

Complementi per il corso
di
“Teoria dei Sistemi e del Controllo”

Andrea Gasparri
Dipartimento di Informatica e Automazione
Università degli Studi “Roma TRE”



17 novembre 2010

Indice

Indice	2
1 Introduzione	4
1.1 Modellazione	4
1.1.1 Rappresentazione Ingresso-Uscita	4
1.1.2 Rappresentazione Ingresso-Stato-Uscita	4
1.2 Variabili di Stato (V.d.S.)	4
2 Sistemi non Lineari e Linearizzazione	6
2.1 Motivazioni	6
2.2 Definizioni	6
2.3 Metodo di Linearizzazione	7
3 Soluzione delle Equazioni Differenziali	10
3.1 Equazioni differenziali lineari omogenee a coefficienti costanti	10
3.2 Equazioni differenziali lineari complete a coefficienti costanti	11
3.3 Riduzione equazione differenziale di ordine n a sistema di n equazioni del primo ordine	12
3.4 Esponenziale di matrice	13
3.5 Evoluzione Libera	15
3.6 Evoluzione Forzata	16
4 Relazioni tra le Rappresentazioni	19
4.1 Passaggio da V.d.S. a $F(s)$	19
4.2 $F(s)$ a V.d.S.	23
4.2.1 Forma Compagna di Controllore	23
4.2.2 Forma Compagna di Osservatore	25
4.2.3 Proprietà di Dualità	28
5 Decomposizione Modale	29
5.1 Autovalori ed Autovettori	29
5.2 Trasformazione di Coordinate	32
5.3 Diagonalizzazione	34
5.4 Jordanizzazione	37

5.5	Calcolo $\exp(A t)$ attraverso Decomposizione Modale	51
5.5.1	Matrici diagonalizzabili	51
5.6	Autovalori ed Autospazi	55
5.6.1	Applicazione al problema del Consenso	58
6	Proprietà Strutturali	64
6.1	Controllabilità	64
6.2	Osservabilità	69
6.3	Dualità	74
6.4	Forma di Kalman per la Controllabilità	76
6.5	Forma di Kalman per l' Osservabilità	82
6.6	Decomposizione Canonica di Kalman	86
7	Forme Canoniche e Realizzazione	89
7.1	Realizzazione	89
7.2	Forme Canoniche per la Realizzazione	94
7.2.1	Forma Compagna di Controllore	94
7.2.2	Forma Compagna di Osservatore	97
8	Luogo delle Radici	100
8.1	Luogo Esatto	100
8.2	Luogo Approssimato	102
8.3	Sintesi	103
A	Richiami di Algebra Lineare	105
A.1	Spazi Vettoriali	105
A.2	Matrici	108
A.3	Trasformazioni Lineari	111
B	Trasformata di Laplace	114
B.1	Trasformata di Laplace	114
B.1.1	Proprietà	115
B.2	Antitrasformata di Laplace	115
B.2.1	Espansione in Fratti Semplici	116
B.3	Trasformate Note	118
C	Matlab per la Teoria dei Sistemi	119
C.1	Rappresentazione di Modelli Lineari	119
C.1.1	Rappresentazione in Spazio di Stato	119
C.1.2	Rappresentazione attraverso Funzione di Trasferimento	121
D	Vecchi Compiti di Esame	122

Notes

Questo piccolo volume raccoglie alcuni degli argomenti presentati al corso di “Teoria dei Sistemi e del Controllo (I° modulo)” presso l’Università degli Studi “Roma Tre”. Esso è da intendersi come integrazione al materiale fornito per il corso “Controlli Automatici” dal Prof. Stefano Panzieri, il quale è fruibile al seguente indirizzo web:

http://www.dia.uniroma3.it/autom/FdAcomm/Ausili/spazio_di_stato.htm

Per chi fosse interessato ad approfondire ulteriormente gli argomenti trattati nel corso un paio di libri che personalmente ritengo estremamente validi sono:

"An Introduction to Linear Control Systems"

By Thomas E. Fortmann, Konrad L. Hitz

Published by Marcel Dekker, 1977

ISBN 0824765125, 9780824765125

"Control System Design:

An introduction to State-Space Methods"

By Bernard Friedland

Published by McGraw-Hill Companies, 1985

ISBN 0070224412, 9780070224414

entrambi disponibili nella biblioteca di area tecnico-scientifica presso il dipartimento di Informatica ed Automazione dell’Università degli Studi “Roma Tre”.

Buon Lavoro,
Andrea

Capitolo 1

Introduzione

1.1 Modellazione

ToDo

1.1.1 Rappresentazione Ingresso-Uscita

ToDo

1.1.2 Rappresentazione Ingresso-Stato-Uscita

ToDo

1.2 Variabili di Stato (V.d.S.)

Si consideri un generico sistema dinamico descritto dal seguente insieme di equazioni differenziali del primo ordine lineari a coefficienti costanti del tipo:

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & a_{11} x_1 + \dots + a_{1n} x_n + b_{11} u_1 + b_{12} u_2 + \dots + b_{1p} u_p \\ \dot{x}_2 & = & a_{21} x_1 + \dots + a_{2n} x_n + b_{21} u_1 + b_{22} u_2 + \dots + b_{2p} u_p \\ & \vdots & \\ \dot{x}_n & = & a_{n1} x_1 + \dots + a_{nn} x_n + b_{n1} u_1 + b_{n2} u_2 + \dots + b_{np} u_p \\ & & \\ y_1 & = & c_{11} x_1 + \dots + c_{1n} x_n + d_{11} u_1 + d_{12} u_2 + \dots + d_{1p} u_p \\ y_2 & = & c_{21} x_1 + \dots + c_{2n} x_n + d_{21} u_1 + d_{22} u_2 + \dots + d_{2p} u_p \\ & \vdots & \\ y_q & = & c_{q1} x_1 + \dots + c_{qn} x_n + d_{q1} u_1 + d_{q2} u_2 + \dots + d_{qp} u_p \end{array} \right. . \quad (1.1)$$

Tali equazioni descrivono un sistema dinamico il cui stato è descritto da n variabili, soggetto a p ingressi e caratterizzato da q uscite.

Per tale sistema è possibile fornire anche una rappresentazione in forma matriciale equivalente caratterizzata da quattro matrici $\{A, B, C, D\}$ come segue:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) + D u(t) \end{cases} \quad (1.2)$$

dove si indica con:

- $u \in \mathbb{R}^p$ vettore delle grandezze di ingresso
- $y \in \mathbb{R}^q$ vettore delle grandezze di uscita
- $x(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbb{R}^n$ stato del sistema con $x_i(t) \in \mathbb{R}$ i -esima variabile di stato
- A matrice dinamica del sistema ($n \times n$)
- B matrice di ingresso ($n \times p$)
- C matrice di uscita ($q \times n$)
- D matrice del legame diretto ingresso-uscita ($q \times p$).

Definizione 1.1 Si definisce **funzione di transizione di stato** $\phi(t, t_0, x(t_0))$ la funzione che fornisce l'andamento di $x(t)$, $t \in [t_0, t_1]$ a partire dallo stato iniziale x_0 e l'ingresso $u(t)$, $t \in [t_0, t_1]$.

Definizione 1.2 Si definisce **traiettoria** di un sistema nello spazio l'insieme \mathcal{T} dei valori di stato che un determinato sistema può assumere in accordo alla funzione di transizione di stato:

$$\mathcal{T} = \{x \in \mathbb{R}^n : x = \phi(t, t_0, x_0, u(t)), t \in [t_0, t_1]\}, \quad (1.3)$$

Definizione 1.3 Si definisce **moto** di un sistema l'insieme così definito.

$$\mathcal{M} = \{(t, x) \in [t_0, t_1] \times \mathbb{R}^n : x = \phi(t, t_0, x_0, u(t)), t \in [t_0, t_1]\}. \quad (1.4)$$

In particolare un moto viene detto **periodico** se $x(t + nT) = x(t)$, con T periodo del moto.

Definizione 1.4 Si definisce spazio di stato lo spazio n -dimensionale \mathbb{R}^n a cui appartiene lo stato $x(t)$.

Capitolo 2

Sistemi non Lineari e Linearizzazione

2.1 Motivazioni

In questo capitolo verrà fornita una motivazione di carattere operativo allo studio dei sistemi lineari. A tal fine verrà introdotto il problema della linearizzazione di un sistema non lineare, mostrando come sotto particolari condizioni operative una rappresentazione (approssimazione) lineare possa fornire informazioni di interesse per il sistema dinamico non lineare di partenza. L'obiettivo è quello di fornire una metodologia sistematica per la derivazione di approssimazioni lineari (a partire da modelli originali non lineari) in grado di descrivere il comportamento dinamico di un sistema fisico a fronte di perturbazioni rispetto ad un dato punto operativo. È importante sottolineare come la validità di tale approssimazione lineare sia limitata all'analisi per piccoli segnali, ovvero deviazioni sufficientemente piccole dai valori di equilibrio o dai punti operativi nominali. Tuttavia nell'area dei controlli automatici tale analisi (per piccoli segnali) ha una grossa rilevanza da un punto di vista applicativo. Infatti, sovente i sistemi di controllo automatici sono pensati per regolare, quindi tenere tanto vicino quanto possibile, un impianto intorno ad un punto operativo.

2.2 Definizioni

Si consideri un sistema dinamico descritto dal seguente sistema di equazioni non lineari:

$$\begin{cases} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t), u(t)) \end{cases} \quad (2.1)$$

dove $x \in \mathbb{R}^n$ è lo stato del sistema, $u \in \mathbb{R}^p$ rappresenta l'ingresso al sistema ed $y \in \mathbb{R}^q$ descrive l'uscita del sistema.

Definizione 2.1 Si definisce punto di equilibrio (pde) di un sistema dinamico descritto dalle equazioni non lineari (2.1) una terna (u_e, x_e, y_e) tale che se $x(0) = x_e$ e $u(t) = u_e$, allora $\forall t \geq 0$ si ha:

$$\begin{cases} \mathbf{0} &= f(x_e, u_e) \\ y_e &= h(x_e, u_e) \end{cases}, \quad (2.2)$$

ovvero lo stato, l'ingresso e l'uscita assumono valori costanti.

Si noti come tale definizione di punto di equilibrio nel caso di un sistema non lineare autonomo $\dot{x}(t) = f(x(t))$ si riduca semplicemente all'individuazione di quel particolare valore dello stato x_e per il quale si azzer la derivata, ovvero:

$$\mathbf{0} = f(x_e) \quad (2.3)$$

Definizione 2.2 *Si definisce sviluppo in serie di Taylor di una generica funzione $f(x)$ la sua rappresentazione attraverso una somma di infiniti termini calcolata a partire dalle derivate di vario ordine di tale funzione $f(x)$ in un dato punto x_0 come segue:*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n. \quad (2.4)$$

Si noti come tale strumento risulti essere di estremo interesse in quanto permette di ottenere una rappresentazione approssimata di accuratezza arbitraria del comportamento della funzione $f(x)$ nell'intorno di un punto operativo x_0 . In particolare, nel caso si decida di troncare lo sviluppo in serie alla derivata di ordine n , l'errore di approssimazione risulta essere dell'ordine di $\mathcal{O}\left(\frac{x^n}{n!}\right)$

2.3 Metodo di Linearizzazione

Si consideri un sistema dinamico descritto dal seguente sistema di equazioni non lineari:

$$\begin{cases} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t), u(t)) \end{cases} \quad (2.5)$$

per il quale si conosce una soluzione nominale del sistema $(u^*(t), x^*(t), y^*(t))$ tale che valga la:

$$\begin{cases} \dot{x}^*(t) &= f(x^*(t), u^*(t)) \\ y^*(t) &= h(x^*(t), u^*(t)) \end{cases} \quad (2.6)$$

Si consideri ora una perturbazione nell'intorno di tale soluzione nominale come segue:

$$\begin{aligned} \bar{x}(t) &= x^*(t) + \epsilon \hat{x}(t) \\ \bar{u}(t) &= u^*(t) + \epsilon \hat{u}(t) \\ \bar{y}(t) &= y^*(t) + \epsilon \hat{y}(t) \end{aligned} \quad (2.7)$$

L'obiettivo e' quello di determinare un sistema di equazioni (lineari) che la $(\hat{x}, \hat{u}, \hat{y})$ dovrebbe rispettare affinché la $(\bar{x}, \bar{u}, \bar{y})$ sia ancora soluzione della (2.5). A tale scopo si consideri la j -esima componente della $f(x, u)$:

$$\dot{x}_j = f_j(x, u) = f_j(x_1, \dots, x_n, u_1, \dots, u_p)$$

si effettui la sostituzione data in (2.7) e si effettui lo sviluppo in serie di Taylor rispetto alla soluzione nominale $(u^*(t), x^*(t), y^*(t))$, in modo da ottenere quanto segue:

$$\begin{aligned} \dot{x}_j^* + \epsilon \dot{\hat{x}}_j &= f_j(x^*, u^*) + \left. \frac{\partial f_j}{\partial x_1} \right|_{(x^*, u^*)} \cdot (x_1^* + \epsilon \hat{x}_1 - x_1^*) + \cdots + \left. \frac{\partial f_j}{\partial x_n} \right|_{(x^*, u^*)} \cdot (x_n^* + \epsilon \hat{x}_n - x_n^*) \\ &+ \left. \frac{\partial f_j}{\partial u_1} \right|_{(x^*, u^*)} \cdot (u_1^* + \epsilon \hat{u}_1 - u_1^*) + \cdots + \left. \frac{\partial f_j}{\partial u_p} \right|_{(x^*, u^*)} \cdot (u_p^* + \epsilon \hat{u}_p - u_p^*) + \mathcal{O}_j(\epsilon^2) \end{aligned} \quad (2.8)$$

È ora possibile definire lo Jacobiano associato al sistema descritto in (2.5) come segue:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla f_1 \\ \vdots \\ \vdots \\ \nabla f_n \end{bmatrix} \quad (2.9)$$

dove $\nabla f_j = \left[\frac{\partial f_j}{\partial x_1} \ \cdots \ \cdots \ \frac{\partial f_j}{\partial x_n} \right]$ è il gradiente rispetto alla j -esima componente della funzione $f(x, u)$. A questo punto, ripetendo lo stesso ragionamento per ogni componente $i \in \{1 \dots, n\}$, e collezionando le n equazioni così ottenute in forma matriciale si ottiene:

$$\dot{x}^* + \epsilon \dot{\hat{x}} = f(x^*, u^*) + \left. \frac{\partial f}{\partial x} \right|_{(x^*, u^*)} \cdot \epsilon \hat{x} + \left. \frac{\partial f}{\partial u} \right|_{(x^*, u^*)} \cdot \epsilon \hat{u} + \mathcal{O}(\epsilon^2) \quad (2.10)$$

dal quale ricordando la (2.6) e dividendo ambo i membri per ϵ si ottiene:

$$\dot{\hat{x}} = \left. \frac{\partial f}{\partial x} \right|_{(x^*, u^*)} \cdot \hat{x} + \left. \frac{\partial f}{\partial u} \right|_{(x^*, u^*)} \cdot \hat{u} + \mathcal{O}(\epsilon). \quad (2.11)$$

Con una procedura del tutto analoga alla precedente rispetto alla funzione $h(x, u)$ si ottiene:

$$\hat{y} = \left. \frac{\partial h}{\partial x} \right|_{(x^*, u^*)} \cdot \hat{x} + \left. \frac{\partial h}{\partial u} \right|_{(x^*, u^*)} \cdot \hat{u} + \mathcal{O}(\epsilon). \quad (2.12)$$

A questo punto raccogliendo la (2.11) e la (2.12) si ottiene:

$$\begin{cases} \dot{\hat{x}}(t) &= A \hat{x}(t) + B \hat{u}(t) \\ \hat{y}(t) &= C \hat{x}(t) + D \hat{u}(t) \end{cases} \quad (2.13)$$

dove:

$$\begin{aligned} A &= \left. \frac{\partial f}{\partial x} \right|_{(x^*, u^*)} & B &= \left. \frac{\partial f}{\partial u} \right|_{(x^*, u^*)} \\ C &= \left. \frac{\partial h}{\partial x} \right|_{(x^*, u^*)} & D &= \left. \frac{\partial h}{\partial u} \right|_{(x^*, u^*)} \end{aligned} \quad (2.14)$$

Si noti che la (2.13) rappresenta il modello lineare approssimato del sistema non lineare originale rispetto rispetto alla soluzione nominale $(u^*(t), x^*(t), y^*(t))$ a fronte di una perturbazione $(\hat{u}(t), \hat{x}(t), \hat{y}(t))$.

Capitolo 3

Soluzione delle Equazioni Differenziali

3.1 Equazioni differenziali lineari omogenee a coefficienti costanti

Definizione 3.1 Si definisce equazione differenziale lineare omogenea a coefficienti costanti, una equazione differenziale del primo ordine del tipo:

$$\dot{x}(t) - ax(t) = 0 \quad (3.1)$$

dove $a \in \mathbb{R}$ è una costante. É noto dall'analisi matematica che tale tipologia di equazioni differenziali ha una famiglia di soluzione del tipo:

$$x(t) = e^{at} c \quad (3.2)$$

dove $c \in \mathbb{R}$ è una costante di integrazione atta a caratterizzare la famiglia di soluzioni.

Dimostrazione: La 3.2 può essere agevolmente ottenuta riarrangiando la 3.1 come segue:

$$\frac{\dot{x}(t)}{x(t)} = a$$

integrando ambo i membri:

$$\begin{aligned} \int \frac{\dot{x}(t)}{x(t)} dt &= \int a dt \\ \int \frac{1}{x(t)} \frac{dx(t)}{dt} dt &= \int a dt \\ \int \frac{dx}{x} &= \int a dt \\ \log(x) + c_1 &= at + c_2 \end{aligned}$$

ed elevando a potenza ambo i membri si ottiene:

$$\begin{aligned} x(t) &= e^{at+(c_2-c_1)} \\ &= e^{at} e^{(c_2-c_1)} \\ &= e^{at} c \end{aligned}$$

Definizione 3.2 Si definisce problema di Cauchy per l'equazione differenziale lineare omogenea a coefficienti costanti definita in 3.1, il seguente problema:

$$\begin{cases} \dot{x}(t) &= a x(t) \\ x(t_0) &= x_0 \end{cases} \quad (3.3)$$

Tale problema ha una soluzione unica che risulta essere:

$$x(t) = e^{a(t-t_0)} x_0 \quad (3.4)$$

Dimostrazione: La dimostrazione è pressochè identica a quella mostrata in 3.1 con la differenza che in questo caso si utilizza l'integrale definito in un intervallo di integrazione $[t_0, t]$.

3.2 Equazioni differenziali lineari complete a coefficienti costanti

Definizione 3.3 Si definisce equazione differenziale lineare completa a coefficienti costanti, una equazione differenziale del primo ordine del tipo:

$$\dot{x}(t) - ax(t) = \beta u(t) \quad (3.5)$$

dove $a \in \mathbb{R}$ è una costante. È noto dall'analisi matematica che la **soluzione generale** di tale equazione differenziale è data dalla **soluzione dell'omogenea** associata più una **soluzione particolare** come segue:

$$x_g(t) = x_o(t) + x_p(t) \quad (3.6)$$

Per il calcolo della soluzione particolare è possibile utilizzare il metodo della variazione delle costanti (o di Lagrange).

Definizione 3.4 (Metodo di Variazione delle Costanti o di Lagrange) Si consideri l'equazione differenziale descritta dalla 3.5. Una soluzione particolare può essere ottenuta a partire dalla soluzione omogenea 3.2 sostituendo la costante c con una funzione $c(t)$ come segue:

$$x_p(t) = e^{at} c(t) \quad (3.7)$$

dove:

$$c(t) = \int e^{-a\tau} \beta u(\tau) d\tau + cost \quad (3.8)$$

quindi:

$$x_p(t) = \int e^{a(t-\tau)} \beta u(\tau) d\tau + cost \quad (3.9)$$

Dimostrazione: Al fine di ottenere la 3.9, si può pensare di sostituire la 3.7 nella 3.5 come segue:

$$\frac{d}{dt} (e^{at} c(t)) - a e^{at} c(t) = \beta u(t)$$

Svolgendo la derivata si ottiene:

$$e^{at} \frac{d}{dt} c(t) + a e^{at} c(t) - a e^{at} c(t) = \beta u(t)$$

$$e^{at} \frac{d}{dt} c(t) = \beta u(t)$$

$$\frac{d}{dt} c(t) = e^{-at} \beta u(t)$$

A questo punto si può pensare di integrare ambo i membri:

$$\int \frac{d}{dt} c(t) dt = \int e^{-a\tau} \beta u(\tau) d\tau$$

$$c(t) = \int e^{-a\tau} \beta u(\tau) d\tau + cost$$

Definizione 3.5 Si definisce problema di Cauchy per l'equazione differenziale lineare completa a coefficienti costanti definita in 3.5, il seguente problema:

$$\begin{cases} \dot{x}(t) = a x(t) + \beta u(t) \\ x(t_0) = x_0 \end{cases} \quad (3.10)$$

Tale problema ha una soluzione unica che risulta essere:

$$x(t) = e^{a(t-t_0)} x_0 + \int_{t_0}^t e^{a(t-\tau)} \beta u(\tau) d\tau \quad (3.11)$$

Dimostrazione: La dimostrazione è pressochè identica a quella mostrata per il metodo delle variazioni delle costanti in 3.4, dove al posto dell'integrale indefinito si fa uso dell'integrale definito nell'intervallo $[t_0, t]$ ponendo $c(t_0) = 0$ senza perdita di generalità.

3.3 Riduzione di una equazione differenziale completa a coefficienti costanti di ordine n ad un sistema di n equazioni differenziali del primo ordine

Si consideri una equazione differenziale di ordine n completa a coefficienti costanti definita come segue:

$$x^{(n)} + a_{n-1} x^{(n-1)} + a_{n-2} x^{(n-2)} + \dots + a_1 x^{(1)} + a_0 x = \beta u(t) \quad (3.12)$$

È sempre possibile riscrivere tale equazione differenziale di ordine n in un sistema di n equazioni differenziali del primo ordine come segue:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \vdots = \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = -a_{n-1} x_n - a_{n-2} x_{n-1} + \dots - a_1 x_2 - a_0 x_1 + \beta u(t) \end{cases} \quad (3.13)$$

3.4 Esponenziale di matrice

Definiamo l'esponenziale di matrice con la seguente sommatoria:

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}, \quad A \in \mathbb{R}^{n \times n}. \quad (3.14)$$

Il quale può essere riscritto in forma estesa come:

$$e^{At} = I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots \quad (3.15)$$

Tale esponenziale di matrice è caratterizzato da una serie di proprietà che andremo ora ad analizzare.

Teorema 3.1 (*Convergenza*) Sia $A \in \mathbb{R}^{n \times n}$ una matrice quadrata, allora la serie esponenziale

$$\sum_{k=0}^{\infty} \frac{(At)^k}{k!}, \quad (3.16)$$

- Converge $\forall t \geq 0, t \in \mathbb{R}$,
- Il limite è una matrice funzione continua del tempo t , nota come e^{At} o $\exp(At)$

Dimostrazione: La convergenza dell'esponenziale di matrice può essere dimostrata come conseguenza della convergenza in norma della serie. Si consideri la serie delle norme:

$$\sum_{k=0}^{\infty} \left\| \frac{(At)^k}{k!} \right\|$$

A questo punto, facendo uso della proprietà di omogeneità di una norma si ha:

$$\|c(At)^k\| \leq |c| \|A\|^k |t|^k$$

dalla quale si ottiene una nuova serie numerica che migliora termine a termine la serie delle norme:

$$\sum_{k=0}^{\infty} \frac{\|A\|^k |t|^k}{k!} = e^{\|A\| |t|}$$

la cui convergenza è garantita dal fatto che questa coincide con lo sviluppo in serie di Taylor della funzione esponenziale di argomento scalare $\|A\| |t|$.

Teorema 3.2 (*Derivazione*) La derivata della 3.14 si può calcolare derivando ogni elemento della sommatoria e vale

$$\frac{d}{dt} e^{At} = A e^{At}. \quad (3.17)$$

Dimostrazione: Per la dimostrazione è sufficiente effettuare la derivazione rispetto allo sviluppo in serie definito in 3.15 come segue:

$$\begin{aligned}
\frac{d}{dt}e^{At} &= \frac{d}{dt} \left(I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots \right) \\
&= \left(\frac{d}{dt}I + \frac{d}{dt}At + \frac{d}{dt} \frac{A^2 t^2}{2!} + \frac{d}{dt} \frac{A^3 t^3}{3!} + \dots \right) \\
&= \left(A + A^2 t + \frac{A^3 t^2}{2!} + \frac{A^4 t^3}{3!} + \dots \right) \\
&= A \left(I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots \right) \\
&= A e^{At}.
\end{aligned}$$

Teorema 3.3 (Composizione) *L'esponenziale di matrice gode della seguente proprietà di composizione:*

$$e^{At}e^{A\tau} = e^{A(t+\tau)} \quad (3.18)$$

Dimostrazione: Per la dimostrazione è sufficiente utilizzare lo sviluppo in serie rispetto alle due variabili temporali t e τ :

$$e^{At}e^{A\tau} = \left(I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots \right) \left(I + A\tau + \frac{A^2 \tau^2}{2!} + \frac{A^3 \tau^3}{3!} + \dots \right)$$

ed effettuando lo sviluppo si ottiene:

$$\begin{aligned}
&= \begin{array}{ccccccc}
I & + & A\tau & + & \frac{A^2 \tau^2}{2!} & + & \frac{A^3 \tau^3}{3!} & + & \dots \\
& & + & At & + & A^2 t \tau & + & \frac{A^3 t \tau^2}{2!} & + & \dots \\
& & & & + & \frac{A^2 t^2}{2!} & + & \frac{A^3 t^2 \tau}{2!} & + & \dots \\
& & & & & & + & \frac{A^3 t^3}{3!} & + & \dots
\end{array} \\
&\hline
&= I + A(t + \tau) + \frac{A^2(t + \tau)^2}{2!} + \frac{A^3(t + \tau)^3}{3!} + \dots \\
&= e^{A(t+\tau)}
\end{aligned}$$

Teorema 3.4 (Invertibilità) *L'esponenziale di matrice gode della seguente proprietà di invertibilità:*

$$(e^{At})^{-1} = e^{-At} \quad (3.19)$$

Tale teorema sull'invertibilità dell'esponenziale di matrice comporta due importanti conseguenze:

- L'esponenziale di matrice ha sempre rango massimo,
- È possibile andare “indietro nel tempo”.

Dimostrazione: Per la dimostrazione è sufficiente far riferimento al precedente teorema di composizione ponendo $\tau = -t$. Infatti, data una generica matrice $A \in \mathbb{R}^{n \times n}$ questa si dice invertibile se $\exists A^{-1}$ tale che:

$$A \cdot A^{-1} = A^{-1} \cdot A = I$$

Nel nostro caso avremmo:

$$e^{At} \cdot e^{-At} = e^{A(t-t)} = e^{A0} = I$$

Teorema 3.5 (*Commutatività*) *L'esponenziale di matrice gode della seguente proprietà di commutatività:*

$$e^{At} \cdot e^{Bt} = e^{(A+B)t} \iff A \cdot B = B \cdot A. \quad (3.20)$$

Dimostrazione: Per la dimostrazione è sufficiente utilizzare lo sviluppo in serie rispetto alle due matrici A e B come segue:

$$e^{At}e^{Bt} = \left(I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \dots \right) \left(I + Bt + \frac{B^2t^2}{2!} + \frac{B^3t^3}{3!} + \dots \right)$$

ed effettuando lo sviluppo si ottiene:

$$\begin{array}{ccccccc} = & I & + & Bt & + & \frac{B^2t^2}{2!} & + \frac{B^3t^3}{3!} + \dots \\ & & & + At & + & \frac{2ABt^2}{2!} & + \dots \\ & & & & + & \frac{A^2t^2}{2!} & + \dots \\ \hline = & I & + & (A+B)t & + & \frac{(A^2 + 2AB + B^2)}{2!}t^2 & + \dots \end{array}$$

dove vale la

$$A^2 + 2AB + B^2 = (A+B)^2 \iff A \cdot B = B \cdot A.$$

3.5 Evoluzione Libera

Sia dato un sistema dinamico descritto dalla seguente equazione differenziale:

$$\begin{cases} \dot{x}(t) &= Ax(t) \\ x(t_0) &= x_0 \end{cases} \quad (3.21)$$

e caratterizzato dalla seguente uscita:

$$y = Cx(t) \quad (3.22)$$

Teorema 3.6 (*Evoluzione Libera*) Si definisce **evoluzione libera** (dello stato) del sistema dinamico la seguente soluzione della 3.21:

$$x(t) = e^{A(t-t_0)}x_0 \quad (3.23)$$

Dimostrazione: Per provare il teorema è sufficiente differenziare la soluzione 3.23 ed applicare il teorema di derivazione per l'esponenziale di matrice 3.2, dove per semplicità abbiamo posto $t_0 = 0$.

$$\frac{d}{dt}x(t) = \frac{d}{dt}(e^{At}x_0)$$

$$\dot{x}(t) = Ae^{At}x_0$$

$$\dot{x}(t) = Ax(t)$$

Inoltre, per la condizione iniziale $t = 0$ si ottiene:

$$x(0) = e^{A \cdot 0}x_0$$

$$x(0) = x_0.$$

il quale prova la tesi.

Definizione 3.6 (*Risposta Libera*) Si definisce **risposta libera** (dell'uscita) del sistema dinamico la seguente soluzione della 3.21 in riferimento alla 3.22:

$$y(t) = Ce^{A(t-t_0)}x_0 \quad (3.24)$$

3.6 Evoluzione Forzata

Sia dato un sistema dinamico descritto dalla seguente equazione differenziale:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (3.25)$$

con $x(t_0) = x_0$.

Definizione 3.7 (*Evoluzione Forzata*) Si definisce **evoluzione forzata** (dello stato) del sistema dinamico la seguente soluzione particolare della 3.28:

$$x(t) = \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau. \quad (3.26)$$

Definizione 3.8 (*Risposta Forzata*) Si definisce **risposta forzata** (dell'uscita) del sistema dinamico la seguente soluzione particolare della 3.25:

$$y(t) = \int_{t_0}^t C e^{A(t-\tau)} B u(\tau) d\tau. \quad (3.27)$$

Teorema 3.7 (*Evoluzione Completa*) Si consideri un sistema dinamico la cui evoluzione dello stato è descritta dalla:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ x(t_0) &= x_0 \end{cases} \quad (3.28)$$

La soluzione di tale sistema è definita **evoluzione completa** (dello stato) ed è data dalla:

$$x(t) = e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-\tau)} B u(\tau) d\tau \quad (3.29)$$

Due importanti caratteristiche possono essere evidenziate:

- Il sistema ha **memoria**: Si può infatti notare come l'uscita al tempo t non dipenda esclusivamente dall'ingresso al tempo t bensì essa è funzione anche degli ingressi $u(\tau)$ passati con $(t_0 \leq \tau \leq t)$.
- Il sistema è **causale** (non anticipatorio): Si può infatti notare come l'uscita al tempo t non dipenda da nessun ingresso $u(\tau)$ con $\tau > t$

Dimostrazione: Per la dimostrazione è sufficiente riarrangiare i termini della 3.28 e moltiplicare per un fattore di integrazione e^{-At} .

$$e^{-At} [\dot{x}(t) - Ax(t)] = e^{-At} B u(t)$$

Notiamo che il termine a sinistra dell'equazione rappresenta la derivata del termine $e^{-At}x(t)$ ed effettuiamo una integrazione ambo i membri:

$$\int_{t_0}^t \frac{d}{d\tau} (e^{-A\tau} x(\tau)) d\tau = \int_{t_0}^t e^{-A\tau} B u(\tau) d\tau$$

Da cui si ottiene:

$$\begin{aligned} e^{-At} x(t) \Big|_{t_0}^t &= \int_{t_0}^t e^{-A\tau} B u(\tau) d\tau \\ e^{-At} x(t) &= e^{-At_0} x(t_0) + \int_{t_0}^t e^{-A\tau} B u(\tau) d\tau \\ x(t) &= e^{A(t-t_0)} x(t_0) + \int_{t_0}^t e^{A(t-\tau)} B u(\tau) d\tau \end{aligned}$$

Definizione 3.9 (*Risposta Completa*) Si definisce **risposta completa** (dell'uscita) del sistema dinamico la seguente soluzione generale della 3.25:

$$y(t) = Ce^{A(t-t_0)}x(t_0) + \int_{t_0}^t Ce^{A(t-\tau)}Bu(\tau)d\tau. \quad (3.30)$$

Dimostrazione: La dimostrazione è una semplice conseguenza del teorema 3.7. Infatti ricordando che l'evoluzione completa dello stato è descritta dallo:

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau$$

e che il legame tra l'uscita e lo stato è dato dalla:

$$y = Cx$$

Si può agevolmente ottenere la risposta completa come:

$$y(t) = Ce^{A(t-t_0)}x(t_0) + \int_{t_0}^t Ce^{A(t-\tau)}Bu(\tau)d\tau.$$

Teorema 3.8 (*Principio di Sovrapposizione degli Effetti*) Si consideri il sistema dinamico descritto dalla 3.25 assumendo di avere $x(t_0) = 0$, e si definisca un ingresso:

$$u(t) = u_1(t) + \beta u_2(t). \quad (3.31)$$

Allora l'uscita del sistema associata a tale ingresso $u(t)$ risulta essere:

$$y(t) = y_1(t) + \beta y_2(t). \quad (3.32)$$

L'applicazione del principio di sovrapposizione degli effetti permette di decomporre il problema in sottoproblemi più semplici. Se l'ingresso è dato dalla somma di più termini, questi possono essere considerati come più ingressi indipendenti tra loro. Quindi la soluzione totale, relativa all'ingresso originale, può essere semplicemente ottenuta come somma delle soluzioni parziali corrispondenti ad ogni singola componente dell'ingresso.

Dimostrazione: Per la dimostrazione è sufficiente far riferimento alla definizione di risposta forzata fornita nella 3.25 e ricordare le proprietà di linearità degli integrali come segue:

$$\begin{aligned} y(t) &= Ce^{A(t-t_0)}x(t_0) + \int_{t_0}^t Ce^{A(t-\tau)}Bu(\tau)d\tau \\ y(t) &= Ce^{A(t-t_0)}0 + \int_{t_0}^t Ce^{A(t-\tau)}B(u_1(\tau) + \beta u_2(\tau))d\tau \\ y(t) &= \int_{t_0}^t Ce^{A(t-\tau)}Bu_1(\tau)d\tau + \beta \int_{t_0}^t Ce^{A(t-\tau)}Bu_2(\tau)d\tau \\ y(t) &= y_1(t) + \beta y_2(t) \end{aligned}$$

Capitolo 4

Relazioni tra le Rappresentazioni

4.1 Passaggio da V.d.S. a F(s)

Dato un sistema dinamico descritto dalla seguente equazione differenziale:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (4.1)$$

con $x(t_0) = x_0$ e $t_0 = 0$ per semplicità.

Trasformiamo l'equazione differenziale secondo Laplace (la quale può essere formalmente ottenuta trattando i vettori come fossero degli scalari):

$$\begin{cases} s X(s) - x(t_0) &= A X(s) + B U(s) \\ Y(s) &= C X(s) \end{cases} \quad (4.2)$$

Riarrangiando opportunamente i termini si ottiene:

$$\begin{cases} X(s) &= (sI - A)^{-1} x_0 + (sI - A)^{-1} B U(s) \\ Y(s) &= C (sI - A)^{-1} x_0 + C (sI - A)^{-1} B U(s) \end{cases} \quad (4.3)$$

dove:

- $(sI - A)^{-1} x_0$ è la trasformata dell'evoluzione libera dello stato
- $(sI - A)^{-1} B U(s)$ è la trasformata dell'evoluzione forzata dello stato.

Teorema 4.1 (*Matrice Risolvente*) La matrice $(sI - A)^{-1}$ è definita **matrice risolvente** ed essa rappresenta la trasformata di Laplace dell'esponenziale di matrice e^{At} :

$$\mathcal{L}[e^{At}] = (sI - A)^{-1}, \quad (4.4)$$

Dimostrazione: Per la dimostrazione si fa riferimento ad un noto sviluppo in serie valido per scalari:

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \dots \quad |z| < 1$$

la cui convergenza è garantita dalla condizione $|z| < 1$. Tale concetto può essere esteso al caso matriciale se vale la $\frac{\|A\|}{|s|} < 1$, la quale risulta sempre verificata per s sufficientemente grande. Effettuando tale espansione in serie per $|s| = +\infty$ si ottiene:

$$\begin{aligned}(sI - A)^{-1} &= \frac{(I - A/s)^{-1}}{s} \\ &= \frac{(I + A/s + A^2/s^2 + A^3/s^3 + \dots)}{s} \\ &= I/s + A/s^2 + A^2/s^3 + \dots\end{aligned}$$

Ora ricordando che $\mathcal{L}^{-1}\left[\frac{1}{(s^{n+1})}\right] = \frac{t^n}{n!}$ e la linearità dell'operatore \mathcal{L}^{-1} possiamo effettuare una antitrasformazione termine a termine come segue:

$$\begin{aligned}\mathcal{L}^{-1}[(sI - A)^{-1}] &= \mathcal{L}^{-1}[I/s] + \mathcal{L}^{-1}[A/s^2] + \mathcal{L}^{-1}[A^2/s^3] + \mathcal{L}^{-1}[A^3/s^4] + \dots \\ &= I + At + \frac{A^2 t^2}{2!} + \frac{A^3 t^3}{3!} + \dots \\ &= e^{At}\end{aligned}$$

Dimostrazione alternativa (Validità limitata alle matrici diagonalizzabili): Nel caso la matrice dinamica del sistema A risulti essere diagonalizzabile, esiste una matrice V tale che la matrice A possa essere espressa come:

$$\begin{aligned}A &= V \operatorname{diag}(\lambda_1, \dots, \lambda_n) V^{-1} \\ &= V \Lambda V^{-1}\end{aligned}$$

ne risulta che l'esponenziale di matrice (a seguito di semplici calcoli che coinvolgono la definizione come sviluppo in serie) può essere riscritto come segue:

$$\begin{aligned}e^{At} &= V e^{\operatorname{diag}(\lambda_1, \dots, \lambda_n) t} V^{-1} \\ &= V e^{\Lambda t} V^{-1}\end{aligned}$$

a questo punto si può pensare di effettuare la trasformata di Laplace sfruttandone la proprietà di linearità:

$$\begin{aligned}\mathcal{L}[e^{At}] &= V \mathcal{L}[e^{\operatorname{diag}(\lambda_1, \dots, \lambda_n)}] V^{-1} \\ &= V \operatorname{diag}\left((s - \lambda_1)^{-1}, \dots, (s - \lambda_n)^{-1}\right) V^{-1} \\ &= V (sI - \Lambda)^{-1} V^{-1} \\ &= (sI - A)^{-1}\end{aligned}$$

il quale prova la tesi. Si noti che nell'ultimo passaggio si é utilizzata la seguente proprietà di invertibilità per il prodotto di matrici:

$$(A \cdot B \cdot C)^{-1} = C^{-1} \cdot B^{-1} \cdot A^{-1}.$$

Teorema 4.2 (*Corrispondenza Uscite*) *L'uscita nel tempo del sistema descritto dalla 4.3 dove si assume $t_0 = 0$ per semplicità è dato dalla:*

$$y(t) = C e^{At} x(0) + \int_0^t C e^{A(t-\tau)} B u(\tau) d\tau. \quad (4.5)$$

Il quale sottolinea, come era lecito aspettarsi, la corrispondenza (sotto condizioni di completa controllabilità e osservabilità) tra la rappresentazione in spazio di stato e la rappresentazione nel dominio di Laplace.

Dimostrazione: Per la dimostrazione si fa riferimento al teorema 4.1 che ci fornisce il legame tra la matrice risolvante e l'esponenziale di matrice. Consideriamo l'uscita del sistema nel dominio di Laplace come da 4.3:

$$\mathcal{L}^{-1}[Y(s)] = \mathcal{L}^{-1}[C(sI - A)^{-1}x_0 + C(sI - A)^{-1}BU(s)]$$

Per la linearità dell'operatore \mathcal{L}^{-1} possiamo riscrivere come:

$$\begin{aligned} \mathcal{L}^{-1}[Y(s)] &= \mathcal{L}^{-1}[C(sI - A)^{-1}x_0] + \mathcal{L}^{-1}[C(sI - A)^{-1}BU(s)] \\ &= C\mathcal{L}^{-1}[(sI - A)^{-1}]x_0 + C\mathcal{L}^{-1}[(sI - A)^{-1}]B * \mathcal{L}^{-1}[U(s)] \end{aligned}$$

La quale per il teorema 4.1 può essere scritta come

$$\begin{aligned} y(t) &= C e^{At} x_0 + C e^{At} B * u(t) \\ &= C e^{At} x(0) + \int_0^t C e^{A(t-\tau)} B u(\tau) d\tau. \end{aligned}$$

dove col simbolo $*$ si vuole indicare l'operatore di convoluzione.

Definizione 4.1 (*Matrice Funzione di Trasferimento*) *Si consideri il sistema descritto dalla 4.1, si definisce **matrice funzione di trasferimento** $H(s)$, la:*

$$H(s) = C(sI - A)^{-1}B \quad (4.6)$$

la quale ha la proprietà di legare l'ingresso e l'uscita del sistema come segue:

$$Y(s) = H(s) \cdot U(s) \quad (4.7)$$

sotto ipotesi di condizioni iniziali nulle per il sistema $x(t_0) = 0$.

Lemma 4.1 *Si consideri il sistema dinamico descritto dalla 4.1, la matrice risolvante associata a tale sistema può sempre essere scritta come un rapporto strettamente proprio di polinomi in s , come segue:*

$$(sI - A)^{-1} = \frac{Q(s)}{p_A(s)} = \frac{Q_m s^m + Q_{m-1} s^{m-1} + \dots + Q_1 s^1 + Q_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (4.8)$$

dove $p_A(s)$ è il **polinomio caratteristico** associato alla matrice A e $Q(s)$ è una matrice polinomiale dove le Q_i sono matrici costanti.

Dimostrazione: Per la dimostrazione si fa riferimento al teorema di inversione di una matrice per la quale si ha:

$$(sI - A)^{-1} = \frac{\Delta^T(sI - A)}{\det(sI - A)}$$

dove $\Delta^T(sI - A)$ è la matrice trasposta dei cofattori di A , e $\det(sI - A)$ è un polinomio associato alla matrice dinamica A noto appunto come polinomio caratteristico.

Due importanti aspetti possono essere messi in evidenza come diretta conseguenza della precedente dimostrazione:

- Il polinomio caratteristico $p_A(s)$ è di ordine n , in quanto risulta essere il determinante della matrice $(sI - A) \in \mathbb{R}^{n \times n}$,
- I termini della matrice dei cofattori $\Delta(sI - A)$ sono dei polinomi che possono avere al massimo grado $n - 1$ essendo ottenuti come determinante di una sottomatrice di ordine $n - 1 \times n - 1$.

Si noti inoltre che il lemma 4.1 fornisce un legame fondamentale tra la funzione di trasferimento $H(s)$ e la matrice dinamica del sistema A . In particolare come verrà evidenziato dal teorema successivo vi è una corrispondenza, a meno di possibili cancellazioni, tra i **poli** della $H(s)$ e gli **autovalori** associati alla matrice A .

Teorema 4.3 (*Relazione Poli/Autovalori*) *Si consideri il sistema dinamico descritto dalla 4.1, i poli della matrice funzione di trasferimento $H(s)$ associata a tale sistema sono, a meno di possibili cancellazioni, gli autovalori della matrice dinamica del sistema A .*

Dimostrazione: Per la dimostrazione è sufficiente ricordare la definizione di matrice funzione di trasferimento fornita nella definizione 4.1 e far riferimento al lemma 4.1.

Infatti, dalla definizione di matrice funzione di trasferimento si ha:

$$H(s) = C (sI - A)^{-1} B$$

sostituendo l'espressione equivalente per la matrice risolvante fornita dal lemma 4.1 si ottiene:

$$\begin{aligned}
H(s) &= C \frac{Q(s)}{p_A(s)} B \\
&= C \frac{\Delta^T(sI - A)}{\det(sI - A)} B
\end{aligned}$$

da cui si evince che i poli del $H(s)$ a meno di cancellazioni sono gli autovalori della matrice dinamica A del sistema, ovvero le radici del polinomio caratteristico $p_A(s)$.

Si noti come il teorema 4.3 evidenzia il fatto che i poli della matrice funzione di trasferimento sono gli autovalori della matrice dinamica del sistema A a meno di eventuali **cancellazioni**. Infatti potrebbe accadere che gli elementi della $Q(s)$ e il polinomio caratteristico $p_A(s)$ abbiano dei fattori in comune da cui deriverebbero delle cancellazioni. Come conseguenza si avrebbe che la matrice funzione di trasferimento $H(s)$ può essere riscritta a valle della semplificazione come:

$$\tilde{H}(s) = \frac{\tilde{Q}(s)}{\tilde{p}(s)} \quad (4.9)$$

dove le radici associate al polinomio $\tilde{p}(s)$ sono dette **poli della matrice di trasferimento** e non sono tutti gli autovalori della matrice dinamica A . (In particolare sono tutti gli autovalori osservabili e raggiungibili).

4.2 F(s) a V.d.S.

Si consideri un sistema dinamico descritto dalla seguente funzione di trasferimento $F(s)$:

$$F(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0} = \frac{Y(s)}{U(s)} \quad (4.10)$$

dove $m < n$. Tale sistema può essere descritto attraverso una rappresentazione equivalente in Spazio di Stato, la quale può essere ricavata in maniera sistematica. In particolare due tecniche verranno presentate, queste permettono di portare il sistema in forma compagna di “controllore” e “osservatore”.

4.2.1 Forma Compagna di Controllore

Il sistema descritto dalla 4.10 può essere visto in termini di variabili di stato, associando il numeratore della $F(s)$ al rapporto $\frac{Y(s)}{X(s)}$ e il denominatore al rapporto $\frac{X(s)}{U(s)}$, cioè

$$F(s) = \frac{Y(s)}{U(s)} = \frac{Y(s)}{X(s)} \cdot \frac{X(s)}{U(s)}, \quad (4.11)$$

dove

$$\frac{X(s)}{U(s)} = \frac{1}{s^n + a_{n-1} s^{n-1} + \dots + a_0} \quad (4.12)$$

$$\frac{Y(s)}{X(s)} = b_m s^m + b_{m-1} s^{m-1} + \dots + b_0. \quad (4.13)$$

Effettuando ora una operazione di antitrasformazione si ottiene:

$$X(s) (s^n + a_{n-1}s^{n-1} + \dots + a_0) = U(s) \Rightarrow \frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \dots + a_0 x = u \quad (4.14)$$

$$Y(s) = (b_m s^m + b_{m-1} s^{m-1} + \dots + b_0) X(s) \Rightarrow y(t) = b_m \frac{d^m x}{dt^m} + \dots + b_0 x, \quad (4.15)$$

ed effettuando la sostituzione

$$\begin{cases} x_1 = x \\ x_2 = x' \\ x_3 = x'' \\ \vdots \\ x_n = x^{(n-1)} \end{cases} \Rightarrow \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 \\ \vdots \\ \dot{x}_{n-1} = x_n \\ \dot{x}_n = -a_0 x_1 - a_1 x_2 - \dots - a_{n-1} x_n + u \end{cases} \quad (4.16)$$

La rappresentazione matriciale in variabili di stato relativa al sistema dinamico descritto dalla $H(s)$ data nella 4.10 sarà quindi:

$$\begin{cases} \dot{x}(t) = A_c x(t) + B_c u(t) \\ y(t) = C_c x(t) \end{cases} \quad (4.17)$$

con le seguenti matrici A_c, B_c, C_c che prendono il nome di forma compagna di controllore

$$A_c = \begin{bmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & \dots & -a_{n-1} \end{bmatrix} \quad (4.18)$$

$$B_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.19)$$

$$C_c = [b_0 \ b_1 \ \dots \ b_m \ 0 \ 0 \ \dots \ 0] \quad (4.20)$$

È interessante osservare come gli elementi dell'ultima riga di una matrice in forma compagna sono, con segno opposto, i coefficienti del polinomio caratteristico.

Inoltre si noti che nel caso in cui sia $m = n$ si esegue la divisione tra il polinomio a numeratore e quello a denominatore: il quoziente rappresenta il legame diretto, il resto è una funzione di trasferimento con $m < n$.

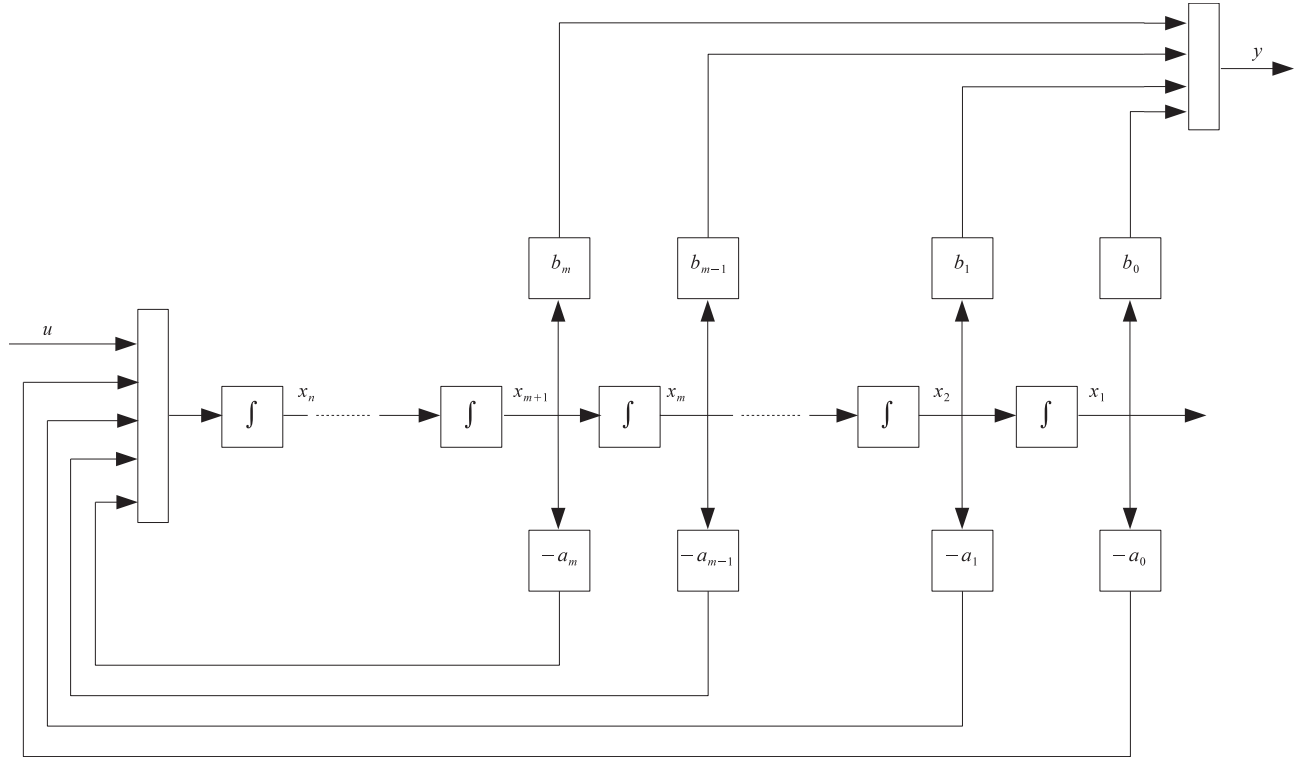


Figura 4.1: Forma Compagna di Controllore

4.2.2 Forma Compagna di Osservatore

Il sistema descritto dalla 4.10 può essere visto diversamente in termini di variabili di stato, effettuando le seguenti manipolazioni. A partire dalla definizione:

$$Y(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0} U(s) \quad (4.21)$$

si moltiplichino ambo i lato per il denominatore $(s^n + a_{n-1} s^{n-1} + \dots + a_0)$ come segue:

$$(s^n + a_{n-1} s^{n-1} + \dots + a_0) Y(s) = (b_m s^m + b_{m-1} s^{m-1} + \dots + b_0) U(s) \quad (4.22)$$

e si metta tutto in funzione della derivata di ordine n rispetto all'uscita:

$$s^n Y(s) = (-a_{n-1} s^{n-1} - \dots - a_1 s - a_0) Y(s) + (b_m s^m + b_{m-1} s^{m-1} + \dots + b_0) U(s) \quad (4.23)$$

si moltiplichino ora ambo i lato per $\frac{1}{s^n}$, in modo da ottenere la:

$$Y(s) = \left(-\frac{a_{n-1}}{s} - \dots - \frac{a_1}{s^{n-1}} - \frac{a_0}{s^n} \right) Y(s) + \left(\frac{b_m}{s^{n-m}} + \frac{b_{m-1}}{s^{n-m+1}} + \dots + \frac{b_0}{s^n} \right) U(s) \quad (4.24)$$

e si raccolgano infine i vari fattori come segue:

$$Y(s) = \frac{1}{s} \left(-a_{n-1}Y(s) \right) + \dots + \frac{1}{s^{n-(m+1)}} \left(-a_{m+1}Y(s) \right) + \quad (4.25)$$

$$+ \frac{1}{s^{n-m}} \left(b_m U(s) - a_m Y(s) \right) + \dots + \frac{1}{s^{n-1}} \left(b_1 U(s) - a_1 Y(s) \right) + \quad (4.26)$$

$$+ \frac{1}{s^n} \left(b_0 U(s) - a_0 Y(s) \right) \quad (4.27)$$

$$(4.28)$$

A questo punto, al fine di passare di ottenere una rappresentazione in spazio di stato si riarrangino i termini come segue:

$$Y(s) = \frac{1}{s} \left[-a_{n-1}Y(s) + \dots + \frac{1}{s} \left[-a_{m+1}Y(s) + \frac{1}{s} \left[\left(b_m U(s) - a_m Y(s) \right) + \right. \right. \right. \quad (4.29)$$

$$\left. \left. \left. + \dots + \frac{1}{s} \left[\left(b_1 U(s) - a_1 Y(s) \right) + \underbrace{\frac{1}{s} \left[\left(b_0 U(s) - a_0 Y(s) \right) \right]}_{x_1} \right] \dots \right] \right] \right] \quad (4.30)$$

x_2

e si definiscano le n variabili di stato (ricordando successivamente il legame $y = x_n$ tra l'uscita e l' n -esima variabile di stato) come segue:

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & -a_0 y + b_0 u \\ \dot{x}_2 & = & x_1 - a_1 y + b_1 u \\ \vdots & & \\ \dot{x}_{m+1} & = & x_m - a_m y + b_m u \\ \dot{x}_{m+2} & = & x_{m+1} - a_{m+1} y \\ \vdots & & \\ \dot{x}_{n-1} & = & x_{n-2} - a_{n-2} y \\ \dot{x}_n & = & x_{n-1} - a_{n-1} y \end{array} \right. \Rightarrow \left\{ \begin{array}{lcl} \dot{x}_1 & = & -a_0 x_n + b_0 u \\ \dot{x}_2 & = & x_1 - a_1 x_n + b_1 u \\ \vdots & & \\ \dot{x}_{m+1} & = & x_m - a_m x_n + b_m u \\ \dot{x}_{m+2} & = & x_{m+1} - a_{m+1} x_n \\ \vdots & & \\ \dot{x}_{n-1} & = & x_{n-2} - a_{n-2} x_n \\ \dot{x}_n & = & x_{n-1} - a_{n-1} x_n \end{array} \right. \quad (4.31)$$

La rappresentazione matriciale in variabili di stato relativa al sistema dinamico descritto dalla $H(s)$ data nella 4.10 sarà quindi:

$$\begin{cases} \dot{x}(t) = A_o x(t) + B_o u(t) \\ y(t) = C_o x(t) \end{cases} \quad (4.32)$$

con le seguenti matrici A_o, B_o, C_o che prendono il nome di forma compagna di osservatore:

$$A_o = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots & \cdots & -a_0 \\ 1 & 0 & 0 & 0 & \cdots & \cdots & -a_1 \\ 0 & 1 & 0 & 0 & \cdots & \cdots & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & 0 & 0 & 1 & -a_{n-1} \end{bmatrix} \quad (4.33)$$

$$B_o = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.34)$$

$$C_o = [0 \ 0 \ \cdots \ 0 \ 0 \ 0 \ \cdots \ 1] \quad (4.35)$$

Si noti che, essendo la funzione di trasferimento $H(s)$ definita come un rapporto strettamente proprio di polinomi (quindi $m < n$), il termine D del legame diretto ingresso/uscita risulta essere sempre nullo.

Diversamente, nel caso in cui il grado del numeratore sia uguale a quello del denominatore, ovvero $m = n$, si esegue la divisione tra il polinomio a numeratore e quello a denominatore: il quoziente rappresenta il legame diretto, il resto è una funzione di trasferimento con $m < n$.

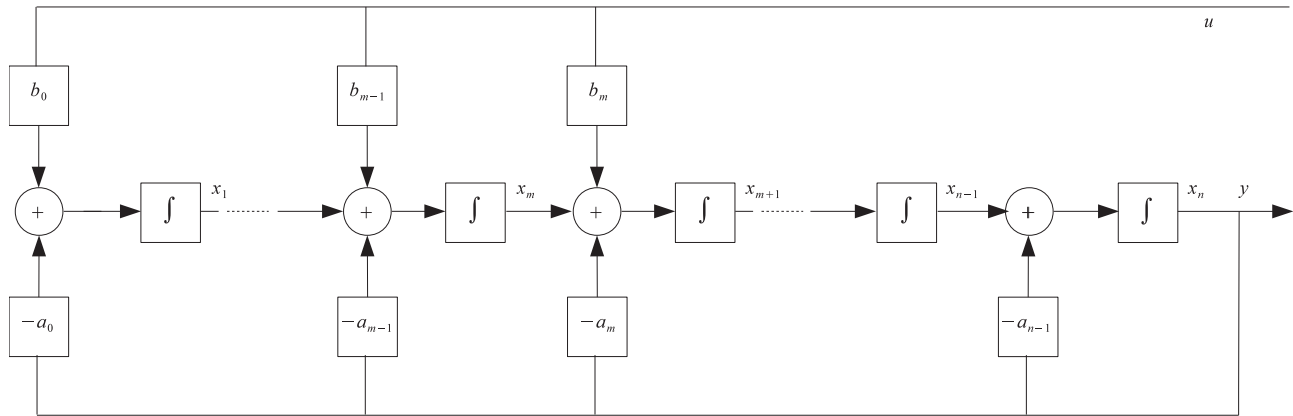


Figura 4.2: Forma Compagna di Osservatore

4.2.3 Proprietà di Dualità

Nella sezione precedente sono state introdotte due tecniche per il passaggio dalla rappresentazione attraverso funzione di trasferimento a quella in spazio di stato. È interessante notare come tali rappresentazioni siano duali.

Teorema 4.4 (*Dualità*) *Si consideri il sistema dinamico descritto dalla 4.10, la forma canonica di controllore descritta dalle matrici (A_c, B_c, C_c) (rispettivamente in 4.18, 4.19, ed 4.20) definisce una rappresentazione duale rispetto alla forma canonica di osservatore descritta dalle matrici $(\hat{A}_o, \hat{B}_o, \hat{C}_o)$ (rispettivamente in 4.33, 4.34, ed 4.35) se si considera un sistema “duale” al sistema di partenza dove gli ingressi vengono scambiati con le uscite, dove $\hat{A} = A^T$, $\hat{B} = C^T$, $\hat{C} = B^T$. In particolare si ha che:*

$$\begin{aligned} A_c &= \hat{A}_o^T & B_c &= \hat{C}_o^T \\ C_c &= \hat{B}_o^T & D_c &= \hat{D}_o \end{aligned} \tag{4.36}$$

Capitolo 5

Decomposizione Modale

5.1 Autovalori ed Autovettori

Sia $A \in \mathbb{R}^{n \times n}$ una applicazione lineare (endomorfismo rispetto ad uno spazio vettoriale):

$$A : \mathbb{R}^n \longrightarrow C \subseteq \mathbb{R}^n \quad (5.1)$$

Definizione 5.1 (*Autovettore destro*) Si definisce **autovettore destro** $v_d \in \mathbb{R}^n$ quel particolare vettore per cui valga la relazione:

$$Av_d = \lambda v_d \quad (5.2)$$

dove $\lambda \in \mathbb{R}$ è uno scalare noto col nome di **autovalore**.

Per il calcolo degli autovettori destri è sufficiente riscrivere la 5.2 come segue:

$$(A - \lambda I)v_d = 0 \quad (5.3)$$

Si noti come al fine di non avere come unica soluzione la soluzione triviale $v_d = 0$ sia necessario che il nucleo della matrice $(A - \lambda I)$ abbia una dimensione diversa da zero. In altri termini, che la matrice $(A - \lambda I)$ sia singolare. Si ricorda che affinché questo accada deve valere la:

$$\det(A - \lambda I) = 0 \quad (5.4)$$

Il termine $\det(A - \lambda I)$ va sotto il nome di **polinomio caratteristico** di grado n in λ e si indica usualmente con $p_A(\lambda)$. È quindi chiaro come gli autovalori della matrice A siano le radici del polinomio caratteristico.

Definizione 5.2 (*Molteplicità Algebrica*) Si definisce **molteplicità algebrica** dell'autovalore λ_i la molteplicità della soluzione $\lambda = \lambda_i$ dell'equazione $p_A(\lambda) = 0$, e si indica col termine *m.a.*(λ_i) o equivalentemente μ_i .

Definizione 5.3 (*Molteplicità Geometrica*) Si definisce **molteplicità geometrica** dell'autovalore λ_i la dimensione dell'autospazio associato all'autovettore v_i , e si indica col termine *m.g.*(λ_i) o equivalentemente ν_i .

Teorema 5.1 (*Indipendenza lineare*) Se gli n autovalori $\{\lambda_1, \dots, \lambda_n\}$ associati alla trasformazione lineare A sono tutti distinti, allora gli autovettori $\{v_1, \dots, v_n\}$ ad essi associati sono linearmente indipendenti e formano una base per l'immagine $C = R(A)$.

Definizione 5.4 (*Autovettore sinistro*) Si definisce **autovettore sinistro** $v_s \in \mathbb{R}^n$ quel particolare vettore per cui valga la relazione:

$$v_s^T A = \lambda v_s^T \quad (5.5)$$

dove $\lambda \in \mathbb{R}$ è uno scalare noto col nome di **autovalore**.

Analogamente al caso degli autovettori destri, anche per il calcolo degli autovettori sinistri è sufficiente riscrivere la 5.5 come segue:

$$v_s^T (A - \lambda I) = 0 \quad (5.6)$$

A questo punto, come nel caso precedente, al fine di non avere come unica soluzione la soluzione triviale $v_s = 0$ occorre verificare che il nucleo della matrice $(A - \lambda I)$ abbia una dimensione diversa da zero. In altri termini, che la matrice $(A - \lambda I)$ sia singolare, ovvero che la condizione 5.4 sia verificata.

È importante a questo punto fare un paio di osservazioni:

- Gli autovalori associati agli autovettori destri e sinistri sono gli stessi,
- Il problema del calcolo degli autovettori sinistri della matrice A equivale al problema del calcolo degli autovettori destri della matrice A^T .

Vediamo ora alcune proprietà che legano gli autovalori destri e sinistri di una applicazione lineare (matrice) A .

Teorema 5.2 (*Ortogonalità*) Si consideri una applicazione lineare $A \in \mathbb{R}^{n \times n}$, ogni autovettore sinistro $v_{s,i}$ associato ad un particolare autovalore λ_i risulta essere ortogonale a tutti gli autovettori destri $v_{d,j}$ associati ai rimanenti autovalori. In altri termini:

$$\langle v_{s,i}, v_{d,j} \rangle \neq 0 \quad \Longleftrightarrow \quad i = j. \quad (5.7)$$

Dimostrazione:

Si considerino 2 autovalori λ_i e λ_j ed i rispettivi autovettori destro e sinistro $v_{s,i}$ e $v_{d,j}$. Applicando la definizione 5.1 e 5.4 si ha:

$$v_{s,i}^T A = \lambda_i v_{s,i}^T \quad A v_{d,j} = \lambda_j v_{d,j}$$

Moltiplicando opportunamente entrambe le equazioni rispettivamente per $v_{d,j}$ e $v_{s,i}^T$ si ottiene:

$$v_{s,i}^T A v_{d,j} = \lambda_i v_{s,i}^T v_{d,j} \quad v_{s,i}^T A v_{d,j} = \lambda_j v_{s,i}^T v_{d,j}$$

Ora, sottraendo membro a membro si ottiene la:

$$v_{s,i}^T A v_{d,j} - v_{s,i}^T A v_{d,j} = (\lambda_i - \lambda_j) v_{s,i}^T v_{d,j}$$

ovvero:

$$(\lambda_i - \lambda_j) v_{s,i}^T v_{d,j} = 0$$

da cui abbiamo che:

$$\lambda_i \neq \lambda_j \implies v_{s,i}^T v_{d,j} = 0$$

Teorema 5.3 (Reciprocità) *Si consideri una applicazione lineare $A \in \mathbb{R}^{n \times n}$ invertibile, e si definisca l'insieme degli autovettori destri associati agli n autovalori distinti della matrice A come $V_d = [v_{d,1}, \dots, v_{d,n}]$. Si definisca allo stesso modo l'insieme degli autovettori sinistri come $V_s = [v_{s,1}, \dots, v_{s,n}]$.*

Allora vale la seguente proprietà:

$$V_s^T \cdot V_d = I \quad (5.8)$$

o equivalentemente:

$$V_s^T = V_d^{-1} \quad (5.9)$$

Dimostrazione: Si considerino 2 autovalori λ_i e λ_j ed i rispettivi autovettori destro e sinistro $v_{s,i}$ e $v_{d,j}$. Applicando la definizione 5.1 e 5.4 si ha:

$$v_{s,i}^T A = \lambda_i v_{s,i}^T \quad A v_{d,j} = \lambda_j v_{d,j}$$

Mettiamo in evidenza come segue:

$$v_{s,i}^T = \lambda_i v_{s,i}^T A^{-1} \quad v_{d,j} = \frac{1}{\lambda_j} A v_{d,j}$$

Moltiplichiamo ora membro a membro:

$$v_{s,i}^T v_{d,j} = \frac{\lambda_i}{\lambda_j} v_{s,i}^T A^{-1} A v_{d,j}$$

Ricordando ora la proprietà di ortogonalità presentata nel teorema 5.2 secondo cui ogni autovettore destro associato all' i -esimo autovalore è ortogonale a tutti gli autovettori sinistri associati ai rimanenti autovalori, si ha:

$$v_{s,i}^T v_{d,j} = v_{s,i}^T v_{d,j} \quad i = j$$

che attraverso una opportuna rinormalizzazione può essere espressa come:

$$v_{s,i}^T v_{d,j} = 1$$

la quale in termini vettoriali diventa:

$$V_s^T V_d = I$$

ovvero:

$$V_s^T = V_d^{-1}$$

È importante osservare come per la dimostrazione si sia fatto uso dell'inversione della matrice dinamica A . La non singolarità di tale matrice è garantita dal fatto di avere n autovalori distinti.

5.2 Trasformazione di Coordinate

Si consideri un sistema dinamico descritto dalle seguenti equazioni differenziali:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (5.10)$$

Si effettui ora il seguente cambiamento di variabili $x = Tz$,

$$\begin{cases} T\dot{z}(t) = ATz(t) + Bu(t) \\ y(t) = CTz(t) + Du(t) \end{cases} \Rightarrow \begin{cases} \dot{z}(t) = T^{-1}ATz(t) + T^{-1}Bu(t) \\ y(t) = CTz(t) + Du(t) \end{cases} \quad (5.11)$$

ovvero, ponendo

$$\begin{aligned} \tilde{A} &= T^{-1}AT \\ \tilde{B} &= T^{-1}B \\ \tilde{C} &= CT \\ \tilde{D} &= D, \end{aligned} \quad (5.12)$$

possiamo riscrivere in forma compatta come:

$$\begin{cases} \dot{z}(t) = \tilde{A}z(t) + \tilde{B}u(t) \\ y(t) = \tilde{C}z(t) \end{cases} \quad (5.13)$$

È interessante notare come le due rappresentazioni siano equivalenti. In particolari le proprietà strutturali quali ad esempio, gli autovalori del sistema dinamico, restano invariati a seguito di una trasformazione di coordinate. A tale proposito verrà ora introdotto un importante teorema dell'algebra che ci permetterà poi di provare quanto appena asserito.

Teorema 5.4 (*Cayley-Hamilton*) Sia $p_A(\lambda)$ il polinomio caratteristico associato alla matrice dinamica A del sistema descritto dalla 5.10, allora vale la:

$$p_A(A) = 0, \quad 0 \in \mathbb{R}^{n \times n} \quad (5.14)$$

ossia ogni matrice quadrata soddisfa la propria equazione caratteristica.

Dimostrazione:

Da inserire.

Lemma 5.1 *Il teorema 5.4 è di estrema importanza in quanto ci dice che ogni potenza di ordine $\geq n$ della matrice A può essere espressa come combinazione lineare delle prime n potenze. Infatti essendo il polinomio caratteristico:*

$$p_A(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 \quad (5.15)$$

per il teorema 5.4 si ha che:

$$p_A(A) = A^n + \alpha_{n-1}A^{n-1} + \dots + \alpha_1A + \alpha_0I = 0 \quad (5.16)$$

da cui si deduce che ogni potenza di ordine $\geq n$ della matrice A può essere espressa come combinazione lineare delle prime n potenze, ovvero:

$$A^n = -\alpha_{n-1}A^{n-1} + \dots - \alpha_1A - \alpha_0I \quad (5.17)$$

Teorema 5.5 (*“Invarianza” Autovalori*) *Una trasformazione di coordinate non cambia gli autovalori associati ad un generico sistema dinamico.*

Dimostrazione: La dimostrazione è una semplice applicazione del teorema di Cayley-Hamilton definito nella 5.4. Dato un sistema dinamico descritto dalla 5.10, si consideri il polinomio caratteristico $p_A(\lambda)$ associato alla matrice dinamica A . Si prenda ora la matrice dinamica \tilde{A} associata ad una qualsiasi rappresentazione ottenuta attraverso una trasformazione di coordinate a partire dalla 5.10, come ad esempio la 5.11. Si prenda ora il polinomio caratteristico $p_A(\lambda)$ calcolato per \tilde{A} :

$$p_A(\tilde{A}) = \tilde{A}^n + \alpha_{n-1}\tilde{A}^{n-1} + \dots + \alpha_1\tilde{A} + \alpha_0I$$

ora ricordando che:

$$\tilde{A}^k = (T^{-1} A T)^k = T^{-1} A^k T$$

il polinomio caratteristico associato alla matrice A calcolato rispetto alla matrice \tilde{A} può essere riscritto come segue:

$$p_A(\tilde{A}) = T^{-1} (A^n + \alpha_{n-1}A^{n-1} + \dots + \alpha_1A + \alpha_0I) T = 0$$

il quale, evidenzia come le due rappresentazioni siano del tutto equivalenti, e quindi gli autovalori siano identici.

Teorema 5.6 (*“Invarianza” Funzione di Trasferimento*) *Una trasformazione di coordinate non cambia la funzione di trasferimento associata ad un generico sistema dinamico. In altre parole, data la funzione di trasferimento $H(s)$ associata al sistema 5.10:*

$$H(s) = C (sI - A)^{-1} B \quad (5.18)$$

e la funzione di trasferimento $\tilde{H}(s)$ associata al sistema 5.11:

$$\tilde{H}(s) = \tilde{C} (sI - \tilde{A})^{-1} \tilde{B} \quad (5.19)$$

si ha che:

$$H(s) = \tilde{H}(s). \quad (5.20)$$

Dimostrazione: Si considerino le matrici A, B, C associate alla rappresentazione 5.10 di un generico sistema dinamico, e si considerino le matrici $\tilde{A}, \tilde{B}, \tilde{C}$ associate ad una rappresentazione alternativa 5.11 relativa allo stesso sistema. Sia ora T la matrice di trasformazione che porta da una rappresentazione all'altra in modo da avere, come già evidenziato nella 5.12, la seguente equivalenza:

$$\begin{aligned} \tilde{A} &= T^{-1} A T \\ \tilde{B} &= T^{-1} B \\ \tilde{C} &= C T \end{aligned}$$

Scriviamo ora la funzione di trasferimento $\tilde{H}(s)$ associata alla rappresentazione 5.11 come:

$$\tilde{H}(s) = \tilde{C} (sI - \tilde{A})^{-1} \tilde{B}$$

La quale può essere tranquillamente riscritta come segue tenendo conto del legame dettato dalla matrice di trasformazione T :

$$\begin{aligned} \tilde{H}(s) &= C T (sI - T^{-1} A T)^{-1} T^{-1} B \\ &= C T (sT^{-1} I T - T^{-1} A T)^{-1} T^{-1} B \\ &= C T T^{-1} (sI - A)^{-1} T T^{-1} B \\ &= C (sI - A)^{-1} B \end{aligned}$$

Il quale prova che:

$$\tilde{H}(s) = H(s)$$

5.3 Diagonalizzazione

Si consideri un sistema dinamico descritto dalle seguenti equazioni:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (5.21)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$.

Un problema estremamente interessante è quello di trovare (e quindi di determinare le condizioni di esistenza) una trasformazione in grado di portare il sistema in forma diagonale. Tale problema è noto come problema della **diagonalizzazione** e la sua rilevanza è diretta conseguenza del fatto che un sistema in forma diagonale comporta una evoluzione “indipendente” e “parallela” di ogni singola dinamica e quindi permette di applicare sforzi di controllo atti alla modifica della singola dinamica.

Tuttavia, non tutti i sistemi possono essere portati in forma diagonale, ed anche quando ciò è possibile, non sempre risulta essere una operazione elementare. Nel seguito andremo ad effettuare una analisi spettrale della matrice dinamica del sistema A al fine di individuare condizioni che permettano il passaggio in forma diagonale e le relative matrici di trasformazione.

Definizione 5.5 (*Diagonalizzazione*) Sia $A \in \mathbb{R}^{n \times n}$ una matrice quadrata di ordine n , il problema della diagonalizzazione consiste nella determinazione di una matrice non singolare P tale che la matrice A risulti essere simile alla matrice diagonale Λ , ovvero:

$$A = P \Lambda P^{-1} \quad (5.22)$$

dove $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ è la matrice con gli autovalori sulla diagonale.

Teorema 5.7 (*Diagonalizzabilità “semplice”*) Sia $A \in \mathbb{R}^{n \times n}$ la matrice dinamica del sistema descritto in 5.21, condizione necessaria e sufficiente per la diagonalizzazione è che tutti gli autovalori $\{\lambda_1, \dots, \lambda_n\}$ associati alla matrice A siano distinti. Inoltre, la matrice P per la trasformazione di coordinate che porta il sistema in forma diagonale è la seguente:

$$P = [v_1, \dots, v_n] \quad (5.23)$$

dove v_i è l'autovettore associato all' i -esimo autovalore λ_i .

Dimostrazione: Per la dimostrazione si farà semplicemente uso della definizione di autovalore destro data in 5.2:

$$A v_i = \lambda v_i$$

In forma matriciale diventa:

$$A [v_1, \dots, v_n] = [v_1, \dots, v_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & \vdots \\ \vdots & \dots & \ddots & 0 \\ 0 & \dots & \dots & \lambda_n \end{bmatrix}$$

il quale può essere riscritto in forma compatta come:

$$A P = P \Lambda$$

da cui si ottiene agevolmente:

$$A = P \Lambda P^{-1}$$

Teorema 5.8 (*Diagonalizzabilità per poli a molteplicità algebrica > 1 , caso $\mu_i = \nu_i$*) Sia $A \in \mathbb{R}^{n \times n}$ la matrice dinamica del sistema descritto in 5.21, condizione necessaria e sufficiente per la diagonalizzazione è che tutti gli autovalori $\{\lambda_1, \dots, \lambda_h\}$ abbiano molteplicità geometrica pari a quella algebrica (in particolare quegli autovalori con molteplicità algebrica > 1). Inoltre, la matrice P per la trasformazione di coordinate che porta il sistema in forma diagonale è la seguente:

$$T = [v_1, v_2, \dots, \{v_{i,1}, \dots, v_{i,\mu_i}\}, \dots, \{v_{j,1}, \dots, v_{j,\mu_j}\}, \dots, v_{h-1}, v_h] \quad (5.24)$$

dove $\{v_{i,1}, \dots, v_{i,\mu_i}\}$ sono gli μ_i autovettori associati all' i -esimo autovalore λ_i di molteplicità algebrica μ_i (i quali peraltro formano una base per il $N(A - \lambda_i I)$).

Dimostrazione: Per la dimostrazione di tale teorema si deve far riferimento alla dimostrazione presentata per il teorema 5.7 ricordando che, nel caso in cui si abbia un autovalore λ_i con molteplicità geometrica pari alla molteplicità algebrica ($\nu_i = \mu_i$) è sempre possibile trovare un insieme di μ_i autovettori indipendenti associato a tale autovalore per la costruzione della matrice P per la trasformazione di coordinate.

Esempio 1 (*Diagonalizzazione, caso $\mu_i = \nu_i = 1$*)

Si consideri la seguente matrice $A \in \mathbb{R}^{4 \times 4}$:

$$A = \begin{bmatrix} 5 & -2 & 0 & 2 \\ 0 & 8 & -2 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 1 & -2 & 7 \end{bmatrix}$$

Il cui polinomio caratteristico risulta essere:

$$\begin{aligned} p_A(\lambda) &= \lambda^4 - 26\lambda^3 + 251\lambda^2 - 1066\lambda + 1680 \\ &= (\lambda - 5)(\lambda - 6)(\lambda - 7)(\lambda - 8) \end{aligned}$$

da cui si evince che vi sono 4 autovalori distinti, rispettivamente $\{\lambda_1 = 5, \lambda_2 = 6, \lambda_3 = 7, \lambda_4 = 8\}$ di molteplicità algebrica $\mu_i = 1 \quad \forall i \in \{1, \dots, 4\}$.

A questo punto ci calcoliamo gli autovettori associati a tali autovalori come segue:

$$\mathcal{B}\{N(A_1)\} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathcal{B}\{N(A_2)\} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \mathcal{B}\{N(A_3)\} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \mathcal{B}\{N(A_4)\} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

dove $A_i = (A - \lambda_i I) \quad \forall i \in \{1, \dots, 4\}$ e $v_i = \mathcal{B}\{N(A_i)\}$ essendo la molteplicità algebrica μ_i unitaria per ogni autovalore. Si ricorda che la molteplicità geometrica ha valori limitati superiormente da quella algebrica, ovvero $\nu_i \leq \mu_i$. Mettiamo ora tutto insieme per ottenere la matrice P per la trasformazione di coordinate:

$$\begin{aligned} P &= [v_1, v_2, v_3, v_4] \\ &= \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \end{aligned}$$

A questo punto per ottenere la forma diagonale è sufficiente calcolare l'inversa della matrice P . È possibile evitare il calcolo diretto dell'inversa P^{-1} in virtù del Teorema 5.3, il quale fa riferimento al calcolo della matrice degli autovettori destri (più semplice).

5.4 Jordanizzazione

Nella sezione precedente sono state evidenziate condizioni necessarie e sufficienti (teoremi 5.7 e 5.8) per portare una matrice in forma diagonale. In accordo con quanto visto precedentemente, in tutti quei casi in cui gli autovalori abbiano una molteplicità algebrica diversa dalla molteplicità geometrica non è possibile trovare una matrice di trasformazione in grado di portare il sistema in forma diagonale.

In ogni caso è sempre possibile portare il sistema in una forma **diagonale a blocchi** nota come forma di Jordan. Il vantaggio è che nonostante la forma diagonale non sia disponibile, è comunque possibile far riferimento a strutture note (blocco a blocco) per il calcolo dell'esponenziale di matrice.

Teorema 5.9 (Forma Canonica di Jordan) Sia $A \in \mathbb{R}^{n \times n}$ la matrice dinamica del sistema descritto in 5.21, si definisce matrice in forma canonica di Jordan una matrice J simile alla matrice A che abbia una struttura di matrice diagonale a blocchi, ovvero:

$$A = P J P^{-1} \quad (5.25)$$

con:

$$J = \begin{bmatrix} J_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & J_{1,\nu_1} & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & J_{q,1} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & J_{q,\nu_q} \end{bmatrix} \quad (5.26)$$

dove q è il numero di autovalori distinti, $J_{i,j} \in \mathbb{R}^{p_{i,j} \times p_{i,j}}$ è il j -esimo **blocco di Jordan** relativo all' i -esimo autovalore λ_i con molteplicità algebrica μ_i e molteplicità geometrica ν_i , ed ha la seguente struttura:

$$J_{i,j} = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \lambda_i \end{bmatrix} \quad (5.27)$$

Inoltre, per tutti i blocchi di Jordan $J_i = \{J_{i,1}, \dots, J_{i,\nu_i}\}$ associati ad uno stesso autovalore λ_i vale:

$$\sum_{j=1}^{\nu_i} p_{i,j} = \mu_i, \quad (5.28)$$

ovvero la somma del numero delle righe degli ν_i blocchi di Jordan associati ad un autovalore λ_i deve essere pari alla sua molteplicità algebrica μ_i .

Per quanto riguarda la matrice di trasformazione P che porta il sistema nella forma canonica di Jordan essa ha la seguente struttura:

$$P = P_1 \oplus \dots \oplus P_q \quad (5.29)$$

con $P_i = \left[\{v_{i,1}^{(1)}, \dots, v_{i,p_{i,1}}^{(1)}\}, \dots, \{v_{i,1}^{(\nu_i)}, \dots, v_{i,p_{i,\nu_i}}^{(\nu_i)}\} \right]$ l'insieme degli **autovettori generalizzati** associati ai vari blocchi di Jordan $J_i = \{J_{i,1}, \dots, J_{i,\nu_i}\}$ relativi all' i -esimo autovalore λ_i di molteplicità algebrica μ_i e molteplicità geometrica ν_i . Inoltre con la notazione $v_{i,j}^{(h)}$ si intende il j -esimo autovettore relativo all' h -esimo blocco di Jordan associato all' i -esimo autovalore λ_i .

Definizione 5.6 (Autovettore generalizzato di ordine k) Sia λ_i un generico autovalore della matrice dinamica del sistema $A \in \mathbb{R}^{n \times n}$ descritto in 5.21 con molteplicità algebrica μ_i , e molteplicità geometrica ν_i . Si definisce autovettore generalizzato v_k di ordine k dell'autovalore λ_i quel particolare vettore $v \in \mathbb{R}^n$ per cui vale la:

$$\left[v \in N(A - \lambda_i I)^k \right] \wedge \left[v \notin N(A - \lambda_i I)^{k-1} \right] \quad (5.30)$$

Inoltre dato un autovettore generalizzato v_k di ordine k , è possibile costruire un autovettore generalizzato v_{k-1} di ordine $k-1$ come segue:

$$v_{k-1} = (A - \lambda_i I) v_k \quad (5.31)$$

È interessante quindi osservare come attraverso tale tecnica si possa sempre costruire delle **catene** di autovettori generalizzati lunghe k .

Definizione 5.7 (Autospazio generalizzato di ordine k) Sia λ_i un generico autovalore della matrice dinamica del sistema $A \in \mathbb{R}^{n \times n}$ descritto in 5.21 con molteplicità algebrica μ_i , e molteplicità geometrica ν_i . Si definisce autospazio generalizzato A_i^k di ordine k il seguente sottospazio:

$$A_i^k = (A - \lambda_i I)^k. \quad (5.32)$$

Teorema 5.10 (Blocco di Jordan, caso $\mu_i \neq \nu_i, \nu_i = 1$) Sia λ_i un generico autovalore della matrice dinamica del sistema $A \in \mathbb{R}^{n \times n}$ descritto in 5.21 con molteplicità algebrica μ_i , e molteplicità geometrica $\nu_i = 1$. Allora per il teorema 5.9 vi è un solo blocco di Jordan J_i e per il calcolo degli μ_i autovettori generalizzati $\{v_{i,1}^{(1)}, \dots, v_{i,\mu_i}^{(1)}\}$ associati all' i -esimo autovalore λ_i di molteplicità algebrica μ_i va risolto il seguente sistema:

$$\begin{cases} A v_{i,1}^{(1)} = \lambda_i v_{i,1}^{(1)} \\ A v_{i,2}^{(1)} = \lambda_i v_{i,2}^{(1)} + v_{i,1}^{(1)} \\ \vdots = \vdots + \vdots \\ A v_{i,\mu_i}^{(1)} = \lambda_i v_{i,\mu_i}^{(1)} + v_{i,\mu_i-1}^{(1)} \end{cases} \Rightarrow \begin{cases} (A - \lambda_i I) v_{i,1}^{(1)} = 0 \\ (A - \lambda_i I) v_{i,2}^{(1)} = v_{i,1}^{(1)} \\ \vdots = \vdots \\ (A - \lambda_i I) v_{i,\mu_i}^{(1)} = v_{i,\mu_i-1}^{(1)} \end{cases} \quad (5.33)$$

dove gli autovettori generalizzati sono scelti in modo da essere indipendenti. Operativamente si può pensare di scegliere $v_{i,\mu_i}^{(1)}$ in maniera tale che esso sia un autovettore generalizzato di ordine μ_i e di costruire a partire da questo una catena lunga μ_i .

Teorema 5.11 (Blocco di Jordan, caso $\mu_i \neq \nu_i, \nu_i > 1$)

Sia λ_i un generico autovalore della matrice dinamica del sistema $A \in \mathbb{R}^{n \times n}$ descritto in 5.21 con molteplicità algebrica μ_i , e molteplicità geometrica ν_i , tale che $\mu_i \neq \nu_i$. Allora per il teorema 5.9 vi sono ν_i blocchi di Jordan J_i associati a tale autovalore per i quali si ha:

$$\sum_{j=1}^{\nu_i} \# \text{row}(J_{i,j}) = \mu_i \quad (5.34)$$

ovvero la somma del numero di righe dei vari blocchi di Jordan deve essere pari alla molteplicità algebrica μ_i dell'autovalore ad essi associato. Per il calcolo della dimensione dei vari blocchi di Jordan si può pensare di utilizzare la seguente tecnica. Per prima cosa si costruisce una successione di elementi d_k dove ogni singolo elemento è definito come segue:

$$d_k = \dim \left(N(A - \lambda_i I)^k \right). \quad (5.35)$$

Successivamente si verifica per quale valore k la successione si stabilizza e sulla base di tale informazione si costruisce una tabella dalla quale si ricava la dimensione dei vari blocchi di Jordan associati all'autovalore λ_i . Si ricorda che una successione si stabilizza per un valore h se vale la:

$$d_{h+1} = d_h = \dim \left(N(A - \lambda_i I)^h \right). \quad (5.36)$$

Si costruisce poi una tabella avente $\#col = h$ e $\#row = \nu_i$. Nella prima colonna si metta una crocetta in ogni riga. Nella seconda colonna si riempiano con una crocetta tante righe quant'è il valore di $d_2 - d_1$ e si lascino vuote le altre, dove si ricorda che $d_1 = \nu_i$ per definizione. Nella terza riga si mettano crocette su un $\#$ di righe pari al valore di $d_3 - d_2$ e così via fino alla colonna $d_h - d_{h-1}$. A questo punto si contano le crocette presenti in ogni riga della tabella: il numero di crocette p_{i,ν_i} sulla prima riga dà la dimensione del blocco J_{i,ν_i} , il numero di crocette p_{i,ν_i-1} della seconda riga dà la dimensione del blocco J_{i,ν_i-1} , e così via.

Una possibile rappresentazione grafica è la seguente:

$d_1 = \nu_i$	$d_2 - d_1$	\dots	$d_h - d_{h-1}$	# Crocette	Dim. blocco
x	x		x	p_{i,ν_i}	$\dim(J_{i,\nu_i}) = p_{i,\nu_i} \times p_{i,\nu_i}$
x	x		\vdots	p_{i,ν_i-1}	$\dim(J_{i,\nu_i-1}) = p_{i,\nu_i-1} \times p_{i,\nu_i-1}$
x	\vdots		x	\vdots	
\vdots	x			p_{i,ν_i-k}	$\dim(J_{i,\nu_i-k}) = p_{i,\nu_i-k} \times p_{i,\nu_i-k}$
\vdots				\vdots	
x				$p_{i,1}$	$\dim(J_{i,1}) = p_{i,1} \times p_{i,1}$

Teorema 5.12 (Blocco di Jordan, caso autovalori complessi e coniugati)

Sia $\lambda_i = \sigma + j\omega$ e $\lambda_{i+1} = \overline{\lambda_i} = \sigma - j\omega$ una generica coppia di autovalori complessi e coniugati della matrice dinamica del sistema $A \in \mathbb{R}^{n \times n}$ descritto in 5.21. Si ha che il blocco di Jordan associato a tale coppia di autovalori è il seguente:

$$J_i = \begin{bmatrix} \sigma + j\omega & 0 \\ 0 & \sigma - j\omega \end{bmatrix} \quad (5.37)$$

e la componente P_i della matrice di trasformazione P che porta il sistema in forma canonica di Jordan ha la seguente struttura:

$$P_i = \begin{bmatrix} v_{i,1}^{(1)} & \overline{v}_{i,1}^{(1)} \end{bmatrix} \quad (5.38)$$

dove:

$$v_{i,1}^{(1)} = q_r + j q_i \quad \overline{v}_{i,1}^{(1)} = q_r - j q_i \quad (5.39)$$

sono la coppia di autovettori complessi e coniugati di parte reale q_r e parte immaginaria q_i associati alla coppia di autovalori complessi e coniugati λ_i e $\overline{\lambda_i}$.

È possibile inoltre riscrivere tale blocco in una forma alternativa (di estrema utilità pratica) in cui viene messa in luce la natura oscillatoria della coppia di autovalori complessi e coniugati come segue:

$$J_i = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} \quad (5.40)$$

la quale può essere ottenuta applicando la seguente trasformazione:

$$\hat{P}_i = P_i \cdot E \quad (5.41)$$

$$= \frac{1}{2} \begin{bmatrix} v_{i,1}^{(1)} & \bar{v}_{i,1}^{(1)} \end{bmatrix} \cdot \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix} \quad (5.42)$$

$$= \begin{bmatrix} \operatorname{Re} [v_{i,1}^{(1)}] & \operatorname{Im} [v_{i,1}^{(1)}] \end{bmatrix} \quad (5.43)$$

$$= \begin{bmatrix} q_r & q_i \end{bmatrix} \quad (5.44)$$

Dimostrazione: Per la dimostrazione faremo riferimento ad una matrice $A \in \mathbb{R}^{2 \times 2}$ caratterizzata da una coppia di autovalori complessi e coniugati, la quale in termini più generali potrebbe essere vista come il blocco di Jordan di dimensione 2×2 associato ad una coppia di autovalori complessi e coniugati. Si definisca quindi la matrice di trasformazione \hat{P} ad essa associata come segue:

$$\begin{aligned} \hat{P} &= \begin{bmatrix} \operatorname{Re} [v_{1,1}^{(1)}] & \operatorname{Im} [v_{1,1}^{(1)}] \end{bmatrix} \\ &= \begin{bmatrix} \frac{(v_{1,1}^{(1)} + \bar{v}_{1,1}^{(1)})}{2} & \frac{(v_{1,1}^{(1)} - \bar{v}_{1,1}^{(1)})}{2j} \end{bmatrix} \end{aligned}$$

dove $v_{i,1}^{(1)}$ e $\bar{v}_{i,1}^{(1)}$ rappresentano la coppia di autovettori complessi coniugati associati alla coppia di autovalori complessi e coniugati. Sia inoltre nel seguito $v = v_{i,1}^{(1)}$ e $\bar{v} = \bar{v}_{i,1}^{(1)}$ per semplicità di notazione. Ora effettuando la moltiplicazione per la matrice A si ottiene:

$$\begin{aligned} A \hat{P} &= A \begin{bmatrix} \frac{(v + \bar{v})}{2} & \frac{(v - \bar{v})}{2j} \end{bmatrix} \\ &= \begin{bmatrix} \frac{A(v + \bar{v})}{2} & \frac{A(v - \bar{v})}{2j} \end{bmatrix} \\ &= \begin{bmatrix} \frac{(A v + A \bar{v})}{2} & \frac{(A v - A \bar{v})}{2j} \end{bmatrix} \\ &= \begin{bmatrix} \frac{(\lambda v + \bar{\lambda} \bar{v})}{2} & \frac{(\lambda v - \bar{\lambda} \bar{v})}{2j} \end{bmatrix} \\ &= \begin{bmatrix} \operatorname{Re}[\lambda v] & \operatorname{Im}[\lambda v] \end{bmatrix} \end{aligned}$$

Sia ora $\lambda = \sigma + j\omega$ da cui si ha:

$$\begin{aligned} \operatorname{Re}[\lambda v] &= \operatorname{Re}[(\sigma + j\omega) v] = \operatorname{Re}[\sigma v + j\omega v] = \sigma \operatorname{Re}[v] - \omega \operatorname{Im}[v] \\ \operatorname{Im}[\lambda v] &= \operatorname{Im}[(\sigma + j\omega) v] = \operatorname{Im}[\sigma v + j\omega v] = \sigma \operatorname{Im}[v] + \omega \operatorname{Re}[v] \end{aligned}$$

il quale sostituito alla precedente e riscritto in forma matriciale risulta essere:

$$A \hat{P} = \hat{P} \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

ovvero:

$$A = \hat{P} \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} \hat{P}^{-1}$$

da cui si evince che:

$$J = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

Esempio 2 (Blocco di Jordan, caso $\mu_i \neq \nu_i, \nu_i = 1$)

Si consideri la seguente matrice $A \in \mathbb{R}^{4 \times 4}$:

$$A = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 0 & 3 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

il cui polinomio caratteristico risulta essere:

$$p_A(\lambda) = (\lambda - 2)^4$$

da cui si evince che vi è un autovalore $\lambda_1 = 2$ di molteplicità algebrica $\mu_1 = 4$.

Andiamo ora ad analizzare la molteplicità geometrica associata a tale autovalore il quale richiede il calcolo della dimensione del nullo della matrice $(A - \lambda_1 I)$, ovvero:

$$\nu_1 = \dim(N(A - \lambda_1 I)) = 1$$

il quale ci informa che vi sarà un solo blocco di Jordan $J_1 = \{J_{1,1}\}$ associato all'unico autovalore λ_1 , e ci fa notare che ci si trova nel caso descritto dal teorema 5.10. Si avrà quindi che la matrice di Jordan ha la seguente struttura:

$$J = J_{1,1} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad (5.45)$$

Mentre per il calcolo degli μ_1 autovettori generalizzati possiamo costruirci la catena di lunghezza 4 a partire da un vettore della base di $N(A - \lambda_1 I)^4$ che non appartenga alla base di $N(A - \lambda_1 I)^3$ come da definizione 5.6. Per tale motivo ci scriviamo le basi per i nuclei relativi al terzo e al quarto ordine di potenza come segue:

$$\mathcal{B}\{N(A_1^3)\} = \begin{Bmatrix} 0 & 0 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix} \quad \mathcal{B}\{N(A_1^4)\} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix}$$

A questo punto notiamo che due possibili candidati possono essere presi in considerazione per la costruzione della catena degli autovettori generalizzati, ovvero:

$$\begin{Bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{Bmatrix}$$

scegliamo di prendere $v_{1,4}$ come il primo dei due candidati e ci costruiamo gli altri tre autovettori della catena come segue:

$$v_{1,4}^{(1)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad v_{1,3}^{(1)} = A_1 v_{1,4}^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad v_{1,2}^{(1)} = A_1 v_{1,3}^{(1)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad v_{1,1}^{(1)} = A_1 v_{1,2}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

Mettiamo ora tutto insieme per ottenere la matrice P per la trasformazione di coordinate:

$$\begin{aligned} P &= [v_{1,1}^{(1)}, v_{1,2}^{(1)}, v_{1,3}^{(1)}, v_{1,4}^{(1)}] \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

la quale ci garantisce che:

$$J = P^{-1} A P.$$

Esempio 3 (Blocco di Jordan, caso $\mu_i \neq \nu_i, \nu_i > 1$)

Si consideri la seguente matrice $A \in \mathbb{R}^{5 \times 5}$:

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

il cui polinomio caratteristico risulta essere:

$$p_A(\lambda) = \lambda (\lambda - 1)^4$$

da cui si evince che vi sono due autovalori $\lambda_1 = 0$ e $\lambda_2 = 1$ rispettivamente con molteplicità algebrica $\mu_1 = 1$ e $\mu_2 = 4$.

Andiamo ora ad analizzare la molteplicità geometrica dei due autovalori. Essendo nota dall'algebra la relazione $1 \leq \nu_i \leq \mu_i$, è immediato verificare che la molteplicità geometrica del primo autovalore è pari ad 1. Diversa è la situazione per il secondo autovalore che richiede il calcolo della dimensione del nullo della matrice $(A - \lambda_2 I)$, ovvero:

$$\nu_2 = \dim(N(A - \lambda_2 I)) = 2$$

il quale ci informa del fatto che vi saranno due blocchi di Jordan associati al secondo autovalore. Siamo a questo punto in grado di conoscere la struttura (a blocchi) della forma canonica di Jordan associata alla matrice A che risulta essere:

$$J = \begin{bmatrix} J_{1,1} & 0 & 0 \\ 0 & J_{2,1} & 0 \\ 0 & 0 & J_{2,2} \end{bmatrix}$$

Ora, poichè la somma del numero delle righe di ogni blocco di Jordan associate allo stesso autovalore deve dare la molteplicità algebrica associata a quell'autovalore, nel caso di λ_2 con $\nu_2 = 2$ e $\mu_2 = 4$ si ha che i due blocchi $\{J_{2,1}, J_{2,2}\}$ possono avere tre possibili configurazioni:

- $p_{2,1} = p_{2,2} = 2$,
- $p_{2,1} = 3, p_{2,2} = 1$,
- $p_{2,1} = 1, p_{2,2} = 3$,

dove $p_{i,j}$ rappresenta il numero di righe (colonne) del j -esimo blocco di Jordan associato all' i -esimo autovalore. È facile notare come il numero delle possibili configurazioni possa crescere in maniera significativa al crescere dei valori della molteplicità algebrica e geometrica associata ad un dato autovalore.

Il calcolo della dimensione dei blocchi $J_i = \{J_{i,1}, \dots, J_{i,\nu_i}\}$ associati ad un determinato autovalore λ_i di molteplicità algebrica μ_i e molteplicità geometrica ν_i si basa sulla tecnica descritta nel teorema 5.11. Definita la matrice $A_2 = (A - \lambda_2 I)$, si deve verificare che la successione degli elementi d_k , dove ogni elemento è definito come:

$$d_k = \dim(N(A - \lambda_2 I)^k) = \dim(N(A_2^k))$$

si stabilizzi, ovvero che sia $d_{h+1} = d_h$. A questo punto, a partire dal calcolo degli elementi della successione si costruisce una tabellina attraverso la quale si determina il grado di ogni blocco di Jordan.

Nel nostro caso si ha che la successione si stabilizza per $k = 3$, e la relativa tabellina è quindi la seguente:

$d_1 = \nu_2 = 2$	$d_2 - d_1 = 1$	$d_3 - d_2 = 1$	# Crocette	Dim. blocco
x	x	x	3	$\dim(J_{2,2}) = 3 \times 3$
x			1	$\dim(J_{2,1}) = 1 \times 1$

dal quale si evince che il primo dei due blocchi ha dimensione 1×1 mentre l'altro ha dimensione 3×3 (si noti l'ordinamento dei blocchi in funzione della dimensione crescente degli stessi). Siamo ora in grado di sviluppare la forma canonica di Jordan associata alla matrice A , la quale risulta essere:

$$J = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

A questo punto va calcolata la matrice di trasformazione P in grado di portare il sistema A nella forma canonica di Jordan J . Sappiamo che vi è un blocco $J_1 = \{J_{1,1}\}$ associato al primo autovalore e due blocchi $J_2 = \{J_{2,1}, J_{2,2}\}$ associati al secondo. Cerchiamo per primo gli autovettori generalizzati $\{v_{2,1}^{(2)}, \dots, v_{2,3}^{(2)}\}$ associati al blocco $J_{2,2}$, a seguire l'autovettore $\{v_{2,1}^{(1)}\}$ associato al blocco $J_{2,1}$ e per finire l'autovettore $\{v_{1,1}^{(1)}\}$ associato al blocco $J_{1,1}$.

Quindi, per il blocco $J_{2,2} \in \mathbb{R}^{3 \times 3}$ ci costruiamo la catena degli autovettori generalizzati a partire dall'autovettore generalizzato di ordine 3 scelto tra gli elementi della base del nucleo di A_2^3 che non appartengono alla base di nessun nucleo della matrice A_2 di potenza 2. Al fine di identificare tale autovettore generalizzato ci scriviamo le basi per i nuclei relativi alle potenze fino al terzo ordine come segue:

$$\mathcal{B}\{N(A_2)\} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathcal{B}\{N(A_2^2)\} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathcal{B}\{N(A_2^3)\} = \begin{bmatrix} -1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In accordo con quanto detto precedentemente possiamo pensare di costruire la catena di autovettori generalizzati prendendo come autovettore generalizzato di ordine 3 il vettore:

$$v_{2,3}^{(2)} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

e costruendo i rimanenti due come segue:

$$v_{2,2}^{(2)} = A_2 v_{2,3}^{(2)} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \quad v_{2,1}^{(2)} = A_2 v_{2,2}^{(2)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Ci calcoliamo ora il vettore $\{v_{2,1}^{(1)}\}$ associato al blocco $J_{2,1}$. Per ottenere questo autovettore possiamo semplicemente prendere un vettore associato al base del nucleo della matrice A_2 che sia linearmente indipendente dalla catena di vettori precedentemente calcolata. Quindi è lecito scegliere:

$$v_{2,1}^{(1)} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Per finire dobbiamo calcolarci l'autovettore relativo al blocco di Jordan $J_{1,1}$. A tale fine possiamo pensare di prendere come autovettore associato a tale blocco uno dei vettori che costituiscono la base del nucleo di A_1 facendo attenzione che tale vettore risulti essere linearmente indipendente dai vettori generalizzati precedentemente calcolati per i due blocchi di Jordan $J_2 = \{J_{2,1}, J_{2,2}\}$ associato al secondo autovalore λ_2 . Ne risulta:

$$v_{1,1}^{(1)} = \mathcal{B}\{N(A_1)\} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Mettiamo ora tutto insieme per ottenere la matrice P per la trasformazione di coordinate:

$$\begin{aligned} P &= \begin{bmatrix} v_{1,1}^{(1)} & v_{2,1}^{(1)} & v_{2,1}^{(2)} & v_{2,2}^{(2)} & v_{2,3}^{(2)} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 \end{bmatrix} \end{aligned}$$

la quale ci garantisce che:

$$J = P^{-1} A P.$$

Esempio 4 Si consideri la seguente matrice $A \in \mathbb{R}^{7 \times 7}$:

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -10 & 2 & -2 & -12 & 0 \\ 0 & 0 & -16 & 2 & -4 & -28 & 0 \\ 0 & 0 & 16 & -4 & 2 & 20 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 9 & -5 \end{bmatrix}$$

il cui polinomio caratteristico risulta essere:

$$p_A(\lambda) = (\lambda - 2)^7$$

da cui si evince che vi è un solo autovalore $\lambda_1 = 2$ con molteplicità algebrica $\mu_1 = 7$.

Andiamo ora ad analizzare la molteplicità geometrica associata a tale autovalore il quale richiede il calcolo della dimensione del nullo della matrice $(A - \lambda_1 I)$, ovvero:

$$\nu_1 = \dim(N(A - \lambda_1 I)) = 4$$

il quale ci informa del fatto che vi saranno quattro blocchi di Jordan associati all'unico autovalore $\lambda_1 = 2$. Siamo a questo punto in grado di conoscere la struttura (a blocchi) della forma canonica di Jordan associata alla matrice A che risulta essere:

$$J = \begin{bmatrix} J_{1,1} & 0 & 0 & 0 \\ 0 & J_{1,2} & 0 & 0 \\ 0 & 0 & J_{1,3} & 0 \\ 0 & 0 & 0 & J_{1,4} \end{bmatrix}$$

A questo punto, poichè la somma del numero delle righe di ogni blocco di Jordan associate allo stesso autovalore deve dare la molteplicità algebrica associata a quell'autovalore, nel caso di λ_1 con $\nu_1 = 4$ e $\mu_1 = 7$ si ha che i quattro blocchi $J_1 = \{J_{1,1}, J_{1,2}, J_{1,3}, J_{1,4}\}$ possono avere un elevato numero di configurazioni.

Tuttavia per risolvere questo problema possiamo far riferimento alla tecnica descritta nel teorema 5.11. Definita la matrice $A_1 = (A - \lambda_1 I)$, si deve verificare che la successione degli elementi d_k , dove ogni elemento è definito come:

$$d_k = \dim(N(A - \lambda_1 I)^k)$$

si stabilizzi, ovvero che sia $d_{h+1} = d_h$. A questo punto, a partire dal calcolo degli elementi della successione si costruisce una tabellina attraverso la quale si determina il grado di ogni blocco di Jordan. Nel nostro caso si ha che la successione si stabilizza per $k = 3$, e la relativa tabellina è quindi la seguente:

$d_1 = \nu_1 = 4$	$d_2 - d_1 = 2$	$d_3 - d_2 = 1$	# Crocette	Dim. blocco
x	x	x	3	$\dim(J_{1,4}) = 3 \times 3$
x	x		2	$\dim(J_{1,3}) = 2 \times 2$
x			1	$\dim(J_{1,2}) = 1 \times 1$
x			1	$\dim(J_{1,1}) = 1 \times 1$

dal quale si evince che i primi due dei quattro blocchi hanno dimensione 1×1 , il terzo ha dimensione 2×2 mentre il quarto ha dimensione 3×3 (si noti l'ordinamento dei blocchi in funzione della dimensione crescente degli stessi). Siamo ora in grado di sviluppare la forma canonica di Jordan associata alla matrice A , la quale risulta essere:

$$J = \begin{bmatrix} -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix}$$

A questo punto va calcolata la matrice di trasformazione P in grado di portare il sistema A nella forma canonica di Jordan J . Sappiamo che vi sono quattro blocchi $J_1 = \{J_{1,1}, J_{1,2}, J_{1,3}, J_{1,4}\}$ associati all'autovalore λ_1 . Per il calcolo degli autovettori generalizzati partiremo dal blocco di dimensione maggiore $J_{1,4}$, ed a seguire determineremo gli autovettori associati agli altri blocchi in ordine decrescente rispetto alla dimensione, quindi $J_{1,3}$, poi $J_{1,2}$ ed infine $J_{1,1}$.

Quindi, per il blocco $J_{1,4} \in \mathbb{R}^{3 \times 3}$ ci costruiamo la catena degli autovettori generalizzati $\{v_{1,1}^{(4)}, v_{1,2}^{(4)}, v_{1,3}^{(4)}\}$ a partire dall'autovettore generalizzato di ordine 3 scelto tra gli elementi della base del nucleo di A_1^3 (matrice A_1 di potenza 3) che non appartengono alla base del nucleo della matrice A_1^2 (matrice A_1 di potenza 3), come dal definizione di autovettore generalizzato (5.6). Per semplificarci le cose scriviamo le basi relative ai nuclei fino alla potenza del terzo ordine della matrice A_1 come segue:

$$\mathcal{B}\{N(A_1)\} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{Bmatrix} \quad \mathcal{B}\{N(A_1^2)\} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{Bmatrix}$$

$$\mathcal{B}\{N(A_1^3)\} = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{Bmatrix}$$

In accordo con la definizione 5.6 possiamo quindi pensare di costruire la catena di autovettori generalizzati prendendo come autovettore generalizzato di ordine 3 relativo al quarto blocco di Jordan il vettore:

$$v_{1,3}^{(4)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

e costruirsi i rimanenti due autovettori della catena come segue:

$$v_{1,2}^{(4)} = A_1 v_{1,3}^{(4)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ -3 \end{bmatrix} \quad v_{1,1}^{(4)} = A_1 v_{1,2}^{(4)} = \begin{bmatrix} 0 \\ 0 \\ 12 \\ 28 \\ -20 \\ 0 \\ 0 \end{bmatrix}$$

Ora, per il blocco $J_{1,3} \in \mathbb{R}^{2 \times 2}$ ci costruiamo la catena degli autovettori generalizzati $\{v_{1,1}^{(3)}, v_{1,2}^{(3)}\}$ a partire dall'autovettore generalizzato di ordine 2 scelto tra gli elementi della base del nucleo di A_1^2 (matrice A_1 di potenza 2) che non appartengono alla base del nucleo della matrice A_1 (matrice A_1 di potenza 1), come dal definizione di autovettore generalizzato (5.6) e che allo stesso tempo sia

linearmente indipendente dalla base associata al blocco $J_{1,4}$ precedentemente individuata. Quindi è lecito scegliere:

$$v_{1,2}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

e calcolarsi l'altro autovettore della catena come segue:

$$v_{1,1}^{(3)} = A_1 v_{1,2}^{(3)} = \begin{bmatrix} 0 \\ 0 \\ -2 \\ -4 \\ 4 \\ 0 \\ 0 \end{bmatrix}$$

A questo punto dobbiamo calcolarci gli autovettori relativi ai blocchi di Jordan $J_{1,2}$, $J_{1,1}$ entrambi di dimensione 1×1 . A tale fine possiamo pensare di prendere come autovettore associato ad ognuno di questi due blocchi uno dei vettori che costituiscono la base del nucleo di A_1 facendo attenzione a che questi risultino essere linearmente indipendenti dai vettori generalizzati precedentemente calcolati per i due blocchi di Jordan $J_{1,3}$, $J_{1,4}$. Per tale ragione, ricordando che:

$$\mathcal{B}\{N(A_1)\} = \left\{ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \right\} \quad [v_{1,1}^{(3)}, v_{1,2}^{(3)}, v_{1,1}^{(4)}, v_{1,2}^{(4)}, v_{1,3}^{(4)}] = \left\{ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 12 & 0 & 0 \\ -4 & 0 & 28 & 0 & 0 \\ 4 & 1 & -20 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -3 & 1 \end{bmatrix} \right\}$$

è lecito scegliere:

$$v_{1,1}^{(2)} = \mathcal{B}\{N(A_1)\} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad v_{1,1}^{(1)} = \mathcal{B}\{N(A_1)\} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Mettiamo ora tutto insieme per ottenere la matrice P per la trasformazione di coordinate:

$$\begin{aligned}
P &= \begin{bmatrix} v_{1,1}^{(1)} & v_{1,1}^{(2)} & v_{1,1}^{(3)} & v_{1,2}^{(3)} & v_{1,1}^{(4)} & v_{1,2}^{(4)} & v_{1,3}^{(4)} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 12 & 0 & 0 \\ 0 & 0 & -4 & 0 & 28 & 0 & 0 \\ 0 & 0 & 4 & 1 & -20 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -3 & 1 \end{bmatrix}
\end{aligned}$$

la quale ci garantisce che:

$$J = P^{-1} A P.$$

5.5 Calcolo $\exp(At)$ attraverso Decomposizione Modale

Nei paragrafi precedenti sono state presentati dei procedimenti per mettere una generica matrice quadrata in forma diagonale o, nel caso questo non fosse possibile, in forma canonica di Jordan. In effetti, avere a disposizione una matrice in forma diagonale (o diagonale a blocchi) semplifica estremamente il calcolo dell'esponenziale di matrice.

5.5.1 Matrici diagonalizzabili

Teorema 5.13 *Si consideri una matrice $A \in \mathbb{R}^{n \times n}$ diagonalizzabile, allora l'esponenziale di matrice ad essa associato può essere scritto come segue:*

$$e^{At} = P e^{\Lambda t} P^{-1} \quad (5.46)$$

dove P è una opportuna matrice di trasformazione di coordinate, e $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ è la matrice con i soli autovalori sulla diagonale.

Dimostrazione: La dimostrazione è una semplice applicazione del Teorema 5.7 e della definizione dell'esponenziale di matrice 3.14. In particolare, poichè la matrice A è diagonalizzabile allora esiste una matrice non singolare P tale che:

$$A = P \Lambda P^{-1}.$$

Si consideri ora l'esponenziale di matrice e^{At} ed il suo sviluppo in serie:

$$e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$$

il quale può essere riscritto come segue:

$$e^{At} = \sum_{k=0}^{\infty} \frac{(P \Lambda P^{-1} t)^k}{k!}$$

ora si noti che per ogni potenza della $P \Lambda P^{-1}$ vale la:

$$(P \Lambda P^{-1})^k = P(\Lambda^k)P^{-1} \quad (5.47)$$

la quale ci permette di riscrivere l'esponenziale di matrice come:

$$\begin{aligned} e^{At} &= \sum_{k=0}^{\infty} \frac{P (\Lambda t)^k P^{-1}}{k!} \\ &= P \left(\sum_{k=0}^{\infty} \frac{(\Lambda t)^k}{k!} \right) P^{-1} \\ &= P e^{\Lambda t} P^{-1} \end{aligned}$$

il quale ci dice che, data una matrice quadrata A simile alla matrice diagonale Λ attraverso la trasformazione P , è sempre possibile effettuare il calcolo del suo esponenziale di matrice in funzione dell'esponenziale della matrice diagonale a cui è simile previa moltiplicazione per la matrice di trasformazione P ad essa associata.

Si ricorda infine che la forma dell'esponenziale di matrice in caso di matrice diagonale è nota e semplice da calcolare come spiegato nel seguente teorema.

Teorema 5.14 (*Esponenziale di matrice per matrici diagonali*) Sia $\Lambda \in \mathbb{R}^{n \times n}$ una matrice diagonale, definita come $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, si ha che l'esponenziale di matrice ad essa associata ha la seguente struttura:

$$e^{\Lambda t} = \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t}) \quad (5.48)$$

Dimostrazione: Per la dimostrazione è sufficiente far riferimento alla definizione di esponenziale di matrice e ricordare lo sviluppo in serie di Taylor dell'esponenziale di uno scalare come segue:

$$\begin{aligned}
e^{\Lambda t} &= \sum_{k=0}^{\infty} \frac{(\Lambda t)^k}{k!} \\
&= \sum_{k=0}^{\infty} \begin{bmatrix} \frac{\lambda_1^k t^k}{k!} & 0 & \dots & \dots & 0 \\ 0 & \frac{\lambda_2^k t^k}{k!} & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & \frac{\lambda_n^k t^k}{k!} \end{bmatrix} \\
&= \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & \dots & 0 \\ 0 & e^{\lambda_2 t} & 0 & \dots & \vdots \\ \vdots & 0 & \ddots & \dots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & e^{\lambda_n t} \end{bmatrix} \\
&= \text{diag} \left(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_n t} \right)
\end{aligned}$$

Teorema 5.15 (*Esponenziale di matrice per coppie di autovalori complessi coniugati*) Sia $A \in \mathbb{R}^{2 \times 2}$ caratterizzata da una coppia di autovalori complessi coniugati del tipo $\lambda = \sigma + j\omega$ e $\bar{\lambda} = \sigma - j\omega$, si ha che l'esponenziale di matrice ad essa associata ha la seguente struttura:

$$e^{At} = e^{\sigma t} \begin{bmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{bmatrix} \quad (5.49)$$

Dimostrazione: Per la dimostrazione è sufficiente far riferimento alla definizione di esponenziale di matrice e ricordare lo sviluppo in serie di Taylor delle funzioni $\sin x$ e $\cos x$. Dal teorema 5.11 sappiamo che la matrice A può essere riscritta (attraverso opportuna trasformazione di coordinate) in una forma che evidenzia la natura oscillatoria della coppia di autovalori complessi e coniugati come segue:

$$\tilde{A} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$

Scomponiamo ora la matrice come segue:

$$\tilde{A} = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} + \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix} = \tilde{A}_1 + \tilde{A}_2$$

al fine di applicare il teorema 3.5 di commutatività dell'esponenziale di matrice, previa verifica che valga la $\tilde{A}_1 \cdot \tilde{A}_2 = \tilde{A}_2 \cdot \tilde{A}_1$. Ora poichè la A_1 è una matrice diagonale la soluzione è data dalla 5.14, mentre per la A_2 effettuiamo lo sviluppo in serie come segue:

$$\begin{aligned}
e^{\tilde{A}_2 t} &= \sum_{k=0}^{\infty} \frac{(\tilde{A}_2 t)^k}{k!} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \omega t \\ -\omega t & 0 \end{bmatrix} + \begin{bmatrix} \frac{-\omega^2 t^2}{2!} & 0 \\ 0 & \frac{-\omega^2 t^2}{2!} \end{bmatrix} + \begin{bmatrix} 0 & \frac{-\omega^3 t^3}{3!} \\ \frac{\omega^3 t^3}{3!} & 0 \end{bmatrix} \\
&+ \begin{bmatrix} \frac{\omega^4 t^4}{4!} & 0 \\ 0 & \frac{\omega^4 t^4}{4!} \end{bmatrix} + \begin{bmatrix} 0 & \frac{\omega^5 t^5}{5!} \\ \frac{-\omega^5 t^5}{5!} & 0 \end{bmatrix} + \dots \\
&= \begin{bmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{bmatrix}
\end{aligned}$$

dove si ricorda che:

$$\begin{aligned}
\sin x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \\
\cos x &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots
\end{aligned}$$

Per cui si ha che:

$$e^{\tilde{A}t} = e^{\tilde{A}_1 t} \cdot e^{\tilde{A}_2 t} = e^{\sigma t} \begin{bmatrix} \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t \end{bmatrix}$$

5.6 Autovalori ed Autospazi

Nelle sezioni precedenti sono stati introdotti i concetti di autovalori ed autovettori, successivamente è stato presentato il problema della diagonalizzazione (e della jordanizzazione) e per finire è stato mostrato il calcolo dell'esponenziale di matrice attraverso la decomposizione modale. In questa sezione, verranno rivisitati in chiave più propriamente sistemistica i concetti di autovalore, autovettore e verrà introdotto il concetto di autospazio.

Si consideri un sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (5.50)$$

con $x(t_0) = x_0$ e $t_0 = 0$ per semplicità, dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$.

Si assuma adesso che esista una trasformazione di coordinate P tale che valga la $A = P \Lambda P^{-1}$. Applicando quindi la trasformazione $x = P z$ il sistema può essere riscritto come segue:

$$\begin{cases} \dot{z}(t) &= \Lambda z(t) + \tilde{B} u(t) \\ y(t) &= \tilde{C} z(t) \end{cases} \quad (5.51)$$

dove $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\tilde{B} = P^{-1}B$ e $\tilde{C} = C P$. Si ricorda che un sistema in forma diagonale presenta delle dinamiche **disaccoppiate**. Inoltre tale forma risulta conveniente anche per analizzare le proprietà strutturali di un sistema quali controllabilità ed osservabilità, come vedremo nel prossimo capitolo. Si assuma ora, per semplicità che il sistema in analisi sia completamente controllabile ed osservabile. La relativa rappresentazione in forma diagonale fornita nella 5.51 ha una corrispettiva rappresentazione a blocchi (Fig. 5.1) di chiara interpretazione.

Tale rappresentazione a blocchi sottolinea la semplicità e allo stesso tempo l'eleganza alla base dei sistemi lineari. Si deve infatti ricordare che qualsiasi sistema interconnesso, anche estremamente complesso, può essere sempre rappresentato attraverso una opportuna trasformazione da un sistema di equazioni differenziali (disaccoppiate) del primo ordine.

La rappresentazione di un sistema attraverso la sua forma diagonale ha inoltre una interpretazione di carattere geometrico estremamente interessante (Fig. 5.2): qualsiasi sia la condizione iniziale delle variabile di stato $\{x_1, x_2, \dots, x_n\}$ (diversa da zero), le dinamiche del sistema evolvono in maniera indipendente su sottospazi, noti col nome di **autospazi**, la cui base è definita dagli autovettori associati ai singoli autovalori. Tali autospazi inoltre sono caratterizzati da una proprietà fondamentale nota col termine di **invarianza**: l'evoluzione di una dinamica x_i associata ad un autovalore λ_i è forzata a rimanere nell'autospazio generato dal corrispettivo autovettore v_i . L' **invarianza** di un autospazio è estremamente importante in quanto, almeno in linea di principio, permette di conoscere a priori l'insieme di tutti i possibili valori che una determinata variabile di stato può assumere. Tale concetto

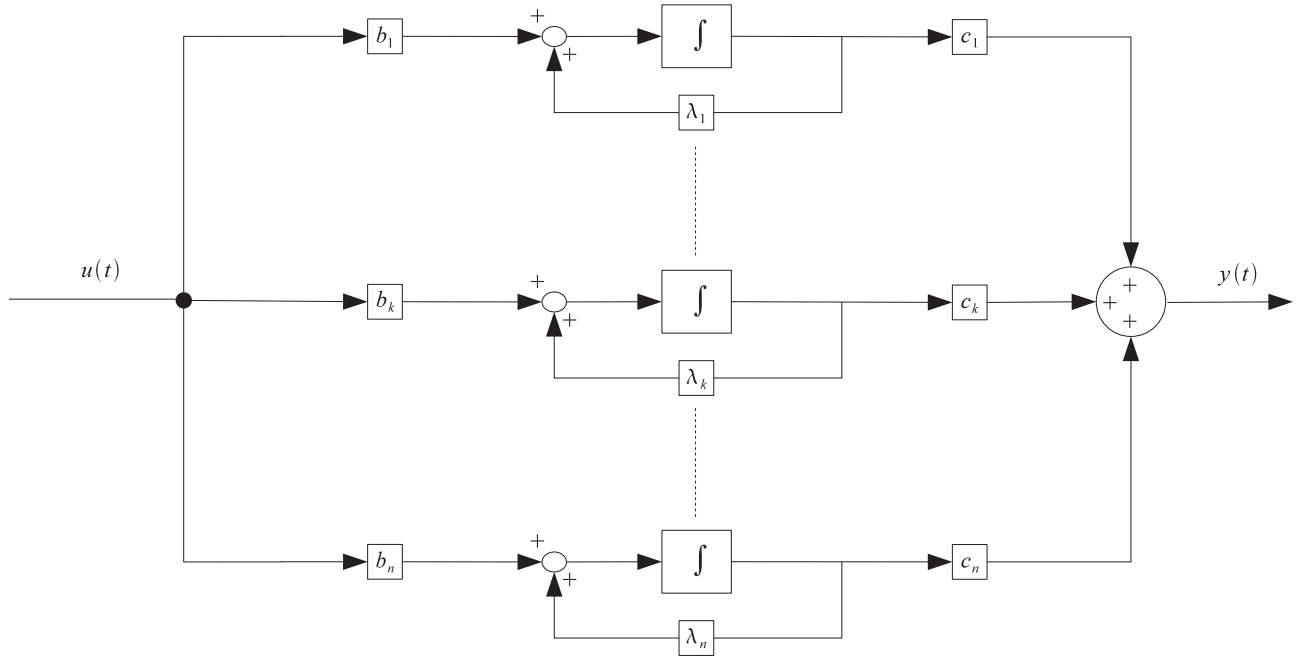


Figura 5.1: Forma Diagonale

può essere generalizzato ad una qualsiasi rappresentazione del sistema dinamico come stabilito dal seguente teorema.

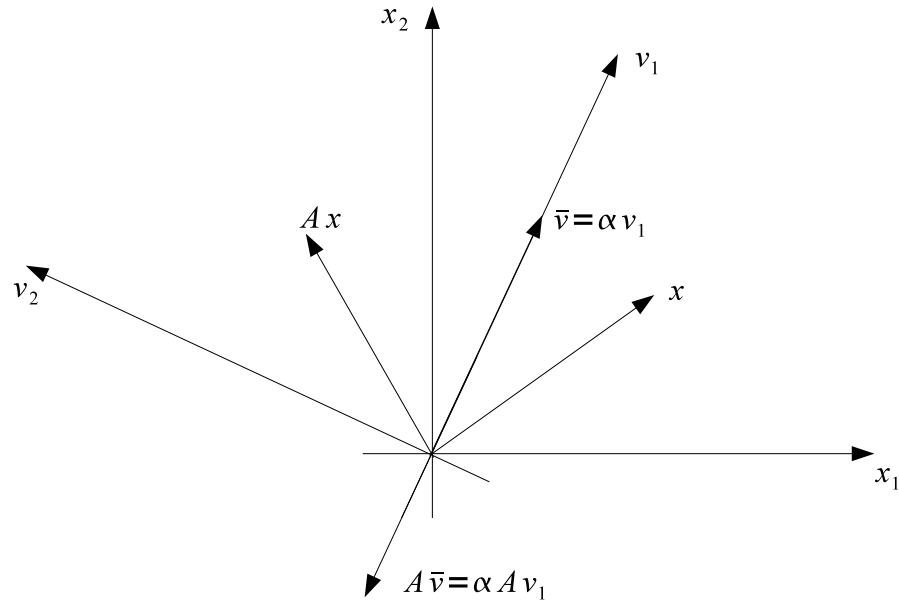


Figura 5.2: Autovettori: Interpretazione geometrica.

Teorema 5.16 (*Invarianza di un Autospatio*) Si consideri un sistema dinamico autonomo del tipo $\dot{x}(t) = Ax(t)$. Siano λ_i e v_i , rispettivamente l' i -esimo autovalore della matrice dinamica A e l'autovettore destro ad esso associato. Se il sistema ha una condizione iniziale allineata all'autovettore v_i ,

allora lo stato $x(t)$ del sistema continuerà ad essere allineato a tale autovettore per la sua evoluzione futura, ovvero:

$$\begin{cases} \dot{x}(t) &= A x(t) \\ x(t_0) &= \alpha v_i \end{cases} \implies x(t) = e^{A(t-t_0)} x(t_0) = \alpha e^{\lambda_i(t-t_0)} v_i \quad (5.52)$$

Dimostrazione: Per la dimostrazione è sufficiente far riferimento alla definizione di esponenziale di matrice e alla definizione di autovettore destro. Si consideri la soluzione del sistema in considerazione:

$$\begin{aligned} x(t) &= e^{A(t-t_0)} x(t_0) \\ &= e^{A(t-t_0)} \alpha v_i \end{aligned}$$

Si assuma per semplicità che $t_0 = 0$ e si applichi la definizione di esponenziale di matrice come segue:

$$\begin{aligned} x(t) &= e^{At} \alpha v_i \\ &= \sum_{k=0}^{\infty} \frac{(At)^k}{k!} \alpha v_i \\ &= \left(I + At + \frac{(At)^2}{2!} + \frac{(At)^3}{3!} + \dots \right) \alpha v_i \\ &= \alpha \left(I v_i + At v_i + \frac{(At)^2}{2!} v_i + \frac{(At)^3}{3!} v_i + \dots \right) \\ &= \alpha \left(v_i + \lambda_i t v_i + \frac{(\lambda_i t)^2}{2!} v_i + \frac{(\lambda_i t)^3}{3!} v_i + \dots \right) \\ &= \alpha e^{\lambda_i t} v_i. \end{aligned}$$

Il teorema 5.16 fornisce una interpretazione dinamica del ruolo degli autovettori destri estremamente interessante: qualunque sia la rappresentazione del sistema dinamico preso in considerazione la scelta di una condizione iniziale $x(t_0)$ che giace nel sottospazio associato ad un dato autovalore vincola l'evoluzione del sistema dinamico a rimanere in tale autospazio.

In maniera del tutto analoga è possibile fornire una interpretazione dinamica del ruolo degli autovettori sinistri come specificato dal seguente teorema.

Teorema 5.17 *Si consideri un sistema dinamico autonomo del tipo $\dot{x}(t) = Ax(t)$. Siano λ_i e v_i , rispettivamente l' i -esimo autovalore della matrice dinamica A e l'autovettore sinistro ad esso associato. Qualunque sia la condizione iniziale $x(t_0)$ vale la:*

$$v_i^T x(t) = e^{\lambda_i t} v_i^T x(t_0) \quad (5.53)$$

Dimostrazione: Per la dimostrazione è sufficiente far riferimento alla definizione di autovettore sinistro ed effettuare una derivazione come segue:

$$\begin{aligned}
\frac{d}{dt} \left(v_i^T x(t) \right) &= v_i^T \dot{x}(t) \\
&= v_i^T A x(t) \\
&= \lambda_i (v_i^T x(t))
\end{aligned}$$

Effettuando ora la sostituzione $z(t) = v_i^T x(t)$ si ottiene la seguente equazione differenziale:

$$\frac{d}{dt} (z(t)) = \lambda_i z(t)$$

la cui soluzione risulta essere:

$$z(t) = e^{\lambda_i t} z(t_0)$$

da cui:

$$v_i^T x(t) = e^{\lambda_i t} v_i^T x(t_0)$$

Si noti che in accordo a quanto enunciato nel Teorema 5.17 si deduce che il ruolo degli autovettori destri è quello di fornire delle funzioni dello stato di semplice computazione qualsiasi sia la condizione iniziale.

5.6.1 Applicazione al problema del Consenso

Lo sviluppo di squadre di agenti autonomi per effettuare le più svariate operazioni da compiti di recupero in seguito a disastri ambientali all'esplorazione distribuita di ambienti ignoti sta negli ultimi anni riscuotendo un elevato interesse da parte della comunità scientifica. Saper raggiungere un accordo su informazioni chiave è un requisito fondamentale per raggiungere la cooperazione in una squadra di agenti autonomi. Tale operazione è resa ancora più complessa dalla necessità di raggiungere tale accordo limitandosi alla sola interazione con i propri vicini, ovvero agenti all'interno del proprio raggio di visibilità. Uno strumento che ha rivoluzionato il modo di sviluppare algoritmi distribuiti è stato l'algoritmo del consenso ("Consensus Algorithm"), introdotto da Richard Murray and Olfati-Saber con la pubblicazione:

R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays", IEEE Trans. Autom. Control, vol. 49, no. 9, pp. 1520–1533, Sep. 2004.

Tale algoritmo si basa su alcuni concetti derivanti dal mondo della Teoria dei Grafi che sono brevemente riportati. Per una trattazione più approfondita di tali concetti il lettore è rimandato ad un qualsiasi testo introduttivo sulla Teoria dei Grafi.

Definizione 5.8 Sia $G = (V, E)$ il grafo indiretto che descrive la topologia della comunicazione per una squadra di n agenti autonomi, dove V è l'insieme dei vertici (agenti) di cardinalità $|V| = n$ e E è l'insieme degli archi (comunicazione tra agenti). Sia inoltre $x = [x_1, \dots, x_n]$ il vettore di stato della squadra di agenti, dove x_i è la variabile di stato associata all' i -esimo agente.

Definizione 5.9 Sia A la matrice di adiacenza associata alla squadra di agenti autonomi descritti attraverso il grafo $G = (V, E)$ dove l'elemento a_{ij} è pari ad uno se la distanza euclidea tra gli agenti i e j è all'interno del reciproco raggio di visibilità, 0 altrimenti.

Definizione 5.10 Sia D la matrice diagonale dei gradi massimi associata alla squadra di agenti autonomi descritti attraverso il grafo $G = (V, E)$ dove l'elemento d_{ii} è pari al numero di vicini associati all'agente i .

Definizione 5.11 Sia M una generica matrice $n \times n$, questa si definisce a diagonale (debolmente) dominante se per ogni riga l'elemento sulla diagonale è maggiore (maggiore uguale) alla somma di tutti gli altri elementi della riga, ovvero:

$$|a_{ii}| > \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{diagonale dominante} \quad (5.54)$$

$$|a_{ii}| \geq \sum_{j=1, j \neq i}^n |a_{ij}| \quad \text{diagonale debolmente dominante} \quad (5.55)$$

Definizione 5.12 Sia $L = D - A$ la matrice laplaciana (Laplaciano) associato al sistema di agenti descritto dal grafo $G = (V, E)$. Si noti che per costruzione il Laplaciano è una matrice simmetrica, a diagonale debolmente dominante le cui righe/colonne sommano a zero. Una diretta conseguenza di tale proprietà è la presenza di almeno un autovalore nullo (il numero di autovalori nulli è pari al numero di componenti connesse) a cui è associato un autovettore destro/sinistro le cui componenti sono tutte pari ad 1.

A questo punto l'algoritmo del consenso può essere enunciato nella sua forma più semplice come segue:

Teorema 5.18 Si consideri una squadra di n agenti autonomi descritta dal grafo indiretto $G = (V, E)$ a cui è associato il Laplaciano L . Si ha che il vettore di stato $x = [x_1, \dots, x_n]$ associato alla dinamica:

$$\dot{x}(t) = -Lx(t) \quad (5.56)$$

raggiunge il consenso, ovvero tutte le variabili di stato x_i dei vari agenti assumono lo stesso valore, se il grafo indiretto G è connesso. In particolare, lo stato di ogni singolo agente converge alla media del vettore degli stati iniziali $x_i = \frac{\sum_{i=1}^n x_i(t_0)}{n}$.

Dimostrazione: La dimostrazione del teorema si divide in due parti e si basa sui risultati presentanti nei Teoremi 5.13 e 5.17. In particolare, è noto dalla Teoria dei Grafi che se un grafo indiretto è (semplicemente) connesso, allora il Laplaciano L ad esso associato ha rango $n - 1$. Inoltre, nel caso di un grafo indiretto connesso, è possibile mostrare attraverso l'impiego del Teorema di Gershgorin, che tutti gli autovalori del Laplaciano sono nel semipiano destro del piano di Gauss. In particolare, vi sarà un solo autovalore nullo a caratterizzare la presenza di una sola componente connessa del grafo.

La prima parte della dimostrazione consiste nel provare la struttura della soluzione $x(t)$:

$$x(t) = \alpha \mathbf{1} \quad t \rightarrow \infty.$$

A tal fine si ricorda che il Laplaciano può essere diagonalizzato come segue:

$$L = P \Lambda P^{-1}$$

dove $\Lambda = \text{diag}(\lambda_1 = 0, \lambda_2, \dots, \lambda_n)$ è la matrice con i soli autovalori sulla diagonale e $P = [v_1, \dots, v_n]$ è la matrice degli autovettori destri di L . A partire da questa relazione in accordo al Teorema 5.13 è possibile scrivere la seguente espressione per l'esponenziale di matrice:

$$e^{-Lt} = P e^{-\Lambda t} P^{-1}$$

dove l'esponenziale di matrice per la matrice diagonale degli autovalori ha la seguente struttura:

$$e^{-\Lambda t} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & e^{-\lambda_2 t} & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & e^{-\lambda_n t} \end{bmatrix}$$

la quale per $t \rightarrow \infty$ diventa:

$$e^{-\Lambda t} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix}$$

a questo punto ricordando la struttura dell'autovettore destro $v_{d,1} = \sqrt{\bar{\alpha}} \mathbf{1}$ e ricordando che ogni matrice simmetrica può essere diagonalizzata attraverso una matrice unitaria¹, si si ottiene:

$$\begin{aligned} e^{-Lt} &= V e^{-\Lambda t} V^{-1} \\ &= \sqrt{\bar{\alpha}} \begin{bmatrix} 1 & * & \dots & * \\ 1 & * & \dots & * \\ \vdots & & \ddots & \vdots \\ 1 & \dots & \dots & * \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & 0 \end{bmatrix} \cdot \sqrt{\bar{\alpha}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ * & * & \dots & * \\ \vdots & & \ddots & \vdots \\ * & \dots & \dots & * \end{bmatrix} \\ &= \bar{\alpha} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & & \ddots & \vdots \\ 1 & \dots & \dots & 1 \end{bmatrix} \\ &= \bar{\alpha} \mathbf{1} \cdot \mathbf{1}^T = \bar{\alpha} v_{d,0} \cdot v_{s,0}^T. \end{aligned}$$

¹Una matrice unitaria è una matrice ortogonale la cui inversa è semplicemente la trasposta, ovvero U tale che $U \cdot U^T = I$.

Ne consegue che qualunque sia la condizione iniziale la soluzione sarà sempre del tipo:

$$x(t) = \bar{\alpha} v_{d,0} \cdot v_{s,0}^T \cdot x(t_0) = \alpha \mathbf{1}.$$

La seconda parte della dimostrazione consiste nel provare che il valore per il quale è raggiunto il consenso è pari alla media del vettore delle condizioni iniziali $x(t_0) = [x_1(t_0), \dots, x_n(t_0)]$. A tal scopo è sufficiente utilizzare il risultato introdotto dal Teorema 5.17, rispetto all'autovalore $\lambda_1 = 0$ ed il corrispettivo autovettore sinistro $v_{s,1} = v_1 = \mathbf{1}$:

$$\begin{aligned} v_1^T x(t) &= e^{\lambda_1 t} v_1^T x(t_0) \\ \mathbf{1}^T \alpha \mathbf{1} &= \mathbf{1}^T x(t_0) \\ \alpha \mathbf{1}^T \mathbf{1} &= \mathbf{1}^T x(t_0) \\ \alpha n &= \sum_{i=1}^n x_i(t_0) \\ \alpha &= \frac{\sum_{i=1}^n x_i(t_0)}{n}. \end{aligned}$$

Da cui si evince che:

$$x(t) = \frac{\sum_{i=1}^n x_i(t_0)}{n} \mathbf{1} \quad t \rightarrow \infty.$$

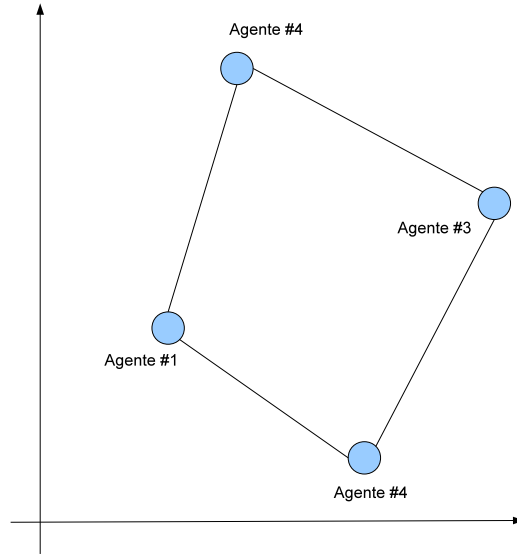


Figura 5.3: Grafo di comunicazione per una squadra di 4 agenti.

Esempio 5 Si consideri una squadra di 4 agenti interconnessi come indicato in Fig. 5.3. Si assuma che gli agenti debbano decidere una locazione nello spazio bidimensionale dove incontrarsi. A tal fine si può pensare che ogni agente proponga la propria locazione come il punto di “rendez-vous”. Tale problema può essere modellato come un consenso su due variabili, ovvero rispetto agli assi x ed y del piano preso in considerazione. Si consideri quindi la seguente matrice $A_0 \in \mathbb{R}^{n \times 2}$ di condizioni iniziali:

$$A_0 = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 5 & 1 \\ 7 & 5 \\ 3 & 7 \end{bmatrix}$$

Si consideri ora la matrice di adiacenza associata al grafo di comunicazione descritto in Fig. 5.3:

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

da cui si deduce facilmente la matrice diagonale dei gradi massimi:

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

da cui si può calcolare il Laplaciano associato al grafo di comunicazione descritto in Fig. 5.3 come segue:

$$L = D - A = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & 2 \end{bmatrix}$$

A questo punto applicando il Teorema del Consensus 5.18 si può costruire una coppia di sistemi dinamici associati ai due assi x e y come segue:

$$\dot{x} = -Lx$$

$$\dot{y} = -Ly$$

Inoltre sempre in accordo al Teorema 5.18 sappiamo che per t sufficientemente grande (teoricamente $t \rightarrow \infty$) i vettori di stato x and y giacciono nel sottospazio $\text{span}\{\mathbf{1}\}$. In particolare, per la conservazione della somma delle condizioni iniziali indotta dal Teorema 5.17 si ha che:

$$\begin{aligned} \alpha_x &= \frac{\sum_{i=1}^4 x_i}{4} = \frac{2+5+7+3}{4} = 4.25 \\ \alpha_y &= \frac{\sum_{i=1}^4 y_i}{4} = \frac{3+1+5+7}{4} = 4 \end{aligned}$$

Da cui si ottiene che il vettore di stato $A(t)$ per t sufficientemente grande è pari ad:

$$A(t) = \begin{bmatrix} x_1(t) & y_1(t) \\ x_2(t) & y_2(t) \\ x_3(t) & y_3(t) \\ x_4(t) & y_4(t) \end{bmatrix} = \begin{bmatrix} 4.25 & 4 \\ 4.25 & 4 \\ 4.25 & 4 \\ 4.25 & 4 \end{bmatrix}.$$

In altri termini, gli agenti hanno raggiunto un consenso su un punto nello spazio bidimensionale dove incontrarsi: il punto di “rendez-vous” è pari a $p_{rv} = [4.25, 4]$. È importante ricordare, che l’algoritmo in questione permette di raggiungere il consenso limitando la collaborazione alla sola interazione locale tra vicini.

Capitolo 6

Proprietà Strutturali

6.1 Controllabilità

Si consideri un sistema dinamico descritto dal seguente sistema di equazioni differenziali lineari a coefficienti costanti:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (6.1)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$.

Definizione 6.1 Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema dinamico descritto in 6.1 si dice **controllabile**, se $\forall T \geq 0 \exists u(t)$ in grado di guidare il sistema dallo stato iniziale $x(0) = \bar{x}$ allo stato finale $x(T) = 0$. Inoltre, se $u(t)$ esiste allora non è unico

Definizione 6.2 Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema dinamico descritto in 6.1 si dice **raggiungibile**, se $\forall T \geq 0 \exists u(t)$ in grado di guidare il sistema dallo stato iniziale $x(0) = 0$ allo stato finale $x(T) = \bar{x}$.

Tali concetti di controllabilità e raggiungibilità definiti per un singolo stato $\bar{x} \in \mathbb{R}^n$ possono essere estesi all'intero sistema come segue.

Definizione 6.3 Un sistema si dice (completamente) controllabile se ogni stato è controllabile.

Definizione 6.4 Un sistema si dice (completamente) raggiungibile se ogni stato è raggiungibile.

È importante sottolineare come tali proprietà siano **strutturali** del sistema. Con il termine strutturale si intende che queste non variano in seguito a manipolazione sul sistema quali ad esempio una trasformazione di coordinate.

Teorema 6.1 Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema dinamico descritto in 6.1 si dice controllabile se e solo se:

$$\bar{x} \in R(W_C(0, T)) \quad (6.2)$$

dove $W_C(0, T)$ (matrice simmetrica $n \times n$) è il **Gramiano di Controllabilità** ed è definito come segue:

$$W_C(0, T) = \int_0^T e^{-A\tau} B B^T e^{-A^T \tau} d\tau \quad (6.3)$$

È ovvio a questo punto vedere che affinché il sistema risulti completamente controllabile il gramiano di controllabilità deve essere non singolare. In altri termini deve essere $\text{rank}\{W_C(0, T)\} = n$.

Dimostrazione sufficienza: Per dimostrare la sufficienza assumiamo di avere $\bar{x} \in R(W_C(0, T))$, e proviamo l'esistenza di un controllo $u(t)$ in grado di guidare il sistema dallo stato $x(0) = \bar{x}$ allo stato $x(T) = 0$.

Ora, poichè $\bar{x} \in R(W_C(0, T))$ esisterà un $z \in \mathbb{R}^n$ tale che:

$$W_C(0, T)z = \bar{x}. \quad (6.4)$$

A questo punto è sufficiente definire un controllo $u(t)$ come segue:

$$u(t) = -B^T e^{-A^T t} z \quad (6.5)$$

per guidare il sistema dallo stato iniziale $x(0) = \bar{x}$ allo stato finale $x(T) = 0$. Infatti si ha:

$$x(T) = e^{AT} \bar{x} + \int_0^T e^{A(T-\tau)} B u(\tau) d\tau \quad (6.6)$$

da cui sostituendo la $u(t)$ come definito in 6.5 si ottiene la:

$$x(T) = e^{AT} \bar{x} - \int_0^T e^{A(T-\tau)} B B^T e^{-A^T \tau} z d\tau \quad (6.7)$$

la quale per la 6.4 dà la seguente:

$$x(T) = e^{AT} \bar{x} - e^{AT} \bar{x} = 0 \quad (6.8)$$

che prova la tesi. ‡

Dimostrazione necessità: Per dimostrare la necessità assumiamo di avere $\bar{x} \notin R(W_C(0, T))$, e di essere in grado di trovare un controllo $u(t)$ tale da guidare il sistema dallo stato $x(0) = \bar{x}$ allo stato $x(T) = 0$. Questo ci permetterà di mostrare l'assurdo per il quale deve essere $\bar{x} \in R(W_C(0, T))$

Infatti, se $\bar{x} \notin R(W_C(0, T))$ vale anche la $\bar{x} \notin N(W_C(0, T))$. Di conseguenza esisterà un $w \in N(W_C(0, T))$ tale che:

$$w^T \bar{x} \neq 0. \quad (6.9)$$

Ora applicando il controllo $u(t)$ si dovrebbe avere:

$$x(T) = e^{AT} \bar{x} + \int_0^T e^{A(T-\tau)} B u(\tau) d\tau = 0 \quad (6.10)$$

Moltiplicando ambo i membri per $w^T e^{-AT}$ si ottiene:

$$0 = w^T x + \int_0^T w^T e^{A(-\tau)} B u(\tau) d\tau \quad (6.11)$$

Dalla 6.9 e 6.11 si deduce che:

$$\int_0^T w^T e^{A(-\tau)} B u(\tau) d\tau \neq 0. \quad (6.12)$$

Tuttavia poichè $w \in N(W_c(0, T))$ si ha che $W_c(0, T)w = 0$, e di conseguenza:

$$0 = w^T W_c(0, T)w = \int_0^T w^T e^{-A\tau} B B^T e^{-A^T \tau} w d\tau \quad (6.13)$$

ovvero:

$$0 = \int_0^T \|(w^T e^{-A\tau} B)\|^2 d\tau \quad (6.14)$$

dal quale si evince che deve essere:

$$w^T e^{-A\tau} B = 0 \quad (0 < \tau \leq T). \quad (6.15)$$

che contraddice la 6.12. Quindi deve essere $\bar{x} \in R(W_C(0, T))$. \ddagger

Lemma 6.1 *Si assuma che $W_C(0, T)$ sia non singolare, allora la funzione di controllo $\bar{u}(t)$ definita come segue:*

$$\bar{u}(t) = -B^T e^{-A^T t} W_C^{-1}(0, T) \bar{x} \quad (6.16)$$

porta il sistem descritto dalla 6.1 dallo stato $x(0) = \bar{x}$ allo stato $x(T) = 0$.

Dimostrazione: Per la dimostrazione è sufficiente scrivere l'evoluzione completa del sistema, sostituire ad essa l'ingresso $\bar{u}(t)$ definito nella 6.16 e calcolarsi il valore dello stato $x(t)$ per $t = T$. Si ha quindi:

$$\begin{aligned} x(T) &= e^{AT} \bar{x} + \int_0^T e^{A(T-\tau)} B \bar{u}(\tau) d\tau \\ &= e^{AT} \bar{x} - \int_0^T e^{A(T-\tau)} B B^T e^{-A^T \tau} W_C^{-1}(0, T) \bar{x} d\tau \\ &= e^{AT} \bar{x} - e^{AT} \int_0^T e^{-A\tau} B B^T e^{-A^T \tau} W_C^{-1}(0, T) d\tau \bar{x} \\ &= e^{AT} \bar{x} - e^{AT} W_C(0, T) W_C^{-1}(0, T) \bar{x} \\ &= e^{AT} \bar{x} - e^{AT} \bar{x} = 0. \end{aligned}$$

Tuttavia da un punto di vista applicativo l'utilizzo del Gramiano di Controllabilità risulta essere poco agevole. Per questo motivo si introduce ora una tecnica alternativa per la verifica della controllabilità di un sistema che si basa sulla definizione della **matrice di controllabilità** P_C e sull'analisi del relativo rango.

Teorema 6.2 *Definiamo la matrice di controllabilità $P_C \in \mathbb{R}^{n \times np}$ associata al sistema 6.1 come segue:*

$$P_C = [B \ AB \ A^2B \ \cdots \ A^{n-1}B] \quad (6.17)$$

Si ha che il sottospazio degli stati controllabili è:

$$R(W_C(0, T)) = R(P_C), \quad (6.18)$$

mente il sottospazio degli stati non controllabili è:

$$N(W_C(0, T)^T) = N(W_C(0, T)) = N(P_C^T) = \begin{bmatrix} B^T \\ B^T A^T \\ B^T (A^T)^2 \\ \vdots \\ B^T (A^T)^{n-1} \end{bmatrix} \quad (6.19)$$

per ogni $T > 0$. Ne risulta che affinché il sistema risulti completamente controllabile, la matrice di controllabilità deve avere rango pieno, ovvero:

$$\text{rank}(P_C) = n. \quad (6.20)$$

Dimostrazione: Ricordando la proprietà di simmetria del Gramiano di Controllabilità $W_c = W_c^T$, e le proprietà di decomposizione delle trasformazioni lineari si ha che:

$$R(W_C)^\perp = N(W_C^T) = N(W_C) \quad R(P_C)^\perp = N(P_C^T)$$

Si deduce che le due relazioni 6.18 e 6.19 sono equivalenti, per questa ragione è sufficiente provare solo la 6.19.

Si dimostrerà dapprima che $N(W_C) \subseteq N(P_C^T)$ e successivamente che $N(W_C) \supseteq N(P_C^T)$ il quale prova l'identità $N(W_C) = N(P_C^T)$.

$$\Rightarrow N(W_C) \subseteq N(P_C^T).$$

Si consideri $z \in N(W_C)$, quindi $W_c z = 0$, e di conseguenza si ha:

$$\begin{aligned} z^T W_c z &= \int_0^T \left(z^T e^{-A\tau} B \right) \left(B^T e^{-A^T \tau} z \right) d\tau \\ &= \int_0^T \left\| B^T e^{-A^T \tau} z \right\|^2 d\tau = 0. \end{aligned}$$

Ora, poichè l'integrale è continuo questo implica che la funzione debba essere identicamente nulla, ovvero:

$$B^T e^{-A^T \tau} z = B^T \left(I - (A^T) \tau + \frac{A^{T^2} \tau^2}{2!} - \frac{A^{T^3} \tau^3}{3!} + \dots \right) z = 0 \quad (0 < \tau \leq T)$$

Ora, poichè lo sviluppo in serie è identicamente nullo su un intervallo finito, ogni termine deve essere nullo, ovvero:

$$B^T z = B^T A^T z = B^T A^{T^2} z = \dots = 0$$

Il quale può essere riscritto come:

$$P_C^T z = \begin{bmatrix} B^T z \\ B^T A^T z \\ B^T (A^T)^2 z \\ \vdots \\ B^T (A^T)^{n-1} z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Il che dimostra che $z \in N(P_C^T)$.

$$\Leftarrow N(W_C) \supseteq N(P_C^T).$$

Si consideri $z \in N(P_C^T)$, quindi $P_C^T z = 0$, e di conseguenza si ha:

$$B^T z = B^T A^T z = B^T A^{T^2} z = \dots = B^T A^{T^{n-1}} z = 0$$

Ora per il teorema di Cayley-Hamilton noi sappiamo che ogni potenza della matrice A di ordine $\geq n$, dove $n \times n$ è la dimensione della matrice, può essere espressa come combinazione lineare delle prime n potenze. Se ne deduce quindi che la precedente relazione vale anche per ogni potenza di ordine $\geq n$, ovvero:

$$B^T A^{T^n} z = B^T A^{T^{n+1}} z = \dots = B^T A^{T^{n+k}} z = \dots = 0$$

Di conseguenza si può scrivere la:

$$\begin{aligned} B^T e^{-A^T \tau} z &= B^T \left[\sum_{i=0}^{\infty} \frac{(-A^T \tau)^i}{i!} \right] z \\ &= \sum_{i=0}^{\infty} \frac{((-I)^i (B^T A^{T^i} z) \tau^i)}{i!} = 0 \end{aligned}$$

da cui si deduce che:

$$W_c z = \int_0^T e^{-A\tau} B (B^T e^{-A^T \tau} z) d\tau = 0$$

Il che dimostra che $z \in N(W_C^T)$.

Con i teoremi 6.1 e 6.2 sono state presentate due tecniche per la verifica della controllabilità di un sistema. In particolare, è stato messo in evidenza come lo spazio degli stati possa essere decomposto nel sottospazio degli stati controllabili e nel sottospazio degli stati non controllabili, ovvero:

$$\mathbb{R}^n = R(W_C(0, T)) \oplus N(W_C(0, T)) \quad (6.21)$$

$$= R(P_C) \oplus N(P_C^T) \quad (6.22)$$

Tale decomposizione comporta che ogni stato $x \in \mathbb{R}^n$ può essere espresso in modo univoco attraverso due componenti:

$$x = x_C + x_{NC} \quad x_C \perp x_{NC} \quad (6.23)$$

dove $x_C \in R(W_C(0, T))$ e $x_{NC} \in N(W_C(0, T))$ sono dette rispettivamente componente controllabile e componente non controllabile.

Lemma 6.2 *Si consideri il sistema descritto dalla 6.1, se lo stato iniziale $x(0) = \bar{x}$ è controllabile, allora lo stato:*

$$x(t) = e^{At}\bar{x} + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau \quad (6.24)$$

resta controllabile per tutto il tempo, a prescindere da quale sia la forzante $u(t)$ applicata in ingresso.

Dimostrazione Da inserire (forse).

6.2 Osservabilità

Si consideri un sistema dinamico descritto dal seguente sistema di equazioni differenziali lineari a coefficienti costanti:

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{cases} \quad (6.25)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$.

Definizione 6.5 *Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema descritto dalla 6.25 si dice **inosservabile** se posto $u(t) = 0$ e $\forall T > 0$, lo stato iniziale $x(0) = \bar{x}$ produce una risposta nulla in uscita:*

$$y(t) = 0 \quad (0 \leq t \leq T). \quad (6.26)$$

Definizione 6.6 Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema descritto dalla 6.25 si dice **non ricostruibile** se l'uscita si annulla per $-T \leq t \leq 0$.

Si noti che il concetto di osservabilità e ricostruibilità sono completamente equivalenti per sistemi lineari tempo invariante. La distinzione tra le due definizioni diviene importante nel caso si vada a considerare sistemi tempo variante.

Definizione 6.7 Un sistema si dice (completamente) osservabile se non esiste nessuno stato, eccetto lo stato identicamente nullo $x(0) = 0$, inosservabile.

Definizione 6.8 Un sistema si dice (completamente) ricostruibile se non esiste nessuno stato, eccetto lo stato identicamente nullo $x(0) = 0$, non ricostruibile.

Teorema 6.3 Un generico stato $\bar{x} \in \mathbb{R}^n$ del sistema descritto dalla 6.25 è inosservabile se e solo se:

$$\bar{x} \in N(W_O(0, T)) \quad (6.27)$$

dove $W_O(0, T)$ (matrice simmetrica $n \times n$) è il **Gramiano di Osservabilità** ed è definito come segue:

$$W_O(0, T) = \int_0^T e^{A^T \tau} C^T C e^{A \tau} d\tau \quad (6.28)$$

È ovvio a questo punto vedere che affinché il sistema risulti completamente osservabile il gramiano di osservabilità deve essere non singolare. In altri termini deve essere $\text{rank}\{W_O(0, T)\} = n$.

Dimostrazione: Si dimostrerà dapprima che se $\bar{x} \in N(W_O(0, T))$ allora lo stato \bar{x} è non osservabile, successivamente si dimostrerà che se lo stato \bar{x} è non osservabile allora $\bar{x} \in N(W_O(0, T))$ il quale prova la tesi.

\Rightarrow

Si assuma $\bar{x} \in N(W_O(0, T))$, questo implica che $W_O(0, T) \bar{x} = 0$, da cui:

$$\begin{aligned} 0 = \bar{x}^T W_O(0, T) \bar{x} &= \int_0^T \bar{x}^T e^{A^T \tau} C^T C e^{A \tau} \bar{x} d\tau \\ 0 &= \int_0^T \|C e^{A \tau} \bar{x}\|^2 d\tau \end{aligned}$$

Il quale per la continuità dell'integrale implica che:

$$y(\tau) = C e^{A \tau} \bar{x} = 0 \quad (0 \leq \tau \leq T)$$

che prova la tesi di inosservabilità per lo stato $\bar{x} \in \mathbb{R}^n$. ‡

\Leftarrow

Si assuma $\bar{x} \in \mathbb{R}^n$ inosservabile. Per l'enunciato del teorema si ha quindi che:

$$y(\tau) = C e^{A\tau} \bar{x} = 0 \quad (0 \leq \tau \leq T)$$

Dalla quale si può facilmente vedere che:

$$W_O(0, T) \bar{x} = \int_0^T e^{A^T \tau} C^T C e^{A\tau} \bar{x} d\tau = 0$$

Il quale prova che $\bar{x} \in N(W_O(0, T))$. ‡

Lemma 6.3 *Si assuma che $W_O(0, T)$ sia non singolare, che l'ingresso del sistema $u(t)$ sia nullo e che lo stato iniziale $x(0) = \bar{x}$ sia sconosciuto. Allora \bar{x} può essere ricostruito processando l'uscita $y(t)$ del sistema come segue:*

$$\bar{x} = \int_0^T W_O^{-1}(0, T) e^{A^T \tau} C^T y(\tau) d\tau \quad (6.29)$$

Dimostrazione Per la dimostrazione basta riscrivere la 6.29 ricordando che l'uscita del sistema può essere espressa come $y(t) = C e^{A^T t} \bar{x}$ per ottenere quanto segue:

$$\begin{aligned} \int_0^T W_O^{-1}(0, T) e^{A^T \tau} C^T C e^{A\tau} \bar{x} d\tau &= W_O^{-1}(0, T) \int_0^T e^{A^T \tau} C^T C e^{A\tau} \bar{x} d\tau \\ &= W_O^{-1}(0, T) W_O(0, T) \bar{x} = \bar{x} \end{aligned}$$

Tuttavia, come nel caso della controllabilità, da un punto di vista applicativo l'utilizzo del Gramiano di Osservabilità risulta essere poco agevole. Per questo motivo si introduce ora una tecnica alternativa per la verifica della osservabilità di un sistema che si basa sulla definizione della **matrice di osservabilità** P_O e sull'analisi del relativo rango.

Teorema 6.4 *Definiamo la matrice di osservabilità $P_O \in \mathbb{R}^{nq \times n}$ associata al sistema 6.25 come segue:*

$$P_O = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (6.30)$$

Si ha che il sottospazio degli stati non osservabili è:

$$N(W_O(0, T)) = N(P_O) \quad (6.31)$$

mentre il sottospazio degli stati osservabili è:

$$R(W_O(0, T)) = R(P_O^T) = R[C^T \ A^T C^T \ (A^T)^2 C^T \ \dots \ (A^T)^{n-1} C^T] \quad (6.32)$$

per ogni $T > 0$. Ne risulta che affinché il sistema risulti completamente osservabile la matrice di osservabilità deve avere rango pieno, ovvero:

$$\text{rank}(P_O) = n \quad (6.33)$$

Si noti che la matrice di osservabilità non dipende dal tempo T . Questo significa, in accordo col lemma 6.3, che almeno in linea di principio lo stato di un sistema osservabile può essere determinando processando l'uscita del sistema per qualsiasi intervallo $[0, T]$. Occorre tuttavia fare attenzione al fatto che la scelta di un intervallo di integrazione troppo piccolo potrebbe portare a **problemi di fisica realizzabilità**.

Dimostrazione Ricordando le proprietà di decomposizione delle trasformazioni lineari si ha che:

$$R(W_O)^\perp = N(W_O^T) = N(W_O) \quad R(P_O^T)^\perp = N(P_O)$$

Si deduce che le due relazioni 6.31 e 6.32 sono equivalenti, per questa ragione è sufficiente provare solo la 6.31.

Si dimostrerà dapprima che $N(W_O) \subseteq N(P_O)$ e successivamente che $N(W_O) \supseteq N(P_O)$ il quale prova l'identità $N(W_O) = N(P_O)$.

$$\Rightarrow N(W_O) \subseteq N(P_O).$$

Si consideri $z \in N(W_O)$, quindi $W_O z = 0$, e di conseguenza si ha:

$$\begin{aligned} z^T W_O z &= \int_0^T \left(z^T e^{A^T \tau} C^T \right) \left(C e^{A \tau} z \right) d\tau \\ &= \int_0^T \left\| C e^{A \tau} z \right\|^2 d\tau = 0. \end{aligned}$$

Ora, poichè l'integrale è continuo questo implica che la funzione debba essere identicamente nulla, ovvero:

$$C e^{A \tau} z = C \left(I + A \tau + \frac{A^2 \tau^2}{2!} + \frac{A^3 \tau^3}{3!} + \dots \right) z = 0 \quad (0 < \tau \leq T)$$

Ora, poich'è lo sviluppo in serie è identicamente nullo su un intervallo finito, ogni termine deve essere nullo, ovvero:

$$C z = C A z = C A^2 z = \dots = C A^{n-1} z = 0$$

Il quale può essere riscritto come:

$$P_O z = \begin{bmatrix} C z \\ CA z \\ CA^2 z \\ \vdots \\ CA^{n-1} z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Il che dimostra che $z \in N(P_O)$.

‡

$\Leftarrow N(W_O) \supseteq N(P_O)$.

Si consideri $z \in N(P_O)$, quindi $P_O z = 0$, e di conseguenza si ha:

$$C z = CA z = CA^2 z = \dots = CA^{n-1} z = 0$$

Ora per il teorema di Cayley-Hamilton noi sappiamo che ogni potenza della matrice A di ordine $\geq n$, dove $n \times n$ è la dimensione della matrice, può essere espressa come combinazione lineare delle prime n potenze. Se ne deduce quindi che la precedente relazione vale anche per ogni potenza di ordine $\geq n$, ovvero:

$$CA^n z = CA^{n+1} z = \dots = CA^{n+k} z = \dots = 0$$

Di conseguenza si può scrivere la:

$$\begin{aligned} C e^{A\tau} z &= C \left[\sum_{i=0}^{\infty} \frac{(A\tau)^i}{i!} \right] z \\ &= \sum_{i=0}^{\infty} \frac{((CA^i z) \tau^i)}{i!} = 0 \end{aligned}$$

da cui si deduce che:

$$W_O z = \int_0^T e^{A^T \tau} C^T (C e^{A\tau} z) d\tau = 0$$

Il che dimostra che $z \in N(W_O^T)$.

‡

Lemma 6.4 *Si consideri il sistema descritto dalla 6.25, se lo stato iniziale $x(0) = \bar{x}$ è inosservabile e l'ingresso $u(t)$ è identicamente nullo, allora lo stato:*

$$x(t) = e^{A^t} \bar{x} \tag{6.34}$$

rimane inosservabile per tutto il tempo.

Dimostrazione Da inserire (forse).

6.3 Dualità

Si può facilmente scorgere una certa similitudine tra le condizioni di controllabilità ed osservabilità mostrate. In effetti, come verrà formalizzato nel teorema a seguire vi è una **dualità** tra questi due concetti.

Teorema 6.5 (*Dualità*) *Il sottospazio di controllabilità del sistema:*

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (6.35)$$

è identico al sottospazio di osservabilità del sistema:

$$\begin{cases} \dot{x}(t) = -A^T x(t) + C^T u(t) \\ y(t) = B^T x(t) \end{cases} \quad (6.36)$$

Si noti che similmente il sottospazio di osservabilità del sistema 6.35 è identico al sottospazio di controllabilità del sistema 6.36. Inoltre, si noti che il teorema vale anche nel caso $-A^T$ sia rimpiazzato con A^T nella 6.36.

Dimostrazione La dimostrazione è una semplice conseguenza dalla definizione dei graminiani di controllabilità e osservabilità, come segue:

$$\begin{aligned} W_{C,1}(0, T) &= \int_0^T e^{-A\tau} B B^T e^{-A^T \tau} d\tau = W_{O,2}(0, T) \\ W_{O,1}(0, T) &= \int_0^T e^{A^T \tau} C^T C e^{A \tau} d\tau = W_{C,2}(0, T) \end{aligned}$$

Un altro aspetto di estremo interesse è il fatto che tali proprietà risultano essere **strutturali** in quanto esse persistono anche quando si decide di adottare una rappresentazione alternativa (ma equivalente) del sistema in esame. D'altronde è logico aspettarsi che il fatto di servirsi di una rappresentazione alternativa non possa modificare la sostanza del sistema in gioco. Tale concetto è formalizzato dal seguente teorema.

Teorema 6.6 *Il sistema:*

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (6.37)$$

è controllabile (osservabile) se e solo se il sistema:

$$\begin{cases} \dot{\hat{x}}(t) = Q A Q^{-1} \hat{x}(t) + Q B u(t) \\ y(t) = C Q^{-1} \hat{x}(t) \end{cases} \quad (6.38)$$

è controllabile (osservabile), dove $Q \in \mathbb{R}^{n \times n}$ è una qualsiasi matrice non singolare tale che $\hat{x} = Qx$. Inoltre, il sottospazio di controllabilità (osservabilità) ha la stessa dimensione per entrambe le rappresentazioni.

Dimostrazione:

I) *Controllabilità*

Per la dimostrazione si fa riferimento alla definizione del gramiano di controllabilità e ad una nota proprietà di trasposizione del prodotto di matrici. Si scriva il gramiano di controllabilità relativo al sistema 6.38:

$$\begin{aligned}
 W_{C,2} &= \int_0^T e^{-(Q A Q^{-1})\tau} Q B (Q B)^T e^{-(Q A Q^{-1})^T \tau} d\tau \\
 &= \int_0^T e^{-(Q A Q^{-1})\tau} Q B (Q B)^T e^{-(Q^{-1T} A^T Q^T)\tau} d\tau \\
 &= \int_0^T Q e^{-A\tau} Q^{-1} Q B B^T Q^T Q^{-1T} e^{-A^T \tau} Q^T d\tau \\
 &= \int_0^T Q e^{-A\tau} B B^T e^{-A^T \tau} Q^T d\tau \\
 &= Q \left(\int_0^T e^{-A\tau} B B^T e^{-A^T \tau} d\tau \right) Q^T \\
 &= Q W_{C,1} Q^T
 \end{aligned}$$

dove $(Q^T)^{-1} = (Q^{-1})^T$ e con $W_{C,1}$ e $W_{C,2}$ si intende rispettivamente il gramiano di controllabilità associata al sistema 6.37 e 6.38. Ora si ricordi che il rango del prodotto di due matrici $A \in \mathbb{R}^{n \times m}$ e $B \in \mathbb{R}^{m \times k}$ rispetta la seguente regola:

$$\text{rank}(AB) \leq \min(\{\text{rank}(A), \text{rank}(B)\})$$

dove in particolare si ha che:

- a) Se B ha rango m allora il prodotto AB ha lo stesso rango di A
- a) Se A ha rango m allora il prodotto AB ha lo stesso rango di B

Nel caso in analisi le matrici in gioco sono quadrate ed in particolare la matrice Q è non singolare (ed il rango di una matrice trasposta è pari al rango della matrice di partenza). Se ne deduce quindi che:

$$\text{rank}(W_{C,2}) = \text{rank}(W_{C,1}).$$

Si noti che in maniera del tutto equivalente si sarebbe potuto effettuare la dimostrazione del teorema facendo uso della matrice di controllabilità P_C .

II) Osservabilità

La dimostrazione è virtualmente identica a quella presentata per il caso della controllabilità e quindi viene lasciata al lettore come esercizio.

6.4 Forma di Kalman per la Controllabilità

Nella sezione 6.1 è stato introdotto il concetto di controllabilità di un sistema ed è stato definito il gramiano di controllabilità. In particolare, dato un sistema dinamico descritto dalla:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (6.39)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$, il teorema 6.2 ha messo in evidenza il fatto che, in caso di controllabilità, un generico stato $x \in \mathbb{R}^n$ del sistema può essere sempre decomposto in:

$$x = x_C + x_{NC}, \quad x_C \perp x_{NC} \quad (6.40)$$

dove x_C e x_{NC} appartengono rispettivamente al sottospazio degli stati controllabili e non controllabili. Si noti che è sempre possibile costruirsi una matrice di trasformazione Q che porta il sistema in **forma di Kalman per la controllabilità**.

Tale matrice di trasformazione Q può essere ottenuta a partire dalla matrice di controllabilità P_C introdotta in 6.2 come segue. Si assuma che il rango della matrice di controllabilità sia k , ovvero $\text{rank}(P_C) = k$, si può allora pensare di trovare un insieme di vettori $\{p_1, \dots, p_k\}$ che formano una base per $R(P_C)$ (il sottospazio degli stati controllabili), ed un insieme di vettori $\{p_{k+1}, \dots, p_n\}$ che formano una base per $R(P_C)^\perp = N(P_C^T)$, (il sottospazio degli stati non controllabili), e di ottenere quindi la matrice di trasformazione Q come concatenazione di tali basi.

Definizione 6.9 Si definisce matrice di trasformazione per la forma di Kalman per la controllabilità, la matrice Q_C ottenuta come segue:

$$Q_C^{-1} = [p_1, \dots, p_k, \quad | \quad p_{k+1}, \dots, p_n] \quad (6.41)$$

dove l'insieme di vettori $\{p_1, \dots, p_k\}$ forma una base per $R(P_C)$ (il sottospazio degli stati controllabili) e l'insieme di vettori $\{p_{k+1}, \dots, p_n\}$ forma una base per $R(P_C)^\perp = N(P_C^T)$, (il sottospazio degli stati non controllabili).

Teorema 6.7 (*Trasformazione per la Forma di Kalman per la Controllabilità*) Si consideri il sistema descritto dalla 6.39 al quale si applica la seguente trasformazione di coordinate:

$$Q_C x = \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_k \\ - - - \\ \hat{x}_{k+1} \\ \vdots \\ \hat{x}_n \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}_1 \\ - - - \\ \hat{\mathbf{x}}_2 \end{bmatrix} \quad (6.42)$$

Il sistema di equazioni risultante è caratterizzato dalla seguente forma:

$$\begin{bmatrix} \dot{\hat{\mathbf{x}}}_1 \\ \dot{\hat{\mathbf{x}}}_2 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} \hat{C}_1 & \hat{C}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} \quad (6.43)$$

dove il sottosistema di dimensione k :

$$\dot{\hat{\mathbf{x}}}_1 = \hat{A}_{11} \hat{\mathbf{x}}_1 + \hat{A}_{12} \hat{\mathbf{x}}_2 + \hat{B}_1 u \quad (6.44)$$

è il sottosistema controllabile, ed il sottosistema di dimensione $n - k$:

$$\dot{\hat{\mathbf{x}}}_2 = \hat{A}_{22} \hat{\mathbf{x}}_2 \quad (6.45)$$

è il sottosistema non controllabile.

Dimostrazione

Per la dimostrazione è sufficiente verificare che il sottospazio $R(P_C)$ è invariante rispetto ad A ovvero:

$$z \in R(P_C) \Rightarrow A z \in R(P_C).$$

A tal fine, si assuma che $\text{rank}\{P_C\} = k$ e si consideri una possibile base dell'immagine di P_C $\mathcal{B}(P_C) = \{p_1, \dots, p_k\}$. Questo implica che ogni vettore $z \in R(P_C)$ può essere espresso come combinazione lineare dei vettori che costituiscono la base $\mathcal{B}(P_C)$ come segue:

$$z = \sum_{i=1}^k \kappa_i p_i.$$

Poichè questo deve anche valere per ogni colonna della matrice P_C , è lecito esprimere il vettore z come combinazione di tali colonne $\{B, AB, \dots, A^{n-1}B\}$ come segue:

$$z = \sum_{i=0}^{n-1} \alpha_i A^i B$$

A questo punto si consideri il vettore Az , il quale può essere scritto come:

$$\begin{aligned} Az &= A \sum_{i=0}^{n-1} \alpha_i A^i B \\ &= A (\alpha_0 B + \alpha_1 A B + \dots + \alpha_{n-1} A^{n-1} B) \\ &= \alpha_0 A B + \alpha_1 A^2 B + \dots + \alpha_{n-1} A^n B. \end{aligned}$$

In particolare, si noti che il termine $A^n B$ per il teorema di Cayley-Hamilton può a sua volta essere espresso come:

$$A^n B = \sum_{i=0}^{n-1} \beta_i A^i B$$

Questo implica che il vettore Az può essere riscritto come:

$$Az = \sum_{i=0}^{n-1} \gamma_i A^i B$$

dove $\gamma_i = \alpha_i + \alpha_{n-1} \beta_i$, il quale prova che $Az \in R(P_C)$.

Una conseguenza diretta del fatto che $Az \in R(P_C)$, $\forall z \in R(P_C)$ è che le prime k colonne della matrice AQ_C^{-1} :

$$AQ_C^{-1} = [Ap_1, \dots, Ap_n]$$

appartengono tutte al $R(P_C)$.

Si consideri ora la matrice Q_C espressa per righe come segue:

$$Q_C = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix}$$

dove q_i è un vettore riga ($q_i \in \mathcal{R}^{1 \times n}$). Poichè vale la $Q_C Q_C^{-1} = I$ si ha che:

$$q_i p_j = 0 \quad \text{for } i \neq j,$$

inoltre essendo $\{p_1 \dots p_k\}$ una base per il $R(P_C)$ si ha che:

$$q_i z = 0 \quad \text{for } i \geq k+1, z \in R(P_C).$$

Questo implica che deve valere anche la:

$$q_i A p_j = 0 \quad \text{for } i \geq k+1, j \leq k.$$

In particolare, essendo $q_i A p_j$ l'espressione dell'elemento \hat{a}_{ij} della matrice $\hat{A} = Q_C A Q_C^{-1}$, se ne deduce che il blocco inferiore sinistro \hat{A}_{21} della matrice dinamica del sistema nella forma di Kalman per la controllabilità deve essere identicamente nullo.

Per concludere, poichè $B \in R(P_C)$, seguendo lo stesso ragionamento, si può dedurre che le ultime $n - k$ righe della matrice $Q_C B$:

$$Q_C B = \begin{bmatrix} q_1 B \\ \vdots \\ q_k B \\ \text{---} \text{---} \text{---} \\ q_{k+1} B \\ \vdots \\ q_n B \end{bmatrix}$$

che rappresentano il blocco \hat{B}_2 sono identicamente nulle.

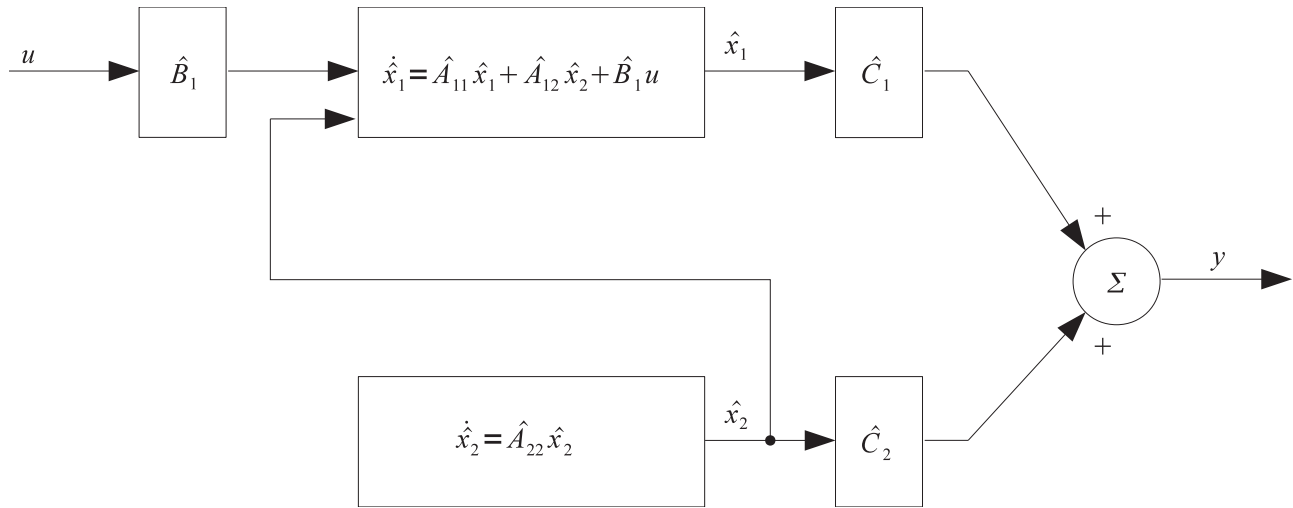


Figura 6.1: Decomposizione in componente controllabile e non controllabile

Esempio 6 Si consideri il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= x_1 - x_2 - 2x_3 + 2u \\ \dot{x}_2 &= 2x_2 + u \\ \dot{x}_3 &= x_2 + 3x_3 - u \\ y &= x_1 + 2x_2 + 2x_3 \end{cases} \quad (6.46)$$

- 1) Si determini la matrice di trasformazione Q_C per il passaggio nella forma di Kalman per la controllabilità,
- 2) Si determini il sottospazio degli stati controllabili ed il sottospazio degli stati non controllabili.

Il sistema descritto dalla 6.46 può essere rappresentato in forma compatta attraverso le seguenti matrici A, B, C :

$$A = \begin{bmatrix} 1 & -1 & -2 \\ 0 & 2 & 0 \\ 0 & 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \quad C = [1 \quad 2 \quad 2]$$

Per prima cosa ci calcoliamo la matrice di controllabilità P_C come segue:

$$P_C = [B \quad AB \quad A^2 B] = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 2 & 4 \\ -1 & -2 & -4 \end{bmatrix}$$

si può facilmente notare che:

$$\text{rank}(P_C) = 2$$

il quale ci informa del fatto che il sottospazio degli stati controllabili ha dimensione 2 mentre quello degli stati non controllabili ha dimensione 1, a supporto di tale ragionamento ricordiamo la seguente relazione valida per ogni applicazione lineare A in \mathbb{R}^n :

$$n = \dim(R(A)) + \dim(N(A^T))$$

Si calcoli ora una base per il sottospazio degli stati controllabili come segue:

$$\mathcal{B}_C = \mathcal{B}(R(P_C)) = [v_{1,C} \ v_{2,C}] = \begin{bmatrix} 2 & 3 \\ 1 & 2 \\ -1 & -2 \end{bmatrix}$$

ovvero si prendano due vettori (colonne) linearmente indipendenti della matrice di controllabilità P_C . Si calcoli ora una base per il sottospazio degli stati non controllabili come segue:

$$\mathcal{B}_{NC} = \mathcal{B}(N(P_C^T)) = [v_{1,NC}] = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

ovvero risolvendo il sistema $Ax = 0$ dove la matrice $A = P_C^T$. Si costruisca a questo punto la matrice di trasformazione (inversa) Q_C^{-1} per il passaggio nella forma di Kalman per la controllabilità come segue:

$$Q_C^{-1} = \mathcal{B}_C \oplus \mathcal{B}_{NC} = [v_{1,C} \ v_{2,C} \ v_{1,NC}] = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 2 & 1 \\ -1 & -2 & 1 \end{bmatrix}$$

Si noti che una semplice “prova del 9” (condizione necessaria ma non sufficiente) per verificare che la matrice Q_C^{-1} sia esatta è verificare che questa abbia rango massimo. A questo punto siamo in grado di calcolarci la matrice Q_C che risulta essere:

$$Q_C = \begin{bmatrix} 2 & \frac{-3}{2} & \frac{3}{2} \\ -1 & 1 & -1 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Effettuiamo ora la trasformazione $\hat{x} = Q_C x$ da cui si ottiene :

$$\hat{A} = Q_C A Q_C^{-1} = \begin{bmatrix} 0 & -2 & -3 \\ 1 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix} \quad \hat{B} = Q_C B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \hat{C} = C Q_C^{-1} = [2 \quad 3 \quad 4]$$

dal quale si possono individuare i seguenti blocchi di interesse:

$$\hat{A}_{11} = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \quad \hat{A}_{22} = [3] \quad \hat{B}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Ovvero il blocco relativo al sottospazio degli stati controllabili \hat{A}_{11} il quale ci dice che gli autovalori controllabili sono $\lambda_1 = 1$ e $\lambda_2 = 2$ (calcolabili come $\det(\hat{A}_{11} - \lambda I)$), il blocco relativo al sottospazio degli stati non controllabili \hat{A}_{22} , il quale ci dice che l'autovalore non controllabile è $\lambda_3 = 3$ (calcolabili come $\det(\hat{A}_{22} - \lambda I)$), ed infine si noti che anche il blocco \hat{B}_1 può avere degli zeri al suo interno.

Nonostante finora il concetto di controllabilità sia stato legato esclusivamente a quello di stato, si può facilmente individuare un legame diretto anche con quello di dinamica. È infatti possibile parlare di controllabilità di un singolo autovalore della matrice dinamica A del sistema. A tal proposito si introduce ora un teorema che ci permette di verificare la controllabilità o meno di una singola dinamica come segue.

Teorema 6.8 (Test PBH (Popov-Belevitch-Hautus) per la Controllabilità) Si consideri il sistema descritto dalla 6.39, la coppia (A, B) è controllabile se e solo se:

$$\text{rank}(A - \lambda I \mid B) = n \quad \forall \text{ autovalore } \lambda \text{ di } A. \quad (6.47)$$

Dimostrazione:

(\Rightarrow) Si assuma che $\text{rank}(A - \lambda I \mid B) < n$ per qualche autovalore. Allora $\exists \bar{v}$:

$$\bar{v}^T (A - \lambda I \mid B) = 0$$

quindi si ha:

$$\bar{v}^T A = \lambda \bar{v}^T, \quad \bar{v}^T B = 0$$

Inoltre si ha che:

$$\begin{aligned} \bar{v}^T A^2 &= \lambda \bar{v}^T A = \lambda^2 \bar{v}^T \\ \bar{v}^T A^3 &= \lambda^2 \bar{v}^T A = \lambda^3 \bar{v}^T \\ &\vdots \\ \bar{v}^T A^k &= \lambda^{k-1} \bar{v}^T A = \lambda^k \bar{v}^T \end{aligned}$$

Da cui si ha che:

$$\bar{v}^T [B, AB, A^2 B, \dots, A^{n-1} B] = [\bar{v}^T B, \bar{v}^T B \lambda, \bar{v}^T B \lambda^2, \dots, \bar{v}^T B \lambda^{n-1}] = 0$$

Il quale prova che la coppia (A, B) è non controllabile.

(\Leftarrow) Si assuma che la coppia (A, B) sia non controllabile. Allora per il teorema 6.9 esiste una matrice di trasformazione Q tale che:

$$\tilde{A} = Q A Q^{-1} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \quad \tilde{B} = Q B \begin{bmatrix} \hat{B}_1 \\ 0 \end{bmatrix}$$

Da cui si evince che:

$$\text{rank}(\tilde{A} - \lambda I \mid \tilde{B}) < n \quad \forall \text{ autovalore } \lambda \text{ di } \hat{A}_{22}.$$

6.5 Forma di Kalman per l'Osservabilità

Nella sezione 6.2 è stato introdotto il concetto di osservabilità di un sistema ed è stato definito il gramiano di osservabilità. In particolare, dato un sistema dinamico descritto dalla:

$$\begin{cases} \dot{x}(t) = A x(t) + B u(t) \\ y(t) = C x(t) \end{cases} \quad (6.48)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$, il teorema 6.4 ha messo in evidenza il fatto che, in caso di osservabilità, un generico stato $x \in \mathbb{R}^n$ del sistema può essere sempre decomposto in:

$$x = x_O + x_{NO}, \quad x_O \perp x_{NO} \quad (6.49)$$

dove x_O e x_{NO} appartengono rispettivamente al sottospazio degli stati osservabili e non osservabili. Si noti che è sempre possibile costruirsi una matrice di trasformazione Q che porta il sistema in **forma di Kalman per l'osservabilità**.

Tale matrice di trasformazione Q può essere ottenuta a partire dalla matrice di osservabilità P_O introdotta in 6.4 come segue. Si assuma che il rango della matrice di osservabilità sia $n - k$, ovvero $\text{rank}(P_O) = n - k$, si può allora pensare di trovare un insieme di vettori $\{p_1, \dots, p_k\}$ che formano una base per $R(P_O)^\perp = N(P_O)$, (il sottospazio degli stati non osservabili), ed un insieme di vettori $\{p_{k+1}, \dots, p_n\}$ che formano una base per $N(P_O)^\perp = R(P_O^T)$ (il sottospazio degli stati osservabili), e di ottenere quindi la matrice di trasformazione Q come concatenazione di tali basi.

Definizione 6.10 Si definisce matrice di trasformazione per la forma di Kalman per l'osservabilità, la matrice Q_O ottenuta come segue:

$$Q_O^{-1} = [p_1, \dots, p_k, \quad | \quad p_{k+1}, \dots, p_n] \quad (6.50)$$

dove l'insieme di vettori $\{p_1, \dots, p_k\}$ forma una base per $R(P_O^T)^\perp = N(P_O)$, (il sottospazio degli stati non osservabili) e l'insieme di vettori $\{p_{k+1}, \dots, p_n\}$ forma una base per $N(P_O)^\perp = R(P_O^T)$ (il sottospazio degli stati osservabili).

Teorema 6.9 (*Trasformazione per la Forma di Kalman per l'Osservabilità*) Si consideri il sistema descritto dalla 6.48 al quale si applica la seguente trasformazione di coordinate:

$$Q_O x = \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \vdots \\ \hat{x}_k \\ - - - \\ \hat{x}_{k+1} \\ \vdots \\ \hat{x}_n \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}_1 \\ - - - \\ \hat{\mathbf{x}}_2 \end{bmatrix} \quad (6.51)$$

Il sistema di equazioni risultante è caratterizzato dalla seguente forma:

$$\begin{bmatrix} \dot{\hat{\mathbf{x}}}_1 \\ \dot{\hat{\mathbf{x}}}_2 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & \hat{C}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_1 \\ \hat{\mathbf{x}}_2 \end{bmatrix} \quad (6.52)$$

dove il sottosistema k -dimensionale

$$\dot{\hat{\mathbf{x}}}_1 = \hat{A}_{11} \hat{\mathbf{x}}_1 + \hat{A}_{12} \hat{\mathbf{x}}_2 + \hat{B}_1 u \quad (6.53)$$

è il sottosistema non osservabile, ed il sottosistema $n - k$ -dimensionale:

$$\dot{\hat{\mathbf{x}}}_2 = \hat{A}_{22} \hat{\mathbf{x}}_2 + \hat{B}_2 u \quad (6.54)$$

$$y = \hat{C}_2 \hat{\mathbf{x}}_2 \quad (6.55)$$

è il sottosistema osservabile.

Esempio 7 Si consideri il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 = 3x_1 + x_2 + 2x_3 + u \\ \dot{x}_2 = 2x_2 + u \\ \dot{x}_3 = -x_2 + x_3 \\ y = 2x_2 + x_3 \end{cases} \quad (6.56)$$

- 1) Si determini la matrice di trasformazione Q_O per il passaggio nella forma di Kalman per la osservabilità,

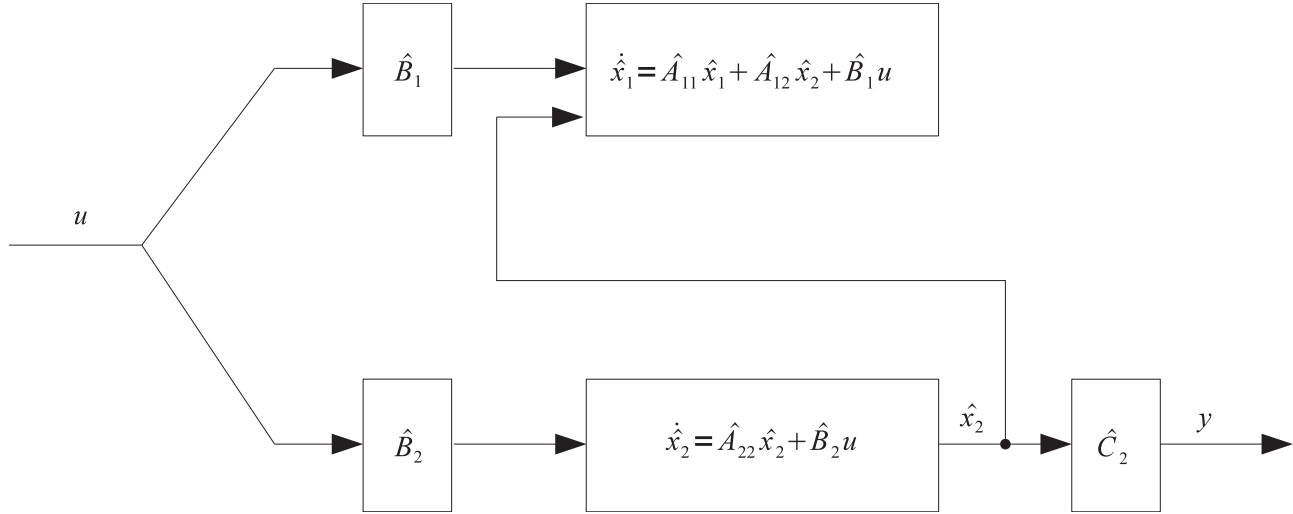


Figura 6.2: Decomposizione in componente osservabile e non osservabile

2) Si determini il sottospazio degli stati osservabili ed il sottospazio degli stati non osservabili.

Il sistema descritto dalla 6.56 può essere rappresentato in forma compatta attraverso le seguenti matrici A, B, C :

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad C = [0 \quad 2 \quad 1]$$

Per prima cosa ci calcoliamo la matrice di osservabilità P_O come segue:

$$P_O = \begin{bmatrix} C \\ C A \\ C A^2 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 3 & 1 \\ 0 & 5 & 1 \end{bmatrix}$$

si può facilmente notare che:

$$\text{nullity}(P_O) = 1$$

il quale ci informa del fatto che il sottospazio degli stati non osservabili ha dimensione 1, mentre quello degli stati osservabili ha dimensione 2 a supporto di tale ragionamento ricordiamo la seguente relazione valida per ogni applicazione lineare A in \mathbb{R}^n :

$$n = \dim(R(A)) + \dim(N(A^T))$$

Si calcoli ora una base per il sottospazio degli stati non osservabili come segue:

$$\mathcal{B}_{NO} = \mathcal{B}(N(P_O)) = [v_{1,NO}] = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

ovvero risolvendo il sistema $Ax = 0$ dove la matrice $A = N(P_O)$. Si calcoli ora una base per il sottospazio degli stati osservabili come segue:

$$\mathcal{B}_O = \mathcal{B}(R(P_O^T)) = [v_{1,O} \ v_{2,O}] = \begin{bmatrix} 0 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}$$

ovvero si prendano due vettori (colonne) linearmente indipendenti della matrice di osservabilità trasposta P_O^T . Si costruisca a questo punto la matrice di trasformazione (inversa) Q_O^{-1} per il passaggio nella forma di Kalman per la osservabilità come segue:

$$Q_O^{-1} = \mathcal{B}_{NO} \oplus \mathcal{B}_O = [v_{1,NO} \ v_{1,O} \ v_{2,O}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 0 & 1 & 1 \end{bmatrix}$$

Si noti che una semplice “prova del 9” (condizione necessaria ma non sufficiente) per verificare che la matrice Q_O^{-1} sia esatta è verificare che questa abbia rango massimo. A questo punto siamo in grado di calcolarci la matrice Q_O che risulta essere:

$$Q_O = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 1 & -2 \end{bmatrix}$$

Effettuiamo ora la trasformazione $\hat{x} = Q_O x$ da cui si ottiene :

$$\hat{A} = Q_O A Q_O^{-1} = \begin{bmatrix} 3 & 4 & 5 \\ 0 & -7 & -12 \\ 0 & 6 & 10 \end{bmatrix} \quad \hat{B} = Q_O B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \quad \hat{C} = C Q_O^{-1} = [0 \ 5 \ 7]$$

dal quale si possono individuare i seguenti blocchi di interesse:

$$\hat{A}_{11} = [3] \quad \hat{A}_{22} = \begin{bmatrix} -7 & -12 \\ 6 & 10 \end{bmatrix}$$

Ovvero il blocco relativo al sottospazio degli stati osservabili \hat{A}_{22} il quale ci dice che gli autovalori osservabili sono $\lambda_1 = 1$ e $\lambda_2 = 2$ (calcolabili come $\det(\hat{A}_{22} - \lambda I)$), il blocco relativo al sottospazio degli stati non osservabili \hat{A}_{11} , il quale ci dice che l'autovalore non osservabile è $\lambda_3 = 3$ (calcolabili come $\det(\hat{A}_{11} - \lambda I)$).

Teorema 6.10 (Test PBH (Popov-Belevitch-Hautus) per l' Osservabilità) Si consideri il sistema descritto dalla 6.48, la coppia (A, C) è osservabile se e solo se:

$$\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n \quad \forall \text{ autovalore } \lambda \text{ di } A. \quad (6.57)$$

Dimostrazione: La dimostrazione è virtualmente identica a quella presentata in 6.8 per il test PBH per la controllabilità, per questo è lasciata al lettore come esercizio.

6.6 Decomposizione Canonica di Kalman

Nelle sezioni precedenti sono state messe in luce due proprietà fondamentali di un sistema lineare, ovvero la controllabilità e l'osservabilità. Sono state successivamente presentate due trasformazioni di coordinate attraverso cui si ottengono due rappresentazioni nella quale tali proprietà sono chiaramente messe in luce, ovvero la forma di Kalman per la controllabilità e la forma di Kalman per la osservabilità. In questa sezione verrà presentata una trasformazione di coordinate alternativa nella quale le due proprietà sono contemporaneamente messe in evidenza: tale operazione va sotto il nome di **decomposizione canonica di Kalman**.

Teorema 6.11 (*Decomposizione Canonica di Kalman*) Dato un sistema dinamico descritto dalle seguenti equazioni:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) \end{cases} \quad (6.58)$$

dove $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$, $C \in \mathbb{R}^{q \times n}$. È sempre possibile trovare una matrice di trasformazione Q che porta il sistema nella forma canonica di Kalman:

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \\ \dot{\hat{x}}_3 \\ \dot{\hat{x}}_4 \end{bmatrix} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} & \hat{A}_{13} & \hat{A}_{14} \\ 0 & \hat{A}_{22} & 0 & \hat{A}_{24} \\ 0 & 0 & \hat{A}_{33} & \hat{A}_{34} \\ 0 & 0 & \hat{0} & \hat{A}_{44} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} + \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \\ 0 \\ 0 \end{bmatrix} u \quad (6.59)$$

$$y = \begin{bmatrix} 0 & \hat{C}_2 & 0 & \hat{C}_4 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix}$$

Teorema 6.12 (*Matrice \hat{Q} per la Decomposizione Canonica di Kalman*) Una matrice di trasformazione \hat{Q} che porta il sistema nella forma canonica di Kalman può essere ottenuta come segue. Si determini una base per il sottospazio degli stati controllabili $\mathcal{B}_C : R(P_C) = \text{span}(\mathcal{B}_C)$ ed una base del sottospazio degli stati inosservabili $\mathcal{B}_{NO} : N(P_O) = \text{span}(\mathcal{B}_{NO})$ e si costruisca la matrice di trasformazione come segue:

$$\hat{Q} = [\chi_1 \quad \chi_2 \quad \chi_3 \quad \chi_4] \quad (6.60)$$

dove le varie basi sono ottenute in accordo alle seguenti regole:

$$I) \chi_1 = \mathcal{B}_C \cap \mathcal{B}_{NO}$$

$$II) \chi_1 \oplus \chi_2 = \mathcal{B}_C$$

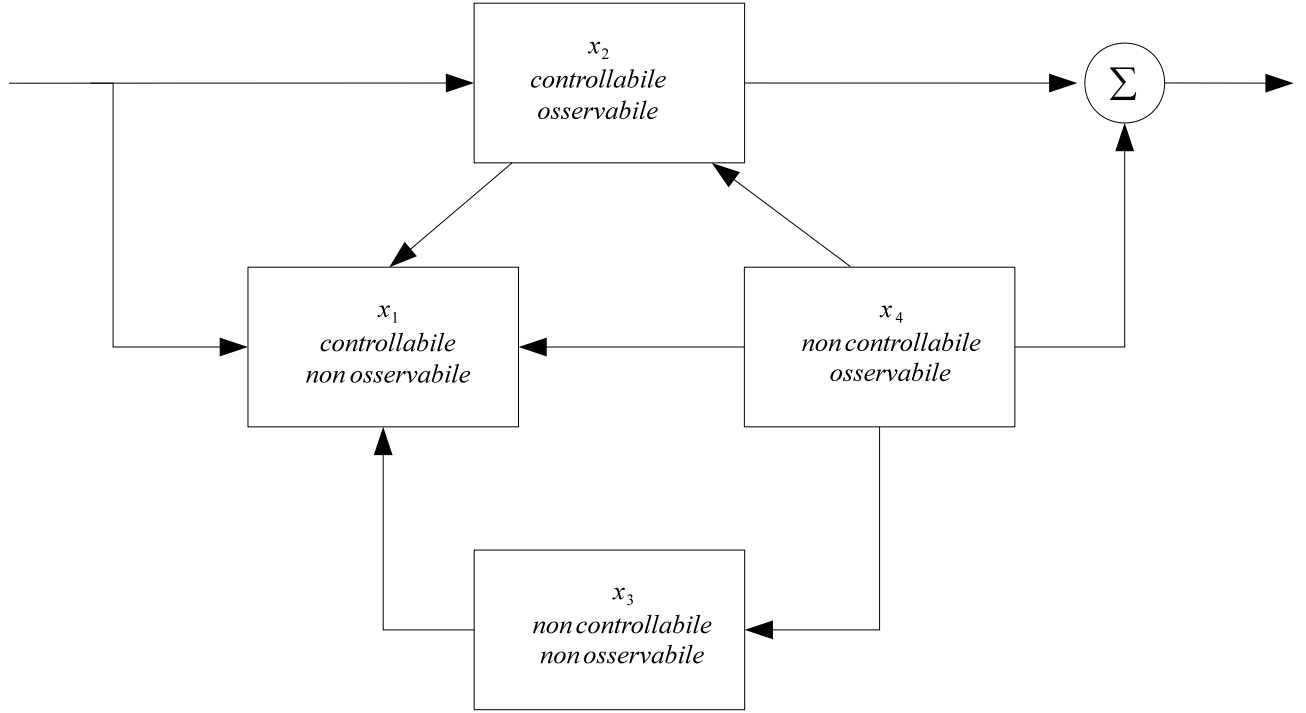


Figura 6.3: Decomposizione Completa

$$III) \chi_1 \oplus \chi_3 = \mathcal{B}_{NO}$$

$$IV) \chi_1 \oplus \chi_2 \oplus \chi_3 \oplus \chi_4 = \mathbb{R}^n$$

dove col simbolo \oplus si intende la composizione di spazi.

Teorema 6.13 (*Relazione V.d.S./ $H(s)$*) Si consideri un sistema dinamico descritto dalla 6.59, la relativa matrice funzione di trasferimento e matrice di risposta impulsiva risultano essere rispettivamente:

$$\hat{H}(s) = \hat{C}_2 (sI - \hat{A}_{22})^{-1} \hat{B}_2 \quad (6.61)$$

$$\hat{h}(t) = \hat{C}_2 e^{\hat{A}_{22} t} \hat{B}_2 \quad (6.62)$$

Dimostrazione: Per la dimostrazione è sufficiente considerare la definizione di matrice risolvante, applicarla al sistema descritto dalla 6.59 e ricordare ovviamente che la 6.62 è semplicemente l'antitrasformata della 6.61:

$$\begin{aligned}
\hat{H}(s) &= \hat{C}(sI - \hat{A})^{-1}\hat{B} \\
&= \begin{bmatrix} 0 & \hat{C}_2 & 0 & \hat{C}_4 \end{bmatrix} \begin{bmatrix} (sI - \hat{A}_{11}) & -\hat{A}_{12} & -\hat{A}_{13} & -\hat{A}_{14} \\ 0 & (sI - \hat{A}_2) & 0 & -\hat{A}_{24} \\ 0 & 0 & (sI - \hat{A}_{33}) & -\hat{A}_{34} \\ 0 & 0 & \hat{0} & (sI - \hat{A}_{44}) \end{bmatrix}^{-1} \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \\ 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & \hat{C}_2 & 0 & \hat{C}_4 \end{bmatrix} \begin{bmatrix} (sI - \hat{A}_{11})^{-1} & \text{x} & \text{x} & \text{x} \\ 0 & (sI - \hat{A}_2)^{-1} & \text{x} & \text{x} \\ 0 & 0 & (sI - \hat{A}_{33})^{-1} & \text{x} \\ 0 & 0 & \hat{0} & (sI - \hat{A}_{44})^{-1} \end{bmatrix} \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \\ 0 \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & \hat{C}_2 & 0 & \hat{C}_4 \end{bmatrix} \begin{bmatrix} \text{x} \\ (sI - \hat{A}_2)^{-1}\hat{B}_2 \\ 0 \\ 0 \end{bmatrix} = \hat{C}_2(sI - \hat{A}_{22})^{-1}\hat{B}_2
\end{aligned}$$

Si osservi che il teorema 6.13 è di estrema importanza in quanto evidenzia come le dinamiche descritte dalla matrice funzione di trasferimento siano solo quelle associate al sottospazio controllabile ed osservabile. Questo concetto tornerà utile, come vedremo nel prossimo capitolo, quando si deve affrontare il problema della realizzazione in quanto fornisce una condizione necessaria e sufficiente per la determinazione di una realizzazione minima associata ad una matrice funzione di trasferimento.

Capitolo 7

Forme Canoniche e Realizzazione

7.1 Realizzazione

Nei capitoli precedenti sono state presentate metodologie per la modellazione dei sistemi dinamici e sono state descritte ed investigate le proprietà strutturali ad esse associate. In questo capitolo, verrà presentato il problema della **realizzazione**, il quale risulta essere estremamente importante da un punto di vista ingegneristico. Infatti, col termine realizzazione si intende il problema di individuare, a partire da una funzione di trasferimento (ottenuta ad esempio per via empirica attraverso misure della risposta in frequenza di un sistema non noto a priori), una rappresentazione in spazio di stato che correttamente rifletta il comportamento ingresso/uscita del sistema in analisi.

Definizione 7.1 (*Realizzazione*) Data una matrice funzione di trasferimento $H(s) \in \mathcal{R}^{q \times p}$ definita come segue:

$$H(s) = \frac{Q_m s^m + Q_{m-1} s^{m-1} + \dots + Q_1 s^1 + Q_0}{s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}, \quad (7.1)$$

si definisce **realizzazione** della $H(s)$ un sistema:

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (7.2)$$

tale che:

$$C (sI - A)^{-1} B = H(s). \quad (7.3)$$

In maniera del tutto equivalente, il sistema 7.2 è detto *realizzazione della risposta impulsiva* se:

$$C e^{At} B = \mathcal{L}^{-1}[H(s)] = h(t). \quad (7.4)$$

Si noti che nel caso la matrice funzione di trasferimento si riduca al caso scalare:

$$F(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0} = \frac{Y(s)}{U(s)} \quad (7.5)$$

il problema della realizzazione è equivalente al problema di individuare una descrizione in spazio di stato per una equazione differenziale di ordine n derivante dalla 7.5 del tipo:

$$y^{(n)} + a_{n-1} y^{(n-1)}(t) + \dots + a_1 \dot{y}(t) + a_0 y(t) = b_m u^{(m)}(t) + b_{m-1} u^{(m-1)}(t) + \dots + b_1 \dot{u}(t) + b_0 u(t) \quad (7.6)$$

In particolare, nel capitolo 4.2 sono state presentate due tecniche per risolvere questo problema in grado di portare il sistema nella forma compagna di controllore ed osservatore. Tale approccio può essere esteso anche al caso multi-input/multi-output (MIMO) per il quale si può pensare di realizzare una forma compagna di controllore come segue:

$$\dot{x}(t) = \begin{bmatrix} 0_p & I_p & 0_p & \dots & \dots & 0_p \\ 0_p & 0_p & I_p & \dots & \dots & 0_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_p & 0_p & 0_p & \dots & \dots & I_p \\ -a_0 I_p & -a_1 I_p & -a_2 I_p & \dots & \dots & -a_{n-1} I_p \end{bmatrix} x(t) + \begin{bmatrix} 0_p \\ 0_p \\ \vdots \\ \vdots \\ 0_p \\ I_p \end{bmatrix} u(t) \quad (7.7)$$

$$y(t) = [Q_0 \quad Q_1 \quad Q_2 \quad \dots \quad \dots \quad Q_{n-2} \quad Q_{n-1}] x(t) \quad (7.8)$$

Analogamente si può pensare di realizzare una forma compagna di osservatore come segue:

$$\dot{x}(t) = \begin{bmatrix} 0_q & \dots & \dots & \dots & \dots & \dots & -a_0 I_q \\ I_q & 0_q & \dots & \dots & \dots & \dots & -a_1 I_q \\ 0_q & I_q & 0_q & \dots & \dots & \dots & -a_2 I_q \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_q & \vdots & \vdots & 0_q & I_q & 0_q & -a_{n-2} I_q \\ 0_q & 0_q & \dots & \dots & 0_q & I_q & -a_{n-1} I_q \end{bmatrix} x(t) + \begin{bmatrix} Q_0 \\ Q_1 \\ Q_2 \\ \vdots \\ \vdots \\ Q_{n-2} \\ Q_{n-1} \end{bmatrix} u(t) \quad (7.9)$$

$$y(t) = [0_q \quad 0_q \quad \dots \quad \dots \quad \dots \quad \dots \quad 0_q \quad I_q] x(t) \quad (7.10)$$

Tuttavia, va precisato che tali rappresentazioni **non** godono della proprietà di dualità presentata per il caso “Single-Input / Single Output” (SISO) nel capitolo 4.2 per ovvie ragioni dimensionali. Infatti nella forma compagna di controllore il sistema ha dimensione np mentre nella forma compagna di osservatore il sistema ha dimensione nq . Inoltre, al fine di giustificare l'utilizzo del termine controllabilità ed osservabilità si fa riferimento al seguente lemma.

Lemma 7.1 *Un sistema espresso in forma compagna di controllore è sempre controllabile ed, analogamente, un sistema espresso in forma compagna di osservatore è sempre osservabile.*

Dimostrazione: La dimostrazione è una semplice conseguenza della struttura delle matrici $(A, B,)$ e (A, C) per i sistemi presi in considerazione. Si ha infatti che la matrice di controllabilità (osservabilità) legata a tale rappresentazione risulta essere strutturalmente sempre di rango massimo. A tal scopo si consideri un generico sistema dinamico SISO di ordine n , la cui rappresentazione in forma compagna di controllore risulta essere:

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

La matrice di controllabilità associata a tale coppia (A, B) risulta essere:

$$P_C = [B \ AB \ A^2B \ \dots \ A^{n-1}B] = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & * \\ \vdots & & & & \vdots \\ 0 & 1 & \dots & * & * \\ 1 & * & \dots & * & * \end{bmatrix}$$

La quale essendo una matrice diagonale inferiore con degli uni sulla anti-diagonale risulta essere strutturalmente di rango massimo. A questo punto l'osservabilità della forma compagna di osservatore è una semplice conseguenza della dualità espressa dal Teorema 4.4.

È importante notare come sia possibile ottenere infinite realizzazione a partire dalla $H(s)$. In particolare, il teorema seguente permette di caratterizzare la realizzazione in forma minima, ovvero la realizzazione di una funzione di trasferimento ottenuta con il minor numero possibile di variabili di stato.

Definizione 7.2 (*Realizzazione minima*) *Si consideri la matrice funzione di trasferimento $H(s)$ descritta dalla 7.1, una sua realizzazione:*

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{cases} \quad (7.11)$$

si dice minima, se non esiste nessun'altra realizzazione della $H(s)$:

$$\begin{cases} \dot{\tilde{x}}(t) &= \tilde{A}\tilde{x}(t) + \tilde{B}u(t) \\ y(t) &= \tilde{C}\tilde{x}(t) \end{cases} \quad (7.12)$$

tale che:

$$\dim(\tilde{x}) < \dim(x). \quad (7.13)$$

In particolare, nel capitolo precedente (6.13) si è mostrato come una funzione di trasferimento associata ad un dato sistema ne riveli la sola parte completamente controllabile ed osservabile. Ciò implica che ogni parte che risulti essere non controllabile o non osservabile può essere eliminata dalla realizzazione senza alterare minimamente la funzione di trasferimento. Se ne deduce quindi che, ai fini della minimalità della realizzazione, tali dinamiche possono essere escluse. Questa relazione tra rappresentazione minima e osservabilità/controllabilità è sottolineata dal seguente teorema.

Teorema 7.1 (*Condizione necessaria e sufficiente per una realizzazione minima*) Una realizzazione della matrice funzione di trasferimento $H(s)$ descritta dalla

$$\begin{cases} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) \end{cases} \quad (7.14)$$

è minima se e solo se essa è completamente controllabile ed osservabile.

(\Rightarrow) Si supponga che il sistema descritto dalla 7.14 non sia completamente controllabile ed osservabile. Allora, in accordo al teorema 6.13, ogni stato non controllabile e/o inosservabile può essere tranquillamente rimosso dalla realizzazione senza alterare la matrice funzione di trasferimento ad essa associata. Di conseguenza la realizzazione così ottenuta è sicuramente di ordine minore di quella di partenza.

(\Leftarrow) Si supponga che il sistema descritto dalla 7.14 sia controllabile e osservabile, e si prenda un'altra realizzazione :

$$\begin{cases} \dot{\tilde{x}}(t) &= \tilde{A} \tilde{x}(t) + \tilde{B} u(t) \\ y(t) &= \tilde{C} \tilde{x}(t) \end{cases}$$

dove $\tilde{A} \in \mathbb{R}^{r \times r}$. Si mostrerà ora che deve essere $r \geq n$. In particolare, poichè le due rappresentazioni sono realizzazioni della stessa matrice funzione di trasferimento $\hat{H}(s)$, esse realizzeranno allo stesso tempo anche la stessa risposta impulsiva $\hat{h}(t) = \mathcal{L}^{-1}(\hat{H}(s))$, quindi si ha che:

$$\begin{aligned} \hat{h}(t - \tau) &= \tilde{C} e^{\tilde{A}(t-\tau)} \tilde{B} = C e^{A(t-\tau)} B \\ &= \tilde{C} e^{\tilde{A}t} e^{-\tilde{A}\tau} \tilde{B} = C e^{At} e^{-A\tau} B \end{aligned}$$

moltiplichiamo ora a destra per $e^{A^T t} C^T$ ed a sinistra per $B^T e^{-A^T \tau}$ ed integriamo ambo i lati. Si ha quindi:

$$\begin{aligned} \int_0^T e^{-A^T t} C^T \tilde{C} e^{\tilde{A}t} dt \int_0^T e^{-\tilde{A}\tau} \tilde{B} B^T e^{-A^T \tau} d\tau &= \\ \int_0^T e^{-A^T t} C^T C e^{At} dt \int_0^T e^{-A\tau} B B^T e^{-A^T \tau} d\tau & \end{aligned}$$

Il quale ricordando la 6.3 e la 6.28 può essere riscritto come:

$$V_O(0, T) V_C(0, T) = W_O(0, T) W_C(0, T).$$

dove $V_O(0, T) \in \mathbb{R}^{n \times r}$, $V_C(0, T) \in \mathbb{R}^{r \times n}$, $W_C(0, T) \in \mathbb{R}^{n \times n}$ e $W_O(0, T) \in \mathbb{R}^{n \times n}$. Si consideri il termine prodotto alla sinistra dell'equazione, poichè il sistema di partenza è controllabile ed osservabile e ricordando le proprietà associate al rango del prodotto di matrici si ha che:

$$\text{rank}(W_O(0, T) W_C(0, T)) = n$$

Da cui si ha che per l'uguaglianza vale anche la:

$$\text{rank}(V_O(0, T) V_C(0, T)) \leq \min(\{n, r\}) = n$$

che ci informa che deve essere $r \geq n$.

È importante sottolineare come il teorema 7.1 fornisca una semplice modalità operativa per determinare una realizzazione minima di una generica matrice funzione di trasferimento $H(s)$. Si può infatti pensare di portare prima il sistema nella forma canonica di osservatore 7.9 e successivamente applicare la trasformazione descritta dal teorema 6.7 per mettere in evidenza il sottospazio non controllabile del sistema. A questo punto, eliminando tale sottospazio la parte rimanente risulta essere completamente controllabile ed osservabile, quindi in altri termini, per il teorema 7.1, una realizzazione minima. Ovviamente in maniera del tutto equivalente, e computazionalmente anche conveniente nel caso la dimensione delle uscite fosse inferiore a quella degli ingressi ($q < p$), si può pensare di portare prima il sistema in forma canonica di controllore e successivamente applicare la trasformazione descritta dal teorema 6.9 per mettere in evidenza il sottospazio non osservabile del sistema e quindi eliminarlo dalla descrizione.

A questo punto, essendo oramai evidente la relazione tra minimalità di una realizzazione e osservabilità e controllabilità associate a tale rappresentazione, introduciamo un teorema il quale stabilisce che tutte le rappresentazioni in forma minima di una stessa matrice funzione di trasferimento sono equivalenti.

Teorema 7.2 (*Equivalenza rappresentazioni minima*) *Se i due sistemi:*

$$\begin{cases} \dot{x}(t) = A x(t) + B u(t) \\ y(t) = C x(t) \end{cases} \quad \begin{cases} \dot{\tilde{x}}(t) = \tilde{A} \tilde{x}(t) + \tilde{B} u(t) \\ y(t) = \tilde{C} \tilde{x}(t) \end{cases} \quad (7.15)$$

sono due realizzazioni minime della stessa funzione di trasferimento $H(s)$, allora esiste una trasformazione di coordinate $\tilde{x} = Q x$ tale che:

$$\tilde{A} = Q A Q^{-1} \quad \tilde{B} = Q B \quad \tilde{C} = C Q^{-1} \quad (7.16)$$

Dimostrazione:

Teorema 7.3 (*Cancellazioni Poli/Zero e Minimalità*) *Il sistema descritto dalla:*

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{cases} \quad (7.17)$$

è controllabile ed osservabile se non vi sono cancellazioni poli-zero tra il numeratore ed il denominatore della matrice funzione di trasferimento:

$$H(s) = C(sI - A)^{-1}B = \frac{Q(s)}{\det(sI - A)} \quad (7.18)$$

dove $Q(s)$ è una matrice polinomiale in s . Si noti che, mentre nel caso dei sistemi MIMO tale condizione è solo sufficiente (ovvero può accadere che un sistema sia ancora controllabile ed osservabile nonostante vi siano delle cancellazioni), per i sistemi SISO tale condizione risulta essere necessaria e sufficiente per garantire la completa controllabilità ed osservabilità del sistema.

Dimostrazione:

7.2 Forme Canoniche per la Realizzazione

Nella sezione 7.1 è stato introdotto il problema della realizzazione, sono state presentate le forme canoniche di controllabilità ed osservabilità per sistemi multi-variabile, e sono state discusse varie proprietà delle realizzazioni come ad esempio la realizzazione minime. In questa sezione, si riprendono in analisi i sistemi SISO e si presentano due tecniche di trasformazione per il passaggio in forma compagna di una qualsiasi rappresentazione in spazio di stato, a fini realizzativi.

7.2.1 Forma Compagna di Controllore

Si consideri un sistema dinamico SISO descritto dal seguente sistema di equazioni differenziali lineari a coefficienti costanti:

$$\begin{cases} \dot{x}(t) &= Ax(t) + bu(t) \\ y(t) &= cx(t) \end{cases} \quad (7.19)$$

dove $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^{n \times 1}$, $c \in \mathbb{R}^{1 \times n}$, in particolare, essendo la tipologia dei sistemi in analisi SISO, si assuma $q = p = 1$.

Teorema 7.4 (*Matrice di trasformazione Q per la forma compagna di controllore*) *Si consideri il sistema descritto dalla 7.19, se la coppia (A, B) è controllabile (ovvero il sistema è controllabile), allora esiste sempre una matrice di trasformazione Q che porta il sistema in forma compagna di controllore:*

$$\begin{cases} \dot{\tilde{x}}(t) &= A_c \tilde{x}(t) + b_c u(t) \\ y(t) &= \tilde{c} \tilde{x}(t) \end{cases} \quad (7.20)$$

dove $A_c = Q A Q^{-1}$, $b_c = Q b$ e $\tilde{c} = c Q^{-1}$.

In particolare, tale matrice di trasformazione Q è ottenuta a partire dalla matrice di controllabilità P_C come segue:

$$Q = \begin{bmatrix} \gamma \\ \gamma A \\ \gamma A^2 \\ \vdots \\ \gamma A^{n-1} \end{bmatrix} \quad (7.21)$$

dove γ è l'ultima riga dell'inverso della matrice di controllabilità P_C :

$$P_C^{-1} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ & & \gamma \end{bmatrix} \quad (7.22)$$

Dimostrazione: La dimostrazione è divisa in due parti: prima si dimostra il fatto che la matrice Q abbia rango pieno (ovvero che sia invertibile), successivamente si prova come tale matrice di trasformazione Q porti il sistema in forma compagna di controllore.

I) Invertibilità matrice Q

Si consideri la matrice di controllabilità

$$P_C = [b \ Ab \ A^2b \ \dots \ A^{n-1}b]$$

per l'assunzione di controllabilità si ha che $P_C^{-1}P_C = I$ ovvero:

$$P_C^{-1}P_C = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ & & \gamma \end{bmatrix} [b \ Ab \ A^2b \ \dots \ A^{n-1}b]$$

da cui si ricava il seguente sistema:

$$\begin{cases} \gamma b & = 0 \\ \gamma Ab & = 0 \\ \gamma A^2b & = 0 \\ \vdots & \\ \gamma A^{n-2}b & = 0 \\ \gamma A^{n-1}b & = 1 \end{cases}$$

questo ci permette di verificare che le righe di Q sono linearmente indipendenti:

$$\sum_{i=0}^{n-1} c_i \gamma A^i = 0 \quad \text{con } c_0 = c_1 = c_2 = \dots = c_{n-1} = 0$$

Si ha infatti che moltiplicando ambo i membri a destra per b si ha:

$$\sum_{i=0}^{n-1} c_i \gamma A^i b = 0 = c_{n-1} = 0$$

A questo punto si può pensare di valutare la sommatoria fino a $n-2$, previa moltiplicazione a destra ambo i membri per Ab come segue:

$$\sum_{i=0}^{n-2} c_i \gamma A^i (Ab) = 0 = c_{n-2} = 0$$

Si può continuare ad operare in questo modo valutando la sommatoria per $n-i$ e moltiplicando a destra ambo i membri per $A^i b$ per provare che il $n-i$ -esimo coefficiente $c_{n-i} = 0$, per arrivare a provare che:

$$c_0 = c_1 = c_2 = \dots = c_{n-1} = 0.$$

Il quale prova l'invertibilità della matrice di trasformazione di coordinate Q .

II) Passaggio in forma compagna

Assumiamo a questo punto di applicare la trasformazione di coordinate $\hat{x} = Qx$ per ottenere \tilde{A} , \tilde{b} , e \tilde{c} come segue:

$$\tilde{A} = Q A Q^{-1} \quad \tilde{b} = Q b \quad \tilde{c} = c Q^{-1}$$

In particolare, si vada ad analizzare la struttura della matrice dinamica del sistema \tilde{A} . Dalla definizione della matrice di trasformazione Q otteniamo:

$$\tilde{A} Q = Q A \quad \Rightarrow \quad \tilde{A} Q = \begin{bmatrix} \gamma \\ \gamma A \\ \gamma A^2 \\ \vdots \\ \gamma A^{n-2} \\ \gamma A^{n-1} \end{bmatrix} A = \begin{bmatrix} \gamma A \\ \gamma A^2 \\ \gamma A^3 \\ \vdots \\ \gamma A^{n-1} \\ \gamma A^n \end{bmatrix}$$

Da cui si evince che:

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{bmatrix}$$

dove per l'ultimo termine si fa riferimento al teorema di Cayley-Hamilton 5.4 grazie al quale si ottiene:

$$-\gamma A^n = a_0 \gamma + a_1 \gamma A + a_2 \gamma A^2 + \dots + a_{n-1} \gamma A^{n-1}$$

In maniera del tutto equivalente si può verificare che:

$$\tilde{b} = Qb \quad \implies \quad \tilde{b} = \begin{bmatrix} \gamma b \\ \gamma A b \\ \gamma A^2 b \\ \vdots \\ \gamma A^{n-2} b \\ \gamma A^{n-1} b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

7.2.2 Forma Compagna di Osservatore

Si consideri un sistema dinamico SISO descritto dal seguente sistema di equazioni differenziali lineari a coefficienti costanti:

$$\begin{cases} \dot{x}(t) = Ax(t) + bu(t) \\ y(t) = cx(t) \end{cases} \quad (7.23)$$

dove $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^{n \times 1}$, $c \in \mathbb{R}^{1 \times n}$, in particolare, essendo la tipologia dei sistemi in analisi SISO, si assuma $q = p = 1$.

Teorema 7.5 (*Matrice di trasformazione Q per la forma compagna di osservatore*) Si consideri il sistema descritto dalla 7.23, se la coppia (A, C) è osservabile (ovvero il sistema è osservabile), allora esiste sempre una matrice di trasformazione Q che porta il sistema in forma compagna di osservabile:

$$\begin{cases} \dot{\tilde{x}}(t) = A_o \tilde{x}(t) + \tilde{b}u(t) \\ y(t) = c_o \tilde{x}(t) \end{cases} \quad (7.24)$$

dove $A_o = Q A Q^{-1}$, $\tilde{b} = Q b$ e $c_o = c Q^{-1}$.

In particolare, tale matrice di trasformazione Q è ottenuta a partire dalla matrice di osservabilità come segue:

$$Q^{-1} = \begin{bmatrix} \xi & A\xi & \dots & A^{n-1}\xi \end{bmatrix} \quad (7.25)$$

dove ξ è l'ultima colonna dell'inverso della matrice di osservabilità P_O :

$$P_O^{-1} = \begin{bmatrix} \vdots & \vdots \\ \vdots & \vdots & \xi \\ \vdots & \vdots \end{bmatrix} \quad (7.26)$$

Dimostrazione: Si basa sul teorema 4.4 di dualità delle forme compagne e sul teorema 7.4 per il calcolo della matrice di trasformazione Q per la forma compagna di controllore. Infatti, in accordo al teorema di dualità 4.4 e ricordando che $A_o = Q A Q^{-1}$ si ha che:

$$\begin{aligned} A_o &= \hat{A}_c^T \\ Q A Q^{-1} &= (\hat{Q} \hat{A} \hat{Q}^{-1})^T \\ Q A Q^{-1} &= \hat{Q}^{-T} \hat{A}^T \hat{Q}^T \end{aligned}$$

da cui si deduce che:

$$Q^{-1} = \hat{Q}^T$$

Inoltre analizzando la struttura di \hat{Q} osserviamo che:

$$\hat{Q} = \begin{bmatrix} \hat{\gamma} \\ \hat{\gamma} \hat{A} \\ \hat{\gamma} \hat{A}^2 \\ \vdots \\ \hat{\gamma} \hat{A}^{n-1} \end{bmatrix} \quad (7.27)$$

dove $\hat{\gamma}$ è l'ultima riga dell'inverso della matrice di controllabilità \hat{P}_C :

$$\hat{P}_C^{-1} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ & & \hat{\gamma} \end{bmatrix} \quad (7.28)$$

Se ne deduce che:

$$Q^{-1} = \begin{bmatrix} \hat{\gamma} \\ \hat{\gamma} \hat{A} \\ \hat{\gamma} \hat{A}^2 \\ \vdots \\ \hat{\gamma} \hat{A}^{n-1} \end{bmatrix}^T = \begin{bmatrix} \hat{\gamma}^T & \hat{A}^T \hat{\gamma}^T & (\hat{A}^T)^2 \hat{\gamma}^T & \dots & (\hat{A}^T)^{n-1} \hat{\gamma}^T \end{bmatrix} \quad (7.29)$$

$$= \begin{bmatrix} \xi & A \xi & A^2 \xi & \dots & A^{n-1} \xi \end{bmatrix} \quad (7.30)$$

dove $\xi = \hat{\gamma}^T$ e si ricorda inoltre che $A = \hat{A}^T$. Vediamo ora come calcolare in maniera diretta ξ . In particolare notiamo che:

$$\begin{bmatrix} \vdots & \vdots & \\ \vdots & \vdots & \xi \\ \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} \dots & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ & & \hat{\gamma} \end{bmatrix}^T = \begin{bmatrix} \hat{P}_C^{-1} \end{bmatrix}^T \quad (7.31)$$

Ora è sufficiente ricordare la seguente proprietà delle matrici:

$$[A^{-1}]^T = [A^T]^{-1} \quad (7.32)$$

per realizzare che vale la:

$$\begin{bmatrix} \hat{P}_C^{-1} \end{bmatrix}^T = \begin{bmatrix} \hat{P}_C^T \end{bmatrix}^{-1} = P_O^{-1} = \begin{bmatrix} \vdots & \vdots & \\ \vdots & \vdots & \xi \\ \vdots & \vdots & \end{bmatrix} \quad (7.33)$$

che prova l'enunciato.

Capitolo 8

Luogo delle Radici

Il luogo delle radici è un diagramma che mostra come si spostano i poli di un sistema dinamico lineare stazionario a singolo ingresso e singola uscita (SISO) al variare di un parametro. È stato ideato nel 1948 da Walter R. Evans.

8.1 Luogo Esatto

Definizione 8.1 *Il luogo delle radici è un procedimento grafico che permette di valutare le prestazioni di un sistema a ciclo chiuso al variare del guadagno K del controllore $C(s)$ in catena diretta. Il ramo positivo è il luogo delle radici ottenuto per $K \in [0, +\infty)$ Il ramo negativo è il luogo delle radici ottenuto per $K \in (-\infty, 0]$ Il ramo completo è il luogo delle radici ottenuto per $K \in (-\infty, +\infty)$ (ovvero l'unione del luogo positivo e negativo)*

Si consideri ad esempio un processo descritto dalla seguente funzione di trasferimento:

$$P(s) = \frac{1}{s \cdot (s - p)} \quad (8.1)$$

La funzione di trasferimento a ciclo chiuso considerando $C(s) = K$ e $H(s) = 1$ risulta:

$$W(s) = \frac{K}{s^2 - p s + K} \quad (8.2)$$

Il polinomio caratteristico è :

$$f(s, K) = s^2 - p s + K \quad (8.3)$$

Di conseguenza poichè il $\Delta = p^2 - 4 \cdot K$ avremo tre casi possibili per le radici al variare di K :

$$\begin{aligned} \Delta = 0 & \rightarrow K = \frac{p^2}{4} \\ \Delta > 0 & \rightarrow K < \frac{p^2}{4} \\ \Delta < 0 & \rightarrow K > \frac{p^2}{4} \end{aligned} \quad (8.4)$$

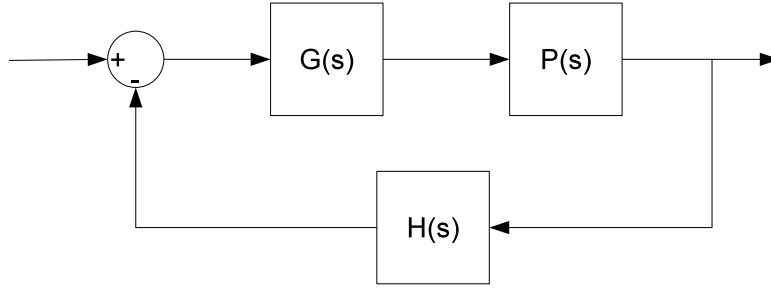


Figura 8.1: Tipico sistema di controllo in retroazione dove $P(s)$ è il processo da controllare, $G(s)$ è il controllore ed $H(s)$ è la funzione di trasferimento del trasduttore.

Si consideri un generico processo descritto dalla seguente funzione di trasferimento:

$$P(s) = \frac{N(s)}{D(s)} = \frac{\prod_{i=1}^m (s - z_i)}{\prod_{i=1}^n (s - p_i)} \quad (8.5)$$

Si costruisca un sistema in retroazione come indicato nella figura 8.1. La funzione di trasferimento a ciclo chiuso considerando $G(s) = K$ and $H(s) = 1$ risulta:

$$W(s) = K \cdot \frac{\prod_{i=1}^m (s - z_i)}{\prod_{i=1}^n (s - p_i) + K \cdot \prod_{i=1}^m (s - z_i)} \quad (8.6)$$

Osservando il polinomio caratteristico:

$$f(s, k) = \prod_{i=1}^n (s - p_i) + K \cdot \prod_{i=1}^m (s - z_i) \quad (8.7)$$

si nota che:

- Poli della $W(s)$ dipendono dal guadagno K , dal numeratore $N(s)$ e dal denominatore $D(s)$ del processo $P(s)$
- Nel caso $K = 0$ i poli della $W(s)$ sono esattamente i poli della $P(s)$

A partire dalla condizione $f(s, k) = 0$ è possibile definire le due equazioni fondamentali per il tracciamento esatto del luogo delle radici.

- Equazione di Tracciamento

$$\angle K = \pi + \sum_{i=1}^n \angle(s - p_i) - \sum_{i=1}^m \angle(s - z_i) + 2h\pi \quad (8.8)$$

- Equazione di Taratura

$$|K| = \frac{\prod_{i=1}^n |(s - p_i)|}{\prod_{i=1}^m |(s - z_i)|} \quad (8.9)$$

Osservazioni:

- L'equazione del tracciamento, eq. (8.8), fornisce una condizione per la verifica della appartenenza dei punti del piano complesso \mathcal{C} al luogo delle radici.
- L'equazione di taratura, eq. (8.9) fornisce una condizione per determinare l'esatto valore del guadagno \bar{K} per cui si ha un determinato polo \bar{s} .

8.2 Luogo Approssimato

Nella sezione precedente sono state introdotte le due equazioni fondamentali per il tracciamento esatto del luogo delle radici, (eq. (8.8) e (8.9)). Un impiego diretto di tali equazioni richiederebbe l'ispezione di tutti i punti del piano complesso limitandone di fatto un possibile utilizzo pratico. Tuttavia a partire da queste equazioni è possibile derivare una serie di regole che permettono il tracciamento del luogo delle radici in maniera qualitativa.

- R1 Il luogo delle radici è costituito da $2n$ rami (n rami positivi e n rami negativi)
- R2 Gli n rami completi passano per $K = 0$ per i poli della $P(s)$
- R3 L'asse reale appartenente al luogo delle radici ad eccezione di eventuali zeri del sistema a ciclo aperto. In particolare:
 - Un generico punto $p \in$ ramo positivo se lascia alla propria destra un numero dispari (ad esclusione del punto improprio) di poli e zeri contati con le loro molteplicità.
 - Un generico punto $p \in$ ramo negativo se lascia alla propria destra un numero pari (ad esclusione del punto improprio) di poli e zeri contati con le loro molteplicità.
- R4 Il luogo delle radici è simmetrico rispetto all'asse reale
- R5 Gli zeri del sistema a ciclo aperto \notin luogo delle radici. In particolare:

$$|K| \rightarrow \inf \begin{cases} m \text{ rami tendono agli zeri della } P(s) \\ n-m \text{ tendono al punto improprio } (\infty) \end{cases}$$

Di conseguenza, il punto improprio può essere interpretato come un particolare zero

- R6 I $2(n-m)$ rami che tendono al punto improprio vi tendono secondo $2(n-m)$ asintoti che dividono l'angolo giro in parti uguali. Le semirette che costituiscono gli asintoti formano una stella la cui origine è:

$$S_0 = \frac{\sum_i^n p_i - \sum_i^m z_i}{n-m} \quad (8.10)$$

Il segno degli asintoti (positivo/negativo) è alternato e la semiretta $[S_0, +\infty)$ è un asintoto del ramo negativo.

- R7 Il luogo delle radici può avere al massimo un numero di punti singolari pari a $n + m - 1$. Tali punti possono essere identificati in forma chiusa attraverso la:

$$\sum_i^m \frac{1}{(s - z_i)} - \sum_i^n \frac{1}{(s - p_i)} = 0 \quad (8.11)$$

- R8 Se p_0 è un punto di singolarità di molteplicità $\nu - 1$, in esso vi confluiscono 2ν segmenti del luogo i quali tendono al punto di singolarità secondo tangenti che spaziano l'angolo giro in parti uguali $\left(\frac{360^\circ}{2\nu}\right)$

- R9 I poli di molteplicità ν sono punti di singolarità di molteplicità $\nu - 1$
- R10 Tra 2 poli (zeri) contigui ci può essere un numero dispari di punti di singolarità, mentre tra 1 polo (zero) e 1 zero (polo) contigui ci può essere un numero pari di punti di singolarità

Procedura sistematica per il tracciamento qualitativo del luogo

Data un processo descritto dalla $P(s) = \frac{\prod_{i=1}^m (s - z_i)}{\prod_{i=1}^n (s - p_i)}$

- Calcolare l'eccesso poli zero ($n - m$) ed il numero massimo ammissibile di punti di singolarità ($n + m - 1$).
- Tracciare i poli $\{p_1, \dots, p_n\}$ e zeri $\{z_1, \dots, z_m\}$ sul piano complesso. Individuare i tratti dell'asse reale che appartengono al luogo (per $K > 0$ e/o $K < 0$).
- Identificare il centro stella S_0 e tracciare gli asintoti.
- Identificare i punti di singolarità
- Effettuare l'analisi per $|K| \rightarrow \infty$
- Orientare il luogo

8.3 Sintesi

Nella sezione precedente è stata presentata una metodologia per il tracciamento qualitativo del luogo delle radici. In questa sezione i concetti precedentemente presentati verranno impiegati ai fini della sintesi di un opportuno controllore in grado di stabilizzare a ciclo chiuso con retroazione unitaria un dato processo $P(s)$.

Si consideri un processo $P(s) = \frac{\prod_{i=1}^m (s - z_i)}{\prod_{i=1}^n (s - p_i)}$ dove $\{z_1, \dots, z_m\} \in \mathcal{C}^-$.

- Se il grado relativo del sistema è pari ad 1, ovvero ($n - m = 1$), allora il sistema è sempre stabilizzabile per valori elevati del guadagno positivo ($K \gg 0$).
- Se il grado relativo del sistema è pari a 2, ovvero ($n - m = 2$), allora vi sono 2 asintoti paralleli all'asse immaginario passanti per il centro stella S_0 . Nel caso il centro stella $S_0 < 0$ è minore di zero, allora per la stabilizzazione è sufficiente un guadagno elevato. Nel caso il centro stella è maggiore di zero $S_0 > 0$ si deve ricorrere ad una funzione compensatrice del tipo:

$$R(s) = \frac{s + z}{s + p} \quad (8.12)$$

la quale non altera il numero di asintoti ma permette di spostare il centro stella come segue:

$$\bar{S}_0 = S_0 - \frac{(p - z)}{2} \quad (8.13)$$

a questo punto la stabilità è garantita da un guadagno K sufficientemente elevato.

- Se il grado relativo del sistema è maggiore di 2, ovvero ($n - m > 2$), allora si aggiungono preliminarmente $n - m - 2$ zeri a parte reale negativa, e si determini una rete compensatrice come nel caso precedente. Si noti che il sistema così ottenuto è, in accordo a quanto detto precedentemente, stabilizzabile. Tuttavia, il controllore:

$$G(s) = K \frac{s + z}{s + p} (s + z_i) \dots (s + s_{n-m-2}) \quad (8.14)$$

risulta essere fisicamente non realizzabile. Per ovviare a tale problema è sufficiente aggiungere $n - m - 2$ poli del tipo $(1 + T_i s)$ con $T_i > 0$ molto piccolo. È infatti possibile dimostrare attraverso il criterio di Nyquist che tale l'aggiunta di tali termini non altera la stabilità del sistema. Infatti, si ha che l'aggiunta di un polo ad alta frequenza altera i diagrammi di Bode solo attorno all'origine quando si è sufficientemente lontani dal punto critico $(-1, 0)$.

Appendice A

Richiami di Algebra Lineare

In questo capitolo sono presentati alcuni concetti di algebra lineare. L'algebra lineare è quella branca della matematica che si occupa dello studio di vettori, spazi vettoriali (o spazi lineari), trasformazioni lineari e sistemi di equazioni lineari. Questi concetti sono di fondamentale importanza nella Teoria dei Sistemi in quanto sono alla base della rappresentazione di un sistema dinamico in spazio di stato. e risultano essere fondamentali

A.1 Spazi Vettoriali

Definizione A.1 Sia G un insieme e sia $*$ operazione binaria su G ($*$: $G \times G \rightarrow G$). Si dice che $(G, *)$ è un gruppo se:

- (a) vale proprietà associativa: $\forall a, b, c \in G$ si ha $(a * b) * c = a * (b * c)$;
- (b) esiste elemento neutro: $\exists e \in G$ tale che $\forall x \in G$ si ha $x * e = e * x = x$;
- (c) esiste inverso/opposto: $\forall x \in G \exists y \in G$ tale che $x * y = y * x = e$.

Se $\forall a, b \in G$ vale $a * b = b * a$, allora G si dice gruppo commutativo.

Estendendo tale concetto a due operazioni si costruisce un campo.

Definizione A.2 Sia K un insieme con due operazioni binarie $+$ e $*$. Si dice che $(K, +, *)$ è un campo se:

- (a) $(K, +)$ è un gruppo commutativo (con elemento neutro 0);
- (b) $(K \setminus \{0\}, *)$ è un gruppo commutativo (con elemento neutro 1);
- (c) vale proprietà distributiva: $\forall a, b, c \in K$ si ha $(a + b) * c = a * c + b * c$;

A questo punto si può introdurre il concetto di spazio vettoriale come segue.

Definizione A.3 Sia V un insieme e sia K un campo (ad esempio il campo dei reali \mathbb{R}). Si dice che l'insieme V è sostegno di uno spazio vettoriale (lineare) sul campo K se in V è definita un'operazione binaria interna $(+)$ per la quale $(V, +)$ è un gruppo commutativo (ossia un gruppo abeliano) ed è inoltre definita una legge di composizione esterna $(*) : K \times V \longrightarrow V$ (detta prodotto esterno o moltiplicazione per uno scalare) per la quale valgono le seguenti proprietà:

- (a) $(\alpha + \beta) * v = \alpha * v + \beta * v$ (Distributività del prodotto esterno rispetto all'addizione di scalari);
- (b) $\alpha * (v + u) = \alpha * v + \alpha * u$ (Distributività del prodotto esterno rispetto all'addizione di vettori);
- (c) $(\alpha * \beta) * v = \alpha * (\beta * v)$ (Associatività del prodotto esterno);
- (d) $1 * v = v$ (Neutralità di 1 rispetto al prodotto esterno).

In questo corso verranno considerati spazi vettoriali Euclidei n -dimensionali ($V \subseteq \mathbb{R}^n$), i cui elementi sono vettori colonna di numeri reali:

$$\vec{x} \in \mathbb{R}^n \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \vec{x} = [x_1, \dots, x_n]^T \quad (\text{A.1})$$

Definizione A.4 Dato uno spazio vettoriale $V \subseteq \mathbb{R}^n$, si definisce combinazione lineare di m vettori $\{x_1, \dots, x_m\} \in V$, la seguente quantità:

$$\sum_{i=1}^m \alpha_i x_i = \alpha_1 x_1 + \dots + \alpha_m x_m \quad (\text{A.2})$$

Definizione A.5 Dati m vettori $\{x_1, \dots, x_m\} \in V \subseteq \mathbb{R}^n$, questi si definiscono linearmente indipendenti se la loro combinazione lineare si annulla solo nel caso in cui tutti i coefficienti sono nulli (caso triviale), ovvero:

$$\sum_{i=1}^m \alpha_i x_i = 0 \quad \Longleftrightarrow \quad \alpha_1 = \dots = \alpha_m = 0 \quad (\text{A.3})$$

Definizione A.6 Sia V un spazio vettoriale in \mathbb{R}^n , dato un insieme $S \subset V$ si dice che S è un sottospazio di V se risulta essere chiuso rispetto ad addizione e moltiplicazione scalare:

$$\forall x, y \in S \quad \Rightarrow \quad \alpha x + \beta y \in S \quad \forall \alpha, \beta \in \mathbb{R} \quad (\text{A.4})$$

Definizione A.7 Dato uno spazio vettoriale V , un insieme $B \subseteq V$ si dice base di V se ogni elemento $u \in V$ si scrive in modo unico come combinazione lineare di elementi di B .

Si noti che la scrittura risulta essere unica solo nel caso in cui gli elementi di B sono linearmente indipendenti.

Definizione A.8 Dato uno spazio vettoriale V , un insieme $B_c \subseteq V$ si dice base canonica di V se questa è costituita da n vettori linearmente indipendenti $\{e_1, \dots, e_n\}$ del tipo:

$$e_i = [0, \dots, \underbrace{1}_i, 0, \dots, 0] \quad (\text{A.5})$$

Definizione A.9 Una norma su uno spazio vettoriale lineare (reale o complesso) X è una funzione

$$\|\cdot\| : X \longrightarrow [0, \infty) \quad (\text{A.6})$$

che verifica le seguenti condizioni:

- $\|x\| \geq 0 \ \forall x \in X$ e $\|x\| = 0 \iff x = 0$, (definita positività),
- $\|\alpha x\| = |\alpha| \|x\|$, (omogeneità),
- $\|x + y\| \leq \|x\| + \|y\|$, (disuguaglianza triangolare).

Definizione A.10 Si definisce norma euclidea, quella particolare norma dove la funzione $\|\cdot\|$ è definita come:

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2} \quad (\text{A.7})$$

Definizione A.11 Sia V uno spazio vettoriale definito in \mathbb{R}^n . Si definisce prodotto scalare $\langle \cdot, \cdot \rangle$ una funzione:

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}, \quad \langle x, y \rangle, \quad (\text{A.8})$$

definita come:

$$\langle x, y \rangle = x^T \cdot y = \sum_{i=1}^n x_i y_i \quad (\text{A.9})$$

che verifica le seguenti condizioni:

- $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in V$ (simmetria),
- $\langle \alpha x + y, z \rangle = \alpha \langle x, z \rangle + \langle y, z \rangle$ (bilinearità),
- $\langle x, x \rangle = \|x\|^2 > 0 \quad \forall x \neq 0$ (definita positività),
- $|\langle x, y \rangle| \leq \|x\| + \|y\|$ (disuguaglianza di Cauchy-Schwarz).

dove con il simbolo $\|\cdot\|$ si indica una opportuna norma.

Definizione A.12 Dati due vettori $x, y \in \mathbb{R}^n$, si dice che questi sono ortogonali se e solo se il loro prodotto scalare è nullo, ovvero:

$$x \perp y \iff \langle x, y \rangle = 0 \quad (\text{A.10})$$

Definizione A.13 Dati due vettori $x, y \in \mathbb{R}^n$, si dice che questi sono ortonormali se sono ortogonali tra loro ed hanno norma unitaria, ovvero:

- $\|x\| = \|y\| = 1$,
- $x \perp y$.

Definizione A.14 Si definisce spazio vettoriale euclideo $V \subseteq \mathbb{R}^n$, quel particolare spazio vettoriale per cui è definito un prodotto scalare Euclideo (ovvero definito rispetto alla norma Euclidea) $\langle \cdot, \cdot \rangle$.

Definizione A.15 Si definisce un vettore funzione del tempo, quel particolare vettore $x(t) \in \mathbb{R}^n$ dove ogni componente è funzione del tempo, ovvero:

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad (\text{A.11})$$

Definizione A.16 Si definisce derivata di un vettore funzione del tempo $x(t) \in \mathbb{R}^n$, quel particolare vettore $\dot{x}(t) \in \mathbb{R}^n$ dove ogni componente è la derivata della corrispondente componente del vettore $x(t)$, ovvero:

$$\dot{x}(t) = \frac{d}{dt}x(t) = \begin{bmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix} \quad (\text{A.12})$$

A.2 Matrici

Definizione A.17 Si definisce matrice A di dimensioni $m \times n$ uno schieramento di elementi organizzati in m righe (orizzontali) ed n colonne (verticali). Una generica matrice di numeri reali $A \in \mathbb{R}^{m \times n}$ è descritta solitamente nel modo seguente:

$$M = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n-1} & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n-1} & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m-11} & a_{m-12} & \dots & a_{m-1n-1} & a_{m-1n} \\ a_{m1} & a_{m2} & \dots & a_{mn-1} & a_{mn} \end{bmatrix} \quad (\text{A.13})$$

Definizione A.18 Sia data una matrice $A \in \mathbb{R}^{n \times m}$. Si definisce polinomio caratteristico $p_A(\cdot)$ associato alla matrice A il polinomio definito nel modo seguente:

$$p_A(\lambda) = \det(A - \lambda I). \quad (\text{A.14})$$

Inoltre, si definisce equazione caratteristica la seguente espressione:

$$p_A(\lambda) = 0. \quad (\text{A.15})$$

Definizione A.19 Sia data una matrice $A \in \mathbb{R}^{n \times m}$, si definisce norma euclidea (indotta) $\|\cdot\|$ la quantità:

$$\|A\| = \max\{\|Ax\|_e : x \in \mathbb{R}^m, \|x\|_e \leq 1\} \quad (\text{A.16})$$

dove $\|\cdot\|_e$ è la norma euclidea per vettori.

Definizione A.20 Una matrice $A \in \mathbb{R}^{n \times n}$ si definisce invertibile se esiste una matrice $A^{-1} \in \mathbb{R}^{n \times n}$ tale che:

$$A \cdot A^{-1} = A^{-1} \cdot A = I \quad (\text{A.17})$$

dove $I \in \mathbb{R}^{n \times n}$ è la matrice identità di dimensione $n \times n$.

Definizione A.21 Siano $A, B \in \mathbb{R}^{n \times n}$ due matrici quadrate $n \times n$. Esse si definiscono simili se esiste una matrice invertibile P tale che:

$$A = P^{-1} \cdot B \cdot P. \quad (\text{A.18})$$

Definizione A.22 Sia $M(n) \in \mathbb{R}^{n \times n}$ lo spazio vettoriale delle matrici quadrate $n \times n$ a valori in un campo K (ad esempio, il campo dei numeri reali o complessi), si definisce determinante di una matrice $A \in M(n)$, una applicazione lineare definita come segue:

$$\det : M(n) \longrightarrow K. \quad (\text{A.19})$$

Inoltre data una generica matrice $A \in M(n)$ definita come segue:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \quad (\text{A.20})$$

si ha che il determinante $\det(A)$ si può indicare come:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \quad (\text{A.21})$$

e può essere calcolato attraverso lo sviluppo di Laplace scegliendo la i -esima riga come segue:

$$\begin{aligned} \det(A) &= \sum_{j=1}^n a_{ij} c_{ij} \\ &= \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(A^{ij}) \end{aligned} \quad (\text{A.22})$$

dove $c_{ij} = (-1)^{i+j} \det(A^{ij})$ è il complemento algebrico della coppia (i, j) .

Teorema A.1 Sia $A_{inv}(n) \in \mathbb{R}^{n \times n}$ lo spazio vettoriale delle matrici quadrate invertibili $n \times n$ a valori in un campo K . Allora $\forall A, B \in M(n)$ e $\forall k \in K$ valgono le seguenti proprietà per il determinante:

- $\det(A B) = \det(A) \det(B)$
- $\det(A^{-1}) = \frac{1}{\det(A)}$
- $\det(k A) = k^n \det(A)$
- $\det(A^T) = \det(A)$
- $\det(\text{diag}(a_1, \dots, a_n)) = \prod_i^n a_i$

Teorema A.2 (Invertibilità di una matrice) Una matrice $A \in \mathbb{R}^{n \times m}$ è invertibile se e solo se il suo determinante è diverso da zero, ovvero:

$$\forall A \in \mathbb{R}^{n \times n} \quad \exists A^{-1} \in \mathbb{R}^{n \times n} \iff \det(A) \neq 0. \quad (\text{A.23})$$

Inoltre se A è invertibile allora la sua matrice inversa A^{-1} si può calcolare come segue:

$$A^{-1} = \frac{1}{\det(A)} \text{Cof}^T(A) \quad (\text{A.24})$$

dove $\text{Cof}(A)$ è la matrice dei cofattori, che si può calcolare come segue:

$$\text{Cof}(A) = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{bmatrix}, \quad c_{ij} = (-1)^{i+j} \det(A^{ij}) \quad (\text{A.25})$$

dove A^{ij} è la matrice $n-1 \times n-1$ ottenuta eliminando dalla matrice A la i -esima riga e la j -esima colonna.

Definizione A.23 Sia data una matrice $A \in \mathbb{R}^{n \times n}$, si definisce traccia $\text{tr}(A)$ della matrice A la somma di tutti gli elementi sulla diagonale principale, ovvero:

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}, \quad (\text{A.26})$$

dove a_{ii} rappresenta l'elemento sulla i -esima e i -esima colonna della matrice A .

Definizione A.24 Sia data una matrice quadrata $A \in \mathbb{R}^{n \times n}$, essa si definisce nilpotente se esiste un intero positivo k tale che:

$$A^k = 0. \quad (\text{A.27})$$

Inoltre, si definisce ordine di nilpotenza il più piccolo k per cui tale condizione è verificata.

Definizione A.25 Si definisce matrice funzione del tempo, quella particolare matrice $A(t) \in \mathbb{R}^{n \times m}$ dove ogni componente è funzione del tempo, ovvero:

$$A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \dots & a_{1n-1}(t) & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \dots & a_{2n-1}(t) & a_{2n}(t) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m-11}(t) & a_{m-12}(t) & \dots & a_{m-1n-1}(t) & a_{m-1n}(t) \\ a_{m1}(t) & a_{m2}(t) & \dots & a_{mn-1}(t) & a_{mn}(t) \end{bmatrix} \quad (\text{A.28})$$

Definizione A.26 Si definisce derivata di una matrice funzione del tempo $A(t) \in \mathbb{R}^{n \times m}$, quella particolare matrice $\dot{A}(t) \in \mathbb{R}^{n \times m}$ dove ogni elemento è la derivata del corrispondente elemento della matrice $A(t)$, ovvero:

$$\dot{A}(t) = \begin{bmatrix} \dot{a}_{11}(t) & \dot{a}_{12}(t) & \dots & \dot{a}_{1n-1}(t) & \dot{a}_{1n}(t) \\ \dot{a}_{21}(t) & \dot{a}_{22}(t) & \dots & \dot{a}_{2n-1}(t) & \dot{a}_{2n}(t) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dot{a}_{m-11}(t) & \dot{a}_{m-12}(t) & \dots & \dot{a}_{m-1n-1}(t) & \dot{a}_{m-1n}(t) \\ \dot{a}_{m1}(t) & \dot{a}_{m2}(t) & \dots & \dot{a}_{mn-1}(t) & \dot{a}_{mn}(t) \end{bmatrix} \quad (\text{A.29})$$

A.3 Trasformazioni Lineari

Definizione A.27 Siano V e W due spazi vettoriali definiti sullo stesso campo K . Una funzione $f : V \rightarrow W$ è una trasformazione lineare (applicazione lineare) se soddisfa le seguenti proprietà:

- $f(x + y) = f(x) + f(y)$ (linearità),
- $f(\alpha x) = \alpha f(x)$ (omogeneità di grado 1),

per ogni coppia di vettori $x, y \in V$ e per ogni scalare $\alpha \in K$.

Definizione A.28 Siano V e W due spazi vettoriali definiti sullo stesso campo K , e $f : V \rightarrow W$ una applicazione lineare. Siano inoltre $B_V = \{v_1, \dots, v_n\}$ e $B_W = \{w_1, \dots, w_m\}$ due basi rispettivamente per V e W . La matrice A associata a f nelle basi B_V e B_W :

$$A_{B_V, B_W} : V \rightarrow W, \quad (\text{A.30})$$

è la matrice $m \times n$ avente nella i -esima colonna le coordinate del vettore $f(v_i)$ rispetto alla base B_W .

Definizione A.29 Siano $V \subseteq \mathbb{R}^n$ e $W \subseteq \mathbb{R}^m$ due spazi vettoriali definiti sullo stesso campo K , e $f : V \rightarrow W$ una applicazione lineare. Sia inoltre A la matrice associata a f . Si definisce immagine di A , quel particolare sottospazio di W tale che:

$$R(A) = \{y \in W : y = Ax \quad \text{per qualche } x\} \quad (\text{A.31})$$

Definizione A.30 Siano $V \subseteq \mathbb{R}^n$ e $W \subseteq \mathbb{R}^m$ due spazi vettoriali definiti sullo stesso campo K , e $f : V \rightarrow W$ una applicazione lineare. Sia inoltre A la matrice associata a f . Si definisce rango di A , la dimensione dell'immagine di A , ovvero:

$$\text{rank}(A) = \dim R(A) \quad (\text{A.32})$$

Definizione A.31 Siano $V \subseteq \mathbb{R}^n$ e $W \subseteq \mathbb{R}^m$ due spazi vettoriali definiti sullo stesso campo K , e $f : V \rightarrow W$ una applicazione lineare. Sia inoltre A la matrice associata a f . Si definisce nullo di A , quel particolare sottospazio di V tale che:

$$N(A) = \{x \in V : Ax = 0\} \quad (\text{A.33})$$

Definizione A.32 Siano $V \subseteq \mathbb{R}^n$ e $W \subseteq \mathbb{R}^m$ due spazi vettoriali definiti sullo stesso campo K , e $f : V \rightarrow W$ una applicazione lineare. Sia inoltre A la matrice associata a f . Si definisce nullità di A , la dimensione del nullo di A , ovvero:

$$\text{nullity}(A) = \dim N(A) \quad (\text{A.34})$$

Teorema A.3 (Teorema fondamentale dell'algebra lineare) Siano \mathbb{R}^n e \mathbb{R}^m due spazi vettoriali definiti sullo stesso campo K . Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ una applicazione lineare con A la matrice ad essa associata. Sia inoltre $f^{-1} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ l'applicazione lineare inversa ed A^T la matrice ad essa associata. Allora per le due applicazioni f e f^{-1} valgono le seguenti relazioni sugli spazi immagine e spazi nulli:

$$R(A)^\perp = N(A^T) \quad (\text{A.35})$$

$$R(A^T)^\perp = N(A) \quad (\text{A.36})$$

ed inoltre valgono la seguente relazione sulle dimensioni:

$$\dim \mathbb{R}^m = \text{rank}(A) + \text{nullity}(A^T) \quad (\text{A.37})$$

$$\dim \mathbb{R}^n = \text{rank}(A^T) + \text{nullity}(A) \quad (\text{A.38})$$

dove $\text{rank}(A) = \text{rank}(A^T)$.

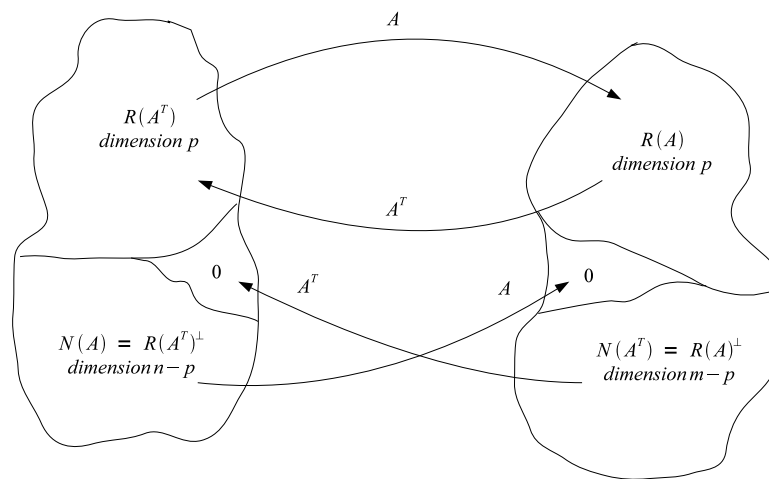


Figura A.1: Decomposizione della matrice A associata all'trasformata lineare f .

Appendice B

Trasformata di Laplace

In questo capitolo verrà illustrato sommariamente il concetto di trasformazione secondo Laplace. In particolare, la trasformata di Laplace è una funzione lineare che permette di passare dallo studio di una variabile temporale (reale) allo studio di una variabile complessa, e viceversa. Tale trasformata integrale ha numerose proprietà che la rendono utile per l'analisi dei sistemi dinamici lineari. Tra i vantaggi più significativi va menzionato il fatto che che l'integrale e la derivata di una funzione diventano rispettivamente una divisione e una moltiplicazione per la variabile complessa, e che l'operazione di convoluzione tra funzioni diventa una semplice operazione di moltiplicazione tra funzioni. Per finire, si ricorda che attraverso l'impiego della trasformata di Laplace le equazioni integrali e le equazioni differenziali vengono trasformate in equazioni polinomiali con ovvi vantaggi in termini risolutivi. Ovviamente tale operazione di trasformazione affinché sia di utilità pratica deve possedere anche una operazione di antitrasformazione che sia biunivoca. Tale operazione di inversione esiste ed è noto sotto il nome di integrale di Bromwich (o di Bromwich-Mellin o anche di Riemann-Fourier) il quale è un integrale complesso.

B.1 Trasformata di Laplace

Definizione B.1 Sia $f(t)$ una funzione del tempo integrabile tale che $f(t) = 0 \forall t \leq 0$. La trasformata di Laplace $F(s) = \mathcal{L}\{f(t)\}$ associata a $f(t)$ è la funzione di variabile complessa $s = \sigma + j\omega$ definita attraverso l'integrale di Laplace come segue:

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^{\infty} f(t)e^{-st}dt \quad (\text{B.1})$$

B.1.1 Proprietà

Teorema B.1 (*Linearità*)

$$\mathcal{L}\{\alpha f_1(t) + f_2(t)\} = \alpha \mathcal{L}\{f_1(t)\} + \mathcal{L}\{f_2(t)\} \quad (\text{B.2})$$

Teorema B.2 (*Ritardo temporale*)

$$\mathcal{L}\{f(t - \Delta)\} = \mathcal{L}\{f(t)\}e^{-s\Delta} \quad (\text{B.3})$$

Teorema B.3 (*Ritardo in frequenza*)

$$\mathcal{L}\{f(t)e^{\alpha t}\} = F(s - \alpha) \quad (\text{B.4})$$

Teorema B.4 (*Derivazione*)

$$\mathcal{L}\{\dot{f}(t)\} = sF(s) - f(0) \quad (\text{B.5})$$

Teorema B.5 (*Integrazione*)

$$\mathcal{L}\left\{\int_0^t f(\tau) d\tau\right\} = \frac{F(s)}{s} \quad (\text{B.6})$$

Teorema B.6 (*Convolvere*)

$$\mathcal{L}\{h(t) * u(t)\} = H(s)U(s) \quad (\text{B.7})$$

dove col simbolo $*$ si indica l'operatore di convoluzione definito come segue:

$$h(t) * u(t) = \int_{-\infty}^{+\infty} h(t - \tau) u(\tau) d\tau \quad (\text{B.8})$$

B.2 Antitrasformata di Laplace

Teorema B.7 Sia $F(s)$ la trasformata di Laplace di una funzione $f(t)$. Allora è possibile calcolare la funzione del tempo $f(t)$ a partire dalla trasformata $F(s)$ attraverso l'operazione di antitrasformazione come segue:

$$f(t) = \mathcal{L}^{-1}\{F(s)\} = \int_{\sigma-j\infty}^{\sigma+j\infty} F(s)e^{st} ds \quad (\text{B.9})$$

Si noti che l'operazione di antitrasformazione fornita dalla (B.9) risulta essere di poca utilità pratica. Infatti, essa richiede la valutazione di un integrale generalmente difficile da calcolare. Per ovviare al calcolo di tale integrale, si può far ricorso ad una tecnica alternativa che si basa sulla espansione della funzione razionale $F(s)$ in fratti semplici, la cui inversione risulta essere estremamente più agevole.

B.2.1 Espansione in Fratti Semplici

Si consideri una funzione razionale $F(s)$ strettamente propria definita come segue:

$$F(s) = \frac{Q(s)}{P(s)} = \frac{q_{n-1}s^{n-1} + \dots + q_1s + q_0}{s^n + p_{n-1}s^{n-1} + \dots + p_1s + p_0} \quad (\text{B.10})$$

$$= \frac{Q(s)}{(s - \lambda_1)(s - \lambda_2) \dots (s - \lambda_n)} \quad (\text{B.11})$$

dove i numeri complessi $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ sono le radici del polinomio $P(s)$ (poli). Allora è possibile scrivere la funzione razionale $F(s)$ attraverso una espansione in fratti semplici come segue:

$$F(s) = \frac{R_1}{(s - \lambda_1)} + \frac{R_2}{(s - \lambda_2)} + \dots + \frac{R_n}{(s - \lambda_n)} \quad (\text{B.12})$$

dove i numeri complessi $R_1, R_2, \dots, R_n \in \mathbb{C}$ sono i residui. A partire da tale scomposizione in fratti semplici l'antitrasformata può essere agevolmente calcolata come segue:

$$f(t) = R_1 e^{\lambda_1 t} + R_2 e^{\lambda_2 t} + \dots + R_n e^{\lambda_n t} \quad (\text{B.13})$$

Per il calcolo dei residui, tre differenti metodi possono essere utilizzare in funzione del fatto che i poli siano reali distinti, reali a molteplicità algebrica maggiore di uno o complessi e coniugati.

Caso poli reali distinti

Nel caso di poli $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ tutti reali e distinti il calcolo dei residui può essere agevolmente effettuato utilizzando la seguente formula:

$$R_i = [(s - \lambda_i) F(s)]_{s=\lambda_i} \quad (\text{B.14})$$

Caso poli reali a molteplicità algebrica maggiore di uno

Nel caso di poli $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ tutti reali con molteplicità algebrica maggiore di uno il calcolo dei residui può essere effettuato come segue. Sia λ_i un polo con molteplicità algebrica pari ad μ_i , il suo contributo nel calcolo della scomposizione in fratti semplici è il seguente:

$$F_i(s) = \sum_j^{\mu_i} \frac{R_{i,j}}{(s - \lambda_i)^j} = \frac{R_{i,1}}{(s - \lambda_i)} + \frac{R_{i,2}}{(s - \lambda_i)^2} + \dots + \frac{R_{i,\mu_i}}{(s - \lambda_i)^{\mu_i}} \quad (\text{B.15})$$

dove i residui associati al polo λ_i possono essere calcolati come segue:

$$R_{i,\mu_i-j} = \frac{1}{j!} \left[\frac{d^j}{ds} \left((s - \lambda_i)^{\mu_i} F(s) \right) \right]_{s=\lambda_i} \quad (\text{B.16})$$

Caso poli complessi coniugati

Nel caso di poli $\lambda_1 \dots, \lambda_n \in \mathbb{C}$ con alcune coppie di poli complesse e coniugate il calcolo dei residui può essere effettuato come segue. Sia λ_i, λ_{i+1} una coppia di poli complessi e coniugati definita come segue:

$$\lambda_i = \sigma_i + j\omega_i \quad (\text{B.17})$$

$$\lambda_{i+1} = \sigma_i - j\omega_i = \bar{\lambda}_i \quad (\text{B.18})$$

con i rispettivi residui:

$$R_i = \frac{1}{2} C_i - \frac{1}{2} j S_i \quad (\text{B.19})$$

$$R_{i+1} = \frac{1}{2} C_i + \frac{1}{2} j S_i = \bar{R}_i \quad (\text{B.20})$$

Si ha che l'espansione in fratti semplici può essere scritta come:

$$\frac{R_i}{s - \lambda_i} + \frac{\bar{R}_i}{s - \bar{\lambda}_i} = \frac{C_i(s - \sigma_i) + S_i\omega_i}{(s - \sigma_i)^2 + \omega_i^2} \quad (\text{B.21})$$

dove $C_i = 2\text{Re}(R_i)$ e $S_i = -2\text{Im}(R_i)$ possono essere calcolati come segue:

$$2j R_i = S_i + j C_i = \frac{1}{\omega_i} \left[(s - \sigma_i)^2 + \omega_i^2 \right] p(s) \Big|_{s=\sigma_i+j\omega_i} \quad (\text{B.22})$$

L'antitrasformata associata a tale coppia di poli complessi e coniugati è la seguente:

$$C_i e^{\sigma_i t} \cos \omega_i t + S_i e^{\sigma_i t} \sin \omega_i t \quad (\text{B.23})$$

il quale può alternativamente essere scritto come:

$$\rho_i e^{\sigma_i t} \sin(\omega_i t + \phi_i) \quad (\text{B.24})$$

dove ρ_i e ϕ_i sono definiti in funzione di $S_i + j C_i = \rho_i e^{j\phi_i}$ come segue:

$$\rho_i = \sqrt{S_i^2 + C_i^2} = 2|R_i| \quad (\text{B.25})$$

$$\phi_i = \tan^{-1} \frac{C_i}{S_i} = 90^\circ + \angle R_i \quad (\text{B.26})$$

B.3 Trasformate Note

Nella Tabella B.3 vengono riportare le trasformate di maggior interesse.

$f(t) = \mathcal{L}^{-1}\{F(s)\}(t)$	$F(s) = \mathcal{L}\{f(t)\} \int_0^\infty e^{-st} f(t) dt$
1	$\frac{1}{s}, \quad s > 0$
$\frac{t^n}{n!}$	$\frac{1}{s^{n+1}}, \quad s > 0$
e^{at}	$\frac{1}{(s-a)}, \quad s > 0$
$\sin(bt)$	$\frac{b}{s^2 + b^2}, \quad s > 0$
$\cos(bt)$	$\frac{s}{s^2 + b^2}, \quad s > 0$
$e^{at} \frac{t^n}{n!}$	$\frac{1}{(s-a)^{n+1}}, \quad s > 0$
$e^{at} \sin(bt)$	$\frac{b}{(s-a)^2 + b^2}, \quad s > 0$
$e^{at} \cos(bt)$	$\frac{(s-a)}{(s-a)^2 + b^2}, \quad s > 0$
$t \sin(bt)$	$\frac{2bs}{(s^2 + b^2)^2}, \quad s > 0$
$t \cos(bt)$	$\frac{s^2 - b^2}{(s^2 + b^2)^2}, \quad s > 0$

Tabella B.1: Tabella delle trasformate di Laplace delle funzioni del tempo di interesse.

Appendice C

Matlab per la Teoria dei Sistemi

In questo capitolo verranno illustrati gli aspetti di interesse per l'applicazione di Matlab alla Teoria dei Sistemi. In particolare siamo interessati alla *Control System Toolbox*. Si ricorda che l'obiettivo di questo capitolo è solo quello di rendere il lettore familiare con tale toolbox, per una versione completa di tutte le funzionalità da essa offerte, si consiglia di far riferimento al manuale di Matlab, la cui versione in formato elettronico può essere acceduta digitando al prompt di Matlab

```
>> doc control
```

C.1 Rappresentazione di Modelli Lineari

La *Control System Toolbox* fornisce vari modelli per la rappresentazione dei sistemi lineari, in particolare è possibile avere:

- Rappresentazione in Spazio di Stato (SS),
- Rappresentazione attraverso Funzione di Trasferimento (TF).

C.1.1 Rappresentazione in Spazio di Stato (State-Space Representation)

La rappresentazione in spazio di stato può essere costruita attraverso l'utilizzo del comando `ss`. In particolare dato il sistema descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= 3x_1 - 1x_2 + 2x_3 + x_4 - u \\ \dot{x}_2 &= x_1 + 5x_2 - x_3 + u \\ \dot{x}_3 &= 3x_3 + u \\ \dot{x}_4 &= x_3 + 4x_4 \\ y &= 2x_1 + 3x_2 + 3x_3 + 3x_4 \end{cases}$$

la relativa rappresentazione in Matlab è la seguente:

```
>> A = [ 3   -1   2   1
```



```

    1     5     -1     0
    0     0     3     0 ;
    0     0     1     4 ]

```

```
A =
```

```

    3     -1     2     1
    1     5     -1     0
    0     0     3     0
    0     0     1     4

```

```
>> B = [ -1  1  1  0 ]'
```

```
B =
```

```

   -1
    1
    1
    0

```

```
>> C = [ 2  3  3  4 ]
```

```
C =
```

```

    2     3     3     4

```

```
>> D = 0
```

```
D =
```

```

    0

```

```
>> sys = ss(A, B, C, D)
```

```
a =
```

```

      x1  x2  x3  x4
x1    3  -1   2   1
x2    1   5  -1   0
x3    0   0   3   0
x4    0   0   1   4

```

```
b =
```

```

      u1
x1   -1
x2    1
x3    1
x4    0

```

```
c =
```

```

      x1  x2  x3  x4
y1    2   3   3   4

```

```
d =
```

```

      u1
y1    0

```

```
Continuous-time model.
```

C.1.2 Rappresentazione attraverso Funzione di Trasferimento (Transfer Function Representation)

La rappresentazione attraverso funzione di trasferimento può essere costruita attraverso l'utilizzo del comando `tf`. In particolare dato il sistema descritto dal seguente insieme di equazioni:

$$H(s) = \frac{4s^2 + 2s + 5}{s^5 + 3s^4 + 2s^3 + 5s^2 + 8s + 1} \quad (\text{C.1})$$

la relativa rappresentazione in Matlab è la seguente:

```
>> num = [ 4 2 5 ]

num =

      4      2      5

>> den = [ 1 3 2 5 8 1 ]

den =

      1      3      2      5      8      1

>> hs = tf(num,den)

Transfer function:
      4 s^2 + 2 s + 5
-----
s^5 + 3 s^4 + 2 s^3 + 5 s^2 + 8 s + 1
```

Per il calcolo dei residui si può far riferimento al comando `residue` come segue.

```
>> num =[ 1 4]

num =

      1      4

>> den = [1 2 10]

den =

      1      2     10

>> [R,P]=residue(num,den)

R =

    0.5000 - 0.5000i
    0.5000 + 0.5000i

P =

   -1.0000 + 3.0000i
   -1.0000 - 3.0000i
```

Appendice D

Vecchi Compiti di Esame

A seguire è riportata una serie di vecchi compiti di esame. Ovviamente per ora il materiale scarseggia ma con gli anni aumenterà ;-)



Laurea Specialistica in
Ingegneria Gestionale
e dell'Automazione
Ing. Andrea Gasparri
3 Novembre 2008

Cognome	Nome	Matricola	E-mail

1. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= x_1 + 3x_2 + x_3 + u \\ \dot{x}_2 &= 3x_2 - 2x_3 + u \\ \dot{x}_3 &= 2x_2 + 3x_3 + u \\ y &= x_1 + 2x_3 \end{cases}$$

dove si assuma come ingresso il segnale $u(t) = 2\sigma_{-1}(t)$ e stato iniziale $x_0 = [1 \ 1 \ 1]^T$,

- Calcolare l'evoluzione libera al tempo $t = 3$,
- Calcolare l'evoluzione forzata al tempo $t = 2$,
- Calcolare la risposta completa al tempo $t = 7$,
- Valutare la stabilità del sistema.

2. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= 2x_1 + 1x_2 - 1x_3 + u \\ \dot{x}_2 &= 3x_2 \\ \dot{x}_3 &= -1x_2 + 3x_3 \\ \dot{x}_4 &= x_1 + x_2 + 3x_4 + x_5 \\ \dot{x}_5 &= x_1 - x_2 + x_3 + x_5 - u \\ y &= 4x_1 + 3x_2 + 3x_3 + 3x_4 + 3x_5 \end{cases}$$

portare il sistema in forma diagonale se possibile, altrimenti in forma canonica di Jordan.

3. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= x_1 - x_2 - 2x_3 - 2x_4 - u \\ \dot{x}_2 &= 2x_2 - 2x_4 - u \\ \dot{x}_3 &= x_2 + 3x_3 + 2x_4 + 2u \\ \dot{x}_4 &= 4x_4 + u \end{cases}$$

$$y = x_2 + 2x_4$$

- Determinare la matrice di trasformazione Q_c per la forma di Kalman di Controllabilità,
 - Determinare la matrice di trasformazione Q_o per la forma di Kalman di Osservabilità,
 - Determinare osservabilità e controllabilità di tutte le dinamiche.
-

4. Dato il sistema dinamico non lineare descritto dalle seguenti equazioni:

$$\begin{cases} \dot{x}_1 &= -x_1 \sqrt{x_2} + 8u \\ \dot{x}_2 &= -x_2 u^2 + 2x_1 \end{cases}$$

$$y = x_1 + x_2$$

si calcoli il modello linearizzato nell'intorno del punto di equilibrio corrispondente ad ingresso costante $u(t) = 2$.



Laurea Specialistica in
Ingegneria Gestionale
e dell'Automazione
Ing. Andrea Gasparri
12 Novembre 2008

Cognome	Nome	Matricola	E-mail

1. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= x_1 + 3x_2 + x_3 + u \\ \dot{x}_2 &= 3x_2 + u \\ \dot{x}_3 &= 2x_3 + u \end{cases}$$

$$y = x_1 + x_2 + 3x_3$$

dove si assuma come ingresso il segnale $u(t) = 3\sigma_{-1}(t)$ e stato iniziale $x_0 = [1 \ 1 \ 1]^T$,

- Calcolare l'evoluzione libera al tempo $t = 3$,
- Calcolare l'evoluzione forzata al tempo $t = 2$,
- Calcolare la risposta completa al tempo $t = 7$,
- Valutare la stabilità del sistema.

2. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= 1x_1 - 1x_2 + 1x_3 + u \\ \dot{x}_2 &= 3x_1 + 4x_2 + 1x_4 \\ \dot{x}_3 &= 2x_1 + 1x_2 + 2x_3 \\ \dot{x}_4 &= -2x_1 - 1x_2 + 2x_4 \end{cases}$$

$$y = 3x_1 + 3x_2 + 2x_3 + 4x_4$$

portare il sistema in forma diagonale se possibile, altrimenti in forma canonica di Jordan.

3. Dato il sistema dinamico descritto dal seguente insieme di equazioni:

$$\begin{cases} \dot{x}_1 &= 5x_1 + 2x_2 + 1x_3 \\ \dot{x}_2 &= 1x_1 + 4x_2 - 1x_3 + u \\ \dot{x}_3 &= -1x_1 - 2x_2 + 3x_3 \end{cases}$$
$$y = 1x_1 + 2x_2 + 1x_3$$

- a) Determinare la matrice di trasformazione Q_c per la forma di Kalman di Controllabilità,
- b) Determinare la matrice di trasformazione Q_o per la forma di Kalman di Osservabilità,
- c) Determinare osservabilità e controllabilità di tutte le dinamiche.

4. Dato il sistema dinamico non lineare descritto dalle seguenti equazioni:

$$\begin{cases} \dot{x}_1 &= x_1 + 2u \\ \dot{x}_2 &= \frac{1}{2}\sqrt{x_2}x_3 + u \\ \dot{x}_3 &= -x_1^2 + x_2 \end{cases}$$
$$y = x_1^2 + x_2 + \log x_3$$

si calcoli il modello linearizzato nell'intorno del punto di equilibrio corrispondente ad ingresso costante $u(t) = 2$.

Bibliografia

- <http://www.matrixanalysis.com/>