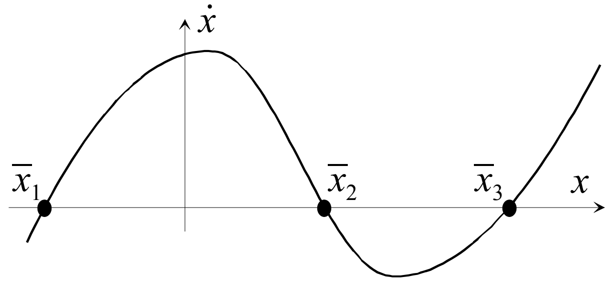

Fig. 6 :Derivata dello stato

Poiché la derivata si annulla per tre valori di x, il sistema ammette tre stati di equilibrio, $\bar{x}_1, \bar{x}_2, \bar{x}_3$.

Non è difficile convincersi che gli stati $\bar{x}_1$ e $\bar{x}_3$ sono instabili, in quanto perturbando lo stato iniziale ad un valore di poco superiore al valore di equilibrio, lo stato, avendo derivata positiva, tenderà ad allontanarsi dallo stato di equilibrio stesso, e viceversa per perturbazioni negative. Con ragionamento analogo si conclude che lo stato $\bar{x}_2$ è invece asintoticamente stabile e la sua regione di attrazione è data dall'insieme $(\bar{x}_1, \bar{x}_3)$.

Stabilità nei sistemi LTI

La stabilità alla Lyapunov discussa nel precedente paragrafo è una proprietà dei singoli movimenti di un sistema e non una proprietà del sistema stesso. In altre parole, per un generico sistema non lineare, l'affermazione "il sistema è stabile (instabile, asintoticamente stabile)" è priva di significato. Le cose stanno diversamente per i sistemi lineari, come una semplice analisi consente di evidenziare. Consideriamo quindi un sistema LTI di equazioni di stato:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t).$$

Assegnato l'ingresso $\overline{\boldsymbol{u}}(t)$ per $t \geq 0$, si considerino un movimento nominale:

$$\boldsymbol{x}_n(t), \quad \boldsymbol{x}_n(0) = \boldsymbol{x}_{n0},$$

ed un movimento perturbato:

$$\boldsymbol{x}_p(t), \quad \boldsymbol{x}_p(0) = \boldsymbol{x}_{p0}.$$

Entrambi i movimenti soddisfano le equazioni di stato del sistema dinamico, per cui risulterà:

$$\dot{\boldsymbol{x}}_n(t) = \boldsymbol{A}\boldsymbol{x}_n(t) + \boldsymbol{B}\overline{\boldsymbol{u}}(t), \quad \boldsymbol{x}_n(0) = \boldsymbol{x}_{n0},$$
$$\dot{\boldsymbol{x}}_p(t) = \boldsymbol{A}\boldsymbol{x}_p(t) + \boldsymbol{B}\overline{\boldsymbol{u}}(t), \quad \boldsymbol{x}_p(0) = \boldsymbol{x}_{p0}.$$

Sottraendo membro a membro (vettorialmente) le due equazioni, e le relative condizioni iniziali, si ottiene:

$$\dot{\boldsymbol{x}}_p(t) - \dot{\boldsymbol{x}}_n(t) = \boldsymbol{A}\left(\boldsymbol{x}_p(t) - \boldsymbol{x}_n(t)\right), \quad \boldsymbol{x}_p(0) - \boldsymbol{x}_n(0) = \boldsymbol{x}_{p0} - \boldsymbol{x}_{n0}.$$

Definiamo ora lo **scostamento** tra il movimento perturbato e quello nominale, e il relativo scostamento tra le condizioni iniziali:

$$\delta\boldsymbol{x}(t) := \boldsymbol{x}_p(t) - \boldsymbol{x}_n(t), \quad \delta\boldsymbol{x}_0 := \boldsymbol{x}_{p0} - \boldsymbol{x}_{n0}.$$

La precedente equazione differenziale potrà quindi essere riscritta in termini della variabile scostamento, come segue:

$$\delta\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\delta\boldsymbol{x}(t), \quad \delta\boldsymbol{x}(0) = \delta\boldsymbol{x}_0.$$

Riprendendo ora la *definizione* di stabilità, e parafrasandola in termini della variabile scostamento, potremo dire che il movimento $\boldsymbol{x}_n$ è stabile se $\forall\varepsilon>0$, esiste in corrispondenza un valore $\delta_\varepsilon>0$, tale che $\forall\ \delta\boldsymbol{x}_0$ che soddisfi la condizione:

$$\|\delta\boldsymbol{x}_0\| \leq \delta_\varepsilon,$$

risulti:

$$\|\delta\boldsymbol{x}(t)\| \leq \varepsilon, \quad \forall t \geq 0.$$

Il movimento è instabile se non soddisfa questa condizione di stabilità, ed è asintoticamente stabile se, oltre ad essere stabile, soddisfa la condizione:

$$\lim_{t\to\infty}\|\delta\boldsymbol{x}(t)\| = 0, \quad \forall\delta\boldsymbol{x}_0 : \|\delta\boldsymbol{x}_0\| < \delta_\varepsilon.$$

Appare del tutto evidente che qualunque sia il movimento nominale $\boldsymbol{x}_n$ preso in considerazione, l'analisi della sua stabilità conduce sempre all'analisi del comportamento delle soluzioni dell'equazione differenziale $\delta\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\delta\boldsymbol{x}(t)$, al variare della condizione iniziale. Poiché il risultato di questa analisi è ovviamente lo stesso qualunque sia il movimento $\boldsymbol{x}_n$ di cui si studia la stabilità, potremo concludere che in un sistema dinamico LTI i movimenti sono tutti stabili, o tutti instabili o tutti asintoticamente stabili. In questo senso si può parlare di **stabilità del sistema**: un sistema dinamico LTI sarà stabile se tutti i suoi movimenti sono stabili (analogamente per l'instabilità e l'asintotica stabilità).

Si osservi inoltre che l'equazione differenziale $\delta\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\delta\boldsymbol{x}(t)$, le cui soluzioni al variare dello stato iniziale definiscono la stabilità di tutti i movimenti del sistema dinamico oggetto di studio, è formalmente identica all'equazione di stato del sistema originario, pur di annullare l'ingresso. Le soluzioni di questa equazione sono quindi i moti liberi del sistema, che si generano al variare della condizione iniziale: in altre parole, l'analisi di stabilità di un sistema LTI coinvolge solo il **moto libero** del sistema.

Ricordando che l'espressione del moto libero del sistema (a partire dall'istante iniziale t=0 con condizione iniziale $\delta\boldsymbol{x}_0$) è la seguente:

$$\delta\boldsymbol{x}_l(t) = e^{\boldsymbol{A}t}\delta\boldsymbol{x}_0\,,$$

si deduce che la stabilità del sistema si analizzerà studiando le proprietà della matrice $\boldsymbol{A}$.

D'altra parte il fatto che il moto libero si scali proporzionalmente alla condizione iniziale $\boldsymbol{x}_0$ porta a concludere che la condizione di **stabilità** è verificata **se tutti i moti liberi** (cioè tutti i movimenti che si generano al variare della condizione iniziale) **sono limitati** nel tempo: in questo caso infatti sarà possibile soddisfare la condizione di stabilità precedentemente enunciata, qualunque sia ε, pur di scegliere in corrispondenza un valore δ_ε sufficientemente piccolo (e quindi limitare la norma dello stato iniziale $\delta\boldsymbol{x}_0$). Se questo non accade, ossia se c'è almeno un moto libero del sistema che non è limitato nel tempo, il sistema è instabile. Infine se tutti i moti liberi del sistema, oltre ad essere limitati, tendono a zero per $t\to\infty$, il sistema è asintoticamente stabile (in questo caso, infatti, lo scostamento tra movimento perturbato e movimento nominale decade a zero, qualunque sia lo stato iniziale perturbato).

Stabilità ed autovalori

Lo studio della stabilità alla Lyapunov di un sistema LTI ha portato a concludere che occorre studiare le soluzioni dell'equazione[1]:

$$\boldsymbol{x}_l(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0$$

(moti liberi del sistema) al variare della condizione iniziale $\boldsymbol{x}_0$. Ci si chiede in particolare se tali soluzioni siano limitate nel tempo e, in caso affermativo, se convergano a zero.

Ci muoveremo nell'ipotesi semplificativa che la matrice $\boldsymbol{A}$ sia diagonalizzabile, ossia che esista una matrice non singolare $\boldsymbol{T}$ tale che:

$$\hat{\boldsymbol{A}} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1} = diag\{\lambda_i, i = 1, \ldots, n\},$$

dove λ_i sono gli autovalori di $\boldsymbol{A}$. Ricordiamo che condizione sufficiente (non necessaria) perché la matrice $\boldsymbol{A}$ sia diagonalizzabile è che gli n autovalori siano distinti, mentre condizione necessaria e sufficiente è che agli n autovalori siano associati n autovettori linearmente indipendenti.

Come già dimostrato, l'espressione del moto libero in termini della matrice diagonale $\hat{\boldsymbol{A}}$ è la seguente:

$$\boldsymbol{x}_l(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0 = e^{\boldsymbol{T}^{-1}\hat{\boldsymbol{A}}\boldsymbol{T}t}\boldsymbol{x}_0 = \boldsymbol{T}^{-1}e^{\hat{\boldsymbol{A}}t}\boldsymbol{T}\boldsymbol{x}_0 = \boldsymbol{T}^{-1}diag\{e^{\lambda_i t}\}\boldsymbol{T}\boldsymbol{x}_0 .$$

Pertanto i moti liberi del sistema sono combinazioni lineari, attraverso gli elementi delle matrici $\boldsymbol{T}$ e $\boldsymbol{T}^{-1}$ e le componenti del vettore $\boldsymbol{x}_0$, degli esponenziali $e^{\lambda_i t}$.

Si ricorda che gli autovalori prendono valore nel campo complesso, e formano un insieme simmetrico rispetto all'asse reale (se esiste un autovalore a parte immaginaria non nulla, esiste anche il complesso coniugato). Dette α_i e β_i rispettivamente la parte reale e immaginaria dell'autovalore λ_i, risulta:

$$e^{\lambda_i t} = e^{\alpha_i t}e^{j\beta_i t} = e^{\alpha_i t}\left(\cos(\beta_i t) + j\sin(\beta_i t)\right).$$

Si osservi però che la combinazione lineare attraverso gli elementi delle matrici $\boldsymbol{T}$ e $\boldsymbol{T}^{-1}$ fa naturalmente elidere tutte le parti immaginarie dei numeri complessi, dal momento che l'esponenziale della matrice $\boldsymbol{A}$ non può che avere elementi reali.

Possiamo a questo punto concludere che se tutti gli autovalori di $\boldsymbol{A}$ hanno parte reale negativa ($\alpha_i < 0, \forall i$), tutti gli esponenziali $e^{\lambda_i t}$ sono limitati nel tempo e tendono a zero, monotonicamente se $\beta_i = 0$, oscillando se $\beta_i \neq 0$. Ne consegue che tutti i moti liberi del sistema (qualunque siano le componenti dello stato iniziale) sono limitati nel tempo e si esauriscono per $t \to \infty$, e quindi che il sistema è asintoticamente stabile.

Se invece tutti gli autovalori di $\boldsymbol{A}$ hanno parte reale negativa o nulla ($\alpha_i \leq 0, \forall i$), e vi è almeno un autovalore a parte reale nulla, tutti gli esponenziali $e^{\lambda_i t}$ sono limitati nel tempo ma quelli associati agli autovalori a parte reale nulla non convergono a zero, rimanendo costanti per $\beta_i = 0$, e oscillando permanentemente se $\beta_i \neq 0$. Essendo sempre possibile individuare

[1] A questo punto, l'indicazione dello scostamento δ non ha ovviamente più alcun rilievo.

almeno una condizione iniziale per cui il moto libero del sistema non converge, in norma, a zero, fermo restando che tutti i moti liberi sono limitati nel tempo, possiamo concludere che il sistema è stabile, non asintoticamente.

Infine se esiste almeno un autovalore di $\boldsymbol{A}$ a parte reale positiva ($\exists i : \alpha_i > 0$), c'è almeno un termine esponenziale $e^{\lambda_i t}$ che diverge (non è limitato nel tempo). Essendo sempre possibile individuare almeno una condizione iniziale per cui il moto libero del sistema non è limitato nel tempo possiamo concludere che il sistema è instabile.

Quanto sopra esposto costituisce dimostrazione del seguente **teorema**.

Un sistema dinamico LTI con matrice $\boldsymbol{A}$ diagonalizzabile è:

asintoticamente stabile:	se e solo se tutti gli autovalori di $\boldsymbol{A}$ hanno parte reale negativa;
stabile:	se e solo se tutti gli autovalori di $\boldsymbol{A}$ hanno parte reale negativa o nulla e ne esistono a parte reale nulla;
instabile:	se e solo se esistono autovalori di $\boldsymbol{A}$ a parte reale positiva.

L'analisi di stabilità di un sistema LTI si conduce quindi studiando la posizione nel piano complesso degli autovalori della matrice $\boldsymbol{A}$.

Osservazioni

1) Ricordiamo che a seguito di un cambiamento di variabili di stato descritto da una matrice di trasformazione $\boldsymbol{T}$, la matrice $\boldsymbol{A}$ del sistema si trasforma secondo una relazione di similitudine ($\hat{\boldsymbol{A}} = \boldsymbol{TAT}^{-1}$). Poiché matrici simili hanno gli stessi autovalori, l'analisi della stabilità è del tutto indipendente dalla scelta delle variabili di stato. In altre parole, la proprietà di stabilità è una **proprietà strutturale** del sistema dinamico.

2) Se la matrice $\boldsymbol{A}$ non è diagonalizzabile, può essere messa in relazione di similitudine con una forma canonica (forma di Jordan). Seguendo questa strada si giunge alla conclusione che il teorema precedentemente enunciato va corredato dalla precisazione che se vi sono autovalori multipli a parte reale nulla (e non vi sono autovalori a parte reale positiva), il sistema è instabile se per almeno uno degli autovalori a parte reale nulla la cosiddetta *molteplicità geometrica* (numero degli autovettori linearmente indipendenti associati all'autovalore) è inferiore alla *molteplicità algebrica* (molteplicità con cui l'autovalore è radice del polinomio caratteristico).

Si consideri a titolo d'esempio il sistema con matrice $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Il polinomio caratteristico della matrice è ovviamente $\chi_A(\lambda) = \lambda^2$, per cui l'autovalore λ=0 ha molteplicità algebrica 2 (è radice doppia del polinomio caratteristico). Per calcolare gli autovettori associati all'autovalore, impostiamo l'equazione:

$$\lambda \boldsymbol{x} = \boldsymbol{Ax}$$

da cui:

$$\begin{cases} x_2 = 0 \\ 0 = 0 \end{cases}.$$

Pertanto tutti gli autovalori hanno seconda componente nulla, ovvero tutti gli autovettori sono linearmente dipendenti dall'autovettore:

$$\boldsymbol{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

La molteplicità geometrica dell'autovalore è quindi pari a 1, inferiore alla molteplicità algebrica: pertanto il sistema è instabile.

A scopo di verifica, calcoliamo l'esponenziale della matrice $\boldsymbol{A}$:

$$e^{\boldsymbol{A}t} = \boldsymbol{I} + \boldsymbol{A}t + \frac{\boldsymbol{A}^2 t^2}{2!} + \frac{\boldsymbol{A}^3 t^3}{3!} + \cdots = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & t \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \cdots = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

L'espressione del moto libero è quindi:

$$\boldsymbol{x}_l(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0 = \begin{cases} x_{01} + x_{02}t \\ x_{02} \end{cases},$$

il che conferma che è sufficiente scegliere lo stato iniziale con componente x_{02} diversa da zero per generare un moto libero non limitato: il sistema è quindi instabile.

3) L'analisi di stabilità condotta sulla base della risposta all'impulso del sistema conduce alle stesse conclusioni della presente analisi. Nello studio della risposta all'impulso si era fatto riferimento ai poli della funzione di trasferimento, i quali d'altra parte coincidono con gli autovalori della matrice $\boldsymbol{A}$ in assenza di cancellazioni di radici tra numeratore e denominatore della funzione di trasferimento. Si osservi che autovalori multipli a parte reale nulla con molteplicità geometrica inferiore alla molteplicità algebrica danno luogo a poli multipli nella funzione di trasferimento che, anche in assenza di poli a parte reale positiva, comportano comunque l'instabilità del sistema.

Stabilità: analisi della matrice *A*

Nello studio della stabilità di un sistema LTI si può ricorrere, in taluni casi particolari, a criteri che, evitando il calcolo diretto degli autovalori, consentono di concludere circa la stabilità del sistema dinamico. Alcuni criteri basati sull'**ispezione della matrice *A*** sono i seguenti:

1) Se $\boldsymbol{A}$ è triangolare, gli autovalori sono sulla diagonale della matrice, per cui l'analisi di stabilità è immediata (si osservi che in questo caso gli autovalori non possono che essere reali).

2) Si ricorda che la traccia di una matrice è la somma degli elementi sulla diagonale principale. Poiché risulta:

$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \mathrm{Re}[\lambda_i],$$

se la traccia di $\boldsymbol{A}$ è positiva o nulla il sistema non è asintoticamente stabile ed è senz'altro instabile se la traccia è positiva (c'è almeno un autovalore a parte reale positiva).

3) Poiché:

$$\det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i ,$$

se il determinante della matrice $\boldsymbol{A}$ è nullo il sistema non è asintoticamente stabile (c'è almeno un autovalore nullo).

Criterio di Routh

Il criterio di Routh fornisce una **condizione necessaria e sufficiente** di asintotica stabilità di un sistema dinamico LTI sulla base di un'analisi dei coefficienti del polinomio caratteristico della sua matrice $\boldsymbol{A}$. In presenza di strumenti di calcolo numerico di diffusione ormai vastissima, in grado di calcolare le radici di polinomi di grado arbitrario, la rilevanza del criterio può apparire modesta, anche in considerazione di una certa laboriosità nella sua applicazione. E' invece indiscutibile il rilievo del criterio nella soluzione di problemi parametrici, ovvero in quei casi in cui i coefficienti del polinomio caratteristico dipendano da uno o più parametri e si vogliano ricavare condizioni di asintotica stabilità del sistema dinamico nello spazio dei parametri.

Sia dunque dato il polinomio caratteristico:

$$\chi_A(\lambda) = \varphi(\lambda) = \varphi_0\lambda^n + \varphi_1\lambda^{n-1} + \varphi_2\lambda^{n-2} + \cdots + \varphi_n$$

E' noto che sussiste una **condizione necessaria** perché tutte le radici del polinomio abbiano parte reale negativa, e quindi perché il sistema sia asintoticamente stabile: se il sistema è asintoticamente stabile, tutti i coefficienti $\varphi_0, \varphi_1, \ldots, \varphi_n$ del polinomio caratteristico sono concordi (hanno lo stesso segno).

Essendo la condizione solo necessaria, non consente di concludere nulla circa l'asintotica stabilità del sistema quando è verificata.

Per l'applicazione del criterio di Routh, occorre preliminarmente formare una tabella (**tabella di Routh**). La tabella è composta di n+1 righe ed è inizializzata con due righe, la prima delle quali è formata dai coefficienti φ_i a partire da φ_0, presi uno sì e uno no, la seconda dai restanti coefficienti φ_i:

$$\begin{array}{cccc}
\varphi_0 & \varphi_2 & \varphi_4 & \cdots \\
\varphi_1 & \varphi_3 & \varphi_5 & \cdots \\
\vdots & \vdots & \vdots & \\
h_1 & h_2 & h_3 & \cdots \\
k_1 & k_2 & k_3 & \cdots \\
l_1 & l_2 & l_3 & \cdots \\
\vdots & \vdots & \vdots &
\end{array}$$

Per formare una generica riga della tabella (la l_sima) ci si serve delle due righe immediatamente precedenti (la k_sima e la h_sima), calcolando ogni elemento della riga secondo la formula:

$$l_i = -\frac{1}{k_1}\det\begin{bmatrix} h_1 & h_{i+1} \\ k_1 & k_{i+1} \end{bmatrix} = h_{i+1} - \frac{h_1 k_{i+1}}{k_1}.$$

Occorre quindi calcolare il determinante della matrice formata dagli elementi in prima colonna delle due righe immediatamente precedenti e dagli elementi di queste due righe nella colonna immediatamente successiva a quella cui appartiene l'elemento che si sta calcolando, cambiare segno al determinante e dividere il risultato per il primo elemento della riga immediatamente precedente. Si osservi che le righe della tabella possono essere completate con zeri laddove questo sia richiesto per il calcolo delle righe successive.

Se il primo elemento di una colonna risulta nullo, la tabella si dice non definita e il procedimento si arresta, altrimenti prosegue fino alla scritture della n+1_sima riga.

Il **criterio di Routh** afferma che il sistema è asintoticamente stabile se e solo se la tabella di Routh è ben definita e tutti gli elementi della prima colonna hanno lo stesso segno.

Esempio

Si consideri il polinomio:

$$\varphi(\lambda) = \lambda^3 + 4\lambda^2 + 5\lambda + 2 .$$

Essendo tutti i coefficienti del polinomio positivi, la condizione necessaria di stabilità non consente di concludere nulla. Costruiamo la tabella di Routh:

$$\begin{array}{ccc} 1 & 5 & 0 \\ 4 & 2 & 0 \\ k_1 & k_2 & \\ h_1 & & \end{array}$$

I coefficienti della tabella indicati simbolicamente si calcolano come illustrato precedentemente:

$$k_1 = -\frac{1}{4}\det\begin{bmatrix} 1 & 5 \\ 4 & 2 \end{bmatrix} = 4.5, \quad k_2 = -\frac{1}{4}\det\begin{bmatrix} 1 & 0 \\ 4 & 0 \end{bmatrix} = 0 .$$

$$h_1 = -\frac{1}{k_1}\det\begin{bmatrix} 4 & 2 \\ k_1 & k_2 \end{bmatrix} = 2$$

La tabella è quindi la seguente:

$$\begin{array}{ccc} 1 & 5 & 0 \\ 4 & 2 & 0 \\ 4.5 & 0 & \\ 2 & & \end{array}$$

Poiché tutti i coefficienti della prima colonna sono positivi, il sistema è asintoticamente stabile.

Esempio

Si consideri il polinomio:

$$\varphi(\lambda) = \lambda^4 + 2\lambda^3 + \lambda^2 + \lambda + a ,$$

dove a è un parametro reale. La condizione necessaria richiede che il parametro a sia positivo per l'asintotica stabilità del sistema. Costruiamo la tabella di Routh:

$$\begin{array}{ccc} 1 & 1 & a \\ 2 & 1 & 0 \\ k_1 & k_2 & \\ h_1 & h_2 & \\ l_1 & & \end{array}$$

I coefficienti della tabella indicati simbolicamente si calcolano come illustrato precedentemente:

$$k_1 = -\frac{1}{2}\det\begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} = 0.5, \quad k_2 = -\frac{1}{2}\det\begin{bmatrix} 1 & a \\ 2 & 0 \end{bmatrix} = a$$

$$h_1 = -\frac{1}{k_1}\det\begin{bmatrix} 2 & 1 \\ k_1 & k_2 \end{bmatrix} = \frac{0.5 - 2a}{0.5} = 1 - 4a, \quad h_2 = -\frac{1}{k_1}\det\begin{bmatrix} 2 & 0 \\ k_1 & 0 \end{bmatrix} = 0\,.$$

$$l_1 = -\frac{1}{h_1}\det\begin{bmatrix} k_1 & k_2 \\ h_1 & h_2 \end{bmatrix} = a$$

La tabella è quindi la seguente:

$$\begin{array}{ccc} 1 & 1 & a \\ 2 & 1 & 0 \\ 0.5 & a & \\ 1-4a & 0 & \\ a & & \end{array}$$

Affinché tutti i coefficienti della prima colonna siano concordi (positivi) devono essere soddisfatte le due condizioni:

$$\begin{cases} 1 - 4a > 0 \\ a > 0 \end{cases},$$

ossia:

$$0 < a < \frac{1}{4}\,.$$

Stabilità degli stati di equilibrio

Discutendo della stabilità alla Lyapunov, si è dato significato al concetto di stabilità di un movimento ed in particolare di uno stato di equilibrio. Mentre la verifica di stabilità di un movimento per un sistema non lineare è decisamente complessa, per l'analisi di stabilità di uno stato di equilibrio esistono delle semplici condizioni deducibili dall'analisi del **sistema linearizzato** nell'intorno dello stato di equilibrio.

Si consideri dunque un sistema non lineare di equazioni di stato:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t)),$$

ed, in corrispondenza dell'ingresso costante $\bar{\boldsymbol{u}}$ esista lo stato di equilibrio $\bar{\boldsymbol{x}}$ (risulta quindi $\boldsymbol{f}(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}) = 0$). Si consideri la matrice $\boldsymbol{A}$ del sistema linearizzato nell'intorno dello stato di equilibrio:

$$\boldsymbol{A} = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}}$$

Si hanno i seguenti risultati:

1) Se la matrice $\boldsymbol{A}$ ha tutti autovalori a parte reale negativa (ossia se il sistema linearizzato è asintoticamente stabile) lo stato di equilibrio $\bar{\boldsymbol{x}}$ è **asintoticamente stabile**.

2) Se la matrice $\boldsymbol{A}$ ha almeno un autovalore a parte reale positiva lo stato di equilibrio $\bar{\boldsymbol{x}}$ è **instabile**.

Resta come caso indecidibile sulla base dell'analisi del sistema linearizzato quello in cui la matrice $\boldsymbol{A}$ non ha autovalori a parte reale positiva ma ne ha a parte reale nulla. In questo caso occorrono approssimazioni del sistema originario non lineare estese a termini di ordine superiore al primo per decidere circa la stabilità dello stato di equilibrio.

Esempio

Si riprenda l'esempio del primo ordine in cui il legame tra x e $\dot{x}$ imposto dalla relazione:

$$\dot{x} = f(x, \bar{u})$$

è rappresentato dalla curva in figura 6.

Si è già osservato che il sistema ammette tre stati di equilibrio, $\bar{x}_1, \bar{x}_2, \bar{x}_3$.

Il sistema linearizzato nell'intorno di questi punti di equilibrio assume espressione:

$$\delta\dot{x}(t) = a\delta x(t) + b\delta u(t),$$

con:

$$a = \left.\frac{\partial f(x,u)}{\partial x}\right|_{\bar{x}_i, \bar{u}} = \left.\frac{df(x,\bar{u})}{dx}\right|_{\bar{x}_i},$$

unico autovalore del sistema. Per gli stati $\bar{x}_1$ e $\bar{x}_3$ risulta a>0, per cui, in base al secondo dei due risultati precedentemente enunciati, i due stati di equilibrio sono instabili. Per lo stato $\bar{x}_2$

risulta invece a<0, per cui, sulla base del primo risultato, lo stato di equilibrio è asintoticamente stabile.

Invarianza della funzione di trasferimento

Dato un sistema dinamico LTI descritto dalle equazioni:

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t)\\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t)\end{aligned}$$

è noto che una rappresentazione alternativa del sistema si ottiene introducendo i vettori $\boldsymbol{U}(s)$ e $\boldsymbol{Y}(s)$, rispettivamente vettori delle trasformate di Laplace degli ingressi e delle uscite del sistema dinamico. Assunto lo stato iniziale del sistema nullo, il legame tra i due vettori è espresso dalla **funzione di trasferimento**:

$$\boldsymbol{G}(s) = \boldsymbol{C}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D},$$

matrice a p righe (quante sono le uscite) e m colonne (quanti sono gli ingressi), funzione della variabile complessa s.

Si supponga ora di effettuare un cambiamento di variabili di stato:

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{T}\boldsymbol{x}(t), \quad \det(\boldsymbol{T}) \neq 0.$$

Il sistema nelle nuove variabili di stato è retto dalle equazioni:

$$\begin{aligned}\dot{\hat{\boldsymbol{x}}}(t) &= \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}(t) + \hat{\boldsymbol{B}}\boldsymbol{u}(t)\\ \boldsymbol{y}(t) &= \hat{\boldsymbol{C}}\hat{\boldsymbol{x}}(t) + \hat{\boldsymbol{D}}\boldsymbol{u}(t)\end{aligned},$$

dove:

$$\hat{\boldsymbol{A}} = \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}, \quad \hat{\boldsymbol{B}} = \boldsymbol{T}\boldsymbol{B}, \quad \hat{\boldsymbol{C}} = \boldsymbol{C}\boldsymbol{T}^{-1}, \quad \hat{\boldsymbol{D}} = \boldsymbol{D}.$$

Calcoliamo la funzione di trasferimento per questo sistema:

$$\begin{aligned}\hat{\boldsymbol{G}}(s) &= \hat{\boldsymbol{C}}\left(s\boldsymbol{I}_n - \hat{\boldsymbol{A}}\right)^{-1}\hat{\boldsymbol{B}} + \hat{\boldsymbol{D}} = \boldsymbol{C}\boldsymbol{T}^{-1}\left(s\boldsymbol{T}\boldsymbol{T}^{-1} - \boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1}\right)^{-1}\boldsymbol{T}\boldsymbol{B} + \boldsymbol{D} =\\ &= \boldsymbol{C}\boldsymbol{T}^{-1}\left[\boldsymbol{T}(s\boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{T}^{-1}\right]^{-1}\boldsymbol{T}\boldsymbol{B} + \boldsymbol{D} = \boldsymbol{C}\boldsymbol{T}^{-1}\boldsymbol{T}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{T}^{-1}\boldsymbol{T}\boldsymbol{B} + \boldsymbol{D} =\\ &= \boldsymbol{C}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D} = \boldsymbol{G}(s)\end{aligned}$$

Pertanto l'espressione della funzione di trasferimento non muta al variare della rappresentazione di stato, ed è quindi **proprietà strutturale** del sistema dinamico.

Dinamiche nascoste

Facendo riferimento ad un sistema SISO, per il quale la funzione di trasferimento è scalare, è noto che la funzione di trasferimento stessa è una funzione razionale in s, ossia esprimibile come il rapporto di due polinomi:

$$G(s) = \frac{N(s)}{D(s)}.$$

Nel formare l'espressione della funzione di trasferimento a partire dalle matrici del sistema dinamico espresso nel dominio del tempo, possono intervenire **cancellazioni** tra termini a numeratore e denominatore. Il risultato netto è che il denominatore della funzione di trasferimento è un polinomio di grado inferiore all'ordine n del sistema.

Vedremo ora alcuni esempi in cui intervengono cancellazioni e l'interpretazione sistemistica che si può dare a questa circostanza.

Esempio 1

Si consideri il sistema del secondo ordine:

$$\begin{cases} \dot{x}_1 = -2x_1 \\ \dot{x}_2 = x_1 - x_2 + u \\ y = 2x_1 + 3x_2 \end{cases} \Rightarrow \quad \boldsymbol{A} = \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 2 & 3 \end{bmatrix}, \quad \boldsymbol{D} = 0,$$

e se ne calcoli la funzione di trasferimento:

$$\boldsymbol{G}(s) = \boldsymbol{C}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{B} + \boldsymbol{D} = \ldots = \frac{3(s+2)}{(s+2)(s+1)} = \frac{3}{s+1}.$$

Si manifesta quindi una cancellazione della radice -2. Si osservi che la variabile di stato x_1 non è in alcun modo influenzata dall'ingresso del sistema: essa influenza l'uscita (sia direttamente che indirettamente attraverso la seconda equazione di stato), ma limitatamente al moto libero. Se lo stato iniziale è nullo, la variabile x_1 è identicamente nulla.

Nel sistema è quindi presente una "dinamica nascosta", associata alla prima variabile di stato, che non partecipa al legame ingresso-uscita e rende il sistema dinamico esternamente di ordine inferiore a quanto è in realtà.

Esempio 2

Si consideri il sistema del secondo ordine:

$$\begin{cases} \dot{x}_1 = -x_1 + x_2 + u \\ \dot{x}_2 = -x_2 + u \\ y = x_2 \end{cases} \Rightarrow \quad \boldsymbol{A} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \boldsymbol{D} = 0,$$

e se ne calcoli la funzione di trasferimento:

$$G(s) = C(sI_n - A)^{-1} B + D = \ldots = \frac{s+1}{(s+1)^2} = \frac{1}{s+1}.$$

Si manifesta quindi una cancellazione della radice −1. Si osservi che la variabile di stato x_1 non influenza l'uscita del sistema, né direttamente, né indirettamente attraverso la seconda variabile di stato. La variabile di stato x_1 è influenzata dall'ingresso, ma la sua evoluzione non ha alcun riscontro nell'uscita y.

Nel sistema è quindi presente una "dinamica nascosta", associata alla prima variabile di stato, che non partecipa al legame ingresso-uscita e rende il sistema dinamico esternamente di ordine 1.

Esempio 3

Si consideri la rete elettrica riportata in figura:

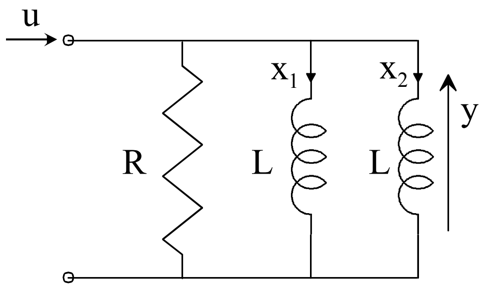

Fig. 7 : Rete elettrica

Dette x_1 e x_2 le correnti nei due induttori, il sistema dinamico è costituito dalle equazioni:

$$\begin{cases} L\dot{x}_1 = R(u - x_1 - x_2) \\ L\dot{x}_2 = R(u - x_1 - x_2) \\ y = R(u - x_1 - x_2) \end{cases} \Rightarrow \quad \boldsymbol{A} = \begin{bmatrix} -\frac{R}{L} & -\frac{R}{L} \\ -\frac{R}{L} & -\frac{R}{L} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \frac{R}{L} \\ \frac{R}{L} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} -R & -R \end{bmatrix}, \quad \boldsymbol{D} = R,$$

Per l'analisi del sistema risulta più comodo effettuare il seguente cambiamento di variabili di stato:

$$\begin{cases} \hat{x}_1 = x_1 + x_2 \\ \hat{x}_2 = x_1 - x_2 \end{cases}$$

Sommando e sottraendo le equazioni si ottiene:

$$\begin{cases} L\dot{\hat{x}}_1 = 2R(u - \hat{x}_1) \\ L\dot{\hat{x}}_2 = 0 \\ y = R(u - \hat{x}_1) \end{cases} \Rightarrow \quad \hat{\boldsymbol{A}} = \begin{bmatrix} -2\frac{R}{L} & 0 \\ 0 & 0 \end{bmatrix}, \quad \hat{\boldsymbol{B}} = \begin{bmatrix} 2\frac{R}{L} \\ 0 \end{bmatrix}, \quad \hat{\boldsymbol{C}} = \begin{bmatrix} -R & 0 \end{bmatrix}, \quad \hat{\boldsymbol{D}} = R$$

e quindi la funzione di trasferimento:

$$\boldsymbol{G}(s) = \hat{\boldsymbol{G}}(s) = \hat{\boldsymbol{C}}\left(s\boldsymbol{I}_n - \hat{\boldsymbol{A}}\right)^{-1} \hat{\boldsymbol{B}} + \hat{\boldsymbol{D}} = \ldots = \frac{RLs}{Ls + 2R}.$$

La variabile di stato $\hat{x}_2$ non è influenzata dall'ingresso e non influenza l'uscita. Pertanto la dinamica ad essa associata (legata alla differenza tra le correnti negli induttori) non partecipa al legame ingresso-uscita e si configura come una "dinamica nascosta". Agli effetti esterni, la rete elettrica originaria è del tutto equivalente alla rete elettrica semplificata con un solo induttore di induttanza $L/2$ riportata in figura:

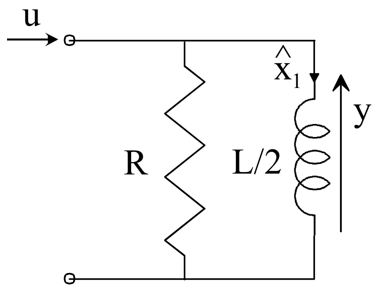

Fig. 8 : Rete elettrica equivalente agli effetti esterni

Raggiungibilità ed osservabilità

La presenza di "dinamiche nascoste" in un sistema, che non si manifestano nel legame ingresso-uscita, può esser formalizzata ricorrendo alla teoria della raggiungibilità e dell'osservabilità, di cui daremo alcuni cenni, anche perché propedeutica alla tecnica di controllo ad assegnamento degli autovalori che verrà trattata nel successivo capitolo.

Diremo che un sistema dinamico lineare tempo invariante è (completamente) **raggiungibile** se è possibile trasferire lo stato del sistema dall'origine a qualsiasi punto in $\Re^n$, con un'opportuna scelta dell'ingresso, in un tempo finito arbitrario.

Un sistema dinamico LTI si dice invece (completamente) **osservabile** se qualunque sia il vettore in $\Re^n$ che costituisce lo stato iniziale, l'uscita libera che origina da tale punto differisce, su qualunque intervallo di tempo finito arbitrario, dall'uscita identicamente nulla.

La raggiungibilità formalizza quindi la possibilità di pilotare il sistema dinamico dall'ingresso in modo da conseguire un'evoluzione arbitraria delle variabili di stato. L'osservabilità concerne invece la possibilità, osservando l'uscita libera del sistema, di distinguere lo stato iniziale dall'origine dello spazio di stato: si vedrà in seguito che di fatto la proprietà di osservabilità consente di ricostruire l'andamento delle variabili di stato del sistema a partire da osservazioni sulle sue variabili di uscita.

Se un sistema di ordine n non è completamente raggiungibile, può tuttavia essere scomposto, previo un opportuno cambiamento di variabili di stato, in due parti, una delle quali è completamente raggiungibile, l'altra completamente non raggiungibile (per la quale cioè non è possibile spostare lo stato dall'origine in alcun punto dello spazio di stato agendo sull'ingresso). Analogamente, un sistema non completamente osservabile può essere scomposto in due parti, una completamente osservabile, l'altra completamente non osservabile (per la quale cioè non è possibile distinguere alcun stato iniziale dall'origine dello spazio di stato sulla base dell'uscita libera).

Se ne deduce che un qualsiasi sistema LTI può essere messo, con un opportuno cambiamento di variabili di stato in una forma canonica, in cui sono poste in evidenza quattro parti, raggiungibili e no, osservabili e no, eventualmente vuote. Il risultato di questo procedimento, che va sotto il nome di **scomposizione canonica di Kalman** e per il quale sono richieste conoscenze di algebra lineare non fornite in questo corso, è rappresentato dallo schema a blocchi in figura, nel quale sono rappresentate le quattro parti[2], con le mutue influenze e le influenze con ingresso e uscita:

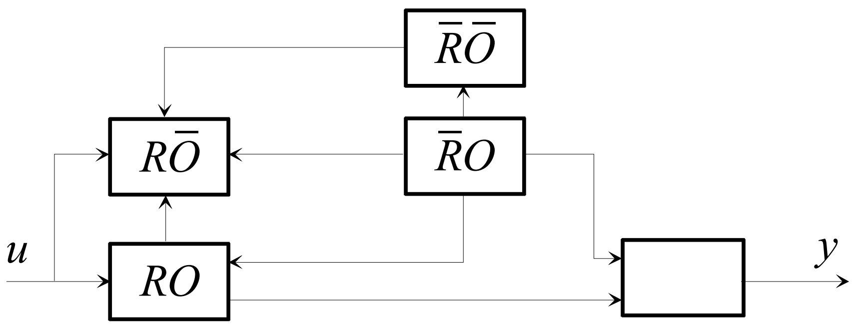

Fig. 9 : Scomposizione canonica di Kalman

---

[2] Le parti non raggiungibili sono indicate con $\overline{R}$, quelle non osservabili con $\overline{O}$.

Si osservi che nel percorso tra ingresso e uscita del sistema interviene solo la parte raggiungibile ed osservabile del sistema, ovvero la **funzione di trasferimento** del sistema coinvolge solo la parte **raggiungibile ed osservabile**. Le altre parti costituiscono dinamiche nascoste, che non partecipano al legame ingresso-uscita.

Esiste un semplice **test di raggiungibilità**, che consente di stabilire se un sistema è completamente raggiungibile. Si costruisca la matrice di raggiungibilità (a n righe e $n \times m$ colonne):

$$\boldsymbol{K}_r = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} & \boldsymbol{A}^2\boldsymbol{B} & \cdots & \boldsymbol{A}^{n-1}\boldsymbol{B} \end{bmatrix}.$$

Il sistema è completamente raggiungibile se e solo se la matrice di raggiungibilità $\boldsymbol{K}_r$ ha rango n, ossia è possibile estrarre n colonne che formino una matrice quadrata di ordine n non singolare. Si osservi che se il sistema è a singolo ingresso (m=1), il test si riduce a verificare che la matrice $\boldsymbol{K}_r$ sia **non singolare**.

Analogamente, esiste un **test di osservabilità**, che consente di stabilire se un sistema è completamente osservabile. Si costruisca la matrice di osservabilità (a n righe e $n \times p$ colonne):

$$\boldsymbol{K}_o = \begin{bmatrix} \boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T & {\boldsymbol{A}^T}^2\boldsymbol{C}^T & \cdots & {\boldsymbol{A}^T}^{n-1}\boldsymbol{C}^T \end{bmatrix}.$$

Il sistema è completamente osservabile se e solo se la matrice di osservabilità $\boldsymbol{K}_o$ ha rango n, ossia è possibile estrarre n colonne che formino una matrice quadrata di ordine n non singolare. Si osservi che se il sistema è a singola uscita (p=1), il test si riduce a verificare che la matrice $\boldsymbol{K}_o$ sia **non singolare**.

Esempi

Si riprendano gli esempi del paragrafo precedente. Nel primo esempio risulta:

$$\boldsymbol{K}_r = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad \det(\boldsymbol{K}_r) = 0$$

$$\boldsymbol{K}_o = \begin{bmatrix} \boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & -3 \end{bmatrix} \quad \Rightarrow \quad \det(\boldsymbol{K}_o) \neq 0$$

Il sistema è quindi completamente osservabile ma non raggiungibile. Sarà quindi scomponibile in una parte raggiungibile ed osservabile ed una parte non raggiungibile e osservabile.

Nel secondo esempio risulta:

$$\boldsymbol{K}_r = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{AB} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad \det(\boldsymbol{K}_r) \neq 0$$

$$\boldsymbol{K}_o = \begin{bmatrix} \boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad \det(\boldsymbol{K}_o) = 0$$

Il sistema è quindi completamente raggiungibile ma non osservabile. Sarà quindi scomponibile in una parte raggiungibile ed osservabile ed una parte raggiungibile e non osservabile.

Per quanto riguarda il terzo esempio, utilizzando le variabili di stato $\hat{x}_1, \hat{x}_2$, risulta:

$$\hat{\boldsymbol{K}}_r = \begin{bmatrix}\hat{\boldsymbol{B}} & \hat{\boldsymbol{A}}\hat{\boldsymbol{B}}\end{bmatrix} = \begin{bmatrix} 2\frac{R}{L} & -4\frac{R^2}{L^2} \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \det\left(\hat{\boldsymbol{K}}_r\right) = 0$$

$$\hat{\boldsymbol{K}}_o = \begin{bmatrix}\hat{\boldsymbol{C}}^T & \hat{\boldsymbol{A}}^T\hat{\boldsymbol{C}}^T\end{bmatrix} = \begin{bmatrix} -R & 2\frac{R^2}{L} \\ 0 & 0 \end{bmatrix} \quad \Rightarrow \quad \det\left(\hat{\boldsymbol{K}}_o\right) = 0$$

Il sistema non è quindi né completamente raggiungibile, né completamente osservabile. Sarà quindi scomponibile in una parte raggiungibile ed osservabile ed una parte non raggiungibile e non osservabile.

Traendo spunto anche dal terzo esempio, ci si può chiedere se le proprietà di raggiungibilità e osservabilità dipendano dalla scelta delle variabili di stato.

Si supponga quindi di effettuare un cambiamento di variabili di stato:

$$\hat{\boldsymbol{x}}(t) = \boldsymbol{T}\boldsymbol{x}(t), \quad \det(\boldsymbol{T}) \neq 0 .$$

Il sistema nelle nuove variabili di stato è caratterizzato dalle matrici:

$$\hat{\boldsymbol{A}} = \boldsymbol{TAT}^{-1}, \quad \hat{\boldsymbol{B}} = \boldsymbol{TB}, \quad \hat{\boldsymbol{C}} = \boldsymbol{CT}^{-1}, \quad \hat{\boldsymbol{D}} = \boldsymbol{D} .$$

La matrice di raggiungibilità per il sistema nelle nuove variabili di stato sarà quindi:

$$\begin{aligned}\hat{\boldsymbol{K}}_r &= \begin{bmatrix}\hat{\boldsymbol{B}} & \hat{\boldsymbol{A}}\hat{\boldsymbol{B}} & \hat{\boldsymbol{A}}^2\hat{\boldsymbol{B}} & \cdots & \hat{\boldsymbol{A}}^{n-1}\hat{\boldsymbol{B}}\end{bmatrix} = \\ &= \begin{bmatrix}\boldsymbol{TB} & \boldsymbol{TAT}^{-1}\boldsymbol{TB} & \boldsymbol{TAT}^{-1}\boldsymbol{TAT}^{-1}\boldsymbol{TB} & \cdots & \left(\boldsymbol{TAT}^{-1}\boldsymbol{TAT}^{-1}\cdots\boldsymbol{TAT}^{-1}\right)\boldsymbol{TB}\end{bmatrix} = . \\ &= \boldsymbol{T}\begin{bmatrix}\boldsymbol{B} & \boldsymbol{AB} & \boldsymbol{A}^2\boldsymbol{B} & \cdots & \boldsymbol{A}^{n-1}\boldsymbol{B}\end{bmatrix} = \boldsymbol{TK}_r\end{aligned}$$

Poiché $\boldsymbol{T}$ è non singolare, il rango di $\hat{\boldsymbol{K}}_r$ coincide con il rango di $\boldsymbol{K}_r$. Pertanto il sistema nelle nuove variabili di stato è completamente raggiungibile se e solo se lo è quello nelle variabili di stato originarie, ovvero la **raggiungibilità è proprietà strutturale**.

Analogamente, la matrice di osservabilità per il sistema nelle nuove variabili di stato sarà:

$$\begin{aligned}\hat{\boldsymbol{K}}_o &= \begin{bmatrix}\hat{\boldsymbol{C}}^T & \hat{\boldsymbol{A}}^T\hat{\boldsymbol{C}}^T & \hat{\boldsymbol{A}}^{T^2}\hat{\boldsymbol{C}}^T & \cdots & \hat{\boldsymbol{A}}^{T^{n-1}}\hat{\boldsymbol{C}}^T\end{bmatrix} = \\ &= \begin{bmatrix}\boldsymbol{T}^{-T}\boldsymbol{C}^T & \boldsymbol{T}^{-T}\boldsymbol{A}^T\boldsymbol{T}^T\boldsymbol{T}^{-T}\boldsymbol{C}^T & \boldsymbol{T}^{-T}\boldsymbol{A}^T\boldsymbol{T}^T\boldsymbol{T}^{-T}\boldsymbol{A}^T\boldsymbol{T}^T\boldsymbol{T}^{-T}\boldsymbol{C}^T & \cdots\end{bmatrix} \\ &= \boldsymbol{T}^{-T}\begin{bmatrix}\boldsymbol{C}^T & \boldsymbol{A}^T\boldsymbol{C}^T & \boldsymbol{A}^{T^2}\boldsymbol{C}^T & \cdots & \boldsymbol{A}^{T^{n-1}}\boldsymbol{C}^T\end{bmatrix} = \boldsymbol{T}^{-T}\boldsymbol{K}_o\end{aligned}$$

Poiché $\boldsymbol{T}$ (e quindi $\boldsymbol{T}^{-T}$) è non singolare, il rango di $\hat{\boldsymbol{K}}_o$ coincide con il rango di $\boldsymbol{K}_o$. Pertanto il sistema nelle nuove variabili di stato è completamente osservabile se e solo se lo è quello nelle variabili di stato originarie, ovvero l'**osservabilità è proprietà strutturale**.

Realizzazione

Data una funzione di trasferimento, si definisce **realizzazione** la scrittura di un sistema dinamico nel dominio del tempo (equazioni di stato e trasformazioni di uscita) che ammetta come funzione di trasferimento quella data. Ci limiteremo, nella trattazione del problema, a sistemi SISO (un ingresso ed un'uscita).

E' evidente che il problema di realizzazione ammette infinite soluzioni. Tra tutte le soluzioni, quelle nelle quali la matrice $\boldsymbol{A}$ ha dimensioni uguali al denominatore della funzione di trasferimento prendono il nome di **realizzazioni minime**. Per quanto discusso precedentemente circa la scomposizione canonica di Kalman, si deduce che le realizzazioni minime sono costituite da sistemi dinamici LTI completamente raggiungibili e osservabili.

Assumono poi particolare rilevanza, tra le realizzazioni minime, alcune **forme canoniche**, immediatamente ricavabili dall'espressione della funzione di trasferimento. Queste forme canoniche semplificano, come si vedrà, la soluzione di alcuni problemi rilevanti nello studio dei sistemi nello spazio di stato, quali l'assegnamento degli autovalori e la stima dello stato.

Consideriamo dunque una generica funzione di trasferimento razionale:

$$G(s)=\frac{N(s)}{D(s)},$$

con $D(s)$ di grado n. Se anche il numeratore è di grado n (sistema non strettamente proprio), sarà sempre possibile riscrivere $G(s)$ come:

$$G(s)=\frac{\tilde{N}(s)}{D(s)}+D,$$

dove D è lo scalare che lega $u(t)$ a $y(t)$ nella trasformazione d'uscita, mentre $\tilde{N}(s)$ è di grado inferiore a n. Possiamo quindi concentrarci sulla seguente espressione di G:

$$G(s)=\frac{b_n s^{n-1}+b_{n-1}s^{n-2}+\ldots+b_2 s+b_1}{s^n+a_n s^{n-1}+a_{n-1}s^{n-2}+\ldots+a_2 s+a_1}.$$

La **forma canonica di raggiungibilità** (o **forma canonica di controllo**) è la seguente realizzazione minima:

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_1 & -a_2 & -a_3 & \cdots & -a_n \end{bmatrix}, \quad \boldsymbol{B}=\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$\boldsymbol{C}=\begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_n \end{bmatrix}$$

E' facile verificare, trasformando secondo Laplace le singole equazioni di stato e procedendo per sostituzione, che la forma canonica di controllo è effettivamente una realizzazione della funzione di trasferimento.

Esempio

Sia:

$$G(s)=\frac{s^2+1}{s^2+4s+3}\,.$$

Scomponendo la frazione, otteniamo:

$$G(s)=\frac{-4s-2}{s^2+4s+3}+1\,.$$

Pertanto la forma canonica di controllo è la seguente:

$$\boldsymbol{A}=\begin{bmatrix}0 & 1\\ -3 & -4\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}0\\ 1\end{bmatrix},$$

$$\boldsymbol{C}=\begin{bmatrix}-2 & -4\end{bmatrix},\quad \boldsymbol{D}=1.$$

La forma canonica di controllo è sempre **completamente raggiungibile** per costruzione. Infatti:

$$\boldsymbol{K}_r=\begin{bmatrix}\boldsymbol{B} & \boldsymbol{AB} & \boldsymbol{A}^2\boldsymbol{B} & \cdots & \boldsymbol{A}^{n-1}\boldsymbol{B}\end{bmatrix}=\begin{bmatrix}0 & 0 & \cdot^{\cdot^{\cdot}} & 0 & 1\\ 0 & \cdot^{\cdot^{\cdot}} & 0 & 1 & *\\ \vdots & 0 & \cdot^{\cdot^{\cdot}} & \vdots & \vdots\\ 0 & 1 & * & \cdots & *\\ 1 & * & * & \cdots & *\end{bmatrix},$$

dove con * si indicano termini il cui valore è inessenziale. Il determinante di $\boldsymbol{K}_r$ vale ± 1, per cui il sistema è completamente raggiungibile. Per quanto osservato prima, il sistema è anche completamente osservabile, purché la funzione di trasferimento sia ben assegnata (non vi siano radici in comune tra numeratore e denominatore).

Un problema di non immediata soluzione è l'individuazione di un **cambiamento di variabili di stato** che porta un sistema LTI completamente raggiungibile in forma canonica di controllo.

Sia ($\boldsymbol{A}$,$\boldsymbol{B}$) la coppia di matrici del sistema di partenza, la cui matrice di raggiungibilità è:

$$\boldsymbol{K}_r=\begin{bmatrix}\boldsymbol{B} & \boldsymbol{AB} & \boldsymbol{A}^2\boldsymbol{B} & \cdots & \boldsymbol{A}^{n-1}\boldsymbol{B}\end{bmatrix},$$

e ($\hat{\boldsymbol{A}},\hat{\boldsymbol{B}}$) la coppia di matrici del sistema di arrivo, la cui matrice di raggiungibilità è:

$$\hat{\boldsymbol{K}}_r=\begin{bmatrix}\hat{\boldsymbol{B}} & \hat{\boldsymbol{A}}\hat{\boldsymbol{B}} & \hat{\boldsymbol{A}}^2\hat{\boldsymbol{B}} & \cdots & \hat{\boldsymbol{A}}^{n-1}\hat{\boldsymbol{B}}\end{bmatrix}=\boldsymbol{TK}_r\,.$$

Pertanto la matrice $\boldsymbol{T}$ di trasformazione è data da:

$\boldsymbol{T} = \hat{\boldsymbol{K}}_r \boldsymbol{K}_r^{-1}$.

Un **algoritmo** per ricavare la matrice $\boldsymbol{T}$ può dunque essere il seguente:

Data ($\boldsymbol{A}$,$\boldsymbol{B}$) coppia raggiungibile:

- Calcolare $\det(s\boldsymbol{I} - \boldsymbol{A}) = s^n + a_n s^{n-1} + a_{n-1}s^{n-2} + \ldots + a_2 s + a_1$
- Ricavare la coppia ($\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}$) dai coefficienti a_i
- Calcolare $\boldsymbol{K}_r$ e $\hat{\boldsymbol{K}}_r$
- Calcolare $\boldsymbol{T} = \hat{\boldsymbol{K}}_r \boldsymbol{K}_r^{-1}$

Una forma canonica utilizzata in alternativa alla forma canonica di controllo è la **forma canonica di osservabilità** (o **forma canonica di ricostruzione**).

Dato un sistema SISO strettamente proprio di matrici ($\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$), la sua funzione di trasferimento, scalare, coinciderà ovviamente con la sua trasposta:

$$G(s) = \boldsymbol{C}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{B} = G^T(s) = \left[\boldsymbol{C}(s\boldsymbol{I}_n - \boldsymbol{A})^{-1}\boldsymbol{B}\right]^T = \boldsymbol{B}^T(s\boldsymbol{I}_n - \boldsymbol{A})^{-T}\boldsymbol{C}^T = \boldsymbol{B}^T\left(s\boldsymbol{I}_n - \boldsymbol{A}^T\right)^{-1}\boldsymbol{C}^T.$$

Posto allora:

$$\tilde{\boldsymbol{A}} = \boldsymbol{A}^T, \quad \tilde{\boldsymbol{B}} = \boldsymbol{C}^T, \quad \tilde{\boldsymbol{C}} = \boldsymbol{B}^T,$$

risulta:

$$G(s) = \tilde{\boldsymbol{C}}\left(s\boldsymbol{I}_n - \tilde{\boldsymbol{A}}\right)^{-1}\tilde{\boldsymbol{B}},$$

per cui il sistema di matrici ($\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}}$), che prende il nome di sistema **duale** del sistema dato, costituisce una realizzazione di $G(s)$. La forma canonica di osservabilità è la duale della forma canonica di raggiungibilità:

$$\tilde{\boldsymbol{A}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_1 \\ 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & -a_n \end{bmatrix} \quad \tilde{\boldsymbol{B}} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix},$$

$$\tilde{\boldsymbol{C}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

E' facile dimostrare che la realizzazione è completamente osservabile per costruzione e completamente raggiungibile se la funzione di trasferimento è assegnata correttamente.

Esempio

Per l'esempio precedente, la forma canonica di osservabilità è la seguente:

$$\boldsymbol{A} = \begin{bmatrix} 0 & -3 \\ 1 & -4 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} -2 \\ -4 \end{bmatrix},$$

$$\boldsymbol{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \boldsymbol{D} = 1.$$