

# Appunti del corso di Model Predictive Control

*versione 0.6*

VITO ANDREA RACANELLI  
vitoandrea.racanelli@poliba.it

17 gennaio 2026



# Indice

<b>Prefazione</b>	<b>III</b>
Changelog . . . . .	III
<b>1 Introduzione ai sistemi</b>	<b>1</b>
1.1 Classificazione dei Sistemi . . . . .	2
<b>2 I diagrammi a blocchi</b>	<b>3</b>
2.1 Elementi fondamentali . . . . .	4
<b>3 Le funzioni di trasferimento</b>	<b>7</b>
3.1 La trasformata di Fourier . . . . .	7
3.2 La trasformata di Laplace . . . . .	8
3.2.1 Trasformate notevoli . . . . .	8
3.2.2 Proprietà della Trasformata di Laplace . . . . .	9
3.2.3 Teoremi fondamentali . . . . .	9
3.3 L'antitrasformata di Laplace . . . . .	9
3.3.1 $m > n$ Funzione razionale fratta impropria . . . . .	9
3.3.2 $m = n$ Funzione razionale fratta propria ma non strettamente . . . . .	10
3.3.3 $m < n$ Funzione razionale fratta strettamente propria . . . . .	10
<b>4 Funzioni di trasferimento</b>	<b>13</b>
4.1 Forme di una funzione di trasferimento . . . . .	13
4.1.1 Forma polinomiale . . . . .	13
4.1.2 Forma fattorizzata . . . . .	13
4.1.3 Forma in costanti di tempo . . . . .	14
4.2 Diagramma poli-zeri . . . . .	15
4.3 Condizione di realizzabilità fisica . . . . .	15
4.4 Funzioni di trasferimento con ritardo . . . . .	16
4.5 Risposte tipiche . . . . .	16
4.5.1 Risposta all'impulso . . . . .	16
4.5.2 Risposta al gradino . . . . .	16
4.5.3 Risposta ad un ingresso generico . . . . .	17
4.6 Sistemi del primo ordine . . . . .	20
4.6.1 Risposta di un sistema del primo ordine . . . . .	20
4.6.2 Tempo di assestamento . . . . .	20
4.6.3 Sistema del primo ordine con uno zero . . . . .	22
4.7 Sistemi del secondo ordine . . . . .	23
4.7.1 $ \delta  > 1$ . . . . .	23
4.7.2 $ \delta  = 1$ . . . . .	23
4.7.3 $ \delta  < 1$ . . . . .	24
4.7.4 Parametri caratteristici della risposta . . . . .	25
4.8 Sistemi del secondo ordine con zero . . . . .	29
4.9 Sistemi di ordine superiore al secondo . . . . .	31
4.10 I modelli in forma <i>spazio di stato</i> . . . . .	33
4.10.1 Sistemi statici e dinamici . . . . .	33
4.11 Sistemi lineari a tempo continuo . . . . .	35
4.11.1 Equazione di Lagrange . . . . .	35

4.12	Sistemi Lineari a Tempo Discreto . . . . .	36
<b>5</b>	<b>Stabilità</b>	<b>39</b>
5.1	Stabilità nei sistemi LTI TC . . . . .	39
5.2	Forma di Jordan . . . . .	39
5.3	Trasformazione per similitudine . . . . .	44
<b>6</b>	<b>Controllabilità e Osservabilità dei sistemi LTI</b>	<b>47</b>
6.1	Controllabilità . . . . .	47
6.1.1	Forma canonica di Controllabilità . . . . .	49
6.2	Osservabilità . . . . .	51
6.2.1	Forma canonica di Osservabilità . . . . .	51
<b>7</b>	<b>Retroazione di stato</b>	<b>55</b>
7.1	Stabizzabilità . . . . .	55
7.2	Progettazione di un sistema in retroazione di stato . . . . .	56
7.3	Progettazione di un regolatore in Matlab . . . . .	59
<b>8</b>	<b>Osservatori o stimatori</b>	<b>61</b>
8.1	Rilevabilità . . . . .	61
8.2	Progettazione di un osservatore . . . . .	61
8.3	Principio di separazione . . . . .	63
8.4	Progettazione di un osservatore in Matlab . . . . .	63
<b>9</b>	<b>Sistemi Non Lineari</b>	<b>65</b>
9.1	Introduzione . . . . .	65
9.2	Linearizzazione . . . . .	65
9.2.1	Linearizzazione di un sistema non lineare . . . . .	66
<b>10</b>	<b>Controllo ottimo</b>	<b>73</b>
10.1	La funzione di costo . . . . .	73
10.1.1	Scelta della funzione di costo . . . . .	77
10.2	Controllo ottimo senza vincoli . . . . .	79
10.3	Metodo dei Moltiplicatori di Lagrange . . . . .	85
10.3.1	Limitazioni . . . . .	85
10.4	Condizioni di Karush–Kuhn–Tucker . . . . .	87
10.5	Condizioni di secondo ordine . . . . .	95
<b>11</b>	<b>MPC in Matlab</b>	<b>99</b>
11.1	MPC Lineare in Matlab . . . . .	99
11.1.1	Comando <code>mpcmove</code> — Calcolo dell'azione di controllo ottimale . . . . .	102
11.2	MPC non lineare in MATLAB . . . . .	104
11.3	Interpretazione dei grafici di simulazione MPC . . . . .	108
11.3.1	Interpretazione dei grafici dell'ingresso di controllo . . . . .	108
11.3.2	Interpretazione dei grafici delle uscite . . . . .	111
11.3.3	Interpretazione dei tempi computazionali . . . . .	114
11.3.4	Interpretazione dell'evoluzione della funzione di costo . . . . .	117

# Prefazione

Questo documento è rivolto principalmente agli studenti del corso di Model Predictive Control del Politecnico di Bari, erogato al secondo anno delle lauree magistrali in Ingegneria dei Sistemi Medicali ed Ingegneria dell'Automazione. Spero che le informazioni contenute possano rendere più agevole e fruttuosa la preparazione dell'esame. Rispetto al programma del corso, gli appunti non sono ancora completi (e potrebbero contenere errori). Per alcuni contenuti, si rimanda ai materiali complementari utilizzati a lezione. Gentilmente, vi chiedo di fare osservazioni o segnalarmi errori via mail.

Vito Andrea Racanelli

## **Changelog**

### **Versione 0.5 – 16 gennaio 2026**

- Sistemati alcuni errori.
- Aggiunta una sezione riguardante il controllo ottimo senza vincoli.
- Aggiunto del materiale aggiuntivo riguardante MPC in matlab con interpretazione dei risultati.

### **Versione 0.6 – 16 gennaio 2026**

- Sistemata definizione 10.5.1.
- MPC in Matlab reso capitolo.



# CAPITOLO 1

## Introduzione ai sistemi

**Definizione 1.0.1** Un **sistema** può essere definito come un insieme di elementi interconnessi che interagiscono tra loro per raggiungere uno scopo comune. Questi elementi possono essere oggetti fisici, concetti astratti o processi, e le loro relazioni determinano il comportamento globale del sistema.

La teoria dei sistemi si occupa di studiare tali interazioni, cercando di individuare principi generali applicabili a diverse discipline, dalla biologia all'ingegneria, dall'economia alla sociologia. Analizzando un sistema, è possibile comprenderne la struttura, prevederne l'evoluzione e ottimizzarne le prestazioni. Questo approccio consente di affrontare problemi complessi in modo integrato, superando la visione frammentaria tipica di metodi più tradizionali. Si potrebbe cercare di ricavare le informazioni che ci interessano facendo degli esperimenti. Una sorta di *trial and error*.

Purtroppo (o per fortuna) non sempre è possibile condurre esperimenti diretti per comprendere il comportamento di un sistema. In molti casi, l'intervento diretto potrebbe essere costoso<sup>1</sup>, rischioso (es. una centrale nucleare) o impraticabile (es. il sistema non è stato costruito).

Qual è la soluzione a questi problemi? La soluzione consiste nel costruire un modello matematico del sistema, rappresentando i fenomeni essenziali attraverso le leggi della fisica, biologia, economia, ecc. Successivamente, si analizzano le equazioni del modello e si simula il sistema risolvendole, manualmente o tramite computer.

Il vantaggio della simulazione è il suo costo quasi nullo, ma la sua efficacia dipende dalla fedeltà del modello al sistema reale. In definitiva, **creare un buon modello è una vera e propria arte!**

È importante distinguere tra **sistema** e **modello**: mentre il sistema è la realtà concreta o astratta composta dagli elementi in interazione, il modello è una rappresentazione semplificata del sistema, costruita per analizzarlo, descriverlo o prevederne il comportamento. Un modello può essere matematico, concettuale o simulativo e serve per comprendere meglio il sistema originale, senza necessariamente replicarlo in ogni dettaglio. La modellizzazione è quindi un passaggio cruciale nella teoria dei sistemi, poiché permette di ridurre la complessità e focalizzarsi sugli aspetti essenziali per l'analisi e la risoluzione di problemi.

---

<sup>1</sup>E' interessante ascoltare l'aneddoto dell'Ing. Andrea Pontremoli, CEO di Dallara Group che racconta in una intervista il caso dei test riguardanti la Bugatti Veyron. <https://www.youtube.com/watch?v=uJa7Y5jF1Lg>

## 1.1 Classificazione dei Sistemi

I sistemi possono essere classificati in base a diversi criteri:

### In base alla dipendenza dalla variabile tempo

- **Sistemi statici** → L'output dipende solo dall'input attuale, senza memoria (es. resistenza elettrica).
- **Sistemi dinamici** → L'output dipende sia dall'input attuale che dagli stati precedenti (es. circuito RLC).

### In base alla natura del tempo

- **Tempo Continuo (TC)** → Descritti da equazioni differenziali (es. circuito RLC).
- **Tempo Discreto (TD)** → Descritti da equazioni alle differenze (es. crescita de popolazione studentesca).

### In base alla linearità

- **Lineari (L)** → Più semplici, risolvibili con metodi analitici, principio di sovrapposizione degli effetti.
- **Non lineari (NL)** → Più complessi da trattare.

### In base alla variazione nel tempo

- **Tempo-invarianti (LTI)** → Le proprietà non cambiano nel tempo; se l'input viene ritardato, l'output subisce lo stesso ritardo (es. circuito RC con componenti fissi).
- **Tempo-varianti (LTV)** → Le proprietà cambiano nel tempo; la risposta del sistema dipende dal momento in cui viene applicato l'input (es. pendolo con lunghezza variabile nel tempo).

### In base alla dipendenza dalle condizioni iniziali

- **Caotici** → Sensibili alle condizioni iniziali, comportamento imprevedibile a lungo termine (es. meteo).
- **Non Caotici** → Le piccole perturbazioni producono variazioni proporzionate.

### In base alla determinazione dell'evoluzione

- **Deterministici** → Lo stato futuro è univocamente determinato dallo stato attuale (es. pendolo senza attrito).
- **Stocastici** → L'evoluzione è influenzata da variabili aleatorie (es. modelli finanziari).

### In base alla dimensionalità dello spazio degli stati

- **Sistemi a bassa dimensionalità** → Descritti con poche variabili di stato (es. oscillatore armonico).
- **Sistemi ad alta dimensionalità** → Necessitano di molte variabili per essere descritti (es. reti neurali).



## I diagrammi a blocchi

Intuitivamente, possiamo considerare un sistema come una scatola nera (black-box) attraverso la quale gli ingressi (ossia le cause) agiscono per fornire un'uscita (ossia gli effetti). Per poter schematizzare il rapporto tra queste variabili, è molto utile fare una rappresentazione grafica attraverso i *diagrammi a blocchi*.

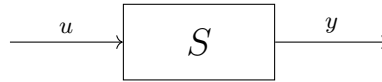
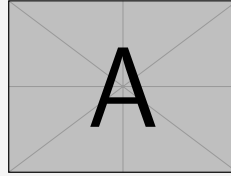


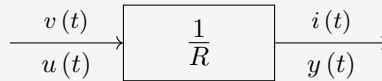
Figura 2.1

La Figura 2.1 è l'esempio più semplice di diagramma a blocchi in cui abbiamo un ingresso  $u$  che entra all'interno di un blocco che rappresenta il sistema  $S$  e fornisce un'uscita  $y$ .

**Esempio 2.0.1** Un esempio di sistema è quello di un resistore di resistenza  $R$  sollecitato con un generatore di tensione  $v(t)$ . Si vuole conoscere la corrente  $i(t)$  che scorre nel resistore.



Iniziamo identificando gli ingressi, le uscite e la relazione tra di esse. In questo caso vi è solo un ingresso  $v(t)$  ed una sola uscita  $i(t)$ . Il legame tra queste due variabili è ben noto come legge di Ohm:  $v(t) = Ri(t)$ .



In tal caso, indicando con  $u(t)$  l'ingresso in tensione  $v(t)$  e con  $y(t)$  l'uscita in corrente  $i(t)$ , l'equazione che lega l'ingresso con l'uscita del sistema è:

$$y(t) = \frac{1}{R}u(t) \quad (2.1)$$

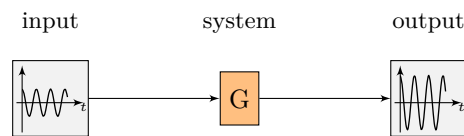
Il valore  $y(t)$  assunto dall'uscita al tempo  $t$  dipende solo dal valore corrente  $u(t)$  assunto dall'ingresso al tempo  $t$ . In tal caso, il sistema è detto **statico**.

### Cos'è il guadagno?

Il **guadagno** è un parametro che indica la relazione tra l'uscita e l'ingresso di un sistema. In termini generali, rappresenta il fattore di amplificazione con cui un segnale in ingresso viene modificato per produrre l'uscita. Matematicamente, il guadagno  $G$  di un sistema può essere espresso come:

$$G = \frac{\text{uscita}}{\text{ingresso}} \quad (2.2)$$

Nel caso dell'esempio 4.10.1 il guadagno è  $G = \frac{1}{R}$ .


 Figura 2.2. Demonstration of system with a gain of  $G = 2$ 

## 2.1 Elementi fondamentali

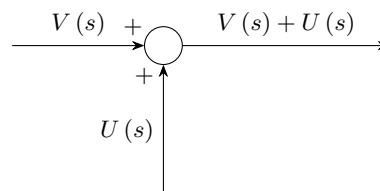
**Blocco** I blocchi rappresentano i singoli elementi o sottosistemi che compongono un sistema complesso. Ciascun blocco descrive una specifica relazione tra un segnale di ingresso e un segnale di uscita, solitamente espressa tramite una funzione di trasferimento nel dominio di Laplace. In questo modo, ogni blocco modella il comportamento dinamico di una parte del sistema, consentendo di analizzare e semplificare la struttura complessiva del sistema.

$$G(s)$$

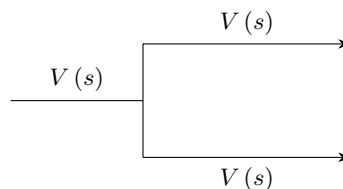
**Freccia** L'elemento freccia rappresenta il flusso del segnale tra i vari blocchi del sistema. Essa indica la direzione di propagazione dell'informazione, ossia il percorso seguito dal segnale dall'uscita di un blocco verso l'ingresso di un altro. Le frecce permettono quindi di comprendere come i diversi sottosistemi siano interconnessi e come le grandezze fisiche o matematiche si trasmettano all'interno del sistema complessivo.

$$\xrightarrow{U(s)}$$

**Nodo Sommatore** Il nodo sommatore (o punto di somma) è l'elemento in cui più segnali si combinano algebricamente per formare un'unica uscita. Ogni segnale in ingresso al nodo può essere additivo o sottrattivo, a seconda del segno indicato accanto alla freccia di ingresso. È quindi fondamentale prestare attenzione ai segni associati a ciascun ramo, poiché da essi dipende la corretta rappresentazione del comportamento del sistema.

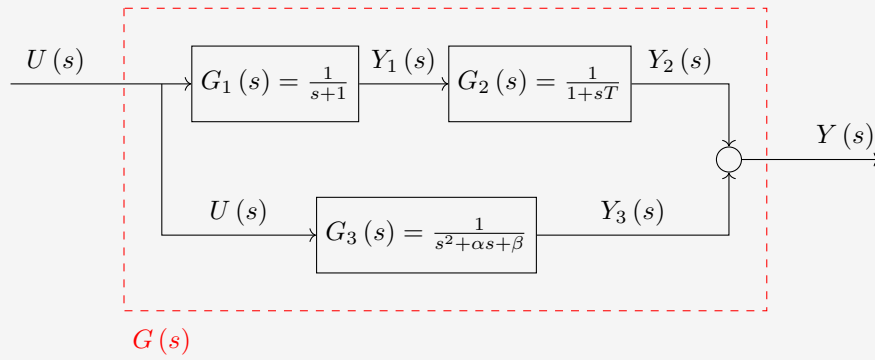


**Punto di diramazione** Il punto di diramazione rappresenta un nodo dal quale un singolo segnale in uscita da un blocco viene suddiviso in più rami, così da poter essere inviato a diversi sottosistemi o nodi sommatore. In altre parole, la diramazione consente a uno stesso segnale di essere utilizzato simultaneamente in più parti del sistema senza alterarne il valore. È importante ricordare che, nei diagrammi a blocchi, la diramazione non modifica né amplifica il segnale, ma ne effettua soltanto una replica verso percorsi differenti.

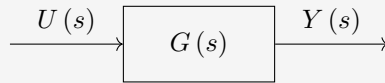


**Teorema 2.1.1** Un sistema costituito da un numero qualunque di sottosistemi fra loro connessi in cascata e/o parallelo è asintoticamente stabile se e solo se sono asintoticamente stabili tutti i sottosistemi che lo compongono.

**Esempio 2.1.2** Si consideri un sistema rappresentato dal seguente diagramma a blocchi:



Considerando il sistema ai capi e facendo uso delle proprietà notevoli dei diagrammi a blocchi già esaminate, è possibile rappresentare il sistema equivalente  $G(s)$  come un unico sistema che ingloba i sottosistemi  $G_1(s)$ ,  $G_2(s)$  e  $G_3(s)$ .



Alla luce del Teorema 2.1.1, il sistema  $G(s)$  è asintoticamente stabile se e solo se ciascuno dei sottosistemi che lo compongono  $G_1(s)$ ,  $G_2(s)$  e  $G_3(s)$  è a sua volta stabile. Tale condizione è soddisfatta per  $T > 0$ ,  $\alpha > 0$ ,  $\beta > 0$ .



## Le funzioni di trasferimento

Abbiamo visto come l'equazioni differenziali siano uno strumento formidabile per modellare il comportamento della realtà che ci circonda. Tuttavia, risolvere un'equazione differenziale potrebbe non essere così semplice.

### 3.1 La trasformata di Fourier

In molte applicazioni si ha a che fare con una grandezza periodica nel tempo (ad esempio un segnale sonoro, luminoso, di trasmissione ecc.) e si vuole risalire alle singole frequenze che la compongono e alle relative ampiezze, *senza conoscerle in anticipo* (si pensi ad esempio a un apparecchio elettronico che, captando attraverso un microfono le vibrazioni dell'aria prodotte da uno strumento musicale, le analizza per riconoscere di quale nota si tratta). La cosiddetta analisi di Fourier che ci accingiamo a studiare ha come obiettivo primario proprio questo: dato un segnale periodico, riuscire a decomporlo come somma (sovrapposizione) di segnali di tipo sinusoidale, ciascuno con una propria frequenza (multipla intera di una frequenza base) e una propria ampiezza.

Fu proprio il matematico francese Joseph Fourier, motivato principalmente dallo studio dell'equazione del calore, il primo a rendersi conto che una funzione  $f(x)$  sostanzialmente arbitraria (purché periodica di periodo  $T$ ) può essere decomposta come somma infinita di seni e coseni (serie trigonometrica).

Si tratta in sostanza di uno sviluppo del tipo:

$$f(x) \sim \sum_{k=0}^{\infty} (a_k \cos(k\omega x) + b_k \sin(k\omega x)), \quad \omega = \frac{2\pi}{T} \quad (3.1)$$

noto appunto come *sviluppo in serie di Fourier*.

Il passaggio dalla descrizione di un segnale (inteso come fenomeno periodico) dal punto di vista del suo andamento temporale alla sua descrizione dal punto di vista delle frequenze in somma di sinusoidi viene spesso chiamato passaggio *dal dominio dei tempi al dominio delle frequenze*.

In matematica, una **trasformata** è un operatore, generalmente lineare, di uno spazio di funzioni su un altro spazio di funzioni. Ovvero trasforma una funzione in un'altra funzione. Tale operatore è di solito applicato ad una funzione per semplificare alcune operazioni o in generale per risolvere più facilmente dei problemi.

Dato da risolvere un problema A, che può essere un calcolo aritmetico o la risoluzione di un'equazione differenziale, uno schema esemplificativo può essere il seguente:

1. si trasforma il problema A in un altro problema B più semplice da risolvere;
2. si risolve il problema B;
3. si antitrasforma la soluzione del problema B nella soluzione del problema A.

La trasformata di Fourier costituisce uno strumento (per certi versi analogo alle serie di Fourier) per analizzare un segnale  $x(t)$  nel dominio delle frequenze. Mentre le serie di Fourier consentono lo studio di un segnale di tipo periodico, la trasformata di Fourier permette l'analisi di un segnale  $x(t)$  il cui andamento nel tempo non sia necessariamente periodico.

**Definizione 3.1.1** Sia  $x(t) : \mathbb{R} \rightarrow \mathbb{C}$  una funzione. Se

$$\int_{-\infty}^{+\infty} x(t) e^{-i\omega t} dt \quad (3.2)$$

esiste finito per ogni  $\omega \in \mathbb{R}$ , allora diciamo che la funzione  $x(t)$  è trasformabile secondo Fourier. In questo caso, la funzione

$$\mathcal{F}[x(t)](\omega) = X(\omega) := \int_{-\infty}^{+\infty} x(t) e^{-i\omega t} dt, \quad \omega \in \mathbb{R} \quad (3.3)$$

viene chiamata **trasformata di Fourier** di  $x(t)$ .

### 3.2 La trasformata di Laplace

Come abbiamo visto nella definizione di trasformata di Fourier (Def. 3.1.1), una funzione  $f(t)$  è trasformabile secondo Fourier se l'integrale in (3.2) esiste finito per ogni  $\omega \in \mathbb{R}$ . Concettualmente, questo significa che è possibile effettuare una trasformazione di Fourier di funzioni che decadono a zero a  $\pm\infty$  come ad esempio una Gaussiana.

Funzioni come  $f(t) = e^{\lambda t}$  o il gradino di Heaviside  $H(t)$  non possono essere trasformate con Fourier. Come fare? Bhe, una soluzione è quella di moltiplicare la funzione che vogliamo trasformare  $f(t)$  con un esponenziale decrescente  $e^{-\gamma t}$  in maniera che  $f(t)e^{-\gamma t} \rightarrow 0$  per  $t \rightarrow \infty$ . Questo risolve il problema per  $+\infty$  ma come fare per  $-\infty$ ? Possiamo moltiplicare la funzione anche per  $H(t)$  il gradino di Heaviside che come sappiamo è nullo per  $t < 0$ .

In questa maniera avremo che la nostra nuova funzione da trasformare è:

$$f_1(t) = f(t) e^{-\gamma t} H(t) = \begin{cases} 0 & t < 0 \\ f(t) e^{-\gamma t} & t \geq 0 \end{cases}$$

Andiamo a scrivere la trasformata di Fourier di questa nuova funzione:

$$\mathcal{F}(f_1(t)) = F(\omega) = \int_{-\infty}^{+\infty} f_1(t) e^{-i\omega t} dt \quad (3.4)$$

Sapendo che  $f_1(t)$  è nulla per  $t < 0$  possiamo riscrivere gli estremi di integrazione come:

$$\int_0^{+\infty} f_1(t) e^{-i\omega t} dt \quad (3.5)$$

Sostituiamo  $f_1(t)$

$$\int_0^{+\infty} f(t) e^{-\gamma t} e^{-i\omega t} dt \quad (3.6)$$

Sfrutto le proprietà degli esponenziali per compattare degli esponenziali:

$$\int_0^{+\infty} f(t) e^{-(\gamma+i\omega)t} dt \quad (3.7)$$

Notiamo che il termine all'esponente dell'esponenziale è una variabile complessa, possiamo quindi dire  $s = \gamma + i\omega$ . Da ciò ne segue la definizione della trasformata di Laplace:

**Definizione 3.2.1** La trasformata di Laplace  $\mathcal{L}$  di  $f$  è una funzione  $F: \mathbb{C} \rightarrow \mathbb{C}$  di una variabile complessa  $s \in \mathbb{C}$  definita da:

$$\mathcal{L}(f(t)) = F(s) = \int_0^{+\infty} e^{-st} f(t) dt$$

per ogni  $s \in \mathbb{C}$  per il quale l'integrale è definito.

#### 3.2.1 Trasformate notevoli

##### Gradino unitario

$$f(t) = 1(t) = \begin{cases} 0 & \text{if } t < 0 \\ 1 & \text{if } t \geq 0 \end{cases} \Rightarrow F(s) = \int_0^{+\infty} e^{-st} dt = -\frac{e^{-st}}{s} \Big|_0^{+\infty} = \frac{1}{s}$$

### 3.2.2 Proprietà della Trasformata di Laplace

#### 3.2.3 Teoremi fondamentali

Nello studio dei sistemi dinamici utilizziamo la trasformata di Laplace per passare dal dominio del tempo a quello complesso, dove le equazioni differenziali diventano semplici relazioni algebriche. Tuttavia, molto spesso il nostro interesse non è conoscere l'intera risposta temporale, bensì prevedere come il sistema si comporta all'inizio di una sollecitazione e quale valore assumerà a regime.

Per ottenere queste informazioni senza calcolare l'intera trasformata inversa, disponiamo di due strumenti fondamentali: il teorema del valore iniziale e il teorema del valore finale.

**Teorema 3.2.2** Teorema del valore iniziale Sia  $F(s)$  la trasformata di Laplace di una funzione  $f(t)$ . Se i limiti esistono, vale la relazione:

$$\lim_{t \rightarrow 0^+} f(t) = \lim_{s \rightarrow \infty} sF(s) \quad (3.8)$$

Questo teorema consente di ricavare immediatamente il valore assunto dalla funzione nel tempo iniziale (subito dopo  $t=0$ ) a partire dalla sua trasformata.

**Teorema 3.2.3** Teorema del valore finale Analogamente, se tutti i poli di  $sF(s)$  hanno parte reale negativa (cioè il sistema è asintoticamente stabile), allora

$$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sF(s) \quad (3.9)$$

Il teorema fornisce il valore a regime della risposta, ed è particolarmente utile in analisi dei controlli per determinare l'errore stazionario senza dover ricostruire l'intera risposta temporale.

### 3.3 L'antitrasformata di Laplace

**Definizione 3.3.1** L'operazione di passaggio inverso dal dominio della frequenza complessa  $s$  al dominio del tempo è detta *antitrasformata* o *trasformazione inversa di Laplace*.

Data una funzione  $F: \mathbb{C} \rightarrow \mathbb{C}$  di una variabile complessa  $s \in \mathbb{C}$ , la sua antitrasformata di Laplace è definita come segue:

$$f(t) = \mathcal{L}^{-1}(F(s)) = \frac{1}{2\pi j} \int_{\gamma-j\infty}^{\gamma+j\infty} F(s) e^{st} ds \quad (3.10)$$

L'espressione integrale utilizzata per definire l'antitrasformata di Laplace è scomoda per antitrasformare delle funzioni complesse e riportarle nel dominio del tempo, per questo motivo si ricorre ad una procedura diversa.

Questa procedura si applica alle funzioni razionali fratte, che individuano le funzioni nel dominio della frequenza complessa  $s$  di interesse nei controlli automatici, essendo le trasformate di Laplace dei segnali tipici di questa disciplina. Una funzione  $F(s)$  di questo tipo si può esprimere come rapporto di due polinomi nella variabile  $s$ :

$$F(s) = \frac{B(s)}{A(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}$$

dove  $m$  è il grado del numeratore  $B(s)$  ed  $n$  quello del denominatore  $A(s)$ .

In base al grado del numeratore  $B(s)$  e quello del denominatore  $A(s)$  è possibile distinguere 3 casi:

#### 3.3.1 $m > n$ Funzione razionale fratta impropria

Effettuando la divisione tra polinomi, si ha

$$F(s) = k_{m-n} s^{m-n} + k_{m-n-1} s^{m-n-1} + \dots + k_0 + \frac{R(s)}{A(s)}$$

con  $R(s)$  polinomio resto di grado  $r < n$ ; pertanto l'antitrasformata vale:

$$f(t) = k_{m-n} \delta^{(m-n+1)}(t) + k_{m-n-1} \delta^{(m-n)}(t) + \dots + k_0 \delta(t) + \mathcal{L}^{-1} \left\{ \frac{R(s)}{A(s)} \right\}$$

dove si è utilizzata la proprietà notevole sulla trasformazione degli impulsi di Dirac di ordine superiore. La funzione antitrasformata presenta dunque degli impulsi di vario ordine. In questo caso la funzione  $f(t)$  si dice anticipativa o non realizzabile (non causale) e la funzione  $F(s)$  è detta funzione razionale fratta impropria.

### 3.3.2 $m = n$ Funzione razionale fratta propria ma non strettamente

Effettuando la divisione tra polinomi si ha in questo caso

$$F(s) = \frac{b_m}{a_n} + \frac{R(s)}{A(s)}$$

con  $R(s)$  polinomio resto di grado  $r < n$ ; pertanto l'antitrasformata vale:

$$f(t) = \frac{b_m}{a_n} \delta(t) + \mathcal{L}^{-1} \left\{ \frac{R(s)}{A(s)} \right\}$$

In questo caso si presenta solo un impulso di Dirac del primo ordine e la funzione  $F(s)$  si dice *razionale fratta propria ma non strettamente*.

### 3.3.3 $m < n$ Funzione razionale fratta strettamente propria

In questo caso la divisione tra polinomi non si può effettuare, quindi l'antitrasformata  $f(t)$  non contiene impulsi di Dirac. La funzione  $F(s)$  si dice *razionale fratta strettamente propria*.

Quindi in ciascuno dei tre casi analizzati è necessario antitrasformare una funzione razionale fratta strettamente propria, pertanto è sufficiente analizzare il caso in cui sia  $m < n$ .

La procedura di antitrasformazione di una funzione razionale fratta strettamente propria  $F(s)$  consiste nell'esprimere la funzione come una somma di altre funzioni razionali fratte strettamente proprie, dette *fratti semplici*, che possono essere facilmente antitrasformate utilizzando le regole di trasformazione e le trasformate notevoli.

I fratti semplici dipendono dai *poli* di  $F(s)$ , cioè dalle radici dell'equazione  $A(s) = 0$ , ossia dagli zeri della funzione  $A(s)$ , che vanno quindi determinati. Tali radici sono in numero pari al grado  $n$  del polinomio  $A(s)$  e possono essere reali o complesse e coniugate, semplici o multiple.

Supponendo che  $A(s)$  abbia  $r$  radici distinte  $p_i$  ( $i = 1, \dots, r$ ), ciascuna di molteplicità  $\alpha_i$ , cioè che sia:

$$A(s) = a_n (s - p_1)^{\alpha_1} (s - p_2)^{\alpha_2} \dots (s - p_r)^{\alpha_r} = a_n \prod_{i=1}^r (s - p_i)^{\alpha_i}$$

con

$$\sum_{i=1}^r \alpha_i = n.$$

È facile dimostrare che lo sviluppo in fratti semplici della funzione da antitrasformare razionale fratta strettamente propria  $F(s)$  è del tipo:

$$\begin{aligned} F(s) = & \frac{k_{11}}{(s - p_1)^{\alpha_1}} + \frac{k_{12}}{(s - p_1)^{\alpha_1-1}} + \dots + \frac{k_{1\alpha_1}}{(s - p_1)} \\ & + \frac{k_{21}}{(s - p_2)^{\alpha_2}} + \frac{k_{22}}{(s - p_2)^{\alpha_2-1}} + \dots + \frac{k_{2\alpha_2}}{(s - p_2)} \\ & \dots \\ & + \frac{k_{r1}}{(s - p_r)^{\alpha_r}} + \frac{k_{r2}}{(s - p_r)^{\alpha_r-1}} + \dots + \frac{k_{r\alpha_r}}{(s - p_r)} \end{aligned}$$

infatti è sufficiente riportare tale espansione in fratti semplici ad un unico membro sommando i diversi termini.



L'espansione in fratti semplici si scrive anche in forma compatta:

$$F(s) = \sum_{i=1}^r \sum_{j=1}^{\alpha_i} \frac{k_{ij}}{(s - p_i)^{\alpha_i - j + 1}}$$

dove tutti e soli i coefficienti  $k_{i\alpha_i}$  dei fratti semplici con denominatori del primo ordine si chiamano *residui* associati ai poli  $p_i$ .

Dalla proprietà della trasformata della moltiplicazione vale:

$$t^n (e^{-at}) \longrightarrow \frac{n!}{(s + a)^{n+1}}$$

e sfruttando tale relazione per  $t^n$  (con  $n = \alpha_i - j$ ) con la proprietà di linearità della trasformata si può antitrasformare l'equazione precedente come segue:

$$f(t) = \left\{ \sum_{i=1}^r \sum_{j=1}^{\alpha_i} \frac{k_{ij}}{(\alpha_i - j)!} t^{\alpha_i - j} e^{p_i t} \right\} \cdot u(t)$$

dove si è introdotto il gradino unitario per rendere biunivoca la trasformazione funzionale.

Il problema è quindi determinare i coefficienti  $k_{ij}$  dell'espansione in fratti.

Di seguito descriviamo i diversi metodi che si utilizzano per individuare i coefficienti dell'espansione in fratti semplici di una funzione  $F(s)$  razionale fratta strettamente propria.



**Definizione 4.0.1** La **funzione di trasferimento** di un sistema lineare tempo-invariante (LTI) è definita come il rapporto tra la trasformata di Laplace dell'uscita e quella dell'ingresso:

$$G(s) = \frac{Y(s)}{X(s)} \quad (4.1)$$

Essa rappresenta il comportamento dinamico del sistema nel dominio complesso di Laplace.

## 4.1 Forme di una funzione di trasferimento

Una funzione di trasferimento può essere espressa in più forme a seconda dell'uso che vogliamo farne.

### 4.1.1 Forma polinomiale

Una delle modalità più comuni per esprimere una funzione di trasferimento  $G(s)$  è la sua **forma polinomiale**, cioè come rapporto tra due polinomi in  $s$ :

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} \quad (4.2)$$

dove:

- $N(s)$  è detto **polinomio del numeratore** e rappresenta gli **zeri** del sistema;
- $D(s)$  è detto **polinomio del denominatore** e rappresenta i **poli** del sistema;
- $m$  e  $n$  indicano i rispettivi gradi dei polinomi del numeratore e denominatore.

In generale l'**ordine di un sistema** è pari al grado del denominatore della funzione di trasferimento.

In questa forma, la funzione di trasferimento evidenzia direttamente la **relazione differenziale** tra ingresso e uscita. Infatti, applicando la trasformata di Laplace all'equazione differenziale che descrive il sistema, si ottiene:

$$a_n s^n Y(s) + a_{n-1} s^{n-1} Y(s) + \dots + a_0 Y(s) = b_m s^m X(s) + b_{m-1} s^{m-1} X(s) + \dots + b_0 X(s), \quad (4.3)$$

da cui, isolando il rapporto  $\frac{Y(s)}{X(s)}$ , si ricava appunto la forma polinomiale di  $G(s)$ .

Questa rappresentazione è molto utile perché:

- consente di collegare direttamente i coefficienti dei polinomi ai parametri fisici del sistema (resistenze, capacità, induttanze, masse, attriti, ecc.);
- permette un'analisi immediata del comportamento del sistema, poiché i poli determinano la **stabilità** e gli zeri influenzano la **forma della risposta**;
- è la base di partenza per ricavare altre rappresentazioni equivalenti, come la forma fattorizzata o quella in termini di poli e zeri.

### 4.1.2 Forma fattorizzata

La **forma fattorizzata** mette in evidenza in modo diretto i **poli** e gli **zeri** del sistema, cioè i valori di  $s$  per i quali la funzione di trasferimento si annulla o diventa infinita.

A partire dalla forma polinomiale:

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0},$$

possiamo fattorizzare numeratore e denominatore in termini delle loro radici:

$$G(s) = K \frac{(s - z_1)(s - z_2) \dots (s - z_m)}{(s - p_1)(s - p_2) \dots (s - p_n)} = K \frac{\prod_{j=1}^m (s - z_j)}{\prod_{i=1}^n (s - p_i)} \quad (4.4)$$

dove:

- $z_i$  sono gli **zeri** del sistema (radici di  $N(s)$ );
- $p_i$  sono i **poli** del sistema (radici di  $D(s)$ );
- $K$  è la **costante di guadagno** (coefficiente direttore del numeratore).

Questa forma è particolarmente utile per l'analisi qualitativa del comportamento del sistema, poiché:

- i poli determinano la **dinamica** e la **stabilità** del sistema (devono avere parte reale negativa per garantire la stabilità asintotica);
- gli zeri influenzano la forma della **risposta in frequenza** e la rapidità della risposta nel dominio del tempo;
- il guadagno  $K$  controlla l'ampiezza complessiva della risposta.

La forma fattorizzata permette inoltre una rappresentazione geometrica immediata nel **piano complesso**  $s$ , attraverso il cosiddetto **diagramma poli-zeri**.

#### 4.1.3 Forma in costanti di tempo

Una versione alternativa della forma fattorizzata è quella in costanti di tempo:

$$\begin{aligned} G(s) &= K \frac{s^\mu (1 + \tau_{z1} s) \dots (1 + \tau_{zm1} s) \left(1 + 2 \frac{\delta_{z1}}{\omega_{nz1}} s + \frac{s^2}{\omega_{nz1}^2}\right) \dots \left(1 + 2 \frac{\delta_{zm2}}{\omega_{nz m2}} s + \frac{s^2}{\omega_{nz m2}^2}\right)}{s^\nu (1 + \tau_{p1} s) \dots (1 + \tau_{pn1} s) \left(1 + 2 \frac{\delta_{p1}}{\omega_{np1}} s + \frac{s^2}{\omega_{np1}^2}\right) \dots \left(1 + 2 \frac{\delta_{pn2}}{\omega_{np n2}} s + \frac{s^2}{\omega_{np n2}^2}\right)} \\ &= K \frac{s^\mu \prod_{j=1}^{m_1} (1 + \tau_{zj} s) \prod_{j=1}^{m_2} \left(1 + 2 \frac{\delta_{zj}}{\omega_{nzj}} s + \frac{s^2}{\omega_{nzj}^2}\right)}{s^\nu \prod_{i=1}^{n_1} (1 + \tau_{pi} s) \prod_{i=1}^{n_2} \left(1 + 2 \frac{\delta_{pi}}{\omega_{npi}} s + \frac{s^2}{\omega_{npi}^2}\right)} \end{aligned} \quad (4.5)$$

In questa forma la funzione è espressa come **prodotto di fattori elementari**, ciascuno dei quali rappresenta un tipo di polo o zero caratteristico:

- $s^\mu$  e  $s^\nu$  rappresentano rispettivamente eventuali zeri e poli nell'origine;
- i termini  $(1 + \tau s)$  rappresentano **zeri o poli reali**;
- i termini  $\left(1 + 2 \frac{\delta}{\omega_n} s + \frac{s^2}{\omega_n^2}\right)$  rappresentano **zeri o poli complessi coniugati**;
- $K_1$  è la costante di guadagno statico del sistema.
- $\omega_n$  = pulsazione naturale;
- $\delta$  = coefficiente di smorzamento;
- $\tau$  = costante di tempo.

È una forma molto usata in analisi e progettazione dei sistemi di controllo, poiché permette di associare immediatamente a ciascun termine un comportamento dinamico ben definito.

## 4.2 Diagramma poli-zeri

Il **diagramma poli-zeri** è una rappresentazione grafica nel **piano complesso**  $s$  che mostra la posizione dei **poli** e degli **zeri** di una funzione di trasferimento  $G(s)$ .

Nel diagramma gli **zeri** vengono rappresentati con il simbolo di una "O" (cerchietto) nel punto  $s = z_i$  mentre i **poli** vengono rappresentati con il simbolo di una "X" nel punto  $s = p_i$ ;

Come vedremo più avanti, il diagramma poli-zeri è uno strumento molto utile per analizzare le **proprietà dinamiche** del sistema:

- la **stabilità** dipende dalla posizione dei poli: un sistema continuo è stabile se tutti i poli si trovano nel **semipiano sinistro** ( $Re(s) < 0$ );
- la **velocità di risposta** è legata alla distanza dei poli dall'origine: poli più vicini all'asse immaginario producono risposte lente;
- la **componente oscillatoria** della risposta è determinata dai poli complessi coniugati;
- gli **zeri** influenzano la forma della risposta in frequenza, introducendo attenuazioni o inversioni di fase.

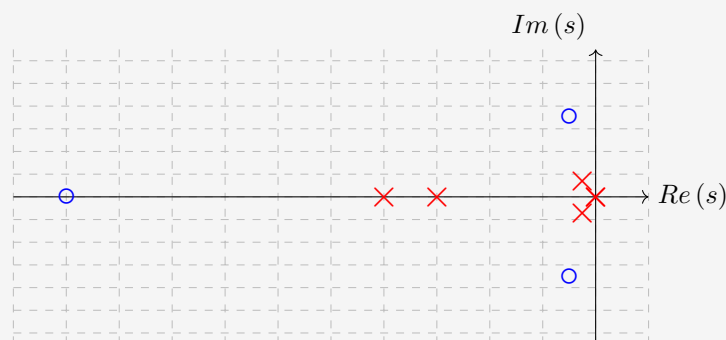
**Esempio 4.2.1** Data la seguente funzione di trasferimento:

$$G(s) = \frac{(s^2 + s + 50)(s + 10)}{s^2(s + 3)(s + 4)(s^2 + 0.5s + 2)}$$

il grado del numeratore è  $m = 3$  e il grado del denominatore è  $n = 6$ . Dato che  $m < n$  il sistema è strettamente proprio. Avremo quindi 3 zeri e 6 poli.

$$\begin{aligned} z_{1,2} &= -\frac{1}{2} \pm 7.05j & z_3 &= -10 \\ p_{1,2} &= 0 & p_3 &= -3 & p_4 &= -4 & p_{5,6} &= -\frac{1}{4} \pm 1.39j \end{aligned}$$

Andiamo a posizionare i poli e gli zeri trovati nel piano complesso in modo da crearci il nostro diagramma poli-zeri:



## 4.3 Condizione di realizzabilità fisica

**Definizione 4.3.1** Affinché un modello rappresenti un sistema fisico reale e sia **fisicamente realizzabile**, è necessario che il grado del numeratore della funzione di trasferimento non superi quello del denominatore, ossia:

$$m \leq n \quad (4.6)$$

Se la funzione di trasferimento ha  $m < n$ , si dice **strettamente propria**, se  $m = n$  si dice **non strettamente propria**. Invece se  $m > n$ , la funzione di trasferimento si dice **impropria**.

**Dimostrazione** Supponiamo per assurdo che sia possibile che  $m > n$ . Allora  $G(s)$  è *impropria*. Per divisione polinomiale si ha

$$G(s) = Q(s) + \frac{R(s)}{D(s)}$$

dove  $Q(s)$  è un polinomio risultante dalla divisione, con grado  $m - n \geq 1$ , e  $R(s)$  è il polinomio resto di grado  $r < n$ .

La risposta all'impulso  $g(t)$  è la trasformata inversa di  $G(s)$ :

$$\begin{aligned} g(t) &= \mathcal{L}^{-1}\{G(s)\} \\ &= \mathcal{L}^{-1}\{Q(s)\} + \mathcal{L}^{-1}\left\{\frac{R(s)}{D(s)}\right\}. \end{aligned}$$

Poiché  $\mathcal{L}^{-1}\{s^k\} = \delta^{(k)}(t)$  (derivata  $k$ -esima dell'impulso), il termine  $\mathcal{L}^{-1}\{Q(s)\}$  contiene combinazioni lineari di  $\delta(t), \delta'(t), \dots, \delta^{(m-n)}(t)$ . In particolare, se  $m > n$  compaiono *derivate* della delta, che equivalgono nel dominio del tempo ad operatori di *derivazione ideale* dell'ingresso, cioè l'uscita è in anticipo sull'ingresso stesso rendendo il sistema **non causale**.  $\square$

## 4.4 Funzioni di trasferimento con ritardo

Abbiamo visto che la funzione di trasferimento di un sistema SISO lineare e tempo-invariante può essere espressa come una funzione razionale fratta nella variabile complessa  $s$ :

Tuttavia, non tutti i sistemi SISO stazionari presentano una funzione di trasferimento di questo tipo. Un esempio importante è quello del **ritardo puro**, ossia un sistema che introduce un ritardo temporale costante  $\tau$  sull'ingresso. Se l'ingresso è indicato con  $x(t)$ , l'uscita sarà:

$$y(t) = x(t - \tau)$$

Questo sistema rimane lineare e tempo-invariante, poiché rispetta il principio di sovrapposizione e le sue caratteristiche non variano nel tempo. Tuttavia, calcolando la sua funzione di trasferimento mediante la trasformata di Laplace, si ottiene:

$$G(s) = \frac{Y(s)}{X(s)} = \frac{\mathcal{L}(x(t - \tau))}{X(s)} = \frac{X(s)e^{-\tau s}}{X(s)} = e^{-\tau s}$$

che non è una funzione razionale fratta.

Il ritardo puro si presenta in molti sistemi di controllo, in particolare in quelli che prevedono trasmissione o elaborazione dei segnali. In tali casi, il ritardo rappresenta il tempo necessario affinché l'effetto di un ingresso si manifesti all'uscita. Esempi tipici sono i **sistemi di controllo remoto**, dove la trasmissione del segnale comporta inevitabilmente un certo ritardo, come accade frequentemente nelle **applicazioni spaziali**.

## 4.5 Risposte tipiche

### 4.5.1 Risposta all'impulso

$$Y(s) = G(s) \cdot X(s) = G(s) \cdot \delta(s) = G(s) \cdot 1 = G(s) \quad (4.7)$$

$$y(t) = \mathcal{L}^{-1}(G(s)) = g(t) \quad (4.8)$$

### 4.5.2 Risposta al gradino

$$Y(s) = G(s) \cdot X(s) = G(s) \cdot \frac{1}{s} \quad (4.9)$$

$$y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}(G(s) \cdot X(s)) = \mathcal{L}^{-1}\left(\frac{G(s)}{s}\right) \quad (4.10)$$

### 4.5.3 Risposta ad un ingresso generico

La risposta di un sistema ad un ingresso generico si determina in modo analogo a quanto visto per le risposte all'impulso e al gradino. Quando la trasformata di Laplace dell'ingresso,  $X(s)$ , può essere ricavata tramite una trasformata notevole, il procedimento risulta diretto: si moltiplica la funzione di trasferimento del sistema  $G(s)$  per  $X(s)$ , ottenendo così  $Y(s)$ , e successivamente si esegue l'antitrasformata per tornare nel dominio del tempo.

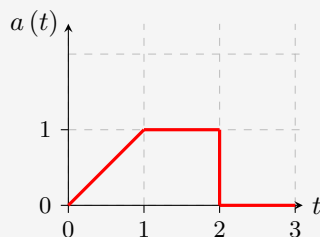
Consideriamo ora un segnale di ingresso come quello mostrato in figura. In questo caso non è possibile esprimere  $X(s)$  mediante una trasformata notevole. Tuttavia, il segnale può essere scomposto in più componenti elementari. Grazie alla **linearità** del sistema e al **principio di sovrapposizione degli effetti**, la risposta complessiva può essere ottenuta come somma delle risposte associate a ciascuna componente del segnale d'ingresso.

Se l'ingresso è costituito da  $N$  componenti elementari, la risposta complessiva del sistema sarà:

$$Y(s) = \sum_{i=0}^N Y_i(s) \quad \Leftrightarrow \quad y(t) = \sum_{i=0}^N y_i(t) \quad (4.11)$$

dove  $Y_i(s)$  e  $y_i(t)$  rappresentano rispettivamente le risposte nel dominio di Laplace e nel dominio del tempo dovute alla  $i$ -esima componente dell'ingresso.

**Esempio 4.5.1** Si consideri un sistema che ha il seguente ingresso:



Osservando il segnale di ingresso, possiamo notare che esso presenta una **rampa** nel tratto  $0 \leq t \leq 1$  e successivamente un **gradino** costante nel tratto  $1 \leq t \leq 2$ . Per  $t \geq 2$ , il segnale torna ad essere nullo.

Per determinare la sua espressione analitica, ricostruiamo il segnale come somma di segnali elementari noti.

Nel primo intervallo ( $0 \leq t \leq 1$ ) è presente una rampa crescente. Sappiamo che una rampa unitaria può essere espressa come:

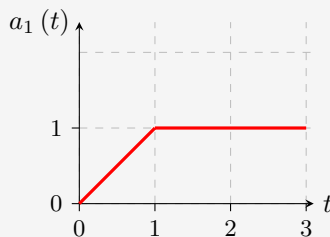
$$r(t) = t, \quad t > 0.$$

Tuttavia, a partire da  $t = 1$ , la crescita termina: la rampa si "blocca". Per rappresentare questa interruzione, sottraiamo una rampa ritardata di un'unità, ossia  $r(t - 1)$ .

Otteniamo quindi:

$$a_1(t) = r(t) - r(t - 1),$$

che rappresenta una rampa crescente tra  $0 \leq t < 1$  con una crescita annullata in  $t = 1$ .



Nel tratto successivo ( $1 \leq t \leq 2$ ) il segnale resta costante a livello 1, cioè presenta un **gradino unitario**. Per troncare anche questo gradino al tempo  $t = 2$ , sottraiamo un gradino ritardato di 2 unità, ossia  $1(t - 2)$ .

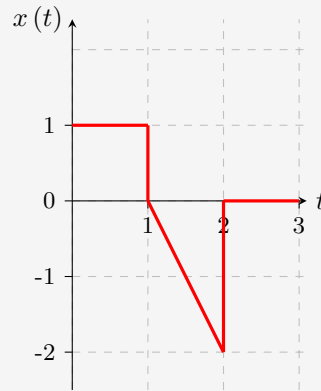
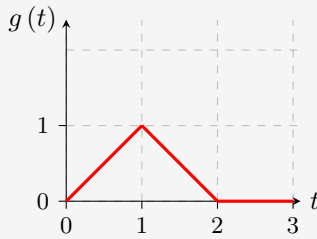
Combinando tutte le componenti otteniamo il segnale completo:

$$a(t) = r(t) - r(t - 1) - 1(t - 2)$$

Il grafico di  $a(t)$  corrisponde esattamente al segnale d'ingresso mostrato in figura. E' molto semplice proseguire ora trasformando il segnale nel dominio di Laplace:

$$A(s) = \mathcal{L}(a(t)) = \frac{1}{s^2} - \frac{1}{s^2}e^{-s} - \frac{1}{s}e^{-2s}$$

**Esempio 4.5.2** Si consideri



$$g(t) = r(t) - 2r(t - 1) + r(t - 2)$$

$$G(s) = \frac{1}{s^2} - \frac{2}{s^2}e^{-s} + \frac{1}{s^2}e^{-2s}$$

$$x(t) = 1(t) - 1(t - 1) - 2r(t - 1) + 2 \cdot 1(t - 2)$$

$$X(s) = \frac{1}{s} - \frac{1}{s}e^{-s} - \frac{2}{s^2}e^{-s} + \frac{2}{s}e^{-2s}$$

$$\begin{aligned} Y(s) &= G(s) \cdot X(s) = \\ &= \left( \frac{1}{s^2} - \frac{2}{s^2}e^{-s} + \frac{1}{s^2}e^{-2s} \right) \cdot \left( \frac{1}{s} - \frac{1}{s}e^{-s} - \frac{2}{s^2}e^{-s} + \frac{2}{s}e^{-2s} \right) = \\ &= \frac{1}{s^2} \frac{1}{s} - \frac{1}{s^2} \frac{1}{s} e^{-s} - \frac{1}{s^2} \frac{2}{s^2} e^{-s} + \frac{1}{s^2} \frac{2}{s^2} e^{-2s} - \frac{2}{s^2} e^{-s} \frac{1}{s} + \frac{2}{s^2} e^{-s} \frac{1}{s} e^{-s} + \frac{2}{s^2} e^{-s} \frac{2}{s^2} e^{-s} \\ &- \frac{2}{s^2} e^{-s} \frac{2}{s^2} e^{-2s} + \frac{1}{s^2} e^{-2s} \frac{1}{s} - \frac{1}{s^2} e^{-2s} \frac{1}{s} e^{-s} - \frac{1}{s^2} e^{-2s} \frac{2}{s^2} e^{-s} + \frac{1}{s^2} e^{-2s} \frac{2}{s^2} e^{-2s} \\ &= \frac{1}{s^3} - \frac{1}{s^3} e^{-s} - \frac{2}{s^4} e^{-s} + \frac{2}{s^3} e^{-2s} - \frac{2}{s^3} e^{-s} + \frac{2}{s^3} e^{-2s} + \frac{4}{s^4} e^{-2s} - \frac{4}{s^3} e^{-3s} \\ &+ \frac{1}{s^3} e^{-2s} - \frac{1}{s^3} e^{-3s} - \frac{2}{s^4} e^{-3s} + \frac{2}{s^3} e^{-4s} \\ &= \frac{2}{s^3} e^{-4s} - \left( \frac{2}{s^4} + \frac{1}{s^3} + \frac{4}{s^3} \right) e^{-3s} + \left( \frac{1}{s^3} + \frac{2}{s^3} + \frac{4}{s^4} + \frac{2}{s^3} \right) e^{-2s} - \left( \frac{1}{s^3} + \frac{2}{s^4} + \frac{2}{s^3} \right) e^{-s} + \frac{1}{s^3} \\ &= \frac{2}{s^3} e^{-4s} - \left( \frac{2}{s^4} + \frac{5}{s^3} \right) e^{-3s} + \left( \frac{4}{s^4} + \frac{5}{s^3} \right) e^{-2s} - \left( \frac{2}{s^4} + \frac{3}{s^3} \right) e^{-s} + \frac{1}{s^3} \end{aligned}$$

$$\begin{aligned} y(t) &= \mathcal{L}^{-1}(Y(s)) \\ &= 2 \cdot \frac{t^2}{2} 1(t - 4) - \left( \frac{2t^3}{6} + 5 \frac{t^2}{2} \right) 1(t - 3) + \left( \frac{4t^3}{6} + 5 \frac{t^2}{2} \right) 1(t - 2) - \left( \frac{2t^3}{6} + 3 \frac{t^2}{2} \right) 1(t - 1) + \frac{t^2}{2} \\ &= t^2 \cdot 1(t - 4) - \left( \frac{t^3}{3} + \frac{5t^2}{2} \right) \cdot 1(t - 3) + \left( \frac{2t^3}{3} + \frac{5t^2}{2} \right) \cdot 1(t - 2) - \left( \frac{t^3}{3} + \frac{3t^2}{2} \right) \cdot 1(t - 1) + \frac{t^2}{2} \end{aligned}$$



Una maniera più semplice è quella di utilizzare il principio di sovrapposizione degli effetti e cercare la risposta delle singole componenti:

$$y(t) = y_1(t) - y_1(t-1) - 2y_p(t-1) + 2y_p(t-2)$$

## 4.6 Sistemi del primo ordine

La funzione di trasferimento di un sistema elementare del primo ordine, si può porre nella forma:

$$G(s) = \frac{1}{1 + \tau s} \quad (4.12)$$

in cui  $\tau$  indica la costante di tempo, cioè un parametro che caratterizza il comportamento dinamico del sistema.

### 4.6.1 Risposta di un sistema del primo ordine

La risposta al gradino unitario è data da:

$$\begin{aligned} y(t) &= \mathcal{L}^{-1} \left[ \frac{1}{(1 + \tau s)s} \right] \\ &= 1 - e^{-\frac{t}{\tau}} \end{aligned} \quad (4.13)$$

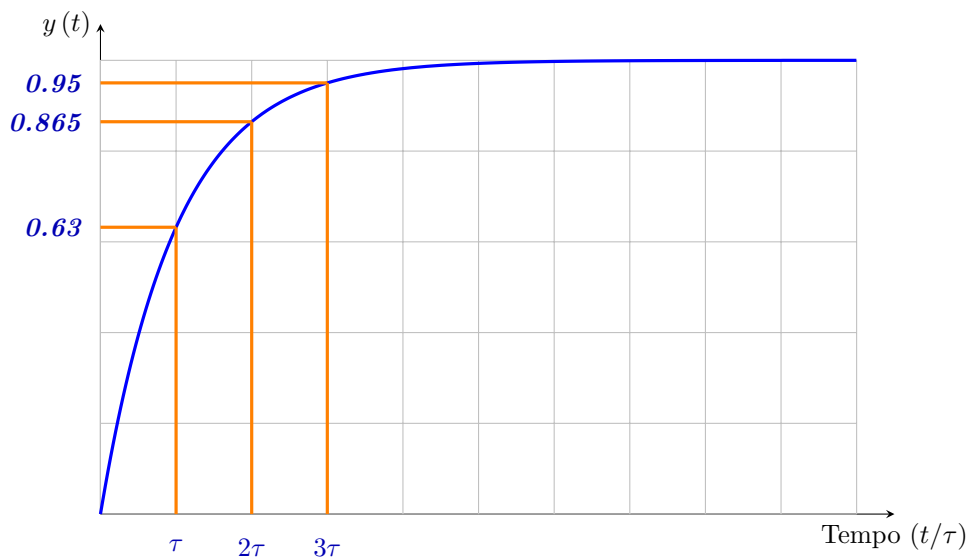
Il valore iniziale della risposta è:

$$y(0) = 0,$$

mentre la **pendenza iniziale** (cioè la derivata della risposta all'istante iniziale) risulta:

$$\frac{dy(0)}{dt} = \frac{1}{\tau}.$$

In altre parole, la risposta parte da zero con una tangente iniziale di pendenza pari a  $1/\tau$ .



**Figura 4.1.** Risposta al gradino unitario di un sistema del primo ordine

- per  $t = \tau$  la risposta raggiunge circa il 63.2% del valore finale di regime.
- per  $t = 2\tau$  il valore corrisponde all'86.5% del valore di regime.
- per  $t = 3\tau$  si arriva a circa il 95% del valore di regime.

### 4.6.2 Tempo di assestamento

**Definizione 4.6.1** Il **tempo di assestamento**  $T_a$  è definito come il tempo necessario affinché la risposta del sistema a un ingresso (tipicamente un gradino unitario) rimanga stabilmente entro una banda di tolleranza prefissata, centrata attorno al valore finale della risposta.

Il **tempo di assestamento ad una data percentuale** rispetto al valore finale si può calcolare, per un sistema del primo ordine, mediante la relazione:

$$T_{aP\%} = -\tau \ln\left(\frac{P}{100}\right)$$

dove  $\tau$  è la costante di tempo del sistema e  $P$  rappresenta la percentuale di errore ammessa (ad esempio  $P = 2$  per una tolleranza del  $\pm 2\%$ , oppure  $P = 5$  per una tolleranza del  $\pm 5\%$ ).

**Dimostrazione** Abbiamo visto come la risposta di un sistema del primo ordine al gradino unitario è:

$$y(t) = 1 - e^{-t/\tau}$$

La risposta di un sistema del primo ordine non raggiunge istantaneamente il valore finale, ma vi si avvicina gradualmente nel tempo. Possiamo quindi definire l'**errore** come la differenza tra il valore finale (che vale 1, se la risposta è normalizzata) e il valore istantaneo della risposta  $y(t)$ :

$$\text{Errore}(t) = 1 - y(t) = 1 - 1 + e^{-t/\tau} = e^{-t/\tau}$$

Questo termine  $e^{-t/\tau}$  rappresenta quanto la risposta è ancora "lontana" dal valore di regime in un certo istante  $t$ . Per dire che il sistema si è **assestato**, vogliamo che questa distanza sia diventata abbastanza piccola, cioè entro una certa tolleranza percentuale stabilita a priori (ad esempio il 2% o il 5%).

In altre parole, imponiamo che l'errore in corrispondenza del **tempo di assestamento** sia uguale al valore della tolleranza. Poiché il tempo di assestamento è un istante ben preciso, lo indichiamo con un simbolo dedicato,  $T_{aP\%}$ , che rappresenta il tempo necessario affinché la risposta rimanga entro la banda di  $\pm P\%$  attorno al valore finale:

$$e^{-T_{aP\%}/\tau} = \frac{P}{100}$$

dove  $P$  indica la **percentuale di errore ammesso** (ad esempio 2 per una tolleranza del  $\pm 2\%$ , oppure 5 per una tolleranza del  $\pm 5\%$ ).

Risolvendo per  $T_{aP\%}$  ottengo:

$$e^{-T_{aP\%}/\tau} = \frac{P_{\text{tol}}}{100} \Rightarrow -\frac{T_a}{\tau} = \ln\left(\frac{P}{100}\right)$$

Quindi:

$$T_{aP\%} = -\tau \ln\left(\frac{P}{100}\right)$$

### Esempi

#### Tempo di assestamento al 5%

$$\begin{aligned} T_{a5\%} &= -\tau \ln\left(\frac{5}{100}\right) \\ &= -\tau \ln(0.05) \\ &= 2.996 \tau \approx 3\tau \end{aligned}$$

#### Tempo di assestamento al 2%

$$\begin{aligned} T_{a2\%} &= -\tau \ln\left(\frac{2}{100}\right) \\ &= -\tau \ln(0.02) \\ &= 3.912 \tau \approx 4\tau \end{aligned}$$

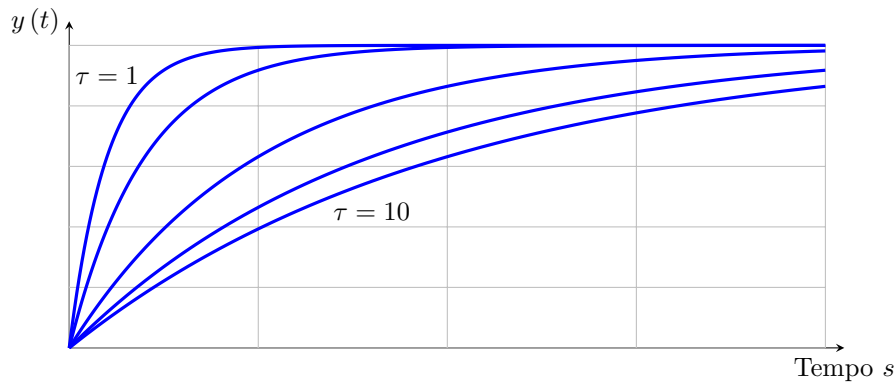
□

Al variare di  $\tau$  varia la velocità di risposta del sistema. Se  $\tau$  aumenta  $\Rightarrow T_a$  aumenta.

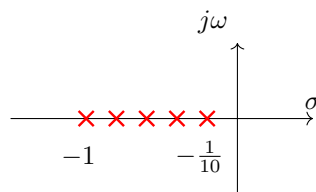
Valutando il polo del sistema:

$$G(s) = \frac{1}{1 + \tau s} \Rightarrow p = -\frac{1}{\tau}. \quad (4.14)$$

Osserviamo che, al diminuire della costante di tempo  $\tau$ , il polo si sposta verso sinistra nel piano complesso. Pertanto, poli più a sinistra (cioè con parte reale più negativa) corrispondono a costanti di tempo più piccole e quindi a **risposte più rapide** nel dominio del tempo.



**Figura 4.2.** Risposta al gradino unitario di un sistema del primo ordine al variare di  $\tau$



### 4.6.3 Sistema del primo ordine con uno zero

Se, oltre al polo, il sistema presenta anche uno zero (sistema proprio), la funzione di trasferimento può essere scritta come:

$$G(s) = \frac{1 + Ts}{1 + \tau s}. \quad (4.15)$$

Scomponendo il termine razionale, si ottiene:

$$G(s) = \frac{T}{\tau} + \frac{1 - T/\tau}{1 + \tau s} = \alpha + \frac{1 - \alpha}{1 + \tau s}, \quad (4.16)$$

dove  $\alpha = \frac{T}{\tau}$  rappresenta il rapporto tra le costanti di tempo associate rispettivamente allo zero e al polo del sistema.

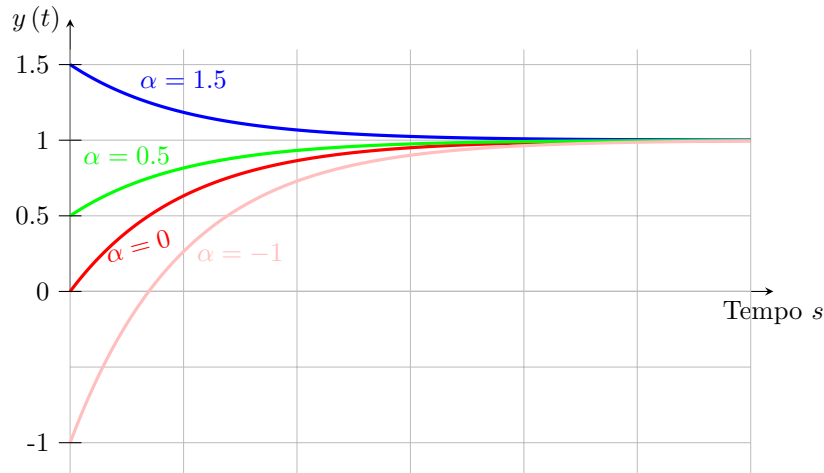
Applicando un ingresso a gradino unitario, la risposta nel dominio del tempo risulta:

$$y(t) = 1 + (\alpha - 1)e^{-t/\tau}. \quad (4.17)$$

Si osserva quindi che la presenza dello zero introduce un termine proporzionale al coefficiente  $\alpha$ , che modifica l'ampiezza della componente esponenziale e può rendere la risposta più rapida o più lenta a seconda che lo zero si trovi più vicino o più lontano dall'origine rispetto al polo ( $p = \alpha z$ ).

Il valore iniziale della risposta è pari a  $\alpha$ , mentre la sua pendenza iniziale risulta

$$\dot{y}(0) = \frac{1 - \alpha}{\tau}.$$



**Figura 4.3.** Risposta al gradino unitario di un sistema del primo ordine con uno zero al variare di  $\alpha$

## 4.7 Sistemi del secondo ordine

Un sistema lineare del secondo ordine può essere descritto, a meno di un fattore costante, dalla seguente funzione di trasferimento:

$$G(s) = \frac{\omega_n^2}{s^2 + 2\delta\omega_n s + \omega_n^2} = \frac{1}{1 + 2\delta\frac{s}{\omega_n} + \frac{s^2}{\omega_n^2}} \quad (4.18)$$

dove

- $\omega_n$  è la **pulsazione naturale** (rad/s).
- $\delta$  è il coefficiente di smorzamento (adimensionale).

La risposta al gradino unitario è data dalla relazione:

$$y(t) = \mathcal{L}^{-1} \left[ \frac{\omega_n^2}{s(s^2 + 2\delta\omega_n s + \omega_n^2)} \right] = 1 - Ae^{-\delta\omega_n t} \sin(\omega t + \varphi) \quad (4.19)$$

dove:

$$\begin{aligned} A &= \frac{1}{\sqrt{1-\delta^2}} \\ \omega &= \omega_n \sqrt{1-\delta^2} \\ \varphi &= \arctan \frac{\sqrt{1-\delta^2}}{\delta} \\ &= \arcsin \sqrt{1-\delta^2} = \arccos \delta \end{aligned}$$

### 4.7.1 $|\delta| > 1$

In questo caso le radici sono reali e distinte

$$\begin{aligned} p_{1,2} &= \sigma \pm j\omega \\ &= -\delta\omega_n \pm \omega_n \sqrt{\delta^2 - 1} \end{aligned} \quad (4.20)$$

con:

$$\omega_n = |p_1| = |p_2| = \sqrt{\sigma^2 + \omega^2} \quad (4.21)$$

Valgono quindi le seguenti relazioni:

$$\begin{aligned} \sigma &= -\delta\omega_n \\ \omega &= \omega_n \sqrt{1-\delta^2} \end{aligned} \quad (4.22)$$

### 4.7.2 $|\delta| = 1$

Le radici sono reali e coincidenti quindi avremo:

$$p_{1,2} = -\delta\omega_n \quad (4.23)$$

**4.7.3**  $|\delta| < 1$ 

In questo caso le radici sono complesse coniugate

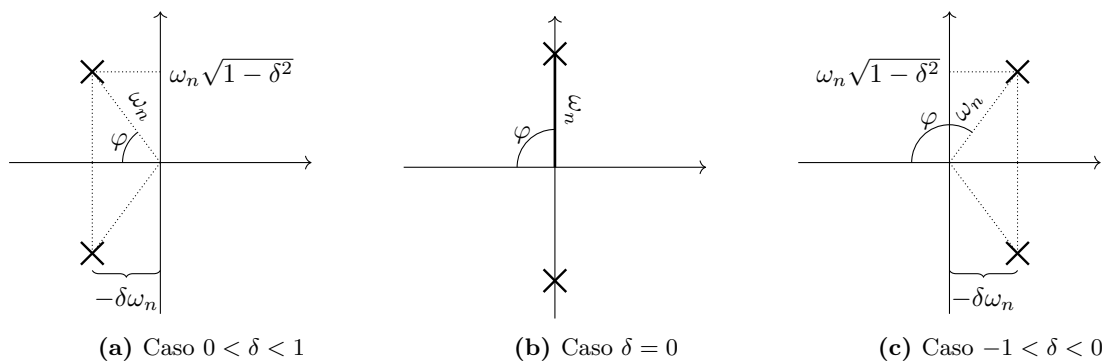
$$\begin{aligned}
 p_{1,2} &= \sigma \pm j\omega \\
 &= -\delta\omega_n \pm j\omega_n\sqrt{1-\delta^2} \\
 &= -\omega_n \cos \varphi \pm j\omega_n \sin \varphi
 \end{aligned} \tag{4.24}$$

con:

$$\begin{aligned}
 \omega_n &= |p_1| = |p_2| = \sqrt{\sigma^2 + \omega^2} \\
 \delta &= \cos \varphi = \frac{-\sigma}{\sqrt{\sigma^2 + \omega^2}}
 \end{aligned} \tag{4.25}$$

Valgono quindi le seguenti relazioni:

$$\begin{aligned}
 \sigma &= -\delta\omega_n \\
 \omega &= \omega_n\sqrt{1-\delta^2}
 \end{aligned} \tag{4.26}$$



**Figura 4.4.** Posizione dei poli al variare di  $\delta$  nel caso di  $|\delta| < 1$

### 4.7.4 Parametri caratteristici della risposta

Per valutare la qualità del transitorio, si definiscono alcune grandezze standard:

**Definizione 4.7.1** Come per i sistemi del primo ordine, il **tempo di assestamento**  $T_a$  è definito come il tempo necessario affinché la risposta del sistema a un ingresso (tipicamente un gradino unitario) rimanga stabilmente entro una banda di tolleranza prefissata, centrata attorno al valore finale della risposta.

Per un sistema del secondo ordine, il **tempo di assestamento ad una data percentuale** rispetto al valore finale si può calcolare, mediante la relazione:

$$T_{aP\%} = \frac{1}{\delta\omega_n} \ln\left(\frac{P}{100}\right) = \tau \ln\left(\frac{P}{100}\right)$$

dove  $\tau = \frac{1}{\delta\omega_n}$  è la costante di tempo e  $P$  rappresenta la percentuale di errore ammessa.

**Dimostrazione** Nel caso di sistema **sottosmorzato** ( $0 < \delta < 1$ ), la risposta al gradino unitario risulta:

$$y(t) = 1 - \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t + \varphi)$$

dove:

$$\omega_d = \omega_n \sqrt{1-\delta^2}, \quad \varphi = \arctan\left(\frac{\sqrt{1-\delta^2}}{\delta}\right)$$

La risposta di un sistema del secondo ordine non raggiunge istantaneamente il valore finale, ma vi si avvicina nel tempo in modo oscillante e smorzato. Possiamo quindi definire l'**errore** come la differenza tra il valore finale (che vale 1, se la risposta è normalizzata) e il valore istantaneo della risposta  $y(t)$ :

$$\text{Errore}(t) = 1 - y(t) = \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t + \varphi)$$

Poiché il termine  $\sin(\omega_d t + \varphi)$  è compreso tra  $-1$  e  $+1$ , l'errore è limitato da un **inviluppo esponenziale**:

$$|\text{Errore}(t)| \leq \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}}$$

Per dire che il sistema si è **assestato**, vogliamo che tale errore sia diventato sufficientemente piccolo, cioè entro una certa tolleranza percentuale prefissata  $P\%$ . In corrispondenza del **tempo di assestamento**  $T_{aP\%}$ , imponiamo dunque:

$$\frac{e^{-\delta\omega_n T_{aP\%}}}{\sqrt{1-\delta^2}} = \frac{P}{100}$$

Risolvendo per  $T_{aP\%}$  otteniamo:

$$-\delta\omega_n T_{aP\%} = \ln\left(\frac{P}{100} \sqrt{1-\delta^2}\right) \Rightarrow T_{aP\%} = -\frac{1}{\delta\omega_n} \ln\left(\frac{100}{P\sqrt{1-\delta^2}}\right)$$

$$T_{aP\%} = \frac{1}{\delta\omega_n} \ln\left(\frac{100}{P\sqrt{1-\delta^2}}\right)$$

Poiché il termine  $\sqrt{1 - \delta^2}$  varia poco per i valori tipici di smorzamento ( $0.4 < \delta < 0.8$ ), si trascura per semplicità ottenendo la **formula approssimata**:

$$T_{a_{P\%}} \approx \frac{1}{\delta\omega_n} \ln\left(\frac{100}{P}\right)$$

### Esempi

**Tempo di assestamento al 5%**

$$\begin{aligned} T_{a_{5\%}} &\approx \frac{1}{\delta\omega_n} \ln\left(\frac{100}{5}\right) \\ &= \frac{1}{\delta\omega_n} \ln(20) \\ &= \frac{2.996}{\delta\omega_n} \approx \frac{3}{\delta\omega_n} \end{aligned}$$

**Tempo di assestamento al 2%**

$$\begin{aligned} T_{a_{2\%}} &\approx \frac{1}{\delta\omega_n} \ln\left(\frac{100}{2}\right) \\ &= \frac{1}{\delta\omega_n} \ln(50) \\ &= \frac{3.912}{\delta\omega_n} \approx \frac{4}{\delta\omega_n} \end{aligned}$$

□

**Definizione 4.7.2** Il **Tempo di ritardo**  $T_r$  è il tempo necessario affinché l'uscita raggiunga il 50% del valore finale.

Il tempo di ritardo si può approssimare mediante la relazione:

$$T_r \approx \frac{1 + 0.7\delta}{\omega_n}$$

Si tratta di un'approssimazione, poiché imponendo  $y(t) = 0.5$  si ottiene un'equazione **trascendente** (il termine incognito  $T_r$  compare sia all'interno del seno che dell'esponenziale).

Per questo motivo non è possibile ricavare una soluzione analitica esplicita e il valore di  $T_r$  viene determinato mediante **metodi numerici** oppure tramite **formule approssimate** ricavate da curve standard di risposta.

**Definizione 4.7.3** Il **Tempo di salita**  $T_s$  è il tempo che serve perché l'uscita passi dal 10% al 90% del valore finale.

Il tempo di salita si può approssimare mediante la relazione:

$$T_s \approx \frac{\pi - \arccos \delta}{\omega_n \sqrt{1 - \delta^2}}$$

Anche in questo caso si tratta di una approssimazione, poiché durante la ricerca della soluzione, si ottiene un'equazione trascendente.

**Definizione 4.7.4** L'**istante di massima sovraelongazione**  $T_m$  o **tempo di picco** è il momento in cui si verifica la massima sovraelongazione.

Il tempo di picco si può calcolare mediante la relazione:

$$T_m = \frac{\pi}{\omega_n \sqrt{1 - \delta^2}}$$

**Dimostrazione** Consideriamo un sistema del secondo ordine con funzione di trasferimento normalizzata

$$G(s) = \frac{\omega_n^2}{s^2 + 2\delta\omega_n s + \omega_n^2},$$



dove  $\omega_n > 0$  è la pulsazione naturale e  $\delta$  è il fattore di smorzamento. Per un ingresso a gradino unitario, la risposta nel tempo (caso sotto smorzato  $0 < \delta < 1$ ) è

$$y(t) = 1 - \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t + \varphi), \quad \omega_d = \omega_n \sqrt{1-\delta^2}, \quad \varphi = \arctan\left(\frac{\sqrt{1-\delta^2}}{\delta}\right).$$

### Ricerca dell'istante di picco

Il tempo di picco  $T_M$  è l'istante del primo massimo locale della risposta. Si impone quindi  $\dot{y}(T_M) = 0$ . Derivando  $y(t)$  si ottiene

$$\begin{aligned} \dot{y}(t) &= -\frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \left[ \omega_d \cos(\omega_d t + \varphi) - \delta\omega_n \sin(\omega_d t + \varphi) \right] \\ &= \omega_n \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \left[ \delta \sin(\omega_d t + \varphi) - \sqrt{1-\delta^2} \cos(\omega_d t + \varphi) \right] \\ &= \omega_n \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \left[ \cos \varphi \sin(\omega_d t + \varphi) - \sin \varphi \cos(\omega_d t + \varphi) \right] \\ &= \omega_n \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t + \varphi - \varphi) \\ &= \omega_n \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t) \end{aligned} \quad (4.27)$$

La condizione  $\dot{y}(T_M) = 0$  implica

$$\omega_n \frac{e^{-\delta\omega_n T_M}}{\sqrt{1-\delta^2}} \sin(\omega_d T_M) = 0 \quad \Leftrightarrow \quad \sin(\omega_d T_M) = 0 \quad \Leftrightarrow \quad \omega_d T_M = k\pi \quad k \in \mathbb{Z} \quad (4.28)$$

segue che

$$T_M = \frac{k\pi}{\omega_d} \quad (4.29)$$

Il primo massimo locale dopo  $t = 0$  corrisponde a  $k = 1$ , dunque

$$T_M = \frac{\pi}{\omega_d} \quad (4.30)$$

EsPLICITANDO in funzione di  $\delta$  e  $\omega_n$ :

$$T_M = \frac{\pi}{\omega_n \sqrt{1-\delta^2}}, \quad 0 < \delta < 1 \quad (4.31)$$

### Osservazioni

- Per  $\delta \geq 1$  la risposta è monotona (nessuna sovra elongazione) e non esiste un picco locale: si pone quindi  $T_M$  non definito.
- La relazione tra tempo di picco e massima sovra elongazione è

$$y(T_M) - 1 = e^{-\delta\omega_n T_M}$$

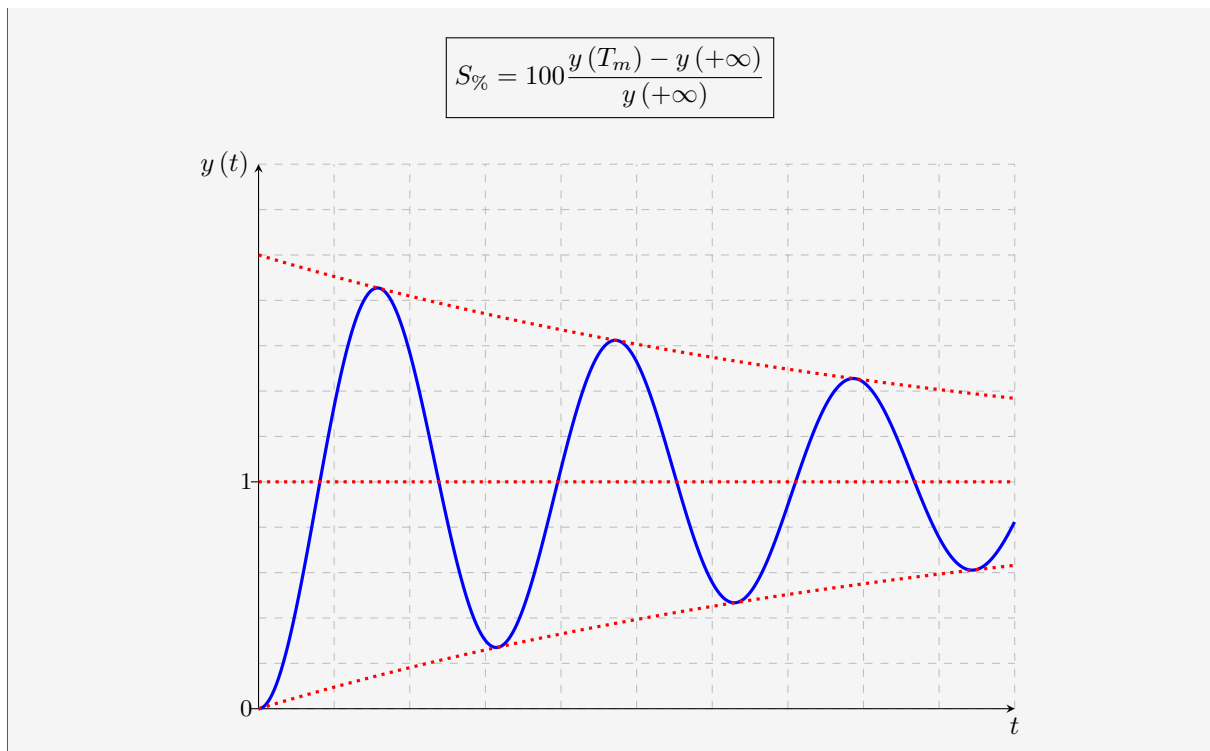
□

**Definizione 4.7.5** **Massima sovraelongazione  $S$ :** è la differenza tra il valore massimo raggiunto dall'uscita e il valore finale.

Noto il tempo di picco, il valore di massima sovra elongazione si può calcolare mediante la relazione:

$$S = 1 + e^{-\frac{\delta\pi}{\sqrt{1-\delta^2}}}$$

Volendo esprimerlo in percentuale rispetto al valore a regime  $y(+\infty)$  abbiamo:



## 4.8 Sistemi del secondo ordine con zero

Consideriamo la funzione di trasferimento di un sistema del secondo ordine con uno zero:

$$G(s) = \frac{\omega_n^2(1 + Ts)}{s^2 + 2\delta\omega_n s + \omega_n^2} \quad (4.32)$$

La stessa può essere riscritta come:

$$G(s) = G_0(s) + TsG_0(s), \quad G_0(s) = \frac{\omega_n^2}{s^2 + 2\delta\omega_n s + \omega_n^2} \quad (4.33)$$

Da cui la risposta al gradino è data da:

$$y(t) = \mathcal{L}^{-1} \left\{ \frac{G_0(s)}{s} \right\} + T \mathcal{L}^{-1} \left\{ s \frac{G_0(s)}{s} \right\} = y_0(t) + T \frac{dy_0(t)}{dt} \quad (4.34)$$

Abbiamo visto in precedenza, nel calcolo del tempo di picco che:

$$\frac{dy(t)}{dt} = \omega_n \frac{e^{-\delta\omega_n t}}{\sqrt{1-\delta^2}} \sin(\omega_d t) \quad (4.35)$$

Poiché questo termine tende a zero per  $t \rightarrow +\infty$ , la risposta al gradino di un sistema del secondo ordine con uno zero si comporta, per tempi lunghi, come quella di un sistema del secondo ordine privo di zeri. Questo risultato è intuitivo: la presenza di uno zero non modifica i modi della funzione di trasferimento, ma influisce solo sui coefficienti dei termini parziali (i cosiddetti fratti semplici).

In altre parole, si ha:

$$y(+\infty) = y_0(+\infty) + T\dot{y}_0(+\infty) = y_0(+\infty) \quad (4.36)$$

il cui valore dipende solo dalla posizione dei poli.

Inoltre, abbiamo che

$$\frac{dy_0(0^+)}{dt} = 0 \Rightarrow y(0^+) = 0 \quad (4.37)$$

Quindi, la risposta al gradino di un sistema del secondo ordine con uno zero inizia a evolversi da un valore nullo, proprio come accade per la risposta al gradino di un sistema del secondo ordine privo di zeri.

**Osservazione 4.8.1** Osserviamo inoltre che per  $t > 0$ , vale:

$$\frac{dy(t)}{dt} = \frac{dy_0(t)}{dt} + T \frac{d^2 y_0(t)}{dt^2} \quad (4.38)$$

che valutato in  $0^+$ :

$$\frac{dy(0^+)}{dt} = \frac{dy_0(0^+)}{dt} + T \frac{d^2 y_0(0^+)}{dt^2} = 0 + T\omega_n^2 = \omega_n^2 \quad (4.39)$$

che assume un valore positivo se lo zero è a **fase minima** ( $T > 0$ ), negativo se è a **fase non minima** ( $T < 0$ ) e nullo per  $T = 0$ .

In definitiva, la presenza di uno zero in un sistema del secondo ordine non modifica né il valore iniziale né quello finale della risposta al gradino. Ciò che cambia è il modo in cui la risposta si stacca dall'asse dei tempi.

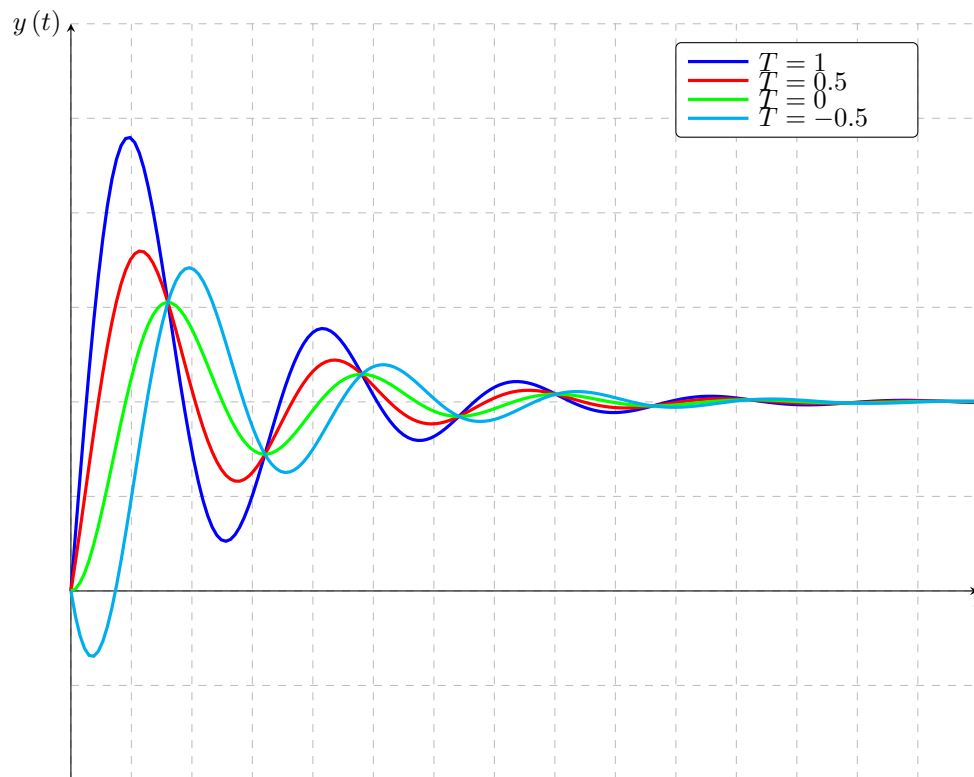
Se lo zero si trova nel semipiano sinistro, la derivata della risposta al gradino nell'origine è positiva; al contrario, in assenza di zeri, tale derivata è nulla. Se invece lo zero si trova nel semipiano destro, la derivata iniziale è negativa: in questo caso la risposta del sistema è più lenta rispetto ai casi precedenti.

In particolare, quando lo zero è a fase non minima, la risposta al gradino inizia evolvendo verso valori negativi, anziché seguire subito l'andamento dell'ingresso (pari a 1) cioè, il sistema inizialmente si allontana dall'ingresso prima di raggiungerlo.

Consideriamo il sistema

$$G(s) = \frac{4(1 + Ts)}{s^2 + 0.8s + 4} \quad (4.40)$$

se andiamo a valutare la risposta a variare di  $T$  otteniamo la situazione in figura



Nel grafico sono rappresentate quattro curve corrispondenti a diversi valori di  $T$ :

$$\begin{cases} T = 1 & \text{(linea blu)} \\ T = 0.5 & \text{(linea rossa)} \\ T = 0 & \text{(linea verde)} \\ T = -0.5 & \text{(linea azzurra)} \end{cases}$$

Si nota che tutte le risposte del sistema presentano un andamento oscillatorio smorzato, caratteristico di un sistema del secondo ordine con poli complessi coniugati. Tuttavia, al variare del parametro  $T$ , la forma della risposta nel tempo cambia sensibilmente, mostrando come questo parametro influenzi la dinamica complessiva del sistema.

Quando  $T$  assume valori positivi, come nei casi  $T = 0.5$  e  $T = 1$ , la risposta tende a mostrare un sovrallungamento più pronunciato, ossia un picco iniziale di maggiore ampiezza rispetto al valore finale di regime. All'aumentare di  $T$ , tale effetto si accentua, indicando che il termine  $(1 + Ts)$  nel numeratore introduce un comportamento più reattivo e amplifica la componente transitoria.

Nel caso  $T = 0$ , lo zero si annulla e quindi abbiamo un semplice sistema del secondo ordine.

Nel caso  $T = -0.5$ , la risposta si comporta in maniera differente: si osserva un'inversione di fase iniziale e un andamento complessivamente più oscillatorio e irregolare. Questo indica che un valore negativo di  $T$  altera la risposta introducendo una componente che tende a contrastare la dinamica naturale del sistema.

## 4.9 Sistemi di ordine superiore al secondo

Quando un sistema ha due poli reali e uno zero, può mostrare un comportamento particolare. Se lo zero ha una costante di tempo simile a quella di uno dei poli, la risposta del sistema diventa molto simile a quella di un sistema del primo ordine. In pratica, polo e zero "quasi coincidenti" si compensano tra loro, e possiamo semplificare il modello eliminandoli.

Lo stesso concetto vale anche per sistemi di ordine superiore. Se un polo e uno zero sono vicini nel piano complesso, possiamo considerarli "accoppiati" e rimuoverli per ottenere un modello più semplice ma con una risposta al gradino molto simile a quella originale. Questa semplificazione è valida solo se le coppie eliminate corrispondono a dinamiche stabili, cioè con parte reale negativa.

Dopo questa riduzione restano alcuni poli che determinano in modo principale l'andamento del sistema: i poli dominanti. Sono quelli più vicini all'origine nel piano complesso, quindi con parte reale minore (in valore assoluto).

La risposta al gradino di un sistema dominato da pochi poli può essere quindi approssimata da quella di un sistema più semplice, che contiene solo i poli dominanti ma mantiene lo stesso guadagno statico. In sintesi, i poli dominanti sono quelli che "governano" la risposta del sistema, mentre gli altri hanno un effetto trascurabile.

**Esempio 4.9.1** Consideriamo la seguente funzione di trasferimento:

$$G(s) = \frac{1100(s+30)}{(s+10)(s^2+0.2s+2.0023)(s^2+5s+123)}$$

E' semplice vedere come la funzione di trasferimento ha uno zero e cinque poli (di cui 4 complessi coniugati).

$$z_1 = -30$$

$$p_1 = -10 \quad p_{2,3} = -0.1 \pm 1.4116j \quad p_{4,5} = -2.5 \pm 10.8051j$$

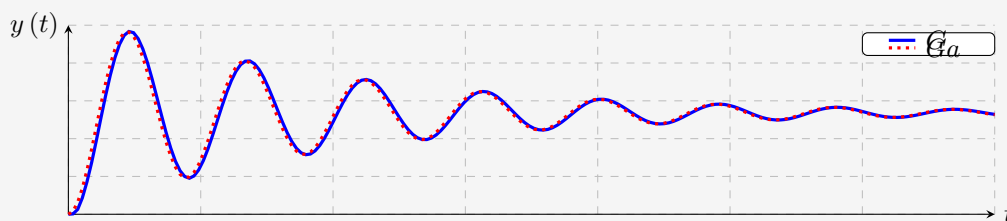
Non è possibile effettuare una cancellazione poli-zeri, tuttavia i poli  $p_{2,3}$  risultano i poli dominanti, quindi è possibile ricondurre il sistema ad uno del secondo ordine.

Cerchiamo il guadagno statico, ponendo  $s = 0$ :

$$G(0) = \frac{1100(30)}{(10)(2.0023)(123)} \approx 13.4$$

Quindi la funzione di trasferimento originaria è approssimabile a:

$$G_a(s) = \frac{13.4}{\left(\frac{s^2}{2.0023} + \frac{0.2}{2.0023}s + 1\right)}$$



**Esempio 4.9.2** Consideriamo la seguente funzione di trasferimento che è un rimaneggiamento della funzione dell'esempio precedente:

$$G(s) = \frac{1100(s+30)}{(s+0.01)(s^2+15s+200)(s^2+5s+123)}$$

Anche in questo caso la funzione di trasferimento ha uno zero e cinque poli (di cui 4 complessi coniugati).

$$z_1 = -30$$

$$p_1 = -0.01 \quad p_{2,3} = -7.5 \pm 11.9896j \quad p_{4,5} = -2.5 \pm 10.8051j$$

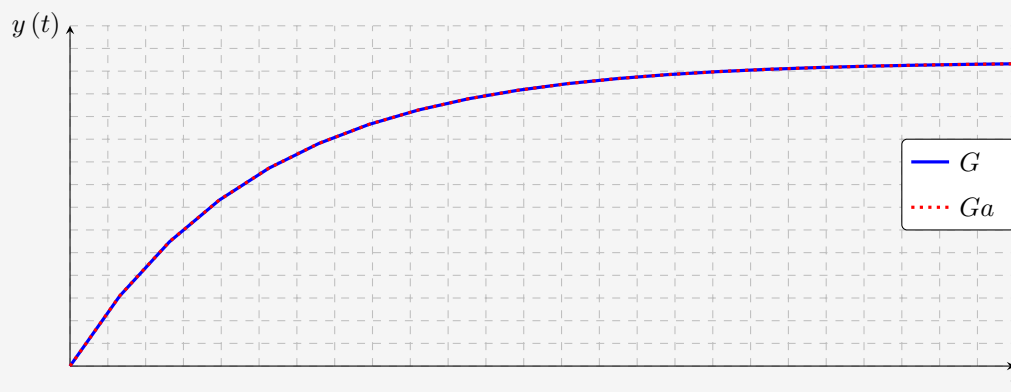
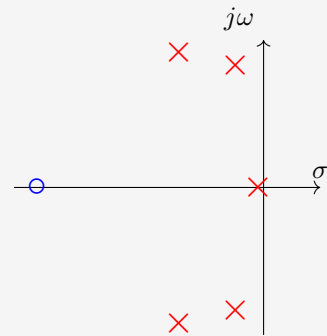
Anche qui non ci sono coppie di poli e zeri sui quali è possibile fare una cancellazione, tuttavia il polo  $p_1$  avendo parte reale molto più piccola degli altri poli, risulta essere un polo dominante. In questo caso è possibile approssimare il sistema addirittura ad uno del primo ordine.

Cerchiamo il guadagno statico, ponendo  $s = 0$ :

$$G(0) = \frac{1100(30)}{(0.01)(200)(123)} \approx 134.15$$

La funzione di trasferimento originaria è approssimabile a:

$$G_a(s) = \frac{134.15}{\left(\frac{s}{0.01}s + 1\right)}$$



**Esempio 4.9.3** Consideriamo la seguente funzione di trasferimento:

$$G(s) = \frac{(s + 0.25)}{(s + 0.2)(s^2 + 10s + 425)}$$

Chiaramente la funzione di trasferimento ha uno zero e tre poli (di cui due complessi coniugati).

$$z_1 = -0.25$$

$$p_1 = -0.2 \quad p_{2,3} = -5 \pm 20j$$

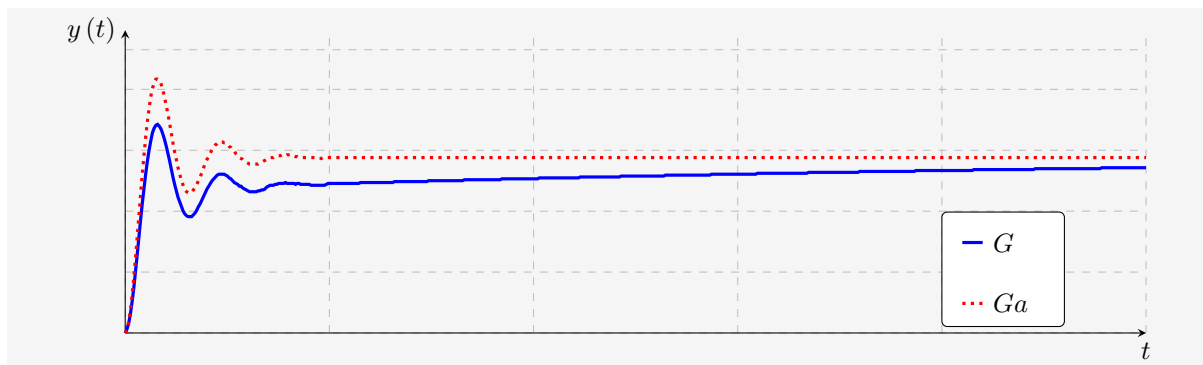
Cerchiamo il guadagno statico, ponendo  $s = 0$ :

$$G(0) = \frac{0.25}{(0.2)(425)} \approx 0.0029$$

Lo zero e il polo semplice, avendo costanti di tempo simili, possono essere eliminate.

La funzione di trasferimento originaria è approssimabile a:

$$G_a(s) = \frac{5.8824}{\left(\frac{s^2}{425} + \frac{10s}{425}s + 1\right)}$$



## 4.10 I modelli in forma *spazio di stato*

Generalmente un sistema  $S$  espresso nella forma *spazio di stato* ha la seguente struttura:

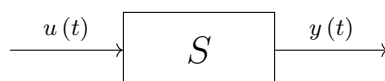
$$S = \begin{cases} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= g(x(t), u(t)) \end{cases} \quad (4.41)$$

dove

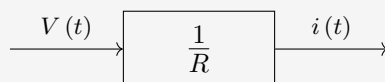
- $u(t)$  è l'**ingresso** del sistema (variabili esogene).
- $y(t)$  è l'**uscita** del sistema (variabili di interesse o misure).
- $x(t)$  è lo **stato** del sistema.
- $f(\cdot)$  è la **funzione di transizione di stato**. Descrive il movimento dello stato, ovvero in che modo avviene il passaggio da uno stato a un altro.
- $g(\cdot)$  è la **funzione di uscita**. Descrive l'uscita del sistema.

### 4.10.1 Sistemi statici e dinamici

Intuitivamente, possiamo considerare un sistema come una scatola nera (black-box) attraverso la quale gli ingressi (ossia le cause) agiscono per fornire un'uscita (ossia gli effetti).



**Esempio 4.10.1** Un esempio di sistema **statico** è quello di un resistore di resistenza  $R$  sollecitato con un generatore di tensione  $V(t)$ . Si vuole conoscere la corrente  $i(t)$  che scorre nel resistore. Iniziamo identificando gli ingressi, le uscite e la relazione tra di esse. In questo caso vi è solo un ingresso  $V(t)$  ed una sola uscita  $i(t)$ . Il legame tra queste due variabili è ben noto come legge di Ohm:  $V(t) = Ri(t)$ .



In tal caso, indicando con  $u(t)$  l'ingresso in tensione ( $V(t)$ ) e con  $y(t)$  l'uscita in corrente ( $i(t)$ ), l'equazione che lega l'ingresso con l'uscita del sistema è:

$$y(t) = \frac{1}{R} u(t) \quad (4.42)$$

Il valore  $y(t)$  assunto dall'uscita al tempo  $t$  dipende solo dal valore corrente  $u(t)$  assunto dall'ingresso al tempo  $t$ . In tal caso, il sistema è detto **statico**.

Si parla invece di sistema **dinamico** quando il valore di un'uscita dipende da tutti i valori assunti dall'ingresso in precedenza. Più in dettaglio, in un sistema dinamico il valore  $y(t)$  assunto dall'uscita

al tempo  $t$  dipende dal valore  $u(\cdot)$  assunto dall'ingresso in tutti gli istanti di tempo nell'intervallo  $[0, t]$  nonché dalle cosiddette condizioni iniziali, che riassumono la storia passata del sistema, ovvero prima dell'applicazione della sollecitazione d'ingresso.

**Esempio 4.10.2** Un esempio di sistema **dinamico** è quello di un condensatore di capacità  $C$  sollecitato con un generatore di corrente  $i(t)$ . Si vuole conoscere la tensione  $v(t)$  ai capi del condensatore.

Iniziamo come visto in precedenza, identificando gli ingressi, le uscite e la relazione tra di esse. In questo caso vi è solo un ingresso  $i(t)$  ed una sola uscita  $v(t)$ . Il legame tra queste due variabili è ben noto ed è:  $i(t) = C \frac{dv(t)}{dt}$ . Invertendo la relazione ed indicando con  $u(t)$  l'ingresso in corrente e con  $y(t)$  l'uscita in tensione, otteniamo:

$$\frac{dy(t)}{dt} = \frac{1}{C} x(t) \quad (4.43)$$

che è una equazione differenziale lineare a coefficienti costanti del primo ordine. Risolvendo l'equazione, si ottiene:

$$y(t) = \frac{1}{C} \int_0^t x(\tau) d\tau + y(0) \quad (4.44)$$

Nel sistema dinamico in questione per conoscere il valore corrente dell'uscita è necessaria la conoscenza della condizione iniziale nonché dell'ingresso in tutto l'intervallo  $[0, t]$ , ossia è necessaria la conoscenza della storia del sistema. Per questo motivo si dice che un sistema dinamico ha memoria.

**Esempio 4.10.3** Modello di un serbatoio.

Consideriamo lo schema di un serbatoio rappresentato in figura. Andiamo a definire le variabili del modello ed i suoi parametri:

altezza serbatoio	$h$	$m$
flusso d'ingresso	$u$	$m^3/s$
flusso d'uscita	$q$	$m^3/s$
velocità d'uscita	$v$	$m/s$
sezione serbatoio	$A$	$m^2$
sezione apertura	$a$	$m^2$
accelerazione di gravità	$g$	$m/s^2$

Fruttando le leggi della fisica:

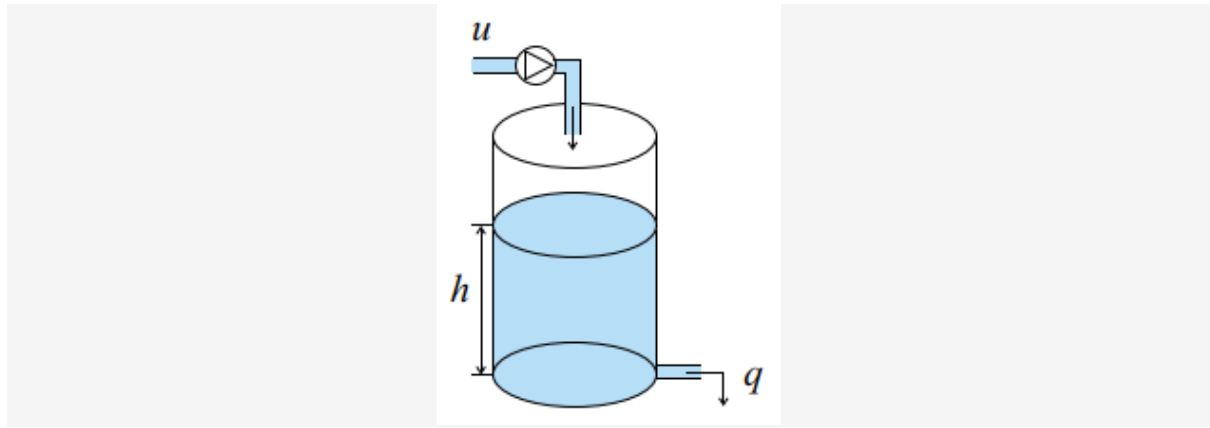
$$\begin{aligned} v(t) &= \sqrt{2gh(t)} && \text{Legge di Torricelli} \\ q(t) &= av(t) && \text{Flusso d'uscita} \\ \frac{d}{dt}[Ah(t)] &= u(t) - q(t) && \text{Bilancio di Massa} \end{aligned}$$

Il modello matematico di un serbatoio è:

$$\begin{cases} \dot{h}(t) = -\frac{a\sqrt{2g}}{A}\sqrt{h(t)} + \frac{1}{A}u(t) \\ q(t) = a\sqrt{2g}\sqrt{h(t)} \end{cases} \quad (4.45)$$

dove  $h(t)$  rappresenta lo stato,  $q(t)$  l'uscita,  $u(t)$  l'ingresso.





## 4.11 Sistemi lineari a tempo continuo

### 4.11.1 Equazione di Lagrange

Consideriamo il seguente problema di Cauchy del primo ordine

$$\begin{cases} \dot{x}(t) &= ax(t) \\ x(0) &= x_0 \end{cases} \quad (4.46)$$

dove  $a$  è uno scalare.

Risolviamo il problema di Cauchy utilizzando la Trasformata di Laplace:

$$\begin{aligned} \mathcal{L}\{\dot{x}(t) = ax(t)\} &\rightarrow sX(s) - X(0) = aX(s) \\ X(s)(s - a) - x_0 &= 0 \\ X(s) &= \frac{x_0}{s - a} \end{aligned} \quad (4.47)$$

Facendo la trasformata inversa, troviamo che esista una e una sola soluzione:  $x(t) = e^{at}x_0$ .

Analogamente, consideriamo il seguente problema di Cauchy del primo ordine in cui includiamo l'ingresso

$$\begin{cases} \dot{x}(t) &= ax(t) + bu(t) \\ x(0) &= x_0 \end{cases} \quad (4.48)$$

dove  $a, b$  sono scalari.

Risolviamo il problema di Cauchy utilizzando la Trasformata di Laplace:

$$\begin{aligned} \mathcal{L}\{\dot{x}(t) = ax(t) + bu(t)\} &\rightarrow sX(s) - X(0) = aX(s) + bU(s) \\ X(s)(s - a) - x_0 &= bU(s) \\ X(s) &= \frac{x_0}{s - a} + \frac{1}{s - a}bU(s) \end{aligned} \quad (4.49)$$

Facendo la trasformata inversa, troviamo che la soluzione esiste ed è unica:

$$\begin{aligned} x(t) &= \underbrace{e^{at}x_0}_{\text{risposta libera}} + \underbrace{\int_0^t e^{a(t-\tau)}bu(\tau)d\tau}_{\text{risposta forzata}} \\ &= x_l(t) + x_f(t) \end{aligned} \quad (4.50)$$

N.B. L'antitrasformata della risposta forzata la si ottiene utilizzando il teorema della convoluzione.

La **risposta libera**  $x_l(t)$  è quella componente influenzata **solo** dalle condizioni iniziali  $x_0$ .

La **risposta forzata**  $x_f(t)$  è quella componente influenzata **solo** dall'ingresso  $u(t)$ .

Nel caso in cui  $a, b$  siano dei vettori, il problema di Cauchy è riscrivibile opportunamente come:

$$\begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{cases} \quad (4.51)$$

la cui soluzione esiste ed è unica ed è:

## Equazione di Lagrange

$$\mathbf{x}(t) = \underbrace{e^{\mathbf{A}t} \mathbf{x}_0}_{\text{risposta libera}} + \underbrace{\int_0^t e^{\mathbf{A}(t-\tau)} \mathbf{B} u(\tau) d\tau}_{\text{risposta forzata}} \quad (4.52)$$

## 4.12 Sistemi Lineari a Tempo Discreto

Nel passato, i plant e i controllori sono stati implementati utilizzando elettronica analogica, che per sua natura è di tipo continuo nel tempo. Al giorno d'oggi, grazie all'avvento dei microprocessori, sia i plant che i controllori possono essere utilizzati all'interno di sistemi digitali. I sistemi a Tempo Discreto (TD) esprimono relazioni fra variabili *campionate* ad intervalli  $T_s$  di tempo.

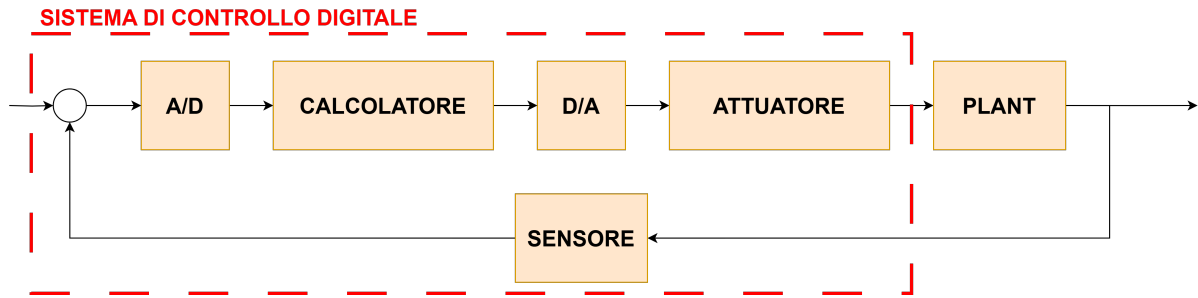


Figura 4.5. Schema di un sistema di controllo digitale.

Un microprocessore non è in grado di leggere un segnale analogico, si rende necessario campionare il segnale.

TODO schema segnale analogico, digitale ecc.

**Esempio 4.12.1 Dinamica degli studenti in un corso di laurea**

Si supponga di voler modellare la dinamica degli studenti del Politecnico iscritti ad un corso di laurea triennale. Consideriamo le seguenti assunzioni:

- La durata del corso è di 3 anni.
- La percentuale di studenti promossi, ripetenti e abbandoni sia costante per ogni anno.
- Non è possibile iscriversi direttamente al II e III anno.
- Non sono ammessi studenti fuori corso.

Il modello della dinamica degli studenti può essere espresso come un sistema di equazioni alle differenze di ordine 3:

$$\begin{cases} x_1(k+1) = \beta_1 x_1(k) + u(k) \\ x_2(k+1) = \alpha_1 x_1(k) + \beta_2 x_2(k) \\ x_3(k+1) = \alpha_2 x_2(k) + \beta_3 x_3(k) \\ y(k) = \alpha_3 x_3(k) \end{cases} \quad (4.53)$$

dove:

- $k$  rappresenta l'anno accademico;
- $x_i(k)$  è il numero di iscritti all' $i$ -esimo anno di corso nell'anno  $k$  ( $i = 1, 2, 3$ );
- $u(k)$  è il numero di matricole nell'anno  $k$ ;
- $y(k)$  è il numero di laureati nell'anno  $k$ ;
- $\alpha_i$  rappresenta il tasso di promossi nell'anno di corso  $i$ -esimo;

- $\beta_i$  rappresenta il tasso di ripetenti nell'anno di corso  $i$ -esimo;

Andando a rappresentare il modello in forma matriciale:

$$\begin{cases} x(k+1) &= \begin{bmatrix} \beta_1 & 0 & 0 \\ \alpha_1 & \beta_2 & 0 \\ 0 & \alpha_2 & \beta_3 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(k) \\ y(k) &= \begin{bmatrix} 0 & 0 & \alpha_3 \end{bmatrix} x(k) \end{cases} \quad (4.54)$$



### 5.1 Stabilità nei sistemi LTI TC

$$\begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \quad (5.1)$$

Sappiamo che l'eq di Lagrange è

$$\mathbf{x}(t) = \underbrace{e^{\mathbf{A}t}\mathbf{x}_0}_{\text{risposta libera}} + \underbrace{\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau) d\tau}_{\text{risposta forzata}} \quad (5.2)$$

Per lo studio della stabilità, abbiamo bisogno di osservare solo la componente libera della risposta, ossia:

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0 \quad (5.3)$$

La stabilità asintotica del sistema dipende da come si comporta  $e^{\mathbf{A}t}$  quanto  $t \rightarrow \infty$ . Quindi: determinare la stabilità equivale a studiare il comportamento dell'esponenziale di una matrice.

Se avessimo un sistema SISO, avremmo:

$$x(t) = e^{at}x_0 \quad (5.4)$$

E quindi per studiare la stabilità del sistema ci basterebbe considerare semplicemente il valore di  $a$  avendo:

- **Instabilità** per  $a > 0$
- **Semplice stabilità** per  $a = 0$
- **Asintotica stabilità** per  $a < 0$

Sarebbe tutto molto semplice se la matrice  $\mathbf{A}$  fosse diagonale in quanto anche  $e^{\mathbf{A}t}$  sarebbe diagonale:

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & \dots & \dots & \lambda_n \end{bmatrix} \Rightarrow e^{\mathbf{A}t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \ddots & \ddots & 0 \\ \vdots & \dots & \dots & e^{\lambda_n t} \end{bmatrix} \quad (5.5)$$

e anche in questo caso, il comportamento di ogni componente di  $x(t)$  sarebbe immediato:

- se  $\text{Re}(\lambda_i) < 0$ , quel modo decade  $\rightarrow$  stabile;
- se  $\text{Re}(\lambda_i) > 0$ , quel modo diverge  $\rightarrow$  instabile;
- se  $\text{Re}(\lambda_i) = 0$ , quel modo è limitato  $\rightarrow$  semplicemente stabile;

Ma la maggior parte delle matrici  $\mathbf{A}$  non è diagonale! Come facciamo allora a capire che succede?

### 5.2 Forma di Jordan

Ricordiamo che l'espansione in serie di Taylor di un esponenziale è definito come:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (5.6)$$

Riscriviamo  $e^{\mathbf{A}t}$  come serie di Taylor:

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2 t^2}{2!} + \frac{\mathbf{A}^3 t^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{\mathbf{A}^n t^n}{n!} \quad (5.7)$$

Notiamo come sia abbastanza complicato calcolare  $e^{\mathbf{A}t}$  per matrici molto grandi e non diagonali. Essendo complicato calcolare  $e^{\mathbf{A}t}$  direttamente, si procede nel trasformare  $\mathbf{A}$  in una matrice  $\mathbf{J}$  denominata **Forma di Jordan della Matrice  $\mathbf{A}$**  sulla quale calcolare  $e^{\mathbf{J}t}$  risulterà molto più agevole e semplice. Come visto in 5.3, effettuare una trasformazione per similitudine, non altera le caratteristiche del sistema ma semplicemente la forma in cui esso viene espresso.

**Definizione 5.2.1** Una matrice  $\mathbf{A}$  si dice **simile** a una matrice  $\mathbf{A}_P$  se esiste una matrice  $\mathbf{P}$  invertibile tale che  $\mathbf{J} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ .

Qualora non sia possibile trasformare  $\mathbf{A}$  in una matrice diagonale, si ricorre alla trasformazione di  $\mathbf{A}$  nella sua forma di Jordan  $\mathbf{J}$ .

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & \dots & \dots & \vdots \\ \vdots & \mathbf{J}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{J}_n \end{bmatrix} \quad (5.8)$$

Le sottomatrici interne alla matrice  $\mathbf{J}$  sono detti **Blocchi di Jordan** e sono pari al numero di autovalori distinti  $n$  della matrice  $\mathbf{A}$ .

Ciascun blocco di Jordan è, a sua volta, una matrice diagonale a blocchi:

$$\mathbf{J}_i = \begin{bmatrix} \mathbf{J}_{i_1} & \dots & \dots & \vdots \\ \vdots & \mathbf{J}_{i_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{J}_{i_{q_i}} \end{bmatrix} \quad (5.9)$$

I termini  $\mathbf{J}_{i_k}$  con  $k = 1, \dots, q_i$ , sono detti **miniblocchi di Jordan**.

Ciascun miniblocco di Jordan è una matrice bidiagonale in cui gli elementi sulla diagonale principale sono uguali all'autovalore  $\lambda_i$  mentre tutti gli elementi sopradiagonale sono uguali a 1 (tutti gli altri elementi sono nulli).

$$\mathbf{J}_{i_j} = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & \ddots & 0 \\ \vdots & 0 & \ddots & 1 \\ \vdots & \vdots & \ddots & \lambda_i \end{bmatrix} \quad (5.10)$$

Un miniblocco ha dimensione uguale a 1 se e solo se  $m_i = q_i$ .

Una matrice quadrata  $\mathbf{A}$  è sempre riconducibile mediante una trasformazione di similitudine a una matrice  $\mathbf{J}$  in forma di Jordan. Tale forma di Jordan è unica, a meno di permutazioni nell'ordine dei blocchi sulla diagonale.

La dimensione di un miniblocco è pari a  $\mu_i$ .

$$m_i = q_i \Rightarrow \mu_i = 1 \quad , \quad m_i \neq q_i \Rightarrow \mu_i \neq 1 \quad (5.11)$$

Essenzialmente,  $\mu_i$  lo si determina cercando il grado massimo di  $\mu_i$  che verifica:

$$n - \text{rank}[(\mathbf{A} - \lambda_i \mathbf{I})^{\mu_i}] = m_i \quad \text{con } 1 \leq \mu_i \leq m_i \quad (5.12)$$

**Esempio 5.2.2** Sia

$$\mathbf{A} = \begin{bmatrix} -6 & 1 & 3 \\ 0 & -4 & 0 \\ 0 & -1 & -4 \end{bmatrix}$$

Gli autovalori sono  $\lambda_1 = -6$  e  $\lambda_2 = -4$ . Abbiamo  $m_1 = 1, q_1 = 1$  e  $m_2 = 2, q_2 = 1$ . Avendo 2 autovalori distinti, avremo 2 blocchi di Jordan  $\mathbf{J}_1, \mathbf{J}_2$ .

Il blocco  $\mathbf{J}_1$  avrà dimensione unitaria, quindi conterrà solo l'autovalore  $\lambda_1 = -6$ , mentre il blocco  $\mathbf{J}_2$  (dato che  $q_2 = 1$ ) conterrà solo un miniblocco  $\mathbf{J}_{21}$

$$\mathbf{J} = \left[ \begin{array}{c|c} \mathbf{J}_1 & \\ \hline & \mathbf{J}_2 \end{array} \right] \quad \mathbf{J}_1 = [-6] \quad \mathbf{J}_2 = [\mathbf{J}_{21}]$$

Di regola, per sapere com'è fatto il miniblocco  $\mathbf{J}_{21}$  dovremmo valutare l'indice  $\mu_2$ . In questo caso, sapendo che  $1 \leq \mu_2 \leq m_2 = 2$  e che *un miniblocco ha dimensione uguale a 1 se e solo se  $m_2 = q_2$* , possiamo dedurre, senza fare ulteriori calcoli, che la dimensione di  $\mathbf{J}_{21}$  è 2, quindi:

$$\mathbf{J}_{21} = \begin{bmatrix} -4 & 1 \\ 0 & -4 \end{bmatrix}$$

Per cui, la nostra matrice in forma di Jordan, sarà:

$$\mathbf{J} = \begin{bmatrix} -6 & 0 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -4 \end{bmatrix} \quad (5.13)$$

**IMPORTANTE:** La forma di Jordan è unica a meno di permutazioni nell'ordine dei blocchi sulla diagonale, quindi scrivere come risultato la 5.13 o 5.14 è equivalente per i nostri scopi.

$$\mathbf{J} = \begin{bmatrix} -4 & 1 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -6 \end{bmatrix} \quad (5.14)$$

**Esempio 5.2.3** Sia

$$\mathbf{A} = \begin{bmatrix} -6 & 0 & 0 & 0 & 0 \\ -2 & -7 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \\ 1 & -1 & 0 & -7 & 0 \\ -3 & 2 & 0 & -1 & -1 \end{bmatrix}$$

Gli autovalori sono  $\lambda_1 = -7, \lambda_2 = -6$  e  $\lambda_3 = -1$ . Abbiamo  $m_1 = 2, q_1 = 1, m_2 = 1, q_2 = 1$  e  $m_3 = 2, q_3 = 2$ . Avendo 3 autovalori distinti, avremo 3 blocchi di Jordan  $\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3$ .

$$\mathbf{J} = \left[ \begin{array}{c|c|c} \mathbf{J}_1 & & \\ \hline & \mathbf{J}_2 & \\ \hline & & \mathbf{J}_3 \end{array} \right]$$

Il blocco  $\mathbf{J}_1$  conterrà solo un miniblocco  $\mathbf{J}_{11}$  in quanto  $q_1 = 1$ , che per esclusione come nell'esempio precedente avrà  $\mu_1 = 2$ .

Il blocco  $\mathbf{J}_2$  sarà di dimensione unitaria in quanto  $m_2 = q_2$  quindi conterrà solo l'autovalore  $\lambda_2 = -6$ .

Il blocco  $\mathbf{J}_3$  avrà due miniblocchi  $\mathbf{J}_{31}, \mathbf{J}_{32}$  in quanto  $q_3 = 2$  entrambi con dimensione unitaria (indice  $\mu = 1$ ) in quanto  $m_3 = q_3$ .

$$\mathbf{J}_1 = [\mathbf{J}_{11}] = \begin{bmatrix} -7 & 1 \\ 0 & -7 \end{bmatrix} \quad \mathbf{J}_2 = [-6] \quad \mathbf{J}_3 = \left[ \begin{array}{c|c} \mathbf{J}_{31} & \\ \hline & \mathbf{J}_{32} \end{array} \right] \quad \mathbf{J}_{31} = [-1] \quad \mathbf{J}_{32} = [-1]$$

Quindi:

$$\mathbf{J} = \begin{bmatrix} -7 & 1 & 0 & 0 & 0 \\ 0 & -7 & 0 & 0 & 0 \\ 0 & 0 & -6 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

**Esempio 5.2.4** Sia

$$\mathbf{A} = \begin{bmatrix} 1 & 5 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 & 0 \\ 4 & -3 & 0 & 0 & 0 & -3 \\ 2 & -1 & 5 & 0 & 5 & -2 \\ -4 & 2 & 0 & 0 & 0 & 2 \\ 5 & -5 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Gli autovalori sono  $\lambda_1 = -2, \lambda_2 = 0$  e  $\lambda_3 = 1$ . Abbiamo  $m_1 = 1, q_1 = 1, m_2 = 3, q_2 = 2$  e  $m_3 = 2, q_3 = 1$ . Avendo 3 autovalori distinti, avremo 3 blocchi di Jordan  $\mathbf{J}_1, \mathbf{J}_2, \mathbf{J}_3$ .

$$\mathbf{J} = \left[ \begin{array}{c|c|c} \mathbf{J}_1 & & \\ \hline & \mathbf{J}_2 & \\ \hline & & \mathbf{J}_3 \end{array} \right]$$

Il blocco  $\mathbf{J}_1$  sarà di dimensione unitaria in quanto  $m_1 = q_1 = 1$  quindi conterrà solo l'autovalore  $\lambda_1 = -2$ .

Il blocco  $\mathbf{J}_2$  avrà due miniblocchi  $\mathbf{J}_{21}, \mathbf{J}_{22}$  in quanto  $q_2 = 2$ .

Il blocco  $\mathbf{J}_3$  avrà un solo miniblocco  $\mathbf{J}_{31}$  in quanto  $q_3 = 1$ . Essendo  $m_3 \neq q_3$  si deduce che  $\mu_3 = 2$ .

$$\mathbf{J}_1 = [-2] \quad \mathbf{J}_2 = \left[ \begin{array}{c|c} \mathbf{J}_{21} & \\ \hline & \mathbf{J}_{22} \end{array} \right] \quad \mathbf{J}_3 = [\mathbf{J}_{31}] = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Ci resta da definire i miniblocchi  $\mathbf{J}_{21}, \mathbf{J}_{22}$ . Senza fare calcoli, possiamo considerare come potrebbero essere formati questi miniblocchi. La somma dei loro ordine deve essere 3. Essendo 2 i miniblocchi da definire, l'unica possibilità è che un blocco sia di ordine 2 e l'altro di ordine 1. Dato che l'ordine dei miniblocchi non ci altera il risultato, non ci interessa chi sia di ordine 2 e chi di ordine 1. Quindi:

$$\mathbf{J}_{21} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \mathbf{J}_{22} = [0]$$

Quindi:

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

**Esempio 5.2.5** Sia

$$\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 & 0 & 5 \\ 0 & -2 & 0 & 0 & 0 \\ 5 & 2 & -2 & -5 & -2 \\ -3 & 2 & 0 & -2 & -3 \\ 0 & -3 & 0 & 0 & -2 \end{bmatrix}$$

Gli autovalori sono  $\lambda_1 = -2, \lambda_2 = 3$ . Abbiamo  $m_1 = 4, q_1 = 2, m_2 = 1, q_2 = 1$ . Avendo 2 autovalori distinti, avremo 2 blocchi di Jordan  $\mathbf{J}_1, \mathbf{J}_2$ .

Il blocco  $\mathbf{J}_1$  conterrà 2 miniblocchi  $\mathbf{J}_{11}, \mathbf{J}_{12}$  in quanto  $q_1 = 2$ .

Il blocco  $\mathbf{J}_2$  sarà unitario e conterrà solo l'autovalore  $\lambda_2 = 3$ .



$$\mathbf{J} = \left[ \begin{array}{c|c} \mathbf{J}_1 & \\ \hline & \mathbf{J}_2 \end{array} \right] \quad \mathbf{J}_1 = \left[ \begin{array}{c|c} \mathbf{J}_{11} & \\ \hline & \mathbf{J}_{12} \end{array} \right] \quad \mathbf{J}_2 = [3]$$

In questo caso, non possiamo ragionare come nell'esempio 5.2.3 in quanto i 2 miniblocchi potrebbero avere dimensione  $2 - 2$  o  $3 - 1$ . Dobbiamo calcolare il valore di  $\mu_1$ .

Usiamo quanto detto in 5.12. Devo trovare un valore di  $\mu_1$  tale che

$$n - \text{rank}[(\mathbf{A} - \lambda_i \mathbf{I})^{\mu_i}] = m_i \quad \Rightarrow \quad 5 - \text{rank}[(\mathbf{A} + 2\mathbf{I})^{\mu_i}] = 4$$

Definisco una matrice  $\mathbf{M}$  (per semplificare la notazione) come:

$$\mathbf{M} = \mathbf{A} - \lambda_i \mathbf{I} = \begin{bmatrix} 5 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 5 & 2 & 0 & -5 & -2 \\ -3 & 2 & 0 & 0 & -3 \\ 0 & -3 & 0 & 0 & 0 \end{bmatrix}$$

Proviamo con  $\mu_1 = 1$ :

$$5 - \text{rank}(\mathbf{M}^{\mu_1}) = 5 - \text{rank}(\mathbf{M}^1) = 5 - 3 = 2$$

Proviamo con  $\mu_1 = 2$ :

$$5 - \text{rank}(\mathbf{M}^2) = 5 - 2 = 3$$

Proviamo con  $\mu_1 = 3$ :

$$5 - \text{rank}(\mathbf{M}^3) = 5 - 1 = 4$$

Abbiamo trovato che  $\mu_1 = 3$ ! Quindi avremo un miniblocco di dimensione 3 e uno di dimensione 1.

$$\mathbf{J}_{11} = \begin{bmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{bmatrix} \quad \mathbf{J}_{12} = [-2]$$

Quindi la forma di Jordan risulta essere:

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

**Definizione 5.2.6** Ogni autovalore  $\lambda_i$  possiede una **Molteplicità Algebrica**  $m_i$ , per cui il polinomio caratteristico può essere riscritto come

$$\rho(\lambda) = \prod_{i=0}^p (\lambda - \lambda_i)^{m_i} \quad (5.15)$$

**Definizione 5.2.7** Ogni autovalore  $\lambda_i$  possiede una **Molteplicità Geometrica**  $q_i$ , definita come il numero di autovettori linearmente indipendenti associati all'autovalore  $\lambda_i$ , ossia il numero di soluzioni della seguente equazione:

$$\mathbf{A}\mathbf{v}_i = \lambda\mathbf{v}_i \Rightarrow (\mathbf{A} - \lambda_i \mathbf{I})\mathbf{v}_i = 0 \quad (5.16)$$

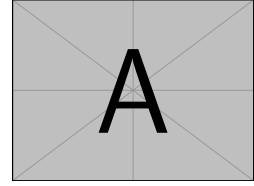
In altri termini, la molteplicità geometrica è pari a:

$$\begin{aligned} q_i &= \dim[\mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})] \\ &= n - \text{rank}(\mathbf{A} - \lambda_i \mathbf{I}) \quad \text{con } 1 \leq q_i \leq m_i \end{aligned} \quad (5.17)$$

### 5.3 Trasformazione per similitudine

Consideriamo il seguente sistema LTI TC:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \quad (5.18)$$



Trasformare per similitudine significa cambiare il sistema di riferimento corrispondente alle variabili di stato per descrivere lo stesso sistema. Adottando la seguente trasformazione:

$$\mathbf{x}(t) = \mathbf{P}\mathbf{z}(t) \Rightarrow \dot{\mathbf{x}}(t) = \mathbf{P}\dot{\mathbf{z}}(t) \quad (5.19)$$

Le equazioni di stato del sistema trasformato, risultano come segue:

$$\begin{cases} \mathbf{P}\dot{\mathbf{z}}(t) = \mathbf{A}\mathbf{P}\mathbf{z}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{P}\mathbf{z}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \Rightarrow \begin{cases} \dot{\mathbf{z}}(t) = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{z}(t) + \mathbf{P}^{-1}\mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{P}\mathbf{z}(t) + \mathbf{D}\mathbf{u}(t) \end{cases} \quad (5.20)$$

La trasformazione di similitudine adottata, la quale è ammissibile soltanto se  $\mathbf{P}$  è una matrice invertibile, comporta quindi una nuova quadrupla di matrici che rappresentano del sistema, passando da  $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$  a  $(\mathbf{A}_P, \mathbf{B}_P, \mathbf{C}_P, \mathbf{D}_P)$  con:

$$\begin{aligned} \mathbf{A}_P &= \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \\ \mathbf{B}_P &= \mathbf{P}^{-1}\mathbf{B} \\ \mathbf{C}_P &= \mathbf{C}\mathbf{P} \\ \mathbf{D}_P &= \mathbf{D} \end{aligned} \quad (5.21)$$

Ossia:

$$\begin{cases} \dot{\mathbf{z}}(t) = \mathbf{A}_P\mathbf{z}(t) + \mathbf{B}_P\mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}_P\mathbf{z}(t) + \mathbf{D}_P\mathbf{u}(t) \end{cases} \quad (5.22)$$

Le proprietà del sistema sono invarianti rispetto alla trasformazione di similitudine. Infatti,  $\mathbf{A}_P$  ha lo stesso polinomio caratteristico di  $\mathbf{A}$ .

#### Dimostrazione

$$\begin{aligned} \rho_{\mathbf{A}_P}(\lambda) &= |\mathbf{A}_P - \lambda\mathbf{I}| \\ &= |\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{I}| \\ &= |\mathbf{P}^{-1}\mathbf{A}\mathbf{P} - \lambda\mathbf{P}^{-1}\mathbf{I}\mathbf{P}| \\ &= |\mathbf{P}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{P}| \quad (\text{dato che } |\mathbf{XYZ}| = |\mathbf{X}||\mathbf{Y}||\mathbf{Z}|) \\ &= |\mathbf{P}^{-1}| |\mathbf{A} - \lambda\mathbf{I}| |\mathbf{P}| \quad (\text{dato che } |\mathbf{P}^{-1}||\mathbf{P}| = 1) \\ &= |\mathbf{A} - \lambda\mathbf{I}| \\ &= \rho_{\mathbf{A}}(\lambda) \end{aligned} \quad (5.23)$$

La matrice  $\mathbf{A}_P$  ha lo stesso polinomio caratteristico della matrice  $\mathbf{A}$ . Di conseguenza, la matrice  $\mathbf{A}_P$  ha gli stessi autovalori, con le medesime molteplicità algebrica e geometrica, della matrice  $\mathbf{A}$ , dunque il sistema trasformato conserva le stesse caratteristiche del sistema di partenza. Solo gli autovettori della matrice trasformata sono differenti.

A partire dal modello del sistema nello spazio di stato, è possibile ottenere il modello ingresso-uscita, ossia il legame tra ingresso e uscita, che è fornito dalla funzione di trasferimento del sistema, ricavabile trasformando secondo Laplace le equazioni di stato e di uscita nell'ipotesi di condizioni iniziali nulle:

$$\begin{cases} s\mathbf{X}(s) = \mathbf{A}\mathbf{X}(s) + \mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) = \mathbf{C}\mathbf{X}(s) + \mathbf{D}\mathbf{U}(s) \end{cases} \quad (5.24)$$

Da cui ricavo:

$$\begin{aligned} (s\mathbf{I} - \mathbf{A})\mathbf{X}(s) &= \mathbf{B}\mathbf{U}(s) \\ \Rightarrow \mathbf{X}(s) &= (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s) \end{aligned} \quad (5.25)$$

Sostituendo in 5.24 ottengo:

$$\begin{cases} \mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{U}(s) \\ \mathbf{Y}(s) = [\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}]\mathbf{U}(s) \end{cases} \quad (5.26)$$

Da cui la matrice di funzioni di trasferimento:

$$\mathbf{W}(s) = \frac{\mathbf{Y}(s)}{\mathbf{U}(s)} = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \quad (5.27)$$

Sfruttando la trasformazione per similitudine, la nuova matrice di funzioni di trasferimento risulta come segue:

$$\begin{aligned} \mathbf{W}_P(s) &= \mathbf{C}_P(s\mathbf{I} - \mathbf{A}_P)^{-1}\mathbf{B}_P + \mathbf{D}_P \\ &= \mathbf{C}\mathbf{P}(s\mathbf{P}^{-1}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P})^{-1}\mathbf{P}^{-1}\mathbf{B} + \mathbf{D} \quad (\text{Sostituendo come in 5.21 e } \mathbf{I} = \mathbf{P}^{-1}\mathbf{P}) \\ &= \mathbf{C}\mathbf{P}\mathbf{P}^{-1}(s\mathbf{I} - \mathbf{A})\mathbf{P}\mathbf{P}^{-1}\mathbf{B} + \mathbf{D} \quad (\text{raccogliendo a sx per } \mathbf{P}^{-1} \text{ e a dx per } \mathbf{P}) \\ &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} \\ &= \mathbf{W}(s) \end{aligned}$$

Si è così dimostrato che anche la matrice di funzioni di trasferimento del sistema trasformato coincide con quella del sistema originario e ciò è dovuto al fatto che la relazione ingresso-uscita non è influenzata dallo stato.

La trasformazione della matrice di stato nella sua forma di Jordan è a sua volta una trasformazione per similitudine.  $\square$

## Tabella di riepilogo

Per determinare la stabilità del sistema LTI, ci basta studiare gli  $n$  autovalori  $\lambda_i$  della matrice  $\mathbf{A}$ .

	Asintotica stabilità	Semplice stabilità	Instabilità
<b>TC</b>	$\operatorname{Re}(\lambda_i) < 0 \quad \forall \lambda_i$	$\operatorname{Re}(\lambda_i) \leq 0 \quad \forall \lambda_i$ e $m_i = q_i \quad \forall \lambda_i : \operatorname{Re}(\lambda_i) = 0$	$\exists \lambda_i : \operatorname{Re}(\lambda_i) > 0$ o $\exists \lambda_i : \operatorname{Re}(\lambda_i) = 0 : m_i \neq q_i$
<b>TD</b>	$ \lambda_i  < 1 \quad \forall \lambda_i$	$ \lambda_i  \leq 1 \quad \forall \lambda_i$ e $m_i = q_i \quad \forall \lambda_i :  \lambda_i  = 1$	$\exists \lambda_i :  \lambda_i  > 1$ o $\exists \lambda_i :  \lambda_i  = 1 : m_i \neq q_i$



## Controllabilità e Osservabilità dei sistemi LTI

### 6.1 Controllabilità

La *controllabilità* di un sistema LTI riguarda la possibilità di modificarne la dinamica, ossia di trasferire il sistema da uno stato iniziale  $x_A$  a uno stato finale  $x_B$ , verificando se esista un ingresso in grado di realizzare tale transizione.

Questo problema si articola in due aspetti principali:

1. **Raggiungibilità:** ci si chiede se sia possibile, partendo dallo stato nullo, raggiungere un generico stato  $x$ ;
2. **Controllabilità:** si indaga se sia possibile, da un dato stato iniziale  $x$ , ricondurre il sistema allo stato nullo.

I due concetti sono tra loro complementari: la *raggiungibilità* è legata all'insieme degli stati che possono essere ottenuti a partire da una condizione iniziale nulla, scegliendo un opportuno segnale di ingresso. La *controllabilità*, invece, riguarda l'insieme degli stati dai quali è possibile condurre il sistema verso uno stato finale assegnato (ad esempio lo stato nullo), mediante una funzione di ingresso adeguata.

Per quanto riguarda i sistemi LTI questi due aspetti coincidono:

$$\text{Raggiungibilità} \Leftrightarrow \text{Controllabilità}$$

**Definizione 6.1.1** Un sistema dinamico LTI  $S(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$  di ordine  $n$  è **controllabile** se e solo se, definita la matrice  $\mathcal{C}$  detta **matrice di controllabilità**:

$$\mathcal{C} = [\mathbf{B} \quad \mathbf{AB} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}] \quad (6.1)$$

essa ha rango pari all'ordine del sistema, cioè  $n$ .

#### Controllabilità in MATLAB

Utilizzando Matlab è semplice verificare la controllabilità di un sistema. Una volta definite le matrici  $\mathbf{A}, \mathbf{B}$  si utilizza:

```
C = ctrb(A,B) % Per calcolare la matrice di controllabilità
rank(C)       % Per verificare il rango della matrice di controllabilità
```

**Esempio 6.1.2** Consideriamo il seguente sistema:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \Rightarrow \mathcal{C} = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 1 \quad (6.2)$$

E' semplice vedere come il sistema non sia controllabile dato che la matrice di controllabilità  $\mathcal{C}$  non ha rango pieno. Anche senza controllare il rango di  $\mathcal{C}$ , è semplice dire che il sistema non è controllabile dato che gli stati  $x_1$  e  $x_2$  sono completamente disaccoppiati e l'ingresso  $u$  influisce solamente sullo stato  $x_2$ .

Modificando l'esempio per includere due ingressi di attuazione, il sistema diventa controllabile grazie a una maggiore capacità di controllo.

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \Rightarrow \mathcal{C} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 2 \quad (6.3)$$

Essendo completamente attuato, questo sistema risulta chiaramente controllabile dato che  $x_1$  e  $x_2$  possono essere controllate indipendentemente con  $u_1$  e  $u_2$ . La controllabilità del sistema inoltre è confermata dal fatto che  $\mathcal{C}$  ha rango pieno.

Consideriamo il sistema

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \Rightarrow \mathcal{C} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 2 \quad (6.4)$$

Poiché gli stati  $x_1$  e  $x_2$  risultano ora dinamicamente accoppiati, il sistema diventa controllabile anche con un unico ingresso di attuazione.

In maniera simile:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \Rightarrow \mathcal{C} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 2 \quad (6.5)$$

è controllabile anche se gli stati  $x_1$  e  $x_2$  sono disaccoppiati. Questo dato che l'attuazione di  $u$  è in grado di influenzare simultaneamente entrambi gli stati.

### 6.1.1 Forma canonica di Controllabilità

La *forma canonica di controllabilità* rappresenta una particolare realizzazione di un sistema dinamico lineare che mette in evidenza, in modo esplicito, le proprietà di controllabilità del sistema stesso. Tale forma è ottenuta mediante un opportuno cambio di coordinate nello spazio degli stati, tale da rendere immediata la verifica della controllabilità.

Per individuare esattamente le componenti controllabili e non dello stato, occorre ricorrere ad una trasformazione per similitudine, che consente di dividere lo stato tra  $n_c$  componenti controllabili  $x_c(t)$  e  $n - n_c$  componenti non controllabili  $x_{nc}(t)$ .

E' semplice dimostrare come la trasformazione per similitudine del sistema, risulti in una semplice trasformazione lineare della matrice di controllabilità.

$$\mathcal{C}_P = P^{-1}\mathcal{C} \quad (6.6)$$

Dalla Eq.6.6 si evince che, poichè il rango di una matrice non varia se essa viene moltiplicata per una matrice a rango pieno, il rango della matrice di controllabilità del sistema trasformato coincide con quello della matrice del sistema iniziale. Ne consegue che la controllabilità non viene modificata dalla trasformazione per similitudine. Se il sistema di partenza è controllabile, lo è anche quello trasformato ed inoltre, lo stesso discorso vale in caso di "controllabilità parziale" del sistema (il numero di componenti controllabili del sistema è invariante rispetto alla trasformazione per similitudine).

Adottiamo una trasformazione di similitudine che ponga il sistema in una forma denominata *forma canonica di controllabilità*, tale da rendere evidenti le proprietà di controllabilità effettuando la già menzionata distinzione fra stati controllabili e non controllabili.

$$\begin{bmatrix} \dot{\mathbf{x}}_c(t) \\ \dot{\mathbf{x}}_{nc}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_{nc} \end{bmatrix} \begin{bmatrix} \mathbf{x}_c(t) \\ \mathbf{x}_{nc}(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_c \\ \mathbf{0} \end{bmatrix} \mathbf{u}(t) \quad (6.7)$$

In questa nuova forma, la matrice di controllabilità assume la seguente espressione:

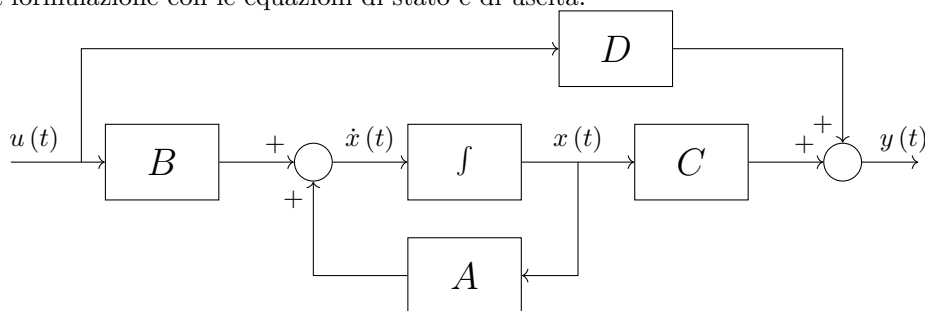
$$\mathcal{C}_P = \begin{bmatrix} \mathbf{B}_c & \mathbf{A}_c\mathbf{B}_c & \mathbf{A}_c^2\mathbf{B}_c & \dots & \mathbf{A}_c^{n-1}\mathbf{B}_c \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathcal{C}_{P1} \\ \mathbf{0} \end{bmatrix} \quad (6.8)$$

e presenta un rango uguale al numero di stati controllabili, nonchè alla dimensione della matrice  $\mathbf{A}_c$ , che minore del numero totale di stati, ovvero:

$$\rho(\mathcal{C}_P) = \dim(\mathbf{A}_c) = n_c < n \quad (6.9)$$

dunque, la sottomatrice  $\mathcal{C}_{P1}$  ha rango pari a  $n_c$ .

Il sistema di partenza può essere rappresentato con il diagramma a blocchi in Fig. TODO, in accordo alla classica formulazione con le equazioni di stato e di uscita.

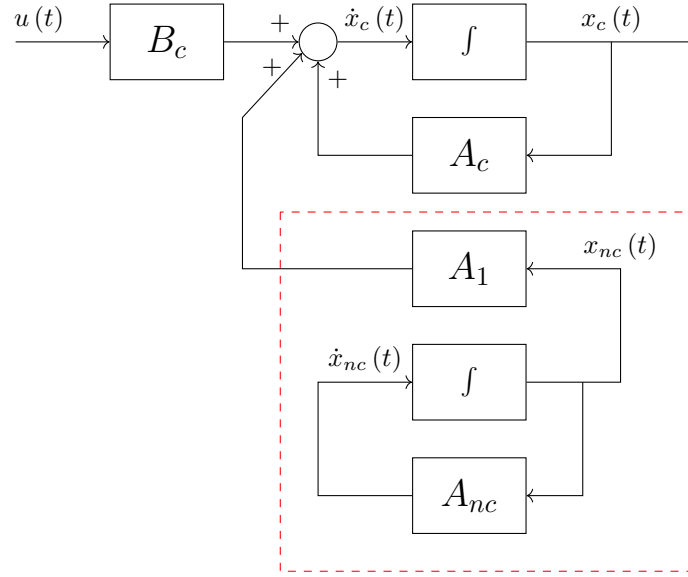


**Figura 6.1.** Caption

Il sistema nella forma di controllabilità è rappresentabile con il diagramma a blocchi in Fig.TODO, in accordo alla seguente equazione di stato:

$$\begin{aligned} \dot{\mathbf{x}}_c(t) &= \mathbf{A}_c\mathbf{x}_c(t) + \mathbf{A}_1\mathbf{x}_{nc}(t) \\ \dot{\mathbf{x}}_{nc}(t) &= \mathbf{A}_{nc}\mathbf{x}_{nc}(t) \end{aligned} \quad (6.10)$$

La parte evidenziata in rosso in Fig. TODO è la parte del sistema non controllabile, completamente distaccata dall'ingresso; Pertanto la sua evoluzione temporale è dovuta alle sole condizioni iniziali, dunque  $\mathbf{x}_{nc}(t)$  può assumere valori non nulli nel tempo se  $\mathbf{x}_{nc}(0) \neq \mathbf{0}$ .



**Figura 6.2.** Schema a blocchi del sistema espresso nella forma di Controllabilità. In rosso è evidenziata la parte del sistema non controllabile.

**Esempio 6.1.3** Vogliamo porre il sistema descritto dalla seguente equazione di stato nella forma canonica di controllabilità:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \mathbf{u}(t)$$

La matrice di controllabilità è data da:

$$\mathcal{C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 2 = n_c < n = 3$$

Per costruire la matrice  $\mathbf{P}$ , che definisce la trasformazione di similitudine operiamo come segue: Le prime  $n_c$  colonne sono costituite da  $n_c$  colonne linearmente indipendenti della matrice di controllabilità, mentre le restanti  $n - n_c$  colonne devono essere scelte in modo da risultare linearmente indipendenti dalle altre e ortogonali. Nel nostro caso si ottiene:

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{P}^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{P}^{-1}\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Il sistema trasformato assume la seguente forma canonica di controllabilità:

$$\begin{bmatrix} \dot{\mathbf{z}}_c(t) \\ \dot{\mathbf{z}}_{nc}(t) \end{bmatrix} = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} \mathbf{z}_c(t) \\ \mathbf{z}_{nc}(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(t)$$

La parte controllabile del sistema è la seguente:

$$\dot{\mathbf{z}}_c(t) = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \mathbf{z}_c(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$



## 6.2 Osservabilità

L'*osservabilità* di un sistema LTI riguarda la possibilità di determinare lo stato interno del sistema a partire dalle misure delle sue uscite e dagli ingressi applicati, in un intervallo di tempo finito.

In altre parole, ci si chiede se, conoscendo l'andamento temporale dell'uscita  $y(t)$  e dell'ingresso  $u(t)$ , sia possibile risalire in modo univoco allo stato iniziale  $x(0)$  del sistema.

Infine, così come la controllabilità si riferisce alla possibilità di *agire* sullo stato del sistema mediante l'ingresso, l'osservabilità si riferisce alla possibilità di *dedurre* lo stato del sistema a partire dalle sue uscite.

**Definizione 6.2.1** Un sistema dinamico LTI  $S(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$  di ordine  $n$  è **osservabile** se e solo se, definita la matrice  $\mathcal{O}$  detta **matrice di osservabilità**:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \quad (6.11)$$

essa ha rango pari all'ordine del sistema, cioè  $n$ .

### Osservabilità in MATLAB

Utilizzando Matlab è semplice verificare l'osservabilità di un sistema. Una volta definite le matrici  $\mathbf{A}, \mathbf{C}$  si utilizza:

```
O = obsv(A,C) % Per calcolare la matrice di osservabilità
rank(O)        % Per verificare il rango della matrice di osservabilità
```

### 6.2.1 Forma canonica di Osservabilità

La *forma canonica di osservabilità* rappresenta una particolare realizzazione di un sistema dinamico lineare che mette in evidenza, in modo esplicito, le proprietà di osservabilità del sistema stesso. Tale forma è ottenuta mediante un opportuno cambio di coordinate nello spazio degli stati, tale da rendere immediata la verifica dell'osservabilità.

Per individuare esattamente le componenti osservabili e non dello stato, occorre ricorrere ad una trasformazione per similitudine, che consente di dividere lo stato tra  $n_o$  componenti osservabili  $x_o(t)$  e  $n - n_o$  componenti non osservabili  $x_{no}(t)$ .

E' semplice dimostrare come la trasformazione per similitudine del sistema, risulti in una semplice trasformazione lineare della matrice di osservabilità.

$$\mathcal{O}_P = \mathcal{O} \mathbf{P} \quad (6.12)$$

Dalla Eq. 6.12 si evince che, poichè il rango di una matrice non varia se essa viene moltiplicata per una matrice a rango pieno, il rango della matrice di osservabilità del sistema trasformato coincide con quello della matrice del sistema iniziale. Ne consegue che l'osservabilità non viene modificata dalla trasformazione per similitudine. Se il sistema di partenza è osservabile, lo è anche quello trasformato ed inoltre, lo stesso discorso vale in caso di "osservabilità parziale" del sistema (il numero di componenti osservabili del sistema è invariante rispetto alla trasformazione per similitudine).

Adottiamo una trasformazione di similitudine che ponga il sistema in una forma denominata *forma canonica di osservabilità*, tale da rendere evidenti le proprietà di osservabilità effettuando la già menzionata distinzione fra stati osservabili e non osservabili.

$$\begin{bmatrix} \dot{\mathbf{x}}_o(t) \\ \dot{\mathbf{x}}_{no}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}_o & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{A}_{no} \end{bmatrix}}_{\mathbf{A}_P} \begin{bmatrix} \mathbf{x}_o(t) \\ \mathbf{x}_{no}(t) \end{bmatrix} + \mathbf{B}_P \mathbf{u}(t) \quad (6.13)$$

$$y(t) = \underbrace{\begin{bmatrix} \mathbf{C}_o & \mathbf{0} \end{bmatrix}}_{\mathbf{C}_P} \begin{bmatrix} \mathbf{x}_o(t) \\ \mathbf{x}_{no}(t) \end{bmatrix} \quad (6.14)$$

In questa nuova forma, la matrice di osservabilità assume la seguente espressione:

$$\mathcal{O}_P = \left[ \begin{array}{c|c} \mathbf{C}_o & \mathbf{0} \\ \mathbf{C}_o \mathbf{A}_o & \mathbf{0} \\ \mathbf{C}_o \mathbf{A}_o^2 & \mathbf{0} \\ \vdots & \vdots \\ \mathbf{C}_o \mathbf{A}_o^{n-1} & \mathbf{0} \end{array} \right] = [\mathbf{C}_{o1} \mid \mathbf{0}] \quad (6.15)$$

e presenta un rango uguale al numero di stati osservabili, nonchè alla dimensione della matrice  $\mathbf{A}_o$ , che minore del numero totale di stati, ovvero:

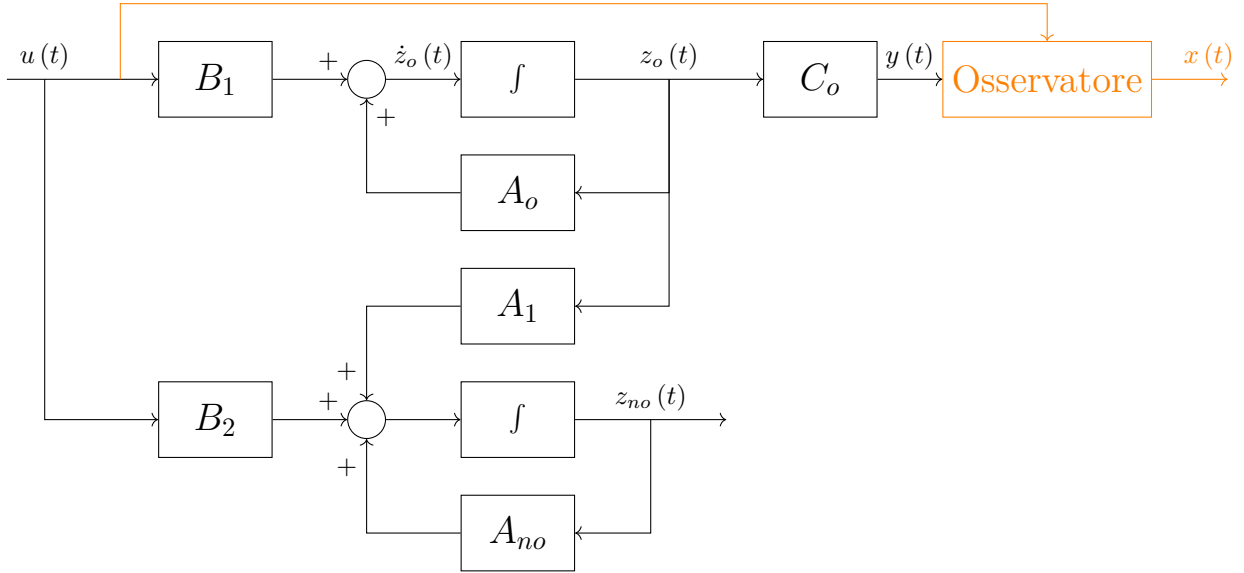
$$\rho(\mathcal{O}_P) = \dim(\mathbf{A}_o) = n_o < n \quad (6.16)$$

dunque, la sottomatrice  $\mathcal{O}_{P1}$  ha rango pari a  $n_c$ .

Il sistema di partenza può essere rappresentato con il diagramma a blocchi in Fig. TODO, in accordo alla classica formulazione con le equazioni di stato e di uscita.

Il sistema nella forma di osservabilità è rappresentabile con il diagramma a blocchi in Fig. TODO, in accordo alla seguente equazione di stato:

$$\begin{cases} \dot{\mathbf{z}}_o(t) = \mathbf{A}_o \mathbf{z}_o(t) + \mathbf{B}_1 \mathbf{u}(t) \\ \dot{\mathbf{z}}_{no}(t) = \mathbf{A}_1 \mathbf{z}_o(t) + \mathbf{A}_{no} \mathbf{z}_{no}(t) + \mathbf{B}_2 \mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C}_o \mathbf{z}_o(t) \end{cases} \quad (6.17)$$



**Figura 6.3.** Schema a blocchi del sistema espresso nella forma canonica di Osservabilità. Non sempre lo stato del sistema è accessibile, potendo trattarsi anche di una grandezza non misurabile, ma quasi sempre si può ricavare l'uscita a partire dall'ingresso; pertanto, si vuole inserire a monte del controllo un **Osservatore**, che consente di **stimare** lo stato noti l'ingresso e l'uscita a condizione che il sistema sia osservabile. Qualora la completa osservabilità non è soddisfatta, tale stima non può essere effettuata, così come la retroazione di stato.

Per la (TODO) la funzione di trasferimento è pari a:

$$\begin{aligned} \mathbf{G}(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1} \mathbf{B} \\ &= [\mathbf{C}_o \mid \mathbf{0}] \begin{bmatrix} s\mathbf{I}_o - \mathbf{A}_o & \mathbf{0} \\ \mathbf{A}_1 & s\mathbf{I}_{no} - \mathbf{A}_{no} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \\ &= [\mathbf{C}_o \mid \mathbf{0}] \begin{bmatrix} (s\mathbf{I}_o - \mathbf{A}_o)^{-1} & \mathbf{0} \\ *1 & (s\mathbf{I}_{no} - \mathbf{A}_{no})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \\ &= [\mathbf{C}_o \mid \mathbf{0}] \begin{bmatrix} (s\mathbf{I}_o - \mathbf{A}_o)^{-1} \mathbf{B}_1 \\ *2 \end{bmatrix} \\ &= \mathbf{C}_o (s\mathbf{I}_o - \mathbf{A}_o)^{-1} \mathbf{B}_1 \end{aligned} \quad (6.18)$$

In definitiva, la funzione di trasferimento è dovuta alla mera parte osservabile del sistema, che la trasformazione sia eseguita o meno. Pertanto si parla di autovalori osservabili, dati dalla matrice  $\mathbf{A}_o$ , e di autovalori non osservabili, di pertinenza della matrice  $\mathbf{A}_{no}$ , e solo i primi di essi caratterizzano la connessione fra l'ingresso e l'uscita, mentre i secondi, come gli autovalori non controllabili, si cancellano con gli zeri del sistema.

**Esempio 6.2.2** Vogliamo porre il sistema descritto dalla seguente equazione di stato nella forma canonica di osservabilità:

$$\begin{cases} \dot{\mathbf{x}}(t) = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -5 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) = \begin{bmatrix} 1 & 0 & 3 \end{bmatrix} \mathbf{x}(t) \end{cases}$$

La matrice di osservabilità è data da:

$$\mathcal{O} = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & -15 \\ 4 & 0 & 75 \end{bmatrix} \Rightarrow \rho(\mathcal{O}) = 2 = n_o < n = 3$$

Per costruire la matrice  $\mathbf{P}$ , che definisce la trasformazione di similitudine operiamo come segue: Le prime  $n_o$  colonne sono costituite da  $n_o$  righe linearmente indipendenti della matrice di osservabilità, mentre le restanti  $n - n_o$  colonne devono essere scelte in modo da risultare linearmente indipendenti dalle altre e ortogonali. Nel nostro caso si ottiene:

$$\mathbf{P} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 3 & -15 & 0 \end{bmatrix} \Rightarrow \mathbf{P}^{-1} = \begin{bmatrix} \frac{5}{7} & 0 & \frac{2}{21} \\ \frac{1}{7} & 0 & -\frac{1}{21} \\ 0 & 1 & 0 \end{bmatrix}$$

Le matrici  $\mathbf{A}_P$  e  $\mathbf{C}_P$  in forma canonica di osservabilità sono allora date da:

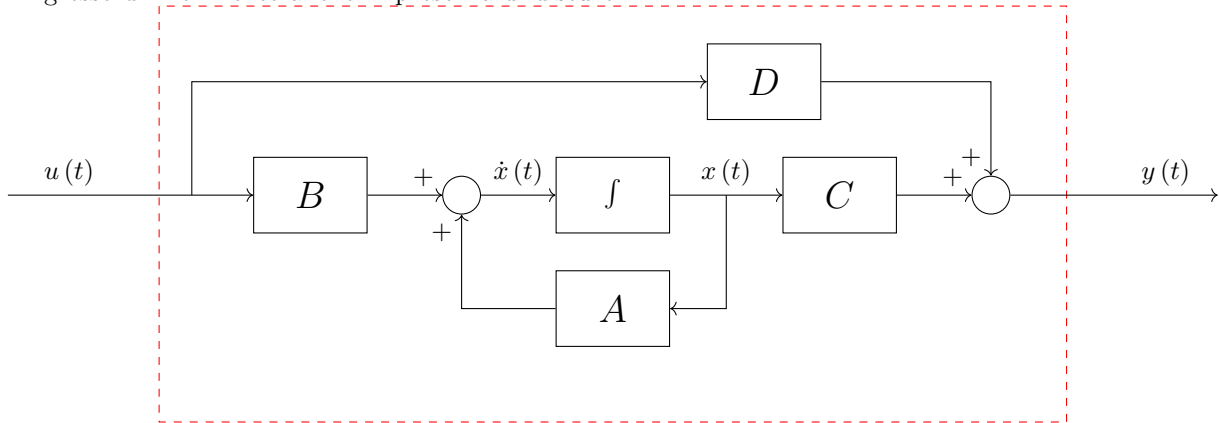
$$\begin{aligned} \mathbf{A}_P = \mathbf{P}^{-1} \mathbf{A} \mathbf{P} &= \begin{bmatrix} \frac{5}{7} & 0 & \frac{2}{21} \\ \frac{1}{7} & 0 & -\frac{1}{21} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 3 & -15 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{5}{7} & 0 & \frac{2}{21} \\ \frac{1}{7} & 0 & -\frac{1}{21} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 \\ 0 & 0 & 4 \\ -15 & 75 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 10 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \\ \mathbf{C}_P = \mathbf{C} \mathbf{P} &= \begin{bmatrix} 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 3 & -15 & 0 \end{bmatrix} = \begin{bmatrix} 10 & -43 & 0 \end{bmatrix} \end{aligned}$$

Il sistema trasformato assume la seguente forma canonica di osservabilità:

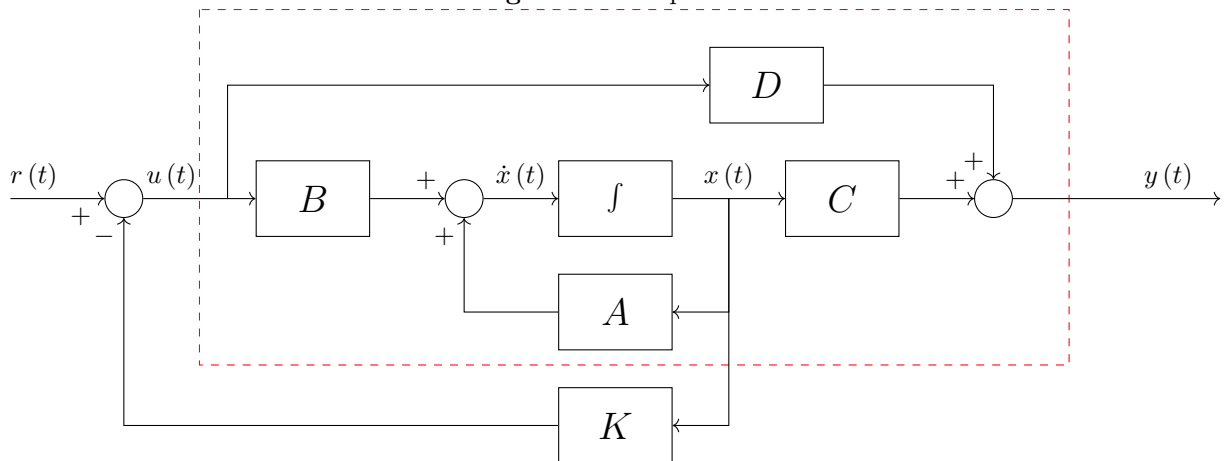
$$\begin{aligned} \begin{bmatrix} \dot{\mathbf{z}}_o(t) \\ \dot{\mathbf{z}}_{no}(\bar{t}) \end{bmatrix} &= \begin{bmatrix} 0 & 10 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} \mathbf{z}_o(t) \\ \mathbf{z}_{no}(\bar{t}) \end{bmatrix} + \mathbf{B}_P \mathbf{u}(t) \\ \mathbf{y}(t) &= \begin{bmatrix} 10 & -43 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{z}_o(t) \\ \mathbf{z}_{no}(\bar{t}) \end{bmatrix} \end{aligned}$$



L'obiettivo di un controllore in retroazione è quello di regolare l'uscite del sistem  $y$  in modo che segua l'ingresso di riferimento anche in presenza di disturbi.



**Figura 7.1.** Caption



**Figura 7.2.** Caption

## 7.1 Stabizzabilità

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} = \mathbf{Cx} + \mathbf{Du} \end{cases} \quad (7.1)$$

$$\mathbf{u} = \mathbf{r} - \mathbf{Kx} \quad (7.2)$$

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \\ \mathbf{y} = \mathbf{Cx} + \mathbf{Du} \end{cases} \Rightarrow \begin{cases} \dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Br} - \mathbf{BKx} \\ \mathbf{y} = \mathbf{Cx} + \mathbf{Dr} - \mathbf{DKx} \end{cases} \Rightarrow \begin{cases} \dot{\mathbf{x}} = (\mathbf{A} - \mathbf{BK})\mathbf{x} + \mathbf{Br} \\ \mathbf{y} = (\mathbf{C} - \mathbf{DK})\mathbf{x} + \mathbf{Dr} \end{cases} \quad (7.3)$$

**Definizione 7.1.1** Un sistema si dice **stabilizzabile** quando è completamente controllabile oppure, se non lo è, quando tutti gli autovalori non controllabili risultano asintoticamente stabili.

## 7.2 Progettazione di un sistema in retroazione di stato

**Teorema 7.2.1** Gli autovalori della matrice di stato in anello chiuso  $\mathbf{A} - \mathbf{BK}$  possono essere assegnati arbitrariamente mediante un'opportuna scelta della matrice  $\mathbf{K}$  se e solo se la matrice di controllabilità  $\mathcal{C}$  ha rango pieno.

**Esempio 7.2.2** Si consideri il sistema dinamico LTI:

$$\mathbf{x}(k+1) = \begin{bmatrix} 0 & -\frac{1}{4} \\ -3 & 1 \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} u(k)$$

al quale vogliamo assegnare gli autovalori  $\frac{1}{2}$  e  $\frac{1}{4}$  in anello chiuso.

Facciamo prima un controllo sulla stabilità del sistema in anello aperto analizzando gli autovalori della matrice  $\mathbf{A}$ .

$$\lambda_1 = -\frac{1}{2} \quad \lambda_2 = \frac{3}{2} \Rightarrow \lambda_2 \text{ è instabile in quanto al di fuori del cerchio unitario}$$

Verifichiamo la controllabilità

$$\mathcal{C} = [\mathbf{B} \quad \mathbf{AB}] = \begin{bmatrix} 1 & \frac{1}{8} \\ -\frac{1}{2} & -\frac{7}{2} \end{bmatrix} \Rightarrow \text{rank}(\mathcal{C}) = 2$$

$\mathcal{C}$  ha rango pieno, quindi il sistema è completamente controllabile, è possibile proseguire con l'assegnazione degli autovalori.

Il polinomio caratteristico desiderato è:

$$\rho_d(\lambda) = \left(\lambda - \frac{1}{2}\right) \left(\lambda - \frac{1}{4}\right) = \lambda^2 - \frac{3}{4}\lambda + \frac{1}{8}$$

Calcoliamo la matrice di stato in anello chiuso  $\mathbf{A} - \mathbf{BK}$ :

$$\begin{aligned} \mathbf{A} - \mathbf{BK} &= \begin{bmatrix} 0 & -\frac{1}{4} \\ -3 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix} \begin{bmatrix} k_1 & k_2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -\frac{1}{4} \\ -3 & 1 \end{bmatrix} - \begin{bmatrix} k_1 & k_2 \\ -\frac{k_1}{2} & -\frac{k_2}{2} \end{bmatrix} \\ &= \begin{bmatrix} -k_1 & -\frac{1}{4} - k_2 \\ -3 + \frac{k_1}{2} & 1 + \frac{k_2}{2} \end{bmatrix} \end{aligned}$$

Il cui polinomio caratteristico è:

$$\lambda^2 + \left(k_1 - \frac{k_2}{2} - 1\right)\lambda + \left(-\frac{7k_1}{8} - \frac{3}{4} - 3k_2\right)$$

Imponiamo i coefficienti del polinomio caratteristico desiderato, risolvendo il sistema:

$$\begin{cases} k_1 - \frac{k_2}{2} - 1 = -\frac{3}{4} \\ -\frac{7k_1}{8} - \frac{3}{4} - 3k_2 = \frac{1}{8} \end{cases} \Rightarrow \begin{cases} k_1 = \frac{1}{11} \\ k_2 = -\frac{7}{22} \end{cases} \Rightarrow \mathbf{K} = \begin{bmatrix} \frac{1}{11} & -\frac{7}{22} \end{bmatrix}$$

La matrice di stato del sistema retro-azionato risulta essere:

$$\mathbf{A} - \mathbf{BK} = \begin{bmatrix} -k_1 & -\frac{1}{4} - k_2 \\ -3 + \frac{k_1}{2} & 1 + \frac{k_2}{2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{11} & \frac{3}{44} \\ -\frac{65}{22} & \frac{37}{44} \end{bmatrix}$$

i cui autovalori sono  $\frac{1}{2}$  e  $\frac{1}{4}$  come da specifica imposta. Si noti come il sistema inizialmente instabile, grazie a questa nostra allocazione è diventato asintoticamente stabile! Abbiamo stabilizzato il sistema!

**Esempio 7.2.3** Si consideri il sistema dinamico LTI:

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

ci viene chiesto come specifica progettuale di assegnare i poli in maniera tale che il tempo di assestamento al 2% ( $t_s^{2\%}$ ) sia 1 secondo e il fattore di smorzamento  $\delta = \frac{\sqrt{2}}{2}$ . Ricordando che  $t_s^{2\%}$  è espresso come  $\frac{4}{\delta\omega_n}$  possiamo riscrivere le specifiche come:

$$\begin{cases} t_s^{2\%} &= \frac{4}{\delta\omega_n} = 1 \Rightarrow \omega_n = \frac{4}{\delta} \\ \delta &= \frac{\sqrt{2}}{2} \end{cases}$$

Sapendo che in un sistema del secondo ordine, una coppia di poli complessi coniugati può essere espressa come:

$$p_{1/2} = -\delta\omega_n \pm j\omega_n\sqrt{1-\delta^2}$$

ricavo i valore dei poli da allocare:

$$\begin{aligned} p_{1/2} &= -\delta\omega_n \pm j\omega_n\sqrt{1-\delta^2} \\ &= -\delta\frac{4}{\delta} \pm j\frac{4}{\delta}\sqrt{1-\left(\frac{\sqrt{2}}{2}\right)^2} \\ &= -4 \pm j\frac{4}{\delta}\sqrt{1-\frac{1}{2}} = -4 \pm j\frac{4}{\delta}\sqrt{\frac{1}{2}} \\ &= -4 \pm j\frac{4}{\delta}\frac{\sqrt{2}}{2} \\ &= -4 \pm 4j \end{aligned}$$

quindi i poli sono  $p_{1/2} = -4 \pm 4j$ .

Verifichiamo la controllabilità

$$\mathcal{C} = [\mathbf{B} \quad \mathbf{AB}] = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \Rightarrow \text{rank}(\mathcal{C}) = 2$$

$\mathcal{C}$  ha rango pieno, quindi il sistema è completamente controllabile, è possibile proseguire con l'assegnazione.

Il polinomio caratteristico desiderato è:

$$\rho_d(\lambda) = (\lambda + 4 - 4j)(\lambda + 4 + 4j) = (\lambda + 4)^2 + 16 = \lambda^2 + 8\lambda + 32$$

Calcolo  $\mathbf{A} - \mathbf{BK}$ :

$$\begin{aligned} \mathbf{A} - \mathbf{BK} &= \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} [k_1 \quad k_2] \\ &= \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ k_1 & k_2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 - k_1 & -1 - k_2 \end{bmatrix} \end{aligned}$$

Il cui polinomio caratteristico è:

$$\lambda^2 + (1 + k_2)\lambda + (-1 + k_1)$$

Imponiamo i coefficienti del polinomio caratteristico desiderato, risolvendo il sistema:

$$\begin{cases} 1 + k_2 &= 8 \\ -1 + k_1 &= 32 \end{cases} \Rightarrow \begin{cases} k_2 &= 7 \\ k_1 &= 33 \end{cases} \Rightarrow \mathbf{K} = [33 \quad 7]$$

La matrice di stato del sistema retroazionato risulta essere:

$$\mathbf{A} - \mathbf{BK} = \begin{bmatrix} 0 & 1 \\ 1 - k_1 & -1 - k_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -32 & -8 \end{bmatrix}$$

i cui autovalori sono quelli imposti.

**Esempio 7.2.4** Consideriamo il seguente sistema:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \mathbf{u}(t)$$

vogliamo allocare i seguenti autovalori  $\lambda_{1,2} = -3$ .

Verifichiamo la controllabilità:

$$\mathcal{C} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \rho(\mathcal{C}) = 2 = n_c < n = 3$$

notiamo che il rango della matrice di controllabilità non è pieno. In particolare abbiamo rango=2, cioè è possibile allocare gli autovalori solo sulla parte controllabile del sistema. Per poter fare la retroazione di stato è necessario trasformare il sistema nella forma canonica di controllabilità.

Costruiamo la matrice  $\mathbf{P}$  che definisce la trasformazione di similitudine e consideriamo le seguenti identità  $\mathbf{A}_P = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ ,  $\mathbf{B}_P = \mathbf{P}^{-1}\mathbf{B}$

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{P}^{-1} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} 0 & -2 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{P}^{-1}\mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Il sistema trasformato assume la seguente forma canonica di controllabilità:

$$\begin{bmatrix} \dot{\mathbf{z}}_c(t) \\ \dot{\mathbf{z}}_{nc}(t) \end{bmatrix} = \begin{bmatrix} 0 & -2 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} \mathbf{z}_c(t) \\ \mathbf{z}_{nc}(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(t)$$

La parte controllabile del sistema è la seguente:

$$\dot{\mathbf{z}}_c(t) = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \mathbf{z}_c(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t)$$

E' possibile procedere all'allocatione degli autovalori sulla parte controllabile nelle medesime modalità viste negli esempi precedenti.

Il polinomio caratteristico desiderato è:

$$\rho_d(\lambda) = (\lambda + 3)^2 = \lambda^2 + 6\lambda + 9$$

Calcolo la matrice di stato in anello chiuso della parte controllabile del sistema  $\mathbf{A}_P - \mathbf{B}_P\mathbf{K}_P$ :

$$\begin{aligned} \mathbf{A}_P - \mathbf{B}_P\mathbf{K}_P &= \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} k_1 & k_2 \end{bmatrix} \\ &= \begin{bmatrix} -k_1 & -k_2 - 2 \\ 1 & 3 \end{bmatrix} \end{aligned}$$

Il cui polinomio caratteristico è:

$$\lambda^2 + (k_1 - 3)\lambda + (2 + k_2 - 3k_1)$$

Imponiamo i coefficienti del polinomio caratteristico desiderato, risolvendo il sistema:



$$\begin{cases} k_1 - 3 &= 6 \\ 2 + k_2 - 3k_1 &= 9 \end{cases} \Rightarrow \begin{cases} k_1 &= 9 \\ k_2 &= 34 \end{cases} \Rightarrow \mathbf{K}_P = \begin{bmatrix} 9 & 34 \end{bmatrix}$$

$\mathbf{K}_P$  è la matrice di retroazione della parte controllabile del sistema. La matrice di retroazione per l'intero sistema risulta:

$$\mathbf{K}_{P2} = [\mathbf{K}_P \quad \mathbf{0}] = [9 \quad 34 \quad 0] \quad (7.4)$$

**NB.**  $n - n_c$  componenti della matrice  $\mathbf{K}_{P2}$  possono essere poste pari a 0 in quanto gli  $n - n_c$  autovalori non controllabili non possono comunque essere spostati.

In ultima istanza, la matrice di retroazione per il sistema di partenza, risulta essere:

$$\mathbf{K}_F = \mathbf{K}_{P2} \mathbf{P}^{-1} = [9 \quad 34 \quad 0] \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -16 & 25 & 0 \end{bmatrix} \quad (7.5)$$

### 7.3 Progettazione di un regolatore in Matlab

In Matlab è molto semplice la progettazione di un regolatore.

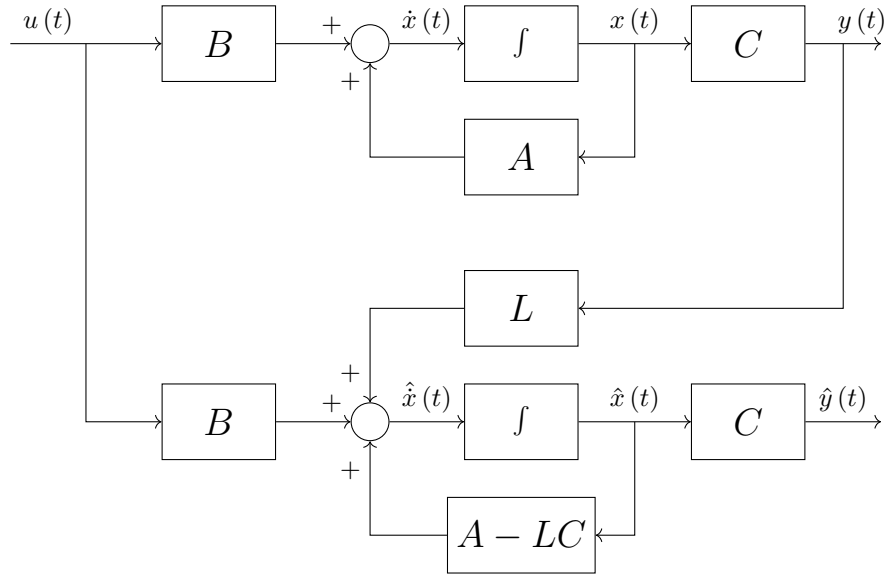
```
K = place(A,B,p)      % Calcola la matrice di retroazione dati i poli in p
K = acker(A,B,p)      % Calcola la matrice di retroazione dati i poli in p.
                       Comando da utilizzare in caso di molteplicità dei poli.
```



Lo stato di un sistema non è sempre direttamente accessibile, in quanto può dipendere da grandezze di difficile misurazione, quali, ad esempio, un flusso. Al contrario, risultano generalmente accessibili e misurabili le variabili di ingresso e di uscita, a partire dalle quali è possibile costruire stimatori di stato volti a ricostruirne l'andamento interno. Tra i principali metodi di stima dello stato si annovera l'**Osservatore di Luenberger**, il quale rappresenta un sistema dinamico descritto dalla seguente formulazione:

$$\begin{cases} \dot{\hat{\mathbf{x}}}(t) = (\mathbf{A} - \mathbf{LC})\hat{\mathbf{x}}(t) + \mathbf{B}u(t) + \mathbf{L}y(t) \\ \hat{\mathbf{y}}(t) = \mathbf{C}\hat{\mathbf{x}}(t) \end{cases} \quad (8.1)$$

dove  $\mathbf{L}$  è la matrice dei guadagni dell'osservatore.



### 8.1 Rilevabilità

**Definizione 8.1.1** Un sistema si dice **rilevabile** quando è completamente osservabile oppure, se non lo è, quando tutti gli autovalori non osservabili risultano asintoticamente stabili.

### 8.2 Progettazione di un osservatore

Il metodo per la progettazione degli osservatori è analogo a quello dei regolatori di stato, dato la natura duale tra controllabilità e osservabilità.

**Lemma 8.2.1** Come regola empirica, i poli dell'osservatore devono essere allocati in maniera da essere  $\approx 10$  volte più veloci dei poli del regolatore.

**Esempio 8.2.2** Consideriamo il seguente sistema:

$$\begin{cases} \dot{\mathbf{x}}(t) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) &= \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \mathbf{x}(t) \end{cases}$$

vogliamo progettare un osservatore che abbia autovalori  $\lambda_1 = 4$  e  $\lambda_{2,3} = -3 \pm 2j$ .  
Per prima cosa verifichiamo l'osservabilità:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & 8 & 9 \end{bmatrix} \Rightarrow \rho(\mathcal{O}) = 3 = n$$

Avendo rango pieno, il sistema risulta completamente osservabile, quindi è progettare un osservatore.  
Il polinomio caratteristico desiderato per l'osservatore è:

$$\rho_d(\lambda) = (\lambda - 4)(\lambda + 3 - 2j)(\lambda + 3 + 2j) = \lambda^3 + 2\lambda^2 - 11\lambda - 52$$

Costruiamo la matrice  $\mathbf{A} - \mathbf{LC}$ :

$$\mathbf{A} - \mathbf{LC} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} - \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1-l_1 & -2l_1 & -l_1 \\ -l_2 & 2-2l_2 & -l_2 \\ -l_3 & -2l_3 & 3-l_3 \end{bmatrix}$$

il cui polinomio caratteristico è:

$$\rho(\lambda) = \lambda^3 + \lambda^2(l_1 + 2l_2 + l_3 - 6) + \lambda(-5l_1 - 8l_2 - 3l_3 + 11) + (6l_1 + 6l_2 + 2l_3 - 6)$$

Imponiamo i coefficienti del polinomio caratteristico desiderato, risolvendo il sistema:

$$\begin{cases} l_1 + 2l_2 + l_3 - 6 &= 2 \\ -5l_1 - 8l_2 - 3l_3 + 11 &= -11 \\ 6l_1 + 6l_2 + 2l_3 - 6 &= -52 \end{cases} \Rightarrow \begin{cases} l_1 &= -30 \\ l_2 &= 29 \\ l_3 &= -20 \end{cases} \Rightarrow \mathbf{L} = \begin{bmatrix} -30 \\ 29 \\ -20 \end{bmatrix}$$

La matrice  $\mathbf{A} - \mathbf{LC}$  risulta essere:

$$\mathbf{A} - \mathbf{LC} = \begin{bmatrix} 1-l_1 & -2l_1 & -l_1 \\ -l_2 & 2-2l_2 & -l_2 \\ -l_3 & -2l_3 & 3-l_3 \end{bmatrix} = \begin{bmatrix} 31 & 60 & 30 \\ -29 & -56 & -29 \\ 20 & 40 & 20 \end{bmatrix}$$

i cui autovalori sono quelli imposti.

**Esempio 8.2.3** Consideriamo il seguente sistema:

$$\begin{cases} \dot{\mathbf{x}}(t) &= \begin{bmatrix} -1 & 0 \\ 1 & -1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 2 \\ 0 \end{bmatrix} \mathbf{u}(t) \\ \mathbf{y}(t) &= \begin{bmatrix} 0 & \frac{1}{2} \end{bmatrix} \mathbf{x}(t) \end{cases}$$

vogliamo progettare un osservatore che abbia autovalori  $\lambda_{1,2} = -4$ .  
Per prima cosa verifichiamo l'osservabilità:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \Rightarrow \rho(\mathcal{O}) = 2 = n$$

Avendo rango pieno, il sistema risulta completamente osservabile, quindi è possibile progettare un osservatore.

Il polinomio caratteristico desiderato per l'osservatore è:

$$\rho_d(\lambda) = (\lambda + 4)^2 = \lambda^2 + 8\lambda + 16$$

Costruiamo la matrice  $\mathbf{A} - \mathbf{LC}$ :

$$\mathbf{A} - \mathbf{LC} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -1 & -\frac{l_1}{2} \\ 1 & -\frac{l_2}{2} - 1 \end{bmatrix}$$

il cui polinomio caratteristico è:

$$\rho(\lambda) = \lambda^2 + \lambda \left( 2 + \frac{l_2}{2} \right) + \left( \frac{\lambda_1}{2} + \frac{\lambda_2}{2} + 1 \right)$$

Imponiamo i coefficienti del polinomio caratteristico desiderato, risolvendo il sistema:

$$\mathbf{L} = \begin{bmatrix} 18 \\ 12 \end{bmatrix}$$

è semplice verifica come gli autovalori di  $\mathbf{A} - \mathbf{LC}$  siano quelli imposti da specifica.

### 8.3 Principio di separazione

**Definizione 8.3.1** La progettazione del regolatore (matrice  $\mathbf{K}$ ) e quella dell'osservatore (matrice  $\mathbf{L}$ ) può essere fatto in maniera indipendente.

**Dimostrazione** TODO

□

### 8.4 Progettazione di un osservatore in Matlab

In Matlab è molto semplice la progettazione di un osservatore. Data la dualità di regolatori e osservatori, si utilizzano gli stessi comandi visti in precedenza.

```
L = place(A',C',p)      % Calcola la matrice di osservazione dati i poli in p
L = acker(A',C',p)      % Calcola la matrice di osservazione dati i poli in p.
                        Comando da utilizzare in caso di molteplicità dei poli.
```



### 9.1 Introduzione

Un sistema non lineare tempo invariante è descritto da una serie di equazioni di stato nella forma:

$$\begin{cases} \dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(t) = g(\mathbf{x}(t), \mathbf{u}(t)) \end{cases} \quad (9.1)$$

**Tempo Continuo**

$$\begin{cases} \mathbf{x}(k+1) = f(\mathbf{x}(k+1), \mathbf{u}(k+1)) \\ \mathbf{y}(k) = g(\mathbf{x}(k+1), \mathbf{u}(k+1)) \end{cases} \quad (9.2)$$

**Tempo Discreto**

dove  $f$  e  $g$  sono due funzioni reali non lineari.

Poiché le osservazioni che faremo sono analoghe sia sui sistemi a tempo continuo che sui sistemi a tempo discreto, nei paragrafi seguenti verrà illustrata la linearizzazione per i sistemi a tempo continuo, sapendo che la stessa procedura si applica in modo del tutto analogo anche ai sistemi a tempo discreto.

### 9.2 Linearizzazione

Iniziamo richiamando un concetto fondamentale: una funzione può essere approssimata localmente da un polinomio.

Sia  $f(x)$  una funzione reale di una variabile reale. e di voler capire come si comporta vicino a un certo punto  $\bar{x}$ . La serie di Taylor ci permette di fare una approssimazione del comportamento della funzione attorno a quel punto. La formula generale è:

$$f(x) = \underbrace{f(\bar{x})}_{\text{valore in } \bar{x}} + \underbrace{\frac{f'(\bar{x})}{1!}(x - \bar{x})}_{\text{contributo derivata prima}} + \underbrace{\frac{f''(\bar{x})}{2!}(x - \bar{x})^2}_{\text{contributo derivata seconda}} + \dots \quad (9.3)$$

Ogni derivata aggiunge un'informazione sulla forma locale della funzione, migliorando l'approssimazione. In forma compatta, la serie di Taylor è:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(\bar{x})}{n!} (x - \bar{x})^n \quad (9.4)$$

Se ci interessa solo un'approssimazione semplice, spesso basta considerare soltanto il primo termine non costante della serie. In questo caso otteniamo lo sviluppo di Taylor al primo ordine:

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) \quad (9.5)$$

Questa espressione rappresenta l'equazione della retta tangente alla funzione nel punto  $x_0$ . In altre parole, in un intorno di  $x_0$  la funzione è indistinguibile dalla sua tangente.

Se la funzione dipende da due variabili,  $f(x_1, x_2)$ , l'idea è la stessa; cambia solo il fatto che dobbiamo considerare le derivate parziali. Lo sviluppo di Taylor al primo ordine diventa:

$$f(x_1, x_2) \approx f(\bar{x}_1, \bar{x}_2) + \left. \frac{\partial f}{\partial x_1} \right|_{\substack{x_1 = \bar{x}_1 \\ x_2 = \bar{x}_2}} (x_1 - \bar{x}_1) + \left. \frac{\partial f}{\partial x_2} \right|_{\substack{x_1 = \bar{x}_1 \\ x_2 = \bar{x}_2}} (x_2 - \bar{x}_2) \quad (9.6)$$

La logica è la seguente:

- si parte dal valore della funzione nel punto centrale  $(\bar{x}_1, \bar{x}_2)$ ;
- si aggiunge quanto cambia la funzione se ci spostiamo leggermente nella direzione di  $x_1$ ;

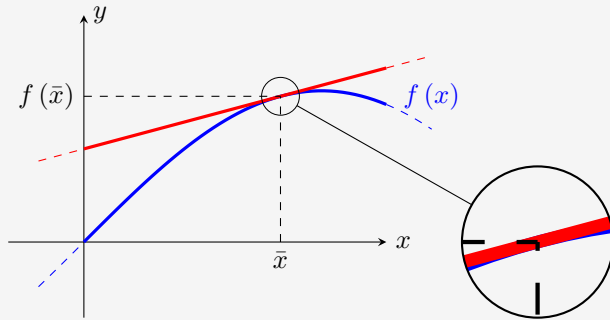
- si aggiunge quanto cambia la funzione se ci spostiamo nella direzione di  $x_2$ .

Le derivate parziali misurano quindi la sensibilità della funzione rispetto a ciascuna direzione.

Da un punto di vista geometrico, questa approssimazione rappresenta il *piano tangente* alla superficie definita da  $f(x_1, x_2)$  nel punto  $(\bar{x}_1, \bar{x}_2)$ .

**FAQ 9.2.1 - Perché nella linearizzazione si utilizza solo il primo termine della serie di Taylor?** La ragione principale è che la linearizzazione non ha lo scopo di ricostruire fedelmente la funzione originale, ma di descriverne il comportamento locale, cioè ciò che accade in un intorno molto piccolo del punto considerato. In questa regione, il contributo dei termini di ordine superiore diventa rapidamente trascurabile: man mano che ci si avvicina al punto di equilibrio, le variazioni quadratiche, cubiche e così via diminuiscono molto più rapidamente della variazione lineare, che finisce quindi per essere l'unica parte significativa dell'espansione. Inoltre, dal punto di vista della dinamica, tutte le proprietà qualitative locali del sistema (come la stabilità e il tipo di equilibrio) dipendono esclusivamente dal termine di primo ordine, mentre i contributi più alti si manifestano solo quando ci si allontana dal punto di equilibrio.

Includerli renderebbe il modello più complesso senza fornire un'informazione utile per l'analisi locale, riportandoci di fatto a un sistema non lineare. La linearizzazione è quindi una descrizione infinitesimale del sistema, una sorta di "zoom estremo" in cui la funzione originale, per quanto complessa, appare indistinguibile dalla retta tangente che ne rappresenta lo sviluppo del primo ordine.



### 9.2.1 Linearizzazione di un sistema non lineare

**Definizione 9.2.2** Dato un sistema non lineare a tempo continuo (9.1), una coppia  $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$  si dice **punto di equilibrio** (o semplicemente equilibrio) se ponendo:

$$\mathbf{x}(0) = \bar{\mathbf{x}}, \quad \mathbf{u}(t) = \bar{\mathbf{u}} \quad \forall t \geq 0$$

la soluzione del sistema rimane costante:

$$\mathbf{x}(t) = \bar{\mathbf{x}} \quad \forall t \geq 0$$

In altre parole, una volta portato il sistema nello stato  $\bar{\mathbf{x}}$  e mantenuto l'ingresso costante a  $\bar{\mathbf{u}}$ , il sistema non si muove più. Dal punto di vista analitico, questa condizione è equivalente a richiedere che il campo vettoriale si annulli in  $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ , cioè:

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

Un punto di equilibrio rappresenta quindi una condizione di funzionamento stazionaria del sistema, attorno alla quale ha senso studiare piccole perturbazioni e applicare la linearizzazione.

Immaginiamo il sistema nel suo punto di equilibrio, dove tutto rimane immobile e nulla varia nel tempo. Anche un lieve cambiamento dell'ingresso o un piccolo spostamento dello stato è però sufficiente per modificare l'evoluzione futura.

Per comprendere come il sistema reagisce a questi disturbi, consideriamo in modo esplicito le perturbazioni attorno all'equilibrio. Questo ci permette di descrivere la risposta a piccoli scostamenti e di ricavare un modello lineare che approssima il comportamento locale.

Per rendere chiaro questo meccanismo, esprimiamo stato e ingresso come somma del valore di equilibrio e di una piccola variazione:

$$x(t) = \bar{x} + \Delta x(t) \quad (9.7a)$$

$$u(t) = \bar{u} + \Delta u(t) \quad (9.7b)$$



L'idea è che, se tali perturbazioni sono sufficientemente piccole, allora anche la dinamica risultante può essere descritta come una "perturbazione" della dinamica d'equilibrio. Quindi dalla Eq. 9.7a abbiamo che:

$$\dot{x}(t) = \frac{d}{dt}(\bar{x} + \Delta x(t)) = \dot{\Delta x}(t) \quad (9.8)$$

Sostituendo queste espressioni nell'equazione di stato e sviluppando  $f(x(t), u(t))$  in serie di Taylor attorno al punto  $(\bar{x}, \bar{u})$  otteniamo:

$$f(x(t), u(t)) = f(\bar{x}, \bar{u}) + \left. \frac{\partial f}{\partial x} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} (x(t) - \bar{x}) + \left. \frac{\partial f}{\partial u} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} (u(t) - \bar{u}) \quad (9.9)$$

Dato che nel punto di equilibrio  $f(\bar{x}, \bar{u}) = 0$ , le piccole perturbazioni corrispondono a  $\Delta x(t) = x(t) - \bar{x}$  e  $\Delta u(t) = u(t) - \bar{u}$  e che  $\dot{x}(t) = f(x(t), u(t))$  abbiamo:

$$\begin{aligned} \Delta \dot{x}(t) &= \left. \frac{\partial f}{\partial x} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} \Delta x(t) + \left. \frac{\partial f}{\partial u} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} \Delta u(t) \\ &= f_x(\bar{x}, \bar{u}) \Delta x(t) + f_u(\bar{x}, \bar{u}) \Delta u(t) \end{aligned} \quad (9.10)$$

Analogamente, abbiamo delle piccole perturbazioni anche nell'equazione d'uscita:

$$y(t) = \bar{y} + \Delta y(t) \quad (9.11)$$

da cui:

$$g(x(t), u(t)) = g(\bar{x}, \bar{u}) + \left. \frac{\partial g}{\partial x} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} (x(t) - \bar{x}) + \left. \frac{\partial g}{\partial u} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} (u(t) - \bar{u}) \quad (9.12)$$

Considerando che  $g(\bar{x}, \bar{u}) = \bar{y}$  perchè siamo all'equilibrio:

$$\begin{aligned} \Delta y(t) &= \left. \frac{\partial g}{\partial x} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} \Delta x(t) + \left. \frac{\partial g}{\partial u} \right|_{\substack{x=\bar{x} \\ u=\bar{u}}} \Delta u(t) \\ &= g_x(\bar{x}, \bar{u}) \Delta x(t) + g_u(\bar{x}, \bar{u}) \Delta u(t) \end{aligned} \quad (9.13)$$

Il sistema linearizzato per un sistema SISO è:

$$\begin{cases} \Delta \dot{x}(t) &= f_x(\bar{x}, \bar{u}) \Delta x(t) + f_u(\bar{x}, \bar{u}) \Delta u(t) \\ \Delta y(t) &= g_x(\bar{x}, \bar{u}) \Delta x(t) + g_u(\bar{x}, \bar{u}) \Delta u(t) \end{cases} \quad (9.14)$$

Nel caso di sistema MIMO avremo:

$$f_x(\bar{x}, \bar{u}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad \begin{matrix} x = \bar{x} \\ u = \bar{u} \end{matrix} \quad f_u(\bar{x}, \bar{u}) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} \\ \vdots \\ \frac{\partial f_n}{\partial u_1} \end{bmatrix} \quad \begin{matrix} x = \bar{x} \\ u = \bar{u} \end{matrix} \quad (9.15)$$

$$g_x(\bar{x}, \bar{u}) = \begin{bmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{bmatrix} \quad \begin{matrix} x = \bar{x} \\ u = \bar{u} \end{matrix} \quad g_u(\bar{x}, \bar{u}) = \begin{bmatrix} \frac{\partial g}{\partial u_1} \end{bmatrix} \quad \begin{matrix} x = \bar{x} \\ u = \bar{u} \end{matrix} \quad (9.16)$$

Il sistema linearizzato viene chiamato **sistema lineare tangente**:

$$\begin{aligned} \Delta \dot{x}(t) &= A \Delta x(t) + B \Delta u(t) \\ \Delta y(t) &= C \Delta x(t) + D \Delta u(t). \end{aligned} \quad (9.17)$$

Questo sistema rappresenta una buona approssimazione del comportamento del sistema non lineare soltanto nelle vicinanze del punto di equilibrio. Man mano che ci si allontana dal punto  $(\bar{x}, \bar{u})$ , gli effetti della non linearità diventano rilevanti e l'approssimazione perde validità.

**Esempio 9.2.3 Modello non lineare glucosio–insulina**

Consideriamo il seguente modello (Bergman minimal model [1]), che descrive l'interazione tra glucosio  $G$ , insulina  $I$  e azione insulinica  $X$ :

$$\dot{G}(t) = -p_1(G(t) - G_b) - X(t)G(t) + M(t), \quad (9.18)$$

$$\dot{X}(t) = -p_2X(t) + p_3(I(t) - I_b), \quad (9.19)$$

$$\dot{I}(t) = -n(I(t) - I_b) + u(t). \quad (9.20)$$

Dove:

- $G(t)$  [mg/dL]: concentrazione di glucosio nel sangue;
- $I(t)$  [ $\mu$ U/mL]: concentrazione di insulina plasmatica;
- $X(t)$  [ $\text{min}^{-1}$ ]: azione insulinica (effetto dell'insulina sul glucosio);
- $G_b, I_b$ : valori basali (a digiuno) di glucosio [mg/dL] e insulina [ $\mu$ U/mL];
- $M(t)$  [mg/dL/min]: apporto esterno di glucosio (es. pasto);
- $u(t)$  [ $\mu$ U/mL/min]: variazione di insulina esogena iniettata (controllo);
- $p_1, p_2, n$  [ $\text{min}^{-1}$ ]: costanti di tempo effettive delle rispettive dinamiche;
- $p_3$  [ $\text{min}^{-2} \cdot \text{mL}/\mu\text{U}$ ] (in pratica: “per unità di insulina per minuto”): guadagno che converte la variazione di insulina in variazione di azione insulinica.

**Punto di equilibrio**

Per linearizzare, scegliamo un punto di equilibrio  $(\bar{x}, \bar{u})$  del sistema, cioè una condizione in cui lo stato è costante nel tempo ( $\dot{x} = 0$ ).

Assumiamo condizioni basali: nessun pasto e nessuna iniezione esterna, cioè:

$$M(t) = 0, \quad u(t) = 0.$$

Cerchiamo una soluzione stazionaria tale che:

$$\dot{G} = 0, \quad \dot{X} = 0, \quad \dot{I} = 0.$$

Dalle equazioni (9.18)–(9.20) si ottiene, in modo naturale, il seguente equilibrio:

$$G^* = G_b, \quad X^* = 0, \quad I^* = I_b, \quad M^* = 0, \quad u^* = 0.$$

**Calcolo delle derivate parziali**

Definiamo le funzioni:

$$f_1(G, X, I, u) = -p_1(G - G_b) - XG + M(t), \quad (9.21)$$

$$f_2(G, X, I, u) = -p_2X + p_3(I - I_b), \quad (9.22)$$

$$f_3(G, X, I, u) = -n(I - I_b) + u. \quad (9.23)$$

**Equazione del glucosio**

$$\frac{\partial f_1}{\partial G} = -p_1 - X, \quad (9.24)$$

$$\frac{\partial f_1}{\partial X} = -G, \quad (9.25)$$

$$\frac{\partial f_1}{\partial I} = 0. \quad (9.26)$$

Valutando al punto di equilibrio  $(G^*, X^*, I^*) = (G_b, 0, I_b)$ :

$$\left. \frac{\partial f_1}{\partial G} \right|_{(x^*, u^*)} = -p_1, \quad (9.27)$$

$$\left. \frac{\partial f_1}{\partial X} \right|_{(x^*, u^*)} = -G_b, \quad (9.28)$$

$$\left. \frac{\partial f_1}{\partial I} \right|_{(x^*, u^*)} = 0. \quad (9.29)$$

**Equazione dell'azione insulinica**

$$\frac{\partial f_2}{\partial G} = 0, \quad (9.30)$$

$$\frac{\partial f_2}{\partial X} = -p_2, \quad (9.31)$$

$$\frac{\partial f_2}{\partial I} = p_3. \quad (9.32)$$

**Equazione dell'insulina plasmatica**

$$\frac{\partial f_3}{\partial G} = 0, \quad (9.33)$$

$$\frac{\partial f_3}{\partial X} = 0, \quad (9.34)$$

$$\frac{\partial f_3}{\partial I} = -n, \quad (9.35)$$

$$\frac{\partial f_3}{\partial u} = 1. \quad (9.36)$$

**Matrici del modello linearizzato**

Raccogliendo i risultati precedenti, otteniamo la matrice  $A$ :

$$A = \begin{bmatrix} -p_1 & -G_b & 0 \\ 0 & -p_2 & p_3 \\ 0 & 0 & -n \end{bmatrix}, \quad (9.37)$$

e la matrice  $B$ :

$$B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (9.38)$$

Definiamo ora le deviazioni rispetto all'equilibrio:

$$\tilde{G} = G - G_b, \quad \tilde{X} = X - 0 = X, \quad \tilde{I} = I - I_b.$$

### Equazioni linearizzate

Il sistema linearizzato (in variabili di deviazione) assume la forma:

$$\dot{\tilde{G}}(t) = -p_1 \tilde{G}(t) - G_b \tilde{X}(t) + M(t), \quad (9.39)$$

$$\dot{\tilde{X}}(t) = -p_2 \tilde{X}(t) + p_3 \tilde{I}(t), \quad (9.40)$$

$$\dot{\tilde{I}}(t) = -n \tilde{I}(t) + u(t). \quad (9.41)$$

Osservazioni:

- Il termine  $-G_b \tilde{X}$  rappresenta l'effetto riduttivo dell'azione insulinica sul glucosio;
- La dinamica  $\tilde{I} \rightarrow \tilde{X} \rightarrow \tilde{G}$  è a cascata: l'insulina influenza l'azione insulinica, che a sua volta influenza il glucosio;
- I parametri  $p_1, p_2, n > 0$  garantiscono stabilità della parte lineare (poli a parte reale negativa).

### Funzione di trasferimento linearizzata

Se trascuriamo l'ingresso di glucosio  $M(t)$  (cioè consideriamo solo l'effetto di  $u(t)$  sulla glicemia), possiamo ricavare la funzione di trasferimento  $\tilde{G}(s)/U(s)$ .

Sotto ipotesi di condizioni iniziali nulle:

$$(s+n)\tilde{I}(s) = U(s) \Rightarrow \tilde{I}(s) = \frac{1}{s+n}U(s), \quad (9.42)$$

$$(s+p_2)\tilde{X}(s) = p_3\tilde{I}(s) = \frac{p_3}{s+n}U(s) \Rightarrow \tilde{X}(s) = \frac{p_3}{(s+p_2)(s+n)}U(s), \quad (9.43)$$

$$(s+p_1)\tilde{G}(s) = -G_b\tilde{X}(s) = -G_b\frac{p_3}{(s+p_2)(s+n)}U(s). \quad (9.44)$$

Ne segue:

$$\frac{\tilde{G}(s)}{U(s)} = \frac{-G_b p_3}{(s+p_1)(s+p_2)(s+n)}. \quad (9.45)$$

Commenti:

- Non ci sono zeri: solo tre poli reali in  $-p_1, -p_2, -n$ ;
- Il *segno negativo* indica che un aumento di insulina esogena  $u(t)$  comporta una *diminuzione* della glicemia  $\tilde{G}(t)$  (effetto ipoglicemizante).

### Interpretazione e limiti della linearizzazione

- La linearizzazione è *valida localmente*, cioè per piccole variazioni attorno al punto di equilibrio  $(G_b, 0, I_b)$ .
- Per variazioni ampie (es. pasti molto grandi, dosi di insulina elevate) la dinamica reale può discostarsi sensibilmente dal modello lineare.
- Nonostante questo, il modello linearizzato è estremamente utile per:
  - studiare la stabilità locale;
  - progettare controllori lineari (PID, LQR, IMC, ...);
  - comprendere qualitativamente l'effetto degli ingressi sulla glicemia.

- [1] Bergman, R. N., Ider, Y. Z., Bowden, C. R., & Cobelli, C. (1979). Quantitative estimation of insulin sensitivity. *American Journal of Physiology*, 236(6), E667–E677.

**Esempio 9.2.4** Si consideri il sistema:

$$\begin{cases} \dot{x}_1(t) &= -x_2(t)u(t) + u(t) \\ \dot{x}_2(t) &= -x_2(t) + x_1(t)u(t) \\ \dot{x}_3(t) &= -x_3(t) - x_2(t)u(t) \\ y(t) &= x_1(t) \end{cases}$$

si determinino lo stato e l'uscita di equilibrio in corrispondenza dell'ingresso costante  $u(t) = \bar{u} = 1 \quad \forall t > 0$  e si linearizzi il sistema intorno all'equilibrio trovato.

Cerchiamo lo stato e l'uscita di equilibrio:

$$\begin{cases} 0 &= -\bar{x}_2 + 1 \\ 0 &= -\bar{x}_2 + \bar{x}_1 \\ 0 &= -\bar{x}_3 - \bar{x}_2 \\ \bar{y} &= \bar{x}_1 \end{cases} \Rightarrow \begin{cases} \bar{x}_2 &= 1 \\ \bar{x}_1 &= \bar{x}_2 = 1 \\ \bar{x}_3 &= -\bar{x}_2 = -1 \\ \bar{y} &= \bar{x}_1 = 1 \end{cases} \Rightarrow \bar{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \quad \bar{y} = 1$$

Una volta trovato il punto di equilibrio, procediamo a linearizzare il sistema:

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

$$\mathbf{A} = \begin{bmatrix} 0 & -\bar{u} & 0 \\ \bar{u} & -1 & 0 \\ 0 & -\bar{u} & -1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & -1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} -\bar{x}_2 + 1 \\ \bar{x}_1 \\ -\bar{x}_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$\mathbf{C} = [1 \quad 0 \quad 0] \quad \mathbf{D} = 0$$

**Esempio 9.2.5** Si consideri il sistema:

$$\begin{cases} \dot{x}_1(t) &= x_1(t)x_2(t) + u(t) \\ \dot{x}_2(t) &= -x_1(t)x_2(t) - x_2(t) \\ y(t) &= \sqrt{x_1(t)} \end{cases}$$

e si determinino i punti di equilibrio e il sistema linearizzato attorno questi ultimi.

**Esempio 9.2.6 Pendolo**

Consideriamo ora il modello non lineare del pendolo:

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = -\frac{g}{l} \sin(x_1) - \frac{h}{Ml}x_2 + \frac{1}{Ml}u.$$

Calcolando le derivate parziali si ottiene la matrice  $A$ , che dipende dal punto di equilibrio scelto. Notiamo un fatto importante:

- Linearizzando attorno a  $x_1 = 0$  (pendolo verso il basso) otteniamo un sistema stabile.
- Linearizzando attorno a  $x_1 = \pi$  (pendolo verso l'alto) otteniamo un sistema instabile.

Questo mostra come la linearizzazione *catturi* la stabilità locale del sistema.



# CAPITOLO 10

## Controllo ottimo

**Definizione 10.0.1** Per **controllo ottimo**, si intende trovare un'azione di controllo  $\mathbf{u}^* \in \mathbf{U}$  tale che il sistema:

$$\begin{cases} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \end{cases}$$

segua la traiettoria  $\mathbf{x}^* \in \mathbf{X}$  che minimizzi l'indice di performance:

$$\mathbf{J} \tag{10.1}$$

### 10.1 La funzione di costo

Vediamo ora alcuni obiettivi di controllo tipici in maniera da fornire le motivazioni specifiche riguardo la scelta di una funzione di costo.

#### Tempo minimo

Nei problemi di **tempo minimo**, ci riferiamo a situazioni in cui vogliamo capire qual è il modo migliore per far evolvere un sistema da uno stato iniziale  $\mathbf{x}_0$  a uno finale  $\mathbf{x}_f$  nel minor tempo possibile, cioè:

$$J = t_f - t_0 = \int_{t_0}^{t_f} dt \tag{10.2}$$

#### Valore a regime

Nei **problemi di valore a regime**, l'obiettivo è minimizzare la deviazione dello stato finale di un sistema  $\mathbf{x}(t_f)$  rispetto al suo valore desiderato  $\mathbf{r}(t_f)$ .

Matematicamente, una sua possibile formulazione è:

$$J = \sum_{i=1}^n [x_i(t_f) - r_i(t_f)]^2 \tag{10.3}$$

in cui  $J$  equivale alla somma quadratica degli errori a regime su tutti gli stati.

Dal momento che un errore positivo o negativo che sia, è considerato una problematica in entrambi i casi, si utilizza il quadrato dell'errore. In teoria si potrebbero impiegare i valori assoluti, ma la forma quadratica adottata nell'Eq. (10.3) risulta molto più comoda da gestire dal punto di vista matematico. Sfruttando la notazione matriciale, si ottiene:

$$\mathbf{J} = [\mathbf{x}(t_f) - \mathbf{r}(t_f)]^T [\mathbf{x}(t_f) - \mathbf{r}(t_f)] \tag{10.4}$$

Per aumentare la flessibilità del modello, è possibile inserire una matrice dei pesi  $\mathbf{H}$  (reale, simmetrica e semidefinita positiva) di dimensione  $n \times n$ . In questo modo si ottiene una formulazione più generale della funzione di costo.

$$\mathbf{J} = [\mathbf{x}(t_f) - \mathbf{r}(t_f)]^T \mathbf{H} [\mathbf{x}(t_f) - \mathbf{r}(t_f)] \tag{10.5}$$

Quando la matrice dei pesi  $H$  è la matrice identità, le equazioni (10.4) e (10.5) coincidono.

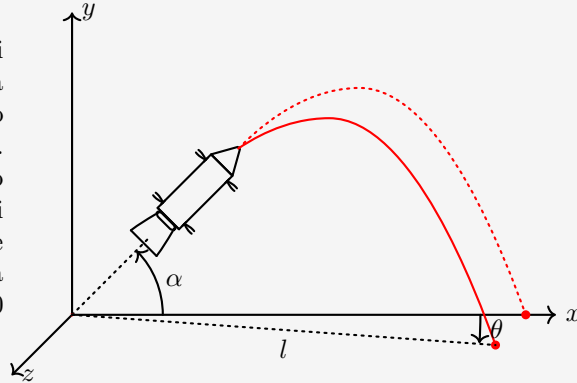
Consideriamo ora il caso in cui  $H$  sia una matrice diagonale. Poiché deve essere semidefinita positiva, i suoi elementi diagonali non possono essere negativi. Questo ci permette di controllare l'importanza relativa delle diverse variabili dello stato finale nella funzione di costo: aumentando un coefficiente  $h_{ii}$  diamo maggiore peso all'errore relativo alla variabile  $x_i(t_f)$ . Ponendolo invece pari a zero, indichiamo che tale variabile non deve contribuire alla misura di prestazione.

**IMPORTANTE:** I valori della matrice  $\mathbf{H}$  devono essere scelti tenendo conto delle diverse scale numeriche delle grandezze coinvolte!



**Esempio 10.1.1 - Importanza delle scale numeriche.**

Consideriamo un razzo della SpaceX la cui posizione sia descritta dai seguenti stati: la distanza dal punto di partenza  $l(t)$ , l'angolo di elevazione  $\alpha(t)$  e l'angolo di azimut  $\theta(t)$ . Supponiamo che questo razzo sia utilizzato per il turismo spaziale e parta dalla base di Cape Canaveral in Florida con destinazione dello spaziorporto di Grottaglie in Puglia. La distanza tra i due spaziorporti è di circa 9000 km.



Supponiamo ora che il razzo presenti un errore molto piccolo nell'angolo di azimut, ad esempio  $\theta(t_f) = 0.01$  rad. Anche se si tratta di un errore angolare minimo lo sfasamento laterale su di una distanza di 9000 km è enorme (circa 90 km). Se ad esempio questo sfasamento fosse verso sud, invece di atterrare a Grottaglie, ci troveremmo nei pressi del comune di Crosia in Calabria! Dunque, un errore angolare apparentemente insignificante produce un errore reale molto grande.

Se adottassimo una funzione di costo semplice del tipo

$$\begin{aligned} J &= [\mathbf{x}(t_f) - \mathbf{r}(t_f)]^T [\mathbf{x}(t_f) - \mathbf{r}(t_f)] \\ &= \begin{bmatrix} l(t_f) - r_l(t_f) \\ \theta(t_f) - r_\theta(t_f) \end{bmatrix}^T \begin{bmatrix} l(t_f) - r_l(t_f) \\ \theta(t_f) - r_\theta(t_f) \end{bmatrix} \\ &= \begin{bmatrix} l(t_f) - 9000 \\ \theta(t_f) - 0 \end{bmatrix}^T \begin{bmatrix} l(t_f) - 9000 \\ \theta(t_f) - 0 \end{bmatrix} \\ &= (l(t_f) - 9000)^2 + \theta^2(t_f) \end{aligned}$$

il termine  $\theta(t_f)^2$  sarebbe molto piccolo (poiché  $0.01^2 = 10^{-4}$ ) e verrebbe praticamente ignorato il che sarebbe un errore dato che al posto di taralli, troveremmo della N'duja all'arrivo.

Per rendere l'effetto dell'errore angolare confrontabile con quello della distanza, è necessario introdurre dei pesi nella funzione di costo. In particolare, bisogna attribuire a  $\theta(t_f)$  un peso maggiore per compensare il fatto che un errore di 0.01 rad corrisponda 90 km di scostamento laterale.

La funzione di costo diventa quindi

$$\begin{aligned} J &= [\mathbf{x}(t_f) - \mathbf{r}(t_f)]^T \mathbf{H} [\mathbf{x}(t_f) - \mathbf{r}(t_f)] \\ &= \begin{bmatrix} l(t_f) - r_l(t_f) \\ \theta(t_f) - r_\theta(t_f) \end{bmatrix}^T \begin{bmatrix} h_{11} & 0 \\ 0 & h_{22} \end{bmatrix} \begin{bmatrix} l(t_f) - r_l(t_f) \\ \theta(t_f) - r_\theta(t_f) \end{bmatrix} \\ &= h_{11} (l(t_f) - 9000)^2 + h_{22} \theta^2(t_f) \end{aligned}$$

Poiché 0.01 rad produce uno scostamento di circa 90 km, il coefficiente  $h_{22}$  deve amplificare opportunamente l'errore angolare. Una scelta adeguata per bilanciare gli effetti potrebbe essere:

$$h_{22} = \left( \frac{90}{0.01} \right)^2 h_{11},$$

così da rendere equivalenti in termini di impatto sulla funzione di costo, le deviazioni nella distanza e quelle nell'angolo di azimut.

In alternativa è possibile normalizzare direttamente le variabili  $l$  e  $\theta$ ; in tal caso una scelta naturale sarebbe porre  $h_{11} = h_{22}$ .

**Minima azione di controllo**

Nei **problemi di controllo ad azione minima**, si vuole passare dallo stato iniziale  $\mathbf{x}_0$  a quello finale  $\mathbf{x}_f$  utilizzando l'azione di controllo minima.

Per rendere l'idea, consideriamo una sonda spaziale impegnata in una missione interplanetaria. Sia  $u(t)$  la spinta del motore, e supponiamo che la spinta sia proporzionale al consumo istantaneo di carburante. Se l'obiettivo è ridurre il più possibile il carburante utilizzato durante la manovra, una scelta naturale per la funzione di costo è la seguente:

$$\mathbf{J} = \int_{t_0}^{t_f} |\mathbf{u}(t)| dt \quad (10.6)$$

Questo integrale rappresenta la quantità totale di spinta (e quindi di carburante) impiegata nell'intervallo di controllo.

Nel caso in cui siano presenti più comandi di controllo, e ciascuno comporti un costo diverso, si può assegnare un peso a ognuno di essi. In tal modo, la quantità totale di sforzo diventa:

$$\mathbf{J} = \int_{t_0}^{t_f} \left[ \sum_{i=1}^m \beta_i |\mathbf{u}_i(t)| \right] dt \quad (10.7)$$

dove i coefficienti  $\beta_i$  sono pesi non negativi che consentono di attribuire maggiore o minore rilevanza allo sforzo associato a ciascun controllo.

Un altro esempio riguarda una sorgente di tensione che alimenta una rete elettrica priva di elementi in grado di immagazzinare energia. Indichiamo con  $u(t)$  la tensione erogata. In questo caso, la corrente fornita dalla sorgente è direttamente proporzionale alla tensione, e l'energia dissipata è proporzionale al quadrato della tensione stessa. Per minimizzare l'energia consumata, la funzione di costo appropriata è:

$$\mathbf{J} = \int_{t_0}^{t_f} \mathbf{u}^2(t) dt \quad (10.8)$$

Se i controlli sono molteplici, la generalizzazione naturale della forma precedente è:

$$\mathbf{J} = \int_{t_0}^{t_f} [\mathbf{u}^T(t) \mathbf{R} \mathbf{u}(t)] dt \quad (10.9)$$

che può essere interpretata anche come una norma pesata:

La matrice  $\mathbf{R}$  è reale, simmetrica e definita positiva, e funge da matrice di pesi che determina quanto sia "costoso" ciascun comando di controllo. Gli elementi della matrice possono dipendere dal tempo, nel caso in cui si desideri modificare la penalizzazione dello sforzo di controllo lungo l'intervallo  $[t_0, t_f]$ .

### Tracciamento

Si desidera mantenere lo stato del sistema  $\mathbf{x}(t)$  il più vicino possibile allo stato desiderato  $\mathbf{r}(t)$  nell'intervallo di tempo  $[t_0, t_f]$ .

Per misurare la qualità del controllo scegliamo come funzione di costo:

$$\mathbf{J} = \int_{t_0}^{t_f} [\mathbf{x}(t) - \mathbf{r}(t)]^T \mathbf{Q} [\mathbf{x}(t) - \mathbf{r}(t)] dt \quad (10.10)$$

dove  $\mathbf{Q}(t)$  è una matrice reale, simmetrica e semidefinita positiva di dimensione  $n \times n$ , definita per ogni istante  $t \in [t_0, t_f]$ . Gli elementi della matrice  $\mathbf{Q}$  vengono scelti in modo da pesare in maniera opportuna l'importanza relativa delle diverse componenti dello stato, oppure per normalizzare le grandezze in gioco. Ad esempio, se  $\mathbf{Q}$  è una matrice diagonale e un certo elemento diagonale  $q_{ii}$  è pari a zero, significa che le deviazioni della componente  $x_i$  non sono rilevanti.

Se l'insieme dei controlli ammissibili è limitato (ad esempio, se  $|u_i(t)| \leq 1$ , per  $i = 1, \dots, m$ ), allora la funzione di costo (2.1-10) è adeguata. Tuttavia, se i controlli non sono limitati, la minimizzazione di questa funzione porta in genere a controlli impulsivi o con derivate impulsive. Per evitare vincoli espliciti sul controllo, oppure quando si desidera penalizzare anche il consumo di energia di controllo, si introduce una funzione di costo modificata:

$$\mathbf{J} = \int_{t_0}^{t_f} [\mathbf{x}(t) - \mathbf{r}(t)]^T \mathbf{Q} [\mathbf{x}(t) - \mathbf{r}(t)] + \mathbf{u}^T(t) \mathbf{R} \mathbf{u}(t) dt \quad (10.11)$$

dove  $\mathbf{R}(t)$  è una matrice reale, simmetrica e definita positiva di dimensione  $m \times m$ , specificata per ogni  $t \in [t_0, t_f]$ . Come vedremo nella Sezione 5.2, se il sistema da controllare è lineare, questa funzione porta a un controllore ottimo che può essere implementato con facilità.

In alcune applicazioni è particolarmente importante che lo stato del sistema sia vicino al valore desiderato proprio nell'istante finale  $t_f$ . In questo caso si può introdurre un termine finale nella funzione di costo, ottenendo:

$$\mathbf{J} = \int_{t_0}^{t_f} [\mathbf{x}(t) - \mathbf{r}(t)]^T \mathbf{Q} [\mathbf{x}(t) - \mathbf{r}(t)] + \mathbf{u}^T(t) \mathbf{R} \mathbf{u}(t) dt + [\mathbf{x}(t_f) - \mathbf{r}(t_f)]^T \mathbf{H} [\mathbf{x}(t_f) - \mathbf{r}(t_f)] \quad (10.12)$$

dove  $\mathbf{H}$  è una matrice reale, simmetrica e semidefinita positiva di dimensione  $n \times n$ . Questo termine aggiuntivo permette di dare un peso specifico all'errore finale, che in molte applicazioni è un requisito cruciale.

### Regolazione

Un problema di regolazione può essere visto come un caso particolare del problema di inseguimento (*tracking*). Tale situazione si verifica quando il valore desiderato dello stato è identicamente nullo per tutto l'intervallo di interesse, ossia:

$$\mathbf{r}(t) = \mathbf{0} \quad \forall t \in [t_0, t_f] \quad (10.13)$$

#### 10.1.1 Scelta della funzione di costo

Quando si progetta un sistema di controllo, è fondamentale disporre di un criterio che permetta di valutare quanto bene il sistema stia funzionando. Questo criterio, chiamato *funzione di costo*, è un'espressione matematica che, quando viene minimizzata, indica che il sistema opera nel modo più desiderabile. In sostanza, definire una misura di prestazione significa tradurre in termini matematici gli obiettivi fisici e funzionali del sistema.

Per comprendere meglio il processo di scelta, si considerino due leggi di controllo ammissibili,  $\mathbf{u}_1(t)$  e  $\mathbf{u}_2(t)$ , che producono entrambe traiettorie di stato accettabili. L'obiettivo è determinare quale dei due controlli offra prestazioni migliori. A tal fine, si può applicare ciascun controllo al sistema e calcolare il corrispondente valore della misura di prestazione, indicato rispettivamente con  $J_1$  e  $J_2$ . Il confronto tra i due valori consente di stabilire quale controllo sia preferibile:

- se  $J_1 < J_2$ , il controllo  $\mathbf{u}_1$  è da considerarsi migliore;
- se  $J_1 > J_2$ , risulta preferibile  $\mathbf{u}_2$ ;
- se  $J_1 = J_2$ , i due controlli sono equivalenti in termini di prestazioni.

Un approccio alternativo consiste nel confrontare direttamente le traiettorie di stato generate da ciascun controllo, valutandole in modo qualitativo o soggettivo. In questo caso, si analizzano i risultati ottenuti e si decide quale traiettoria rappresenti meglio il comportamento desiderato del sistema. Se la misura di prestazione scelta riflette correttamente gli obiettivi del progetto, la traiettoria ritenuta "più soddisfacente" dal progettista dovrebbe corrispondere a quella con il valore di  $J$  più basso.

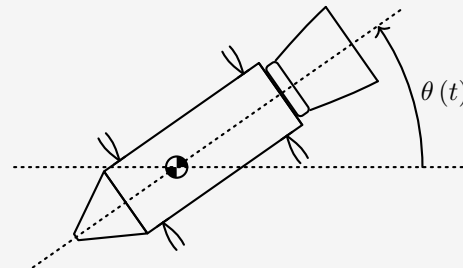
Se invece la valutazione soggettiva e il risultato numerico non coincidono, è segno che la funzione di costo adottata non rappresenta adeguatamente gli obiettivi desiderati. In tal caso, può essere necessario modificare la funzione di prestazione o rivedere i vincoli imposti al sistema, in modo che il criterio matematico esprima in modo più fedele le reali preferenze del progettista.

**Esempio 10.1.2** Si supponga di voler controllare una navicella spaziale che per semplicità modelleremo per controllare solo l'angolo di pitch  $\theta(t)$ .

L'equazione che descrive il moto è:

$$I\ddot{\theta}(t) = T(t)$$

dove  $I$  è il momento angolare d'inerzia e  $T(t)$  è la coppia prodotta dai getti di gas.



Se consideriamo  $x_1(t) = \theta(t)$  e  $x_2(t) = \dot{\theta}(t)$  come variabili di stato e  $u(t) = T(t)/I$  come ingresso, otteniamo il seguente sistema nello spazio di stato:

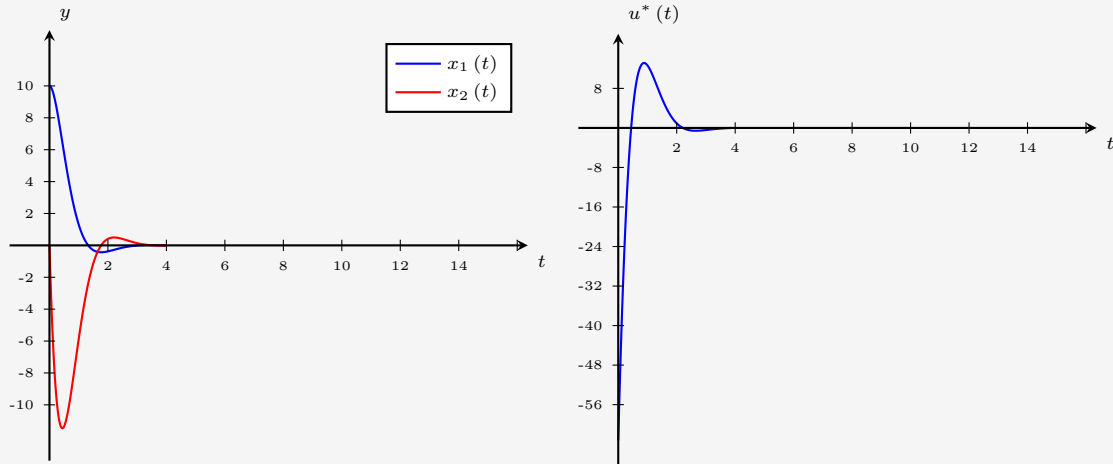
$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = u(t) \end{cases}$$

L'obiettivo del sistema di controllo è quello di mantenere quanto più possibile la posizione angolare a zero. Questo obiettivo deve essere raggiunto utilizzando piccole accelerazioni.

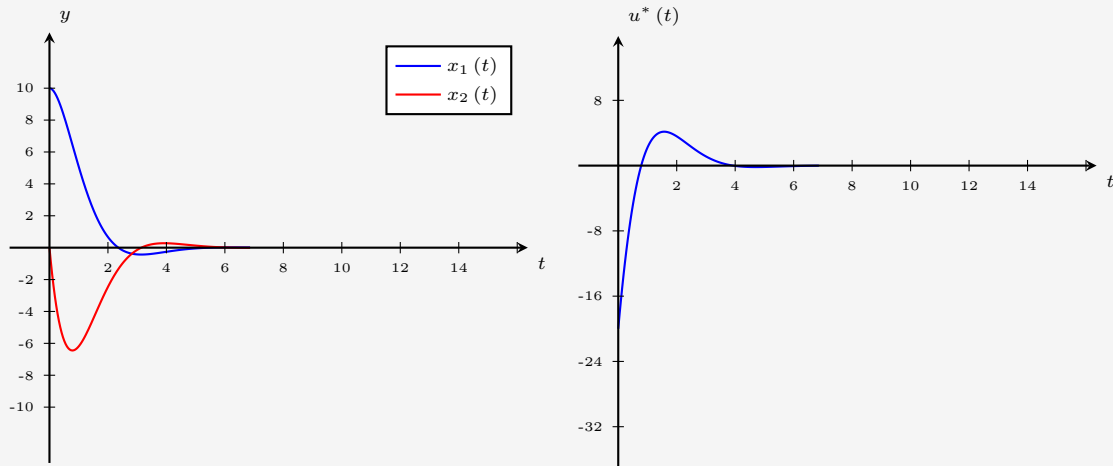
Dall'obiettivo, capiamo che si tratta di un problema di regolazione e controllo minimo. Scriviamo la funzione di costo:

$$J = \int_0^\infty [\mathbf{x}^T \mathbf{Q} \mathbf{x} + Ru] dt$$

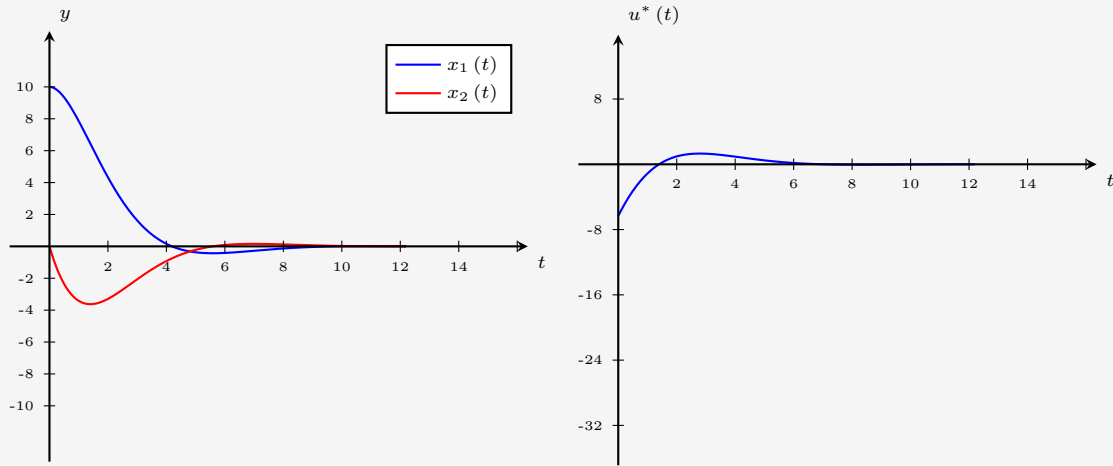
Nelle figure che seguono, vediamo l'effetto della variazione dei pesi sulle traiettorie ottimali:



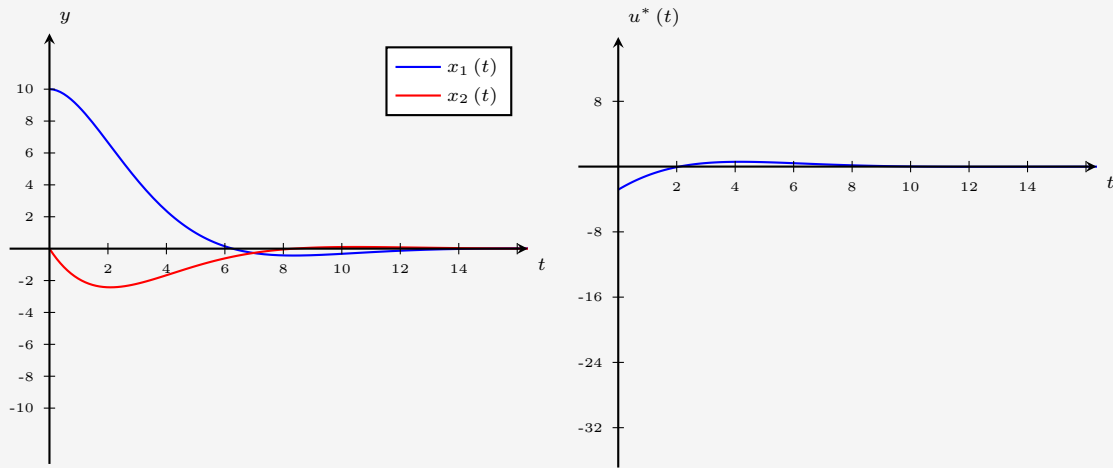
**Figura 10.1.** Andamento di uscita e ingresso con  $\mathbf{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $R = 0.1$ ,  $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$



**Figura 10.2.** Andamento di uscita e ingresso con  $\mathbf{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $R = 1$ ,  $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$



**Figura 10.3.** Andamento di uscita e ingresso con  $\mathbf{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $R = 10$ ,  $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$



**Figura 10.4.** Andamento di uscita e ingresso con  $\mathbf{Q} = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$ ,  $R = 50$ ,  $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 0 \end{bmatrix}$

Aumentare il valore di  $R$  comporta una penalizzazione maggiore sull'accelerazione e sull'impegno richiesto al controllore. Tutte queste traiettorie sono ottimali, ciascuna rispetto ad una diversa funzione di costo. Se il nostro obiettivo principale è ridurre rapidamente lo spostamento angolare a zero, allora la traiettoria mostrata in Figura 10.1 sarebbe la più indicata. Gli astronauti, tuttavia, probabilmente preferirebbero la traiettoria della Figura 10.4, poiché comporta accelerazioni molto più ridotte.

Dobbiamo essere molto cauti nell'interpretare il valore numerico della misura di prestazione minima. Infatti, se moltiplichiamo tutti i pesi presenti nella funzione di costo per una costante positiva  $K$ , il valore complessivo della misura sarà moltiplicato per  $K$ , ma il controllo ottimale e la traiettoria rimarranno esattamente gli stessi. In realtà, può essere persino possibile modificare i diversi fattori di ponderazione in misura diversa, mantenendo comunque lo stesso controllo e la stessa traiettoria ottimali.

Anche l'interpretazione fisica del valore della funzione di costo è un aspetto importante. Di solito, i valore presenti all'interno di una funzione di costo come il tempo impiegato o il carburante consumato hanno un chiaro significato fisico. Nei casi in cui la funzione di costo sia una combinazione ponderata di diverse grandezze fisiche (come in questo esempio) il valore numerico della funzione di costo non rappresenta una quantità fisicamente significativa.

## 10.2 Controllo ottimo senza vincoli

**Esempio 10.2.1** Esempio guidato: costo e controllo ottimo (orizzonte  $H_p = H_w = 2$ )

Si consideri il sistema discreto (SISO) con stato  $x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$ :

$$x_1(k+1) = u(k), \quad (10.14)$$

$$x_2(k+1) = x_1(k), \quad (10.15)$$

$$y(k) = x(k). \quad (10.16)$$

In forma matriciale:

$$x(k+1) = A x(k) + B u(k), \quad y(k) = C x(k) + D u(k), \quad (10.17)$$

con

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = I_2, \quad D = 0. \quad (10.18)$$

Il riferimento è costante

$$r(k) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \forall k. \quad (10.19)$$

*Funzione di costo.* La funzione di costo (tracking pesato) è

$$V(k) = \sum_{i=H_w}^{H_p} [r(k+1) - x(k+1)]^\top Q [r(k+1) - x(k+1)], \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad (10.20)$$

con  $H_p = 2$  e  $H_w = 2$ .

Poiché  $H_p = H_w$ , la sommatoria contiene *un solo termine*: il costo penalizza l'errore predetto al tempo  $k+1$ . Inoltre, in questo sistema  $x_1(k+1)$  dipende direttamente da  $u(k)$ , mentre  $x_2(k+1)$  dipende solo dallo stato corrente  $x_1(k)$ ; quindi una parte del costo risulta indipendente dalla variabile di decisione  $u(k)$ .

**Calcolare  $V(k)$**

Dalle dinamiche:

$$x(k+1) = \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} u(k) \\ x_1(k) \end{bmatrix}. \quad (10.21)$$

Con  $r(k+1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$  si ha

$$r(k+1) - x(k+1) = \begin{bmatrix} 1 - u(k) \\ 1 - x_1(k) \end{bmatrix}. \quad (10.22)$$

Quindi:

$$V(k) = \begin{bmatrix} 1 - u(k) \\ 1 - x_1(k) \end{bmatrix}^\top \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 - u(k) \\ 1 - x_1(k) \end{bmatrix} \quad (10.23)$$

$$= (1 - u(k))^2 + 2(1 - x_1(k))^2. \quad (10.24)$$

Il secondo addendo  $2(1 - x_1(k))^2$  non dipende da  $u(k)$ : è un termine costante rispetto alla minimizzazione in  $u(k)$ .

**Calcolare  $u(k)^*$** 

Per trovare  $u(k)^*$  si minimizza  $V(k)$  rispetto a  $u(k)$ :

$$\frac{\partial V(k)}{\partial u(k)} = \frac{\partial}{\partial u(k)} \left[ (1 - u(k))^2 + 2(1 - x_1(k))^2 \right] \quad (10.25)$$

$$= 2(1 - u(k))(-1) \quad (10.26)$$

$$= 2u(k) - 2. \quad (10.27)$$

Imponendo la condizione di ottimalità del primo ordine:

$$2u(k) - 2 = 0 \implies u(k)^* = 1. \quad (10.28)$$

Il controllo ottimo forza  $x_1(k+1) = u(k)$  a seguire il riferimento  $r_1 = 1$  al passo successivo. La componente  $x_2(k+1)$  non è influenzabile da  $u(k)$  in un solo passo, quindi non altera la scelta ottima.

**Calcolare  $u(k)^*$  quando  $r(k) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$** 

Se  $r(k+1) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ , allora:

$$r(k+1) - x(k+1) = \begin{bmatrix} 2 - u(k) \\ 2 - x_1(k) \end{bmatrix}, \quad (10.29)$$

e

$$V(k) = (2 - u(k))^2 + 2(2 - x_1(k))^2. \quad (10.30)$$

Derivando rispetto a  $u(k)$ :

$$\frac{\partial V(k)}{\partial u(k)} = 2(2 - u(k))(-1) = 2u(k) - 4. \quad (10.31)$$

Condizione del primo ordine:

$$2u(k) - 4 = 0 \implies u(k)^* = 2. \quad (10.32)$$

Con lo stesso ragionamento, il controllo ottimo impone  $x_1(k+1)$  uguale alla prima componente del riferimento (qui 2). Il peso maggiore sulla seconda componente (fattore 2) non modifica  $u^*$  in questo setting, perché quella componente non dipende da  $u(k)$  a un solo passo.

**Esempio 10.2.2** Esempio guidato: funzione di costo e controllo ottimo

Si consideri il seguente sistema discreto con stato  $x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$ :

$$x_1(k+1) = u(k) + x_2(k), \quad (10.33)$$

$$x_2(k+1) = x_1(k) - x_2(k), \quad (10.34)$$

$$y(k) = x(k). \quad (10.35)$$

In forma matriciale:

$$x(k+1) = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(k), \quad y(k) = x(k). \quad (10.36)$$

**Riferimento.** Il riferimento è costante:

$$r(k) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \forall k. \quad (10.37)$$

**Funzione di costo.** La funzione di costo di tipo tracking è definita come:

$$V(k) = \sum_{i=H_w}^{H_p} [r(k+i) - x(k+i)]^\top Q [r(k+i) - x(k+i)], \quad Q = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, \quad (10.38)$$

con orizzonti:

$$H_p = 2, \quad H_w = 1.$$

Poiché  $H_w = 1$  e  $H_p = 2$ , il costo include *due istanti di predizione*:  $k+1$  e  $k+2$ . A differenza di esempi a un solo passo, il controllo  $u(k)$  influenza entrambe le predizioni tramite la dinamica del sistema.

#### **Calcolare la funzione di costo $V(k)$**

Dalle equazioni di stato:

$$x(k+1) = \begin{bmatrix} u(k) + x_2(k) \\ x_1(k) - x_2(k) \end{bmatrix}. \quad (10.39)$$

Incrementando di un passo:

$$x_1(k+2) = u(k+1) + x_2(k+1), \quad (10.40)$$

$$x_2(k+2) = x_1(k+1) - x_2(k+1). \quad (10.41)$$

Sostituendo le espressioni di  $x(k+1)$ :

$$x_1(k+2) = u(k+1) - x_1(k) + x_2(k), \quad (10.42)$$

$$x_2(k+2) = u(k) - x_1(k) + 2x_2(k). \quad (10.43)$$

Il costo risulta quindi:

$$\begin{aligned} V(k) &= 2(1 - x_1(k+1))^2 + 3(1 - x_2(k+1))^2 \\ &\quad + 2(1 - x_1(k+2))^2 + 3(1 - x_2(k+2))^2, \end{aligned} \quad (10.44)$$

ossia, sostituendo:

$$\begin{aligned} V(k) &= 2(1 - u(k) - x_2(k))^2 + 3(1 - x_1(k) + x_2(k))^2 \\ &\quad + 2(1 - u(k+1) + x_1(k) - x_2(k))^2 + 3(1 - u(k) + x_1(k) - 2x_2(k))^2. \end{aligned} \quad (10.45)$$

È perfettamente lecito fermarsi a questa forma compatta. L'espansione completa dei quadrati serve solo se si vuole evidenziare esplicitamente la dipendenza quadratica in  $u(k)$ .

#### **Calcolare il controllo ottimo $u(k)$**

Considerando  $u(k+1)$  fissato (strategia MPC a un passo), si calcola la derivata parziale rispetto a  $u(k)$ :

$$\frac{\partial V(k)}{\partial u(k)} = -4(1 - u(k) - x_2(k)) - 6(1 - u(k) + x_1(k) - 2x_2(k)) \quad (10.46)$$

$$= -10 + 10u(k) - 6x_1(k) + 16x_2(k). \quad (10.47)$$



Imponendo la condizione di ottimalità del primo ordine:

$$\frac{\partial V(k)}{\partial u(k)} = 0 \implies u(k)^* = \frac{6x_1(k) - 16x_2(k) + 10}{10}. \quad (10.48)$$

Il controllo ottimo è una legge di feedback lineare nello stato più un termine costante legato al riferimento. Questo è tipico di problemi LQ di tracking senza vincoli.

**Calcolare il controllo ottimo**  $u(k+1)$

Il procedimento è del tutto analogo, derivando  $V(k)$  rispetto a  $u(k+1)$ :

$$\frac{\partial V(k)}{\partial u(k+1)} = -4(1 - u(k+1) + x_1(k) - x_2(k)). \quad (10.49)$$

Ponendo la derivata nulla:

$$-4 + 4u(k+1) + 4x_1(k) - 4x_2(k) = 0, \quad (10.50)$$

da cui:

$$u(k+1)^* = -x_1(k) + x_2(k) + 1. \quad (10.51)$$

In un contesto MPC reale,  $u(k+1)$  non verrebbe applicato ma solo calcolato come parte della sequenza ottima; al passo successivo il problema verrebbe risolto nuovamente con lo stato aggiornato (principio del *receding horizon*).

**Esempio 10.2.3** Esempio guidato: calcolo di  $V(k)$ ,  $u(k)$ ,  $u(k+1)$  e variante con costo su input

Si consideri il sistema discreto con stato  $x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$ :

$$x_1(k+1) = -x_1(k) + x_2(k), \quad (10.52)$$

$$x_2(k+1) = x_1(k) + u(k), \quad (10.53)$$

$$y(k) = x(k). \quad (10.54)$$

Il riferimento è costante:

$$r(k) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \forall k. \quad (10.55)$$

Con orizzonti  $H_p = 2$  e  $H_w = 1$ :

$$V(k) = \sum_{i=H_w}^{H_p} [r(k+i) - x(k+i)]^\top Q [r(k+i) - x(k+i)], \quad Q = \begin{bmatrix} 4 & 0 \\ 0 & -3 \end{bmatrix}. \quad (10.56)$$

Qui  $Q$  non è definita positiva (compare il peso  $-3$  sulla seconda componente): quindi  $V(k)$  non è, in generale, un' "energia" da minimizzare in senso standard LQ. Tuttavia, il calcolo analitico di punti stazionari (derivata nulla) resta fattibile e produce le espressioni richieste dall'esercizio.

**Calcolare**  $V(k)$

Dalle dinamiche (10.52)–(10.53) si ottiene:

$$x_1(k+1) = -x_1(k) + x_2(k), \quad (10.57)$$

$$x_2(k+1) = x_1(k) + u(k). \quad (10.58)$$

Per  $k + 2$ :

$$x_1(k+2) = -x_1(k+1) + x_2(k+1), \quad (10.59)$$

$$x_2(k+2) = x_1(k+1) + u(k+1). \quad (10.60)$$

Sostituendo le espressioni a  $k + 1$ :

$$x_1(k+2) = -(-x_1(k) + x_2(k)) + (x_1(k) + u(k)) = 2x_1(k) - x_2(k) + u(k), \quad (10.61)$$

$$x_2(k+2) = (-x_1(k) + x_2(k)) + u(k+1) = -x_1(k) + x_2(k) + u(k+1). \quad (10.62)$$

Poiché  $r(k+i) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ , il costo (10.56) diventa:

$$\begin{aligned} V(k) &= 4(1 - x_1(k+1))^2 - 3(1 - x_2(k+1))^2 \\ &\quad + 4(1 - x_1(k+2))^2 - 3(1 - x_2(k+2))^2. \end{aligned} \quad (10.63)$$

Sostituendo  $x_1(k+1), x_2(k+1), x_1(k+2), x_2(k+2)$ :

$$\begin{aligned} V(k) &= 4(1 + x_1(k) - x_2(k))^2 - 3(1 - x_1(k) - u(k))^2 \\ &\quad + 4(1 - 2x_1(k) + x_2(k) - u(k))^2 - 3(1 + x_1(k) - x_2(k) - u(k+1))^2. \end{aligned} \quad (10.64)$$

È consigliabile lasciare  $V(k)$  in forma “quadratica” come in (10.64): l’espansione completa è lunga e non aggiunge informazione strutturale.

#### **Calcolare $u(k)$ (caso base)**

Per trovare il valore stazionario rispetto a  $u(k)$ , deriviamo (10.64):

$$\frac{\partial V(k)}{\partial u(k)} = -3 \cdot 2(1 - x_1(k) - u(k)) \cdot (-1) + 4 \cdot 2(1 - 2x_1(k) + x_2(k) - u(k)) \cdot (-1) \quad (10.65)$$

$$= 6(1 - x_1(k) - u(k)) - 8(1 - 2x_1(k) + x_2(k) - u(k)) \quad (10.66)$$

$$= -2 + 10x_1(k) - 8x_2(k) + 2u(k). \quad (10.67)$$

Imponendo  $\frac{\partial V}{\partial u(k)} = 0$ :

$$-2 + 10x_1(k) - 8x_2(k) + 2u(k) = 0 \implies u(k) = -5x_1(k) + 4x_2(k) + 1. \quad (10.68)$$

#### **Calcolare $u(k+1)$ (caso base)**

Derivando (10.64) rispetto a  $u(k+1)$  (compare solo nell’ultimo termine):

$$\frac{\partial V(k)}{\partial u(k+1)} = -3 \cdot 2(1 + x_1(k) - x_2(k) - u(k+1)) \cdot (-1) \quad (10.69)$$

$$= 6(1 + x_1(k) - x_2(k) - u(k+1)). \quad (10.70)$$

Ponendo la derivata uguale a zero:

$$6(1 + x_1(k) - x_2(k) - u(k+1)) = 0 \implies u(k+1) = x_1(k) - x_2(k) + 1. \quad (10.71)$$

In logica MPC (receding horizon) si applica solo  $u(k)$ ;  $u(k+1)$  è parte della sequenza predetta e verrà ricalcolato al passo successivo usando lo stato aggiornato.

**Aggiunta di un termine su input e calcolo di  $V(k)$ ,  $u(k)$ ,  $u(k+1)$** 

Ora si considera la funzione di costo modificata:

$$V_1(k) = \sum_{i=H_w}^{H_p} [r(k+i) - x(k+i)]^\top Q [r(k+i) - x(k+i)] + \sum_{i=0}^{H_c-1} 2u(k+i)^2, \quad (10.72)$$

con  $H_c = 1$ . Poiché  $H_c = 1$ , il secondo termine vale semplicemente  $2u(k)^2$ , quindi:

$$V_1(k) = V(k) + 2u(k)^2. \quad (10.73)$$

*Calcolo di  $u(k)$  (caso con costo su input).* Derivando (10.73) rispetto a  $u(k)$ :

$$\frac{\partial V_1(k)}{\partial u(k)} = \frac{\partial V(k)}{\partial u(k)} + 4u(k). \quad (10.74)$$

Usando  $\frac{\partial V}{\partial u(k)} = -2 + 10x_1(k) - 8x_2(k) + 2u(k)$ :

$$\frac{\partial V_1(k)}{\partial u(k)} = -2 + 10x_1(k) - 8x_2(k) + 6u(k). \quad (10.75)$$

Imponendo la condizione di stazionarietà:

$$-2 + 10x_1(k) - 8x_2(k) + 6u(k) = 0 \implies u(k) = \frac{-10x_1(k) + 8x_2(k) + 2}{6} = \frac{-5x_1(k) + 4x_2(k) + 1}{3}. \quad (10.76)$$

*Calcolo di  $u(k+1)$  (caso con costo su input).* Poiché il termine aggiuntivo è solo  $2u(k)^2$  (con  $H_c = 1$ ), non dipende da  $u(k+1)$ :

$$\frac{\partial V_1(k)}{\partial u(k+1)} = \frac{\partial V(k)}{\partial u(k+1)}. \quad (10.77)$$

Quindi  $u(k+1)$  resta invariato rispetto al caso base:

$$u(k+1) = x_1(k) - x_2(k) + 1. \quad (10.78)$$

Aggiungere un termine  $\rho u^2$  (qui  $\rho = 2$ ) “regolarizza” l’azione di controllo, riducendo l’ampiezza di  $u(k)$  (si passa da (10.68) a (10.76)). Con  $H_c = 1$  la penalizzazione agisce solo sul primo controllo, quindi non altera  $u(k+1)$ .

## 10.3 Metodo dei Moltiplicatori di Lagrange

Il metodo dei moltiplicatori di Lagrange è una tecnica che consente di determinare i punti stazionari (massimi e minimi) di una funzione  $f(\mathbf{x})$  soggetta a uno o più vincoli nella forma  $g(\mathbf{x}) = 0$ .

Invece che considerare la funzione e i vincoli separatamente, si considera una funzione aumentata detta Lagrangiana:

$$\mathcal{L}(\mathbf{x}) = f(\mathbf{x}) + \lambda g(\mathbf{x}) \quad (10.79)$$

dove  $\lambda$  è un vettore di scalari detti **moltiplicatori di Lagrange**.

Se  $\mathbf{x}^*$  è stazionario, allora esiste un  $\lambda^*$  tale che  $(\mathbf{x}^*, \lambda^*)$  è stazionario.

### 10.3.1 Limitazioni

Non tutti i punti stazionari portano ad una soluzione del problema originario, quindi il metodo dei moltiplicatori di Lagrange fornisce una **condizione necessaria ma non sufficiente** per l’ottimizzazione dei problemi vincolati.

- **Condizione necessaria ma non sufficiente.** Il fatto che valga

$$\nabla f = \lambda \nabla g \quad \text{e} \quad g = 0 \quad (10.80)$$

non garantisce che il punto trovato sia effettivamente un massimo o un minimo vincolato: potrebbe trattarsi di un *punto di sella* o di un estremo locale non rilevante.

- **Regolarità del vincolo.** Il metodo richiede che il gradiente del vincolo (o dei vincoli, nel caso multiplo) sia non nullo nel punto considerato, ovvero che  $\nabla g(x^*) \neq 0$ . In caso contrario, la condizione di Lagrange non è applicabile. Questa ipotesi è nota come *condizione di qualificazione del vincolo*.
- **Vincoli di disuguaglianza.** Il metodo classico si applica solo a vincoli di uguaglianza  $g(x) = 0$ . Se sono presenti vincoli di disuguaglianza  $h(x) \leq 0$ , occorre ricorrere a una generalizzazione: le *condizioni di Karush–Kuhn–Tucker (KKT)*.
- **Globalità dell'ottimo.** Anche quando le condizioni di Lagrange sono soddisfatte, il punto trovato è solo un candidato a estremo *locale*. Per verificare che si tratti di un massimo o minimo *globale*, è necessario confrontare i valori della funzione nei punti candidati oppure analizzare la natura (convessa o concava) della funzione.
- **Differenziabilità.** Il metodo presuppone che le funzioni coinvolte,  $f$  e  $g$ , siano almeno una volta derivabili (classe  $C^1$ ). In presenza di punti angolosi, cuspidi o discontinuità, il metodo non è direttamente applicabile.

**Esempio 10.3.1** Consideriamo il problema dell'esempio precedente per trovare il punto sulla retta  $x_1 + x_2 = 5$  che sia più vicino all'origine.

Scrivendo matematicamente il nostro obiettivo, ossia quello trovare il punto  $(x_1^*, x_2^*)$  che minimizzi una distanza, possiamo scrivere:

$$J(x_1, x_2) = f(x_1, x_2) = \sqrt{x_1^2 + x_2^2} = x_1^2 + x_2^2$$

$$g(x_1, x_2) = x_1 + x_2 - 5$$

Scriviamo la nostra funzione Lagrangiana:

$$\begin{aligned} \mathcal{L}(x_1, x_2) &= f(x_1, x_2) + \lambda g(x_1, x_2) \\ &= x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 5) \end{aligned}$$

Cerchiamo i punti stazionari  $(x_1^*, x_2^*)$  calcolando il gradiente di  $\mathcal{L}$ :

$$\nabla \mathcal{L} = (2x_1^* + \lambda) \Delta x_1 + (2x_2^* + \lambda) \Delta x_2 + (x_1^* + x_2^* - 5) \Delta \lambda = 0$$

Questo risolvere il seguente sistema a 3 equazioni e 3 incognite  $x_1^*, x_2^*, \lambda$ :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x_1} = 2x_1^* + \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} = 2x_2^* + \lambda = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} = x_1^* + x_2^* - 5 = 0 \end{cases}$$

Il sistema può essere risolto con uno qualsiasi dei metodi diretti che conosciamo (es. per sostituzione, eliminazione di Gauss, Cramer o Gauss-Jordan).

Utilizzo il metodo della sostituzione isolando  $\lambda$  dalla prima e seconda equazione ed eguagliando le due equazioni ottenute:

$$\begin{cases} \lambda = -2x_1^* \\ \lambda = -2x_2^* \end{cases} \Rightarrow -2x_1^* = -2x_2^* \Rightarrow x_1^* = x_2^*$$

Sostituisco quanto trovato nella terza equazione:

$$x_1^* + x_1^* = 5 \quad \Rightarrow \quad 2x_1^* = 5 \quad \Rightarrow \quad x_1^* = \frac{5}{2} = 2.5$$

Quindi  $x_2^* = 2.5$ . Cerco  $\lambda$  usando una delle equazioni ricavate in precedenza:

$$\lambda = -2x_1^* = -2 \cdot \frac{5}{2} = -5$$

Quindi il punto di minimo del nostro problema è  $\mathbf{x}^* = (x_1^*, x_2^*) = (2.5, 2.5)$ ,  $\lambda = -5$ .

**Esempio 10.3.2** Trovare il punto in  $\mathbb{R}^3$  che sia quanto più vicino all'origine e soddisfi i seguenti vincoli:

$$g(x_1, x_2, x_3) = \begin{cases} x_3 = x_1 x_2 + 5 \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

La funzione da minimizzare è:

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

La Lagrangiana è:

$$\mathcal{L}(x_1, x_2, x_3, \lambda_1, \lambda_2) = x_1^2 + x_2^2 + x_3^2 + \lambda_1 (x_1 x_2 - x_3 + 5) + \lambda_2 (x_1 + x_2 + x_3 - 1)$$

Il gradiente della Lagrangiana è:

$$\begin{aligned} \nabla \mathcal{L} &= (2x_1 + \lambda_1 x_2 + \lambda_2) \Delta x_1 \\ &+ (2x_2 + \lambda_1 x_1 + \lambda_2) \Delta x_2 \\ &+ (2x_3 - \lambda_1 + \lambda_2) \Delta x_3 \\ &+ (x_1 x_2 - x_3 + 5) \Delta \lambda_1 \\ &+ (x_1 + x_2 + x_3 - 1) \Delta \lambda_2 \end{aligned}$$

Imponendo  $\nabla \mathcal{L} = 0$  per la ricerca dei punti stazionari, otteniamo il sistema:

$$\begin{cases} 2x_1^* + \lambda_1^* x_2^* + \lambda_2^* = 0 \\ 2x_2^* + \lambda_1^* x_1^* + \lambda_2^* = 0 \\ 2x_3^* - \lambda_1^* + \lambda_2^* = 0 \\ x_1^* x_2^* - x_3^* + 5 = 0 \\ x_1^* + x_2^* + x_3^* - 1 = 0 \end{cases}$$

Abbiamo sempre un sistema a 5 equazioni e 5 incognite. Risolvendolo per sostituzione:

$$\begin{aligned} 2x_1^* + \lambda_1^* x_2^* + \lambda_2^* - 2x_2^* + \lambda_1^* x_1^* + \lambda_2^* &= 0 \\ (2x_1^* - 2x_2^*) + \lambda_1^* x_2^* + \lambda_2^* + \lambda_1^* x_1^* + \lambda_2^* &= 0 \end{aligned}$$

Risolvendolo, troviamo due punti stazionari che soddisfano i vincoli:

$$(x_1^*, x_2^*, x_3^*) = \begin{cases} (2, -2, 1) \\ o \\ (-2, 2, 1) \end{cases}$$

con  $\lambda_1^* = 2, \lambda_2^* = 0$ .

Inoltre,  $f_{min} = 9$  quindi la distanza minima è 3.

## 10.4 Condizioni di Karush–Kuhn–Tucker

Il metodo dei moltiplicatori di Lagrange consente di trattare problemi di ottimizzazione vincolata in presenza di vincoli di uguaglianza. Tuttavia, in molti problemi di interesse pratico, e in particolare nel controllo ottimo e nel Model Predictive Control, i vincoli sono spesso espressi sotto forma di *disugua-*

*glanze*. In questi casi è necessario ricorrere a una generalizzazione del metodo di Lagrange, nota come *condizioni di Karush–Kuhn–Tucker* (KKT).

Consideriamo il seguente problema di ottimizzazione vincolata:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_j(x) = 0, \quad j = 1, \dots, p \end{aligned} \quad (10.81)$$

dove  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  è la funzione obiettivo,  $g_i$  sono i vincoli di disuguaglianza e  $h_j$  sono i vincoli di uguaglianza.

Analogamente al caso dei moltiplicatori di Lagrange, introduciamo la *Lagrangiana estesa*:

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{j=1}^p \lambda_j h_j(x) + \sum_{i=1}^m \mu_i g_i(x) \quad (10.82)$$

dove:

- $\lambda_j \in \mathbb{R}$  sono i moltiplicatori associati ai vincoli di uguaglianza;
- $\mu_i \geq 0$  sono i moltiplicatori associati ai vincoli di disuguaglianza.

Le condizioni di Karush–Kuhn–Tucker forniscono un insieme di condizioni *necessarie* per l'ottimalità (di tipo primo ordine) che generalizzano le condizioni di Lagrange. Un punto  $x^*$  è candidato a soluzione ottima se esistono moltiplicatori  $\lambda^*$  e  $\mu^*$  tali che siano soddisfatte le seguenti condizioni.

1. *Stazionarietà* Il gradiente della Lagrangiana rispetto alle variabili decisionali deve annullarsi:

$$\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = \nabla f(x^*) + \sum_{j=1}^p \lambda_j^* \nabla h_j(x^*) + \sum_{i=1}^m \mu_i^* \nabla g_i(x^*) = 0 \quad (10.83)$$

2. *Fattibilità primale* Il punto candidato deve soddisfare tutti i vincoli del problema:

$$\begin{aligned} g_i(x^*) &\leq 0, \quad i = 1, \dots, m \\ h_j(x^*) &= 0, \quad j = 1, \dots, p \end{aligned} \quad (10.84)$$

3. *Fattibilità duale* I moltiplicatori associati ai vincoli di disuguaglianza devono essere non negativi:

$$\mu_i^* \geq 0, \quad i = 1, \dots, m \quad (10.85)$$

4. *Complementarità* Per ciascun vincolo di disuguaglianza deve valere la condizione di complementarità:

$$\mu_i^* g_i(x^*) = 0, \quad i = 1, \dots, m \quad (10.86)$$

La condizione di complementarità ha una chiara interpretazione fisica:

- se un vincolo è *attivo* ( $g_i(x^*) = 0$ ), il corrispondente moltiplicatore  $\mu_i^*$  può essere strettamente positivo;
- se un vincolo è *inattivo* ( $g_i(x^*) < 0$ ), il moltiplicatore associato deve essere nullo.

#### Osservazioni

- In assenza di vincoli di disuguaglianza, le condizioni KKT si riducono esattamente alle condizioni dei moltiplicatori di Lagrange.
- Le condizioni KKT sono condizioni necessarie per l'ottimalità locale; diventano anche sufficienti se il problema è convesso, ossia se  $f$  è convessa e l'insieme dei vincoli è convesso.
- Le KKT giocano un ruolo fondamentale nella formulazione e risoluzione numerica dei problemi di controllo ottimo, in quanto le condizioni di ottimalità dei problemi MPC sono spesso espresse proprio in forma KKT.

**Esempio 10.4.1** Consideriamo il seguente problema di ottimizzazione vincolata:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x_1, x_2) = x_1^2 + x_2^2 \\ \text{subject to} \quad & g(x_1, x_2) = x_1 + x_2 - 1 \leq 0 \end{aligned} \quad (10.87)$$

Dal punto di vista geometrico, il problema consiste nel trovare il punto del semipiano  $x_1 + x_2 \leq 1$  che sia più vicino all'origine.

Introduciamo la Lagrangiana associata al problema:

$$\mathcal{L}(x_1, x_2, \mu) = x_1^2 + x_2^2 + \mu(x_1 + x_2 - 1) \quad (10.88)$$

dove  $\mu \geq 0$  è il moltiplicatore associato al vincolo di disuguaglianza.

*Condizione di stazionarietà* Calcoliamo il gradiente della Lagrangiana rispetto alle variabili decisionali:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 2x_1 + \mu = 0 \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 + \mu = 0 \end{aligned} \quad (10.89)$$

*Fattibilità primale* Il punto candidato deve soddisfare il vincolo:

$$x_1 + x_2 - 1 \leq 0 \quad (10.90)$$

*Fattibilità duale* Il moltiplicatore deve essere non negativo:

$$\mu \geq 0 \quad (10.91)$$

*Complementarità* Deve inoltre valere:

$$\mu(x_1 + x_2 - 1) = 0 \quad (10.92)$$

*Analisi dei casi* Dalla condizione di complementarità emergono due possibili situazioni.

*Caso 1: vincolo inattivo* Se  $x_1 + x_2 - 1 < 0$ , allora necessariamente  $\mu = 0$ . In questo caso, dalle condizioni di stazionarietà otteniamo:

$$2x_1 = 0, \quad 2x_2 = 0$$

da cui:

$$x_1^* = 0, \quad x_2^* = 0$$

Verifichiamo la fattibilità:

$$0 + 0 - 1 = -1 \leq 0$$

Il punto è ammissibile e quindi candidato ottimo.

*Caso 2: vincolo attivo* Se il vincolo è attivo, allora:

$$x_1 + x_2 - 1 = 0$$

con  $\mu \geq 0$ .

Dalle condizioni di stazionarietà si ottiene:

$$x_1 = x_2 = -\frac{\mu}{2}$$

Sostituendo nel vincolo attivo:

$$-\frac{\mu}{2} - \frac{\mu}{2} = 1 \Rightarrow -\mu = 1 \Rightarrow \mu = -1$$

Poiché  $\mu < 0$ , questa soluzione viola la condizione di fattibilità duale e deve essere scartata.

*Soluzione ottima* L'unico punto che soddisfa tutte le condizioni KKT è:

$$x^* = (0, 0), \quad \mu^* = 0$$

Il valore minimo della funzione obiettivo è:

$$f(x^*) = 0$$

Questo esempio evidenzia come le condizioni KKT permettano di gestire correttamente i vincoli di disuguaglianza e come il concetto di vincolo attivo o inattivo emerga in modo naturale attraverso la condizione di complementarità.

**Esempio 10.4.2** Consideriamo ora un esempio in cui il vincolo di disuguaglianza risulta effettivamente *attivo* nel punto ottimo. Vogliamo minimizzare la distanza dall'origine imponendo che il punto appartenga a un semipiano che *non* contiene l'origine:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x_1, x_2) = x_1^2 + x_2^2 \\ \text{subject to} \quad & g(x_1, x_2) = 1 - x_1 - x_2 \leq 0 \end{aligned} \tag{10.93}$$

Il vincolo  $1 - x_1 - x_2 \leq 0$  è equivalente a  $x_1 + x_2 \geq 1$ . Geometricamente stiamo cercando il punto del semipiano  $x_1 + x_2 \geq 1$  più vicino all'origine: in questo caso è intuitivo aspettarsi che la soluzione giaccia sul bordo  $x_1 + x_2 = 1$ , cioè con vincolo attivo.

Introduciamo la Lagrangiana:

$$\mathcal{L}(x_1, x_2, \mu) = x_1^2 + x_2^2 + \mu(1 - x_1 - x_2), \tag{10.94}$$

con  $\mu \geq 0$ .

*Condizioni KKT* (i) **Stazionarietà.** Calcoliamo le derivate parziali:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 2x_1 - \mu = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 - \mu = 0. \end{aligned} \tag{10.95}$$

Da cui otteniamo:

$$x_1 = \frac{\mu}{2}, \quad x_2 = \frac{\mu}{2}. \tag{10.96}$$

(ii) **Fattibilità primale.**

$$1 - x_1 - x_2 \leq 0. \tag{10.97}$$

(iii) **Fattibilità duale.**

$$\mu \geq 0. \tag{10.98}$$



**(iv) Complementarità.**

$$\mu(1 - x_1 - x_2) = 0. \quad (10.99)$$

*Analisi dei casi* Anche qui, la complementarità ci impone di distinguere tra vincolo inattivo e vincolo attivo.

*Caso 1: vincolo inattivo* Se  $1 - x_1 - x_2 < 0$ , allora  $\mu = 0$ . Dalla stazionarietà segue  $x_1 = x_2 = 0$ . Tuttavia, tale punto non è ammissibile, perché:

$$1 - 0 - 0 = 1 \not\leq 0.$$

Quindi il caso di vincolo inattivo non può fornire la soluzione.

*Caso 2: vincolo attivo* Deve allora valere:

$$1 - x_1 - x_2 = 0 \quad \Rightarrow \quad x_1 + x_2 = 1.$$

Sostituendo (10.96):

$$\frac{\mu}{2} + \frac{\mu}{2} = 1 \quad \Rightarrow \quad \mu = 1.$$

Quindi:

$$x_1^* = \frac{1}{2}, \quad x_2^* = \frac{1}{2}, \quad \mu^* = 1.$$

Verifichiamo la fattibilità duale:  $\mu^* = 1 \geq 0$  (ok). La complementarità è soddisfatta perché il vincolo è attivo ( $1 - x_1^* - x_2^* = 0$ ).

*Soluzione ottima* La soluzione del problema è:

$$x^* = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \mu^* = 1.$$

Il valore minimo della funzione obiettivo è:

$$f(x^*) = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2}.$$

Questo esempio mette in evidenza il significato operativo delle KKT: poiché l'ottimo non può cadere all'interno della regione ammissibile senza violare il vincolo, la soluzione si colloca sul bordo  $x_1 + x_2 = 1$  e il corrispondente moltiplicatore risulta strettamente positivo ( $\mu^* > 0$ ), segnalando che il vincolo è effettivamente *attivo* all'ottimo.

**Esempio 10.4.3** Consideriamo il seguente problema di ottimizzazione vincolata:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x_1, x_2) = x_1^2 + x_2^2 \\ \text{subject to} \quad & h(x_1, x_2) = x_1 + x_2 - 1 = 0, \\ & g(x_1, x_2) = x_1 \geq 0. \end{aligned} \quad (10.100)$$

Il problema consiste nel trovare il punto della retta  $x_1 + x_2 = 1$  che sia più vicino all'origine, imponendo inoltre che la prima coordinata sia non negativa.

*Lagrangiana* Il vincolo di disuguaglianza viene riscritto nella forma standard richiesta dalle condizioni KKT:

$$g(x_1, x_2) = -x_1 \leq 0.$$

Introduciamo quindi la Lagrangiana:

$$\mathcal{L}(x_1, x_2, \lambda, \mu) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1) + \mu(-x_1), \quad (10.101)$$

con  $\mu \geq 0$ .

**Condizioni KKT (i) Stazionarietà.**

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x_1} &= 2x_1 + \lambda - \mu = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 + \lambda = 0.\end{aligned}\tag{10.102}$$

**(ii) Fattibilità primale.**

$$x_1 + x_2 - 1 = 0, \quad -x_1 \leq 0.\tag{10.103}$$

**(iii) Fattibilità duale.**

$$\mu \geq 0.\tag{10.104}$$

**(iv) Complementarità.**

$$\mu(-x_1) = 0.\tag{10.105}$$

**Analisi dei casi** Dalla condizione di complementarità emergono due possibili situazioni.

**Caso 1: vincolo di disuguaglianza inattivo** Se  $x_1 > 0$ , allora  $-x_1 < 0$  e quindi  $\mu = 0$ . Il sistema di stazionarietà diventa:

$$2x_1 + \lambda = 0, \quad 2x_2 + \lambda = 0.$$

Da cui segue:

$$x_1 = x_2.$$

Imponendo il vincolo di uguaglianza:

$$x_1 + x_2 = 1 \quad \Rightarrow \quad x_1 = x_2 = \frac{1}{2}.$$

Poiché  $x_1 = \frac{1}{2} > 0$ , il vincolo di disuguaglianza è effettivamente inattivo e il punto è ammissibile.

**Caso 2: vincolo di disuguaglianza attivo** Se il vincolo è attivo, allora:

$$x_1 = 0.$$

Dal vincolo di uguaglianza segue:

$$x_2 = 1.$$

La stazionarietà fornisce:

$$2 \cdot 0 + \lambda - \mu = 0, \quad 2 \cdot 1 + \lambda = 0.$$

Da cui:

$$\lambda = -2, \quad \mu = -2.$$

Poiché  $\mu < 0$ , la fattibilità duale è violata e questo caso deve essere scartato.

**Soluzione ottima** L'unica soluzione che soddisfa tutte le condizioni KKT è:

$$x^* = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \lambda^* = -1, \quad \mu^* = 0.$$

Il valore minimo della funzione obiettivo è:

$$f(x^*) = \frac{1}{2}.$$

In questo esempio il vincolo di uguaglianza è sempre attivo per definizione, mentre il vincolo di disuguaglianza risulta inattivo all'ottimo.

**Esempio 10.4.4** Vogliamo risolvere il problema

$$\begin{aligned} \max_{x,y} \quad & f(x, y) = x \\ \text{subject to} \quad & y \leq (1-x)^3. \end{aligned} \quad (10.106)$$

Per applicare le KKT conviene riscrivere il vincolo nella forma standard  $g(x, y) \leq 0$ :

$$g(x, y) = y - (1-x)^3 \leq 0.$$

La Lagrangiana (per un problema di massimo) si scrive come

$$\mathcal{L}(x, y, \mu) = x - \mu(y - (1-x)^3), \quad (10.107)$$

dove  $\mu \geq 0$  è il moltiplicatore associato al vincolo di disuguaglianza.

**Condizioni KKT (i) Stazionarietà.** Impongo l'annullamento del gradiente rispetto a  $x$  e  $y$ :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= 1 - \mu \frac{\partial}{\partial x} (-(1-x)^3) = 1 - \mu (-3(1-x)^2) = 1 + 3\mu(1-x)^2 = 0, \\ \frac{\partial \mathcal{L}}{\partial y} &= -\mu = 0. \end{aligned} \quad (10.108)$$

**(ii) Fattibilità primale.**

$$y - (1-x)^3 \leq 0. \quad (10.109)$$

**(iii) Fattibilità duale.**

$$\mu \geq 0. \quad (10.110)$$

**(iv) Complementarità.**

$$\mu(y - (1-x)^3) = 0. \quad (10.111)$$

**Discussione** Dalla seconda equazione di stazionarietà in (10.108) si ottiene immediatamente

$$-\mu = 0 \quad \Rightarrow \quad \mu = 0.$$

Sostituendo  $\mu = 0$  nella prima equazione di stazionarietà si ottiene

$$1 + 3\mu(1-x)^2 = 1 = 0,$$

che è impossibile. Questo significa che *non esistono punti che soddisfano le condizioni KKT*.

L'assenza di punti KKT, in questo caso, non è un “fallimento” del metodo: indica che il problema (10.106) *non ammette un massimo* (è non limitato superiormente). Infatti il vincolo impone solo un limite superiore a  $y$ , ma non pone alcun limite su  $x$ . Per ogni valore arbitrariamente grande di  $x$  si può scegliere, ad esempio,

$$y = (1-x)^3$$

che soddisfa il vincolo con uguaglianza, ottenendo un valore dell'obiettivo pari a  $f = x$  sempre più grande. Quindi:

$$\sup f = +\infty \quad \text{e non esiste un massimo finito.}$$

**NB (vincolo aggiuntivo spesso usato)** Se si aggiunge anche il vincolo  $y \geq 0$ , cioè  $-y \leq 0$ , il problema diventa

$$\max x \quad \text{s.t.} \quad 0 \leq y \leq (1-x)^3.$$

In tal caso l'ammissibilità richiede  $(1-x)^3 \geq 0 \Rightarrow x \leq 1$ , e il massimo esiste ed è  $x^* = 1$  (con  $y^* = 0$ ). In questo caso le KKT producono effettivamente una soluzione.

**Esempio 10.4.5** Costruiamo un esempio in continuità con i precedenti: minimizziamo la distanza dall'origine, imponendo due vincoli di uguaglianza (che definiscono un insieme affine) e due vincoli

di disuguaglianza (che selezionano una porzione ammissibile di tale insieme). Consideriamo il problema:

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & f(x) = x_1^2 + x_2^2 + x_3^2 \\ \text{subject to} \quad & h_1(x) = x_1 + x_2 - 1 = 0, \\ & h_2(x) = x_3 - 1 = 0, \\ & g_1(x) = -x_1 \leq 0 \quad (\Leftrightarrow x_1 \geq 0), \\ & g_2(x) = -x_2 \leq 0 \quad (\Leftrightarrow x_2 \geq 0). \end{aligned} \quad (10.112)$$

I due vincoli di uguaglianza impongono che  $(x_1, x_2)$  appartenga alla retta  $x_1 + x_2 = 1$  e che  $x_3$  sia fissato a 1. I due vincoli di disuguaglianza richiedono inoltre che  $x_1$  e  $x_2$  siano non negativi.

**Lagrangiana** Introduciamo i moltiplicatori  $\lambda_1, \lambda_2 \in \mathbb{R}$  per i vincoli di uguaglianza e  $\mu_1, \mu_2 \geq 0$  per i vincoli di disuguaglianza. La Lagrangiana è:

$$\mathcal{L}(x, \lambda, \mu) = x_1^2 + x_2^2 + x_3^2 + \lambda_1(x_1 + x_2 - 1) + \lambda_2(x_3 - 1) + \mu_1(-x_1) + \mu_2(-x_2). \quad (10.113)$$

**Condizioni KKT (i) Stazionarietà.** Impongo  $\nabla_x \mathcal{L} = 0$ :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 2x_1 + \lambda_1 - \mu_1 = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 + \lambda_1 - \mu_2 = 0, \\ \frac{\partial \mathcal{L}}{\partial x_3} &= 2x_3 + \lambda_2 = 0. \end{aligned} \quad (10.114)$$

**(ii) Fattibilità primale.**

$$x_1 + x_2 - 1 = 0, \quad x_3 - 1 = 0, \quad -x_1 \leq 0, \quad -x_2 \leq 0. \quad (10.115)$$

**(iii) Fattibilità duale.**

$$\mu_1 \geq 0, \quad \mu_2 \geq 0. \quad (10.116)$$

**(iv) Complementarità.**

$$\mu_1(-x_1) = 0, \quad \mu_2(-x_2) = 0. \quad (10.117)$$

**Risoluzione** Dai vincoli di uguaglianza in (10.115) otteniamo subito:

$$x_3^* = 1, \quad x_2^* = 1 - x_1^*.$$

Sostituendo  $x_3^* = 1$  nella terza equazione di stazionarietà in (10.114):

$$2 \cdot 1 + \lambda_2^* = 0 \quad \Rightarrow \quad \lambda_2^* = -2.$$

Per quanto riguarda  $x_1$  e  $x_2$ , notiamo che, se all'ottimo valgono  $x_1^* > 0$  e  $x_2^* > 0$ , allora entrambi i vincoli di disuguaglianza sono inattivi e, per complementarità, deve valere:

$$\mu_1^* = \mu_2^* = 0.$$

In tal caso, le prime due equazioni di stazionarietà diventano:

$$2x_1^* + \lambda_1^* = 0, \quad 2x_2^* + \lambda_1^* = 0 \quad \Rightarrow \quad x_1^* = x_2^*.$$

Imponendo il vincolo di uguaglianza  $x_1^* + x_2^* = 1$  otteniamo:

$$x_1^* = x_2^* = \frac{1}{2}.$$

Verifichiamo la fattibilità dei vincoli di disuguaglianza:  $x_1^* = \frac{1}{2} \geq 0$ ,  $x_2^* = \frac{1}{2} \geq 0$ , quindi l'ipotesi di vincoli inattivi è coerente.

A questo punto si ricava  $\lambda_1^*$  da una delle condizioni di stazionarietà:

$$2 \cdot \frac{1}{2} + \lambda_1^* = 0 \quad \Rightarrow \quad \lambda_1^* = -1.$$

**Soluzione ottima** La soluzione che soddisfa tutte le condizioni KKT è quindi:

$$x^* = \left( \frac{1}{2}, \frac{1}{2}, 1 \right), \quad \lambda^* = (\lambda_1^*, \lambda_2^*) = (-1, -2), \quad \mu^* = (\mu_1^*, \mu_2^*) = (0, 0).$$

Il valore minimo della funzione obiettivo è:

$$f(x^*) = \left( \frac{1}{2} \right)^2 + \left( \frac{1}{2} \right)^2 + 1^2 = \frac{1}{4} + \frac{1}{4} + 1 = \frac{3}{2}.$$

**Osservazione** In questo esempio i due vincoli di uguaglianza sono (per definizione) sempre attivi. I due vincoli di disuguaglianza risultano invece inattivi all'ottimo, come indicato dal fatto che i relativi moltiplicatori sono nulli. Se la soluzione non rispettasse  $x_1 \geq 0$  o  $x_2 \geq 0$ , le condizioni di complementarità avrebbero forzato l'attivazione del vincolo corrispondente e l'ottimo si sarebbe spostato sul bordo dell'insieme ammissibile.

## 10.5 Condizioni di secondo ordine

Le condizioni di Karush–Kuhn–Tucker forniscono condizioni di ottimalità di primo ordine e permettono di individuare punti *candidati* all'ottimo. Tuttavia, come nel caso dei problemi non vincolati, esse non consentono in generale di stabilire se un punto stazionario sia di minimo, di massimo o di sella.

Per distinguere la natura dell'estremo è necessario introdurre condizioni di secondo ordine, basate sull'Hessiano della Lagrangiana.

Nel caso vincolato, tale Hessiano deve essere analizzato non su tutto lo spazio  $\mathbb{R}^n$ , ma ristretto alle direzioni ammissibili, ossia alle direzioni tangenti all'insieme dei vincoli attivi nel punto considerato.

Nei casi in cui il problema sia convesso (funzione obiettivo convessa e insieme ammissibile convesso), le condizioni KKT diventano anche sufficienti per garantire l'ottimalità globale, rendendo superflua l'analisi di secondo ordine.

**Definizione 10.5.1** Dato un problema nella forma:

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq 0 \\ & h_j(x) = 0 \end{aligned} \tag{10.118}$$

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^m \mu_i (0 - g_i(x)) + \sum_{j=1}^k \lambda_j (0 - h_j(x))$$

Hessiana:

$$\mathbf{H}(x, \lambda) = \begin{bmatrix} \mathcal{L}_{xx} & \mathcal{L}_{xy} & \mathcal{L}_{x\lambda} \\ \mathcal{L}_{yx} & \mathcal{L}_{yy} & \mathcal{L}_{y\lambda} \\ \mathcal{L}_{\lambda x} & \mathcal{L}_{\lambda y} & \mathcal{L}_{\lambda\lambda} \end{bmatrix} \tag{10.119}$$

Hessiana bordata:

$$\mathbf{H}(x, \lambda) = \begin{bmatrix} 0 & g_x & g_y \\ g_x & \mathcal{L}_{xx} & \mathcal{L}_{xy} \\ g_y & \mathcal{L}_{yx} & \mathcal{L}_{yy} \end{bmatrix} \tag{10.120}$$

$$\text{Se } \mathbf{x}^*, \lambda \text{ è un punto stazionario di } \mathcal{L}: \begin{cases} \mathbf{H}(\mathbf{x}^*, \lambda) \succ 0 & \Rightarrow \mathbf{x}^* \text{ è un minimo locale vincolato,} \\ \mathbf{H}(\mathbf{x}^*, \lambda) \prec 0 & \Rightarrow \mathbf{x}^* \text{ è un massimo locale vincolato,} \\ \det(\mathbf{H}(\mathbf{x}^*, \lambda)) < 0 & \Rightarrow \mathbf{x}^* \text{ è un punto di sella,} \\ \det(\mathbf{H}(\mathbf{x}^*, \lambda)) = 0 & \Rightarrow \text{nessuna conclusione possibile.} \end{cases}$$

**Esempio 10.5.2** Per vedere in modo netto come l'informazione di secondo ordine (Hessiano) permetta di distinguere tra minimo, massimo e sella, consideriamo un esempio semplice in  $\mathbb{R}^2$ .

*Caso 1: punto di minimo*

$$f_1(x, y) = x^2 + y^2.$$

I punti stazionari si ottengono imponendo  $\nabla f_1 = 0$ :

$$\nabla f_1(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} = 0 \quad \Rightarrow \quad (x^*, y^*) = (0, 0).$$

Calcoliamo l'Hessiano:

$$\nabla^2 f_1(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Questa matrice è *definita positiva* (autovalori  $2, 2 > 0$ ), quindi  $(0, 0)$  è un *minimo locale*. In questo caso, essendo  $f_1$  convessa, il minimo è anche globale.

*Caso 2: punto di massimo*

$$f_2(x, y) = -(x^2 + y^2).$$

Anche qui:

$$\nabla f_2(x, y) = \begin{bmatrix} -2x \\ -2y \end{bmatrix} = 0 \quad \Rightarrow \quad (x^*, y^*) = (0, 0).$$

L'Hessiano vale:

$$\nabla^2 f_2(x, y) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Questa matrice è *definita negativa* (autovalori  $-2, -2 < 0$ ), quindi  $(0, 0)$  è un *massimo locale* (e anche globale).

*Caso 3: punto di sella (nessuna conclusione con solo il gradiente)*

$$f_3(x, y) = x^2 - y^2.$$

Il gradiente è:

$$\nabla f_3(x, y) = \begin{bmatrix} 2x \\ -2y \end{bmatrix} = 0 \quad \Rightarrow \quad (x^*, y^*) = (0, 0).$$

L'Hessiano è:

$$\nabla^2 f_3(x, y) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Questa matrice è *indefinita* (un autovalore positivo e uno negativo), quindi  $(0, 0)$  è un *punto di sella*: muovendosi lungo la direzione  $x$  la funzione cresce, mentre lungo la direzione  $y$  decresce.

**Esempio 10.5.3** Consideriamo il seguente problema di ottimizzazione vincolata:

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & f(x_1, x_2) = x_1^2 + x_2^2 \\ \text{subject to} \quad & h(x_1, x_2) = x_1 + x_2 - 1 = 0, \\ & g(x_1, x_2) = -x_1 \leq 0 \quad (\Leftrightarrow x_1 \geq 0). \end{aligned} \tag{10.121}$$

Il vincolo di uguaglianza impone che il punto appartenga alla retta  $x_1 + x_2 = 1$ , mentre il vincolo di disuguaglianza seleziona la semiretta con  $x_1 \geq 0$ .

**Lagrangiana** Introduciamo il moltiplicatore  $\lambda \in \mathbb{R}$  per il vincolo di uguaglianza e  $\mu \geq 0$  per il vincolo di disuguaglianza:

$$\mathcal{L}(x_1, x_2, \lambda, \mu) = x_1^2 + x_2^2 + \lambda(x_1 + x_2 - 1) + \mu(-x_1). \quad (10.122)$$

**Condizioni KKT di primo ordine** **Stazionarietà:**

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_1} &= 2x_1 + \lambda - \mu = 0, \\ \frac{\partial \mathcal{L}}{\partial x_2} &= 2x_2 + \lambda = 0. \end{aligned} \quad (10.123)$$

**Fattibilità primale:**

$$x_1 + x_2 - 1 = 0, \quad x_1 \geq 0.$$

**Fattibilità duale:**

$$\mu \geq 0.$$

**Complementarità:**

$$\mu(-x_1) = 0.$$

**Soluzione KKT** Supponiamo inizialmente  $x_1 > 0$ . Allora il vincolo di disuguaglianza è inattivo e, per complementarità,  $\mu = 0$ . Le equazioni di stazionarietà diventano:

$$2x_1 + \lambda = 0, \quad 2x_2 + \lambda = 0 \quad \Rightarrow \quad x_1 = x_2.$$

Imponendo il vincolo di uguaglianza:

$$x_1 + x_2 = 1 \quad \Rightarrow \quad x_1 = x_2 = \frac{1}{2}.$$

Poiché  $x_1^* = \frac{1}{2} > 0$ , l'ipotesi iniziale è coerente e il punto è ammissibile. La soluzione KKT è quindi:

$$x^* = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \lambda^* = -1, \quad \mu^* = 0.$$

**Analisi di secondo ordine** L'Hessiano della Lagrangiana rispetto alle variabili decisionali è:

$$\nabla_{xx}^2 \mathcal{L}(x, \lambda, \mu) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Nel problema vincolato, l'Hessiano va analizzato lungo le *direzioni ammissibili*. Poiché il vincolo di uguaglianza è attivo, le direzioni ammissibili  $d$  devono soddisfare:

$$\nabla h(x^*)^T d = [1 \ 1] d = 0 \quad \Rightarrow \quad d_1 + d_2 = 0.$$

Una direzione ammissibile è quindi  $d = (1, -1)$ .

Valutiamo la forma quadratica:

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) d = [1 \ -1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 4 > 0.$$

Poiché l'Hessiano della Lagrangiana è positivo sulle direzioni ammissibili, il punto  $x^*$  è un *minimo locale vincolato*.





## 11.1 MPC Lineare in Matlab

### Comando `ss` — Modello in spazio di stato

Il comando `ss` viene utilizzato per creare un modello di sistema nella forma spazio di stato.

La rappresentazione in spazio di stato è centrale perché fornisce il modello utilizzato internamente dal controllore MPC per predire l'evoluzione futura del sistema a partire dallo stato attuale.

La documentazione ufficiale è consultabile cliccando [qui](#).

Un sistema in spazio di stato continuo è descritto da:

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{cases} \quad (11.1)$$

dove:

- $x(t)$  è il vettore degli stati,
- $u(t)$  è il vettore degli ingressi,
- $y(t)$  è il vettore delle uscite,
- $A, B, C, D$  sono le matrici del modello.

In maniera analoga, un modello discreto, con tempo di campionamento  $T_s$ , ha forma:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{cases} \quad (11.2)$$

La sintassi di base del comando è la seguente:

```
1 sys = ss(A,B,C,D)           % Caso Tempo Continuo
2 sys = ss(A,B,C,D,Ts)       % Caso Tempo Discreto
```

con la quale si crea un oggetto di tipo `ss` che rappresenta un modello a tempo continuo dato dalle matrici  $A, B, C, D$ . Analogamente, un sistema a tempo discreto con periodo di campionamento  $T_s$ .

#### Esempio 11.1.1 Definizione di un sistema TD

```
1 A = [0 1; -5 -2];
2 B = [0; 3];
3 C = [0 1];
4 D = 0;
5 Ts = 0.1;
6
7 % modello spazio di stato discreto
8 plant = ss(A,B,C,D,Ts);
```

In questo esempio, con il comando `ss` è stato definito un modello discreto a due stati, un ingresso e un'uscita.

**NB:** E' possibile utilizzare `ss` anche per convertire modelli rappresentati in altre forme (ad esempio funzione di trasferimento) in spazio di stato.

### Comando `tf` — Funzione di trasferimento

Il comando `tf` consente di creare un modello dinamico nella forma di **funzione di trasferimento**. Una funzione di trasferimento rappresenta matematicamente la relazione tra ingresso e uscita di un sistema

lineare tempo-invariante (LTI) nel dominio di Laplace (per sistemi continui) o nel dominio  $z$  (per sistemi discreti).

La documentazione ufficiale è consultabile cliccando [qui](#).

Una funzione di trasferimento ha la forma:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_0 + b_1s + \dots + b_ms^m}{a_0 + a_1s + \dots + a_ns^n}$$

dove  $s$  è la variabile complessa del dominio di Laplace per sistemi continui.

Nel caso discreto, la variabile è  $z$  e la forma diventa:

$$G(z) = \frac{Y(z)}{U(z)} = \frac{b_0 + b_1z^{-1} + \dots + b_mz^{-m}}{a_0 + a_1z^{-1} + \dots + a_nz^{-n}}$$

Per un sistema continuo definito dai vettori numeratore e denominatore:

```
1 sys = tf(num, den)
```

crea una funzione di trasferimento:

$$G(s) = \frac{\text{num}(s)}{\text{den}(s)}$$

Per un sistema discreto con tempo di campionamento  $T_s$ :

```
1 sys = tf(num, den, Ts)
```

dove  $T_s$  rappresenta il periodo di campionamento.

#### Esempio 11.1.2 Come usare tf.

```
1 % funzione di trasferimento continua G(s)=1/(s^2+3s+2)
2 num = [1];
3 den = [1 3 2]; % s^2 + 3s + 2
4 sys = tf(num, den);
5
6 % funzione di trasferimento discreta z^{-1} con Ts = 0.1
7 Ts = 0.1;
8 numd = [0.1 0.05];
9 dend = [1 -0.8];
10 sysd = tf(numd, dend, Ts);
11
12 % funzione di trasferimento continua utilizzando variabili simboliche
13 s = tf('s') % Definisce s come variabile simbolica tf
14 sys = 1/(s^2 + 3*s + 1); % sys e' gia' un oggetto di tipo tf
15
```

### Comando mpc — Creazione di un controllore predittivo

Il comando `mpc` serve per creare un oggetto di tipo MPC. Questo oggetto contiene tutte le informazioni necessarie affinché il controllore possa prevedere l'evoluzione futura del sistema e risolvere iterativamente un problema di ottimizzazione vincolata per calcolare l'azione di controllo ottimale.

Quando si crea un oggetto `mpc`, MATLAB utilizza il modello del sistema specificato, insieme a eventuali disturbi o rumore, per costruire internamente il modello di predizione e lo stato del controllore. In ciascun intervallo di controllo viene calcolata una sequenza futura di ingressi che minimizza una funzione di costo quadratica, tenendo conto dei vincoli su ingressi, uscite e variazioni di controllo.

La documentazione ufficiale è consultabile cliccando [qui](#).

La sintassi è la seguente:

```
1 mpcobj = mpc(plant, Ts)
2 mpcobj = mpc(plant, Ts, P, M, W, MV, OV, DV)
```

- **plant**: modello di predizione del processo, in genere un modello lineare in spazio di stato o funzione di trasferimento discretizzato.
- **Ts**: tempo di campionamento desiderato per il controllore.

- P: orizzonte di predizione (`PredictionHorizon`).
- M: orizzonte di controllo (`ControlHorizon`).
- W: struttura dei pesi (`Weights`) per la funzione di costo.
- MV: struttura di vincoli sulle variabili manipolate (`ManipulatedVariables`).
- OV: struttura di vincoli sulle variabili di uscita (`OutputVariables`).
- DV: struttura di vincoli sulle variabili di disturbo (`DisturbanceVariables`).

Se alcuni parametri vengono omessi o lasciati vuoti, vengono adottati i valori predefiniti.

#### Comportamento rispetto al modello

- Se il modello `plant` è continuo, l'MPC lo discretizza internamente usando il tempo di campionamento `Ts`.
- Se il modello è già discreto ma ha tempo di campionamento non specificato (`plant.Ts = -1`), allora eredita il valore di `Ts`.

#### Esempio 11.1.3 Esempio di utilizzo:

```

1  % definizione del modello discretizzato
2  plant = ss(A,B,C,D,Ts);
3
4  % creazione di un MPC con orizzonte di predizione 10 e controllo 2
5  mpcobj = mpc(plant, Ts, 10, 2);
6

```

In questo esempio l'oggetto `mpcobj` eredita dal modello `plant` il tempo di campionamento e utilizza orizzonti di predizione e controllo definiti dall'utente.

#### Proprietà principali dell'oggetto `mpc`

Dopo la creazione, molte proprietà dell'oggetto possono essere modificate tramite notazione a punti, ad esempio:

```

1 mpcobj.Ts = 0.1;
2 mpcobj.PredictionHorizon = 12;
3 mpcobj.ControlHorizon = 3;
4 mpcobj.Weights.OutputVariables = 1;
5 mpcobj.MV.Min = -5;
6 mpcobj.MV.Max = 5;

```

In questo modo si definiscono i principali parametri di progetto del controllore.

#### Note:

- L'oggetto `mpc` non esegue azioni di controllo da solo: esso definisce il controllore che poi viene applicato tramite comandi come `mpcmove` o `sim`.
- È possibile anche includere modelli di disturbi misurati o non misurati durante la creazione dell'oggetto.

#### Comando `mpcstate` — Stato del controllore MPC

Il comando `mpcstate` in MATLAB viene utilizzato per creare e inizializzare un oggetto dello stato interno associato a un controllore MPC. Questo oggetto rappresenta lo “stato” del controllore, ovvero contiene tutte le variabili interne necessarie affinché l'MPC possa operare correttamente durante la simulazione o durante l'applicazione reale.

La documentazione ufficiale è consultabile cliccando [qui](#).

Quando si utilizza un controllore MPC, non si lavora soltanto con il modello del sistema fisico, ma anche con variabili interne al controllore che memorizzano informazioni sullo stato corrente del processo, errori passati, condizioni dinamiche, predizioni, ecc. L'oggetto restituito da `mpcstate` serve quindi per:

- inizializzare lo stato interno del controllore prima di una simulazione o di un'applicazione in anello chiuso,
- mantenere queste informazioni tra una chiamata e la successiva del comando `mpcmove`,
- permettere un aggiornamento coerente della predizione e della soluzione ottimale nel tempo.

La sintassi è la seguente:

```
1 xk = mpcstate(mpcobj)
```

- `mpcobj`: oggetto MPC creato con il comando `mpc`.
- `xk`: oggetto che rappresenta lo stato interno del controllore.

L'oggetto `xk` viene tipicamente creato all'inizio di una simulazione o di un loop di controllo, e poi passato come argomento in input alla funzione `mpcmove`. Ad ogni passo di campionamento, lo stato interno viene aggiornato automaticamente.

#### Esempio 11.1.4 Esempio tipico di uso:

```
1 % creazione del controllore MPC
2 mpcobj = mpc(plant, Ts, Hp, Hc);
3
4 % inizializzazione dello stato interno
5 xk = mpcstate(mpcobj);
6
7 % loop di controllo
8 for k = 1:N
9     % misura dell'uscita
10    yk = C*xk.Plant;
11
12    % calcolo dell'azione di controllo ottimale
13    uk = mpcmove(mpcobj, xk, yk, ref);
14 end
15
```

In questo schema, la variabile `xk` viene aggiornata automaticamente da `mpcmove` ad ogni iterazione.

#### Considerazioni pratiche

- L'oggetto `mpcstate` non contiene lo stato fisico del processo (che viene gestito dal modello stesso), ma lo stato interno del controllore MPC.
- Se si utilizzano più canali di uscita o disturbi misurati, l'oggetto `xk` si adatta automaticamente alla dimensione appropriata dello spazio degli stati.
- L'inizializzazione corretta di `xk` è fondamentale per evitare comportamenti instabili o inaspettati durante l'avvio del controllo predittivo.

**NB:** Se si utilizza la funzione `sim` per simulare l'intero sistema con MPC, MATLAB può gestire automaticamente lo stato interno senza richiedere una variabile `xk` esplicita.

### 11.1.1 Comando `mpcmove` — Calcolo dell'azione di controllo ottimale

Il comando `mpcmove` è il principale strumento per utilizzare un controllore MPC in anello chiuso. In un tipico schema di controllo predittivo, questo comando calcola l'azione di controllo ottimale in base alle misure attuali del sistema, aggiorna lo stato interno del controllore e restituisce la variabile manipolata da applicare al processo.

La documentazione ufficiale è consultabile cliccando [qui](#).

Ad ogni passo di campionamento, `mpcmove` risolve un problema di ottimizzazione quadratica basato sul modello interno e sulle informazioni correnti:

- lo stato interno del controllore (`xk`),
- le uscite misurate del processo (`ym`),

- il riferimento o setpoint desiderato ( $\mathbf{r}$ ),
- eventuali disturbi misurati ( $\mathbf{v}$ ), se presenti.

L'uscita del comando è la variabile manipolata ottimale ( $\mathbf{mv}$ ) da applicare al processo all'istante corrente e lo stato interno viene automaticamente aggiornato per il passo successivo.

La sintassi è la seguente:

```
1 mv = mpcmove(mpcobj, xk, ym, r)
2 mv = mpcmove(mpcobj, xk, ym, r, v)
```

I principali argomenti di input sono:

- `mpcobj` — oggetto MPC creato con `mpc`;
- `xk` — oggetto stato interno del controllore, ottenuto con `mpcstate`;
- `ym` — vettore delle uscite misurate all'istante corrente;
- `r` — vettore dei riferimenti o setpoint;
- `v` — (opzionale) vettore dei disturbi misurati.

*Comportamento a runtime* Se uno degli argomenti `ym`, `r` o `v` è specificato come `[]` oppure omesso, `mpcmove` utilizza i valori di default dell'oggetto `mpcobj.Model.Nominal`. Il comando aggiorna inoltre internamente lo stato di stima (ad esempio tramite filtro di Kalman integrato) se si utilizza la stima di stato predefinita.

#### Esempio 11.1.5 Esempio di utilizzo:

```
1 % mpcobj: controllore MPC gi creato
2 % xk: stato interno del controllore
3 % yk: uscita misurata al tempo k
4 % ref: riferimento desiderato
5 mv = mpcmove(mpcobj, xk, yk, ref);
6
```

In caso di presenza di disturbi misurabili, la sintassi diventa:

```
1 mv = mpcmove(mpcobj, xk, yk, refk, vk);
```

#### Note importanti

- Il comando `mpcmove` va chiamato sequenzialmente all'interno di un loop cioè di una simulazione step-by-step per riprodurre il comportamento in tempo reale del controllore.
- L'oggetto `xk` deve essere passato come riferimento: il suo stato interno viene modificato dal comando e queste modifiche sono utilizzate nel passo successivo.

## 11.2 MPC non lineare in MATLAB

Tutta la documentazione riguardante i sistemi non lineari è disponibile [qui](#).

### Comando `nlpmpc` — Controllore predittivo non lineare

Il comando `nlpmpc` consente di creare un oggetto controllore predittivo non lineare. A differenza del controllore lineare (`mpc`), l'oggetto `nlpmpc` è progettato per sistemi non lineari in cui la dinamica del processo e/o l'output non possono essere rappresentati esattamente tramite un modello lineare. La documentazione ufficiale è disponibile [qui](#).

Un controllore NMPC utilizza un modello non lineare del processo per prevedere l'evoluzione futura degli stati e dell'uscita, risolvendo ad ogni istante un problema di ottimizzazione non lineare con vincoli. Questo permette di gestire direttamente non linearità significative del processo, saturazioni, effetti dinamici non lineari, ecc.

La sintassi è la seguente:

```
1 nlobj = nlpmpc(nx,ny,nu)
2 nlobj = nlpmpc(nx,ny,'MV',mvIndex,'MD',mdIndex)
```

dove:

- `nx`: numero di stati del modello non lineare,
- `ny`: numero di uscite,
- `nu`: numero di ingressi (variabili manipolate),
- l'opzione `'MV'`, `mvIndex` indica quali variabili del modello sono considerate come variabili manipolate,
- l'opzione `'MD'`, `mdIndex` indica quali variabili del modello sono considerate come disturbi misurati.

Dopo la creazione dell'oggetto `nlobj`, è necessario specificare le funzioni che descrivono la dinamica non lineare e l'uscita del sistema:

```
1 nlobj.Model.StateFcn = 'stateFcn';
2 nlobj.Model.OutputFcn = 'outputFcn';
3 nlobj.Model.IsContinuousTime = true;
```

- `stateFcn`: funzione che calcola  $\dot{x} = f(x, u)$  (o  $x[k+1] = f(x[k], u[k])$ , se tempo discreto),
- `outputFcn`: funzione che calcola l'uscita ( $y = h(x, u)$ ),
- `IsContinuousTime`: specifica se il modello è in tempo continuo o discreto.

### Specifica della funzione di stato e di Uscita

In un controllore `nlpmpc`, è necessario definire esplicitamente il modello di predizione non lineare. Questo modello consiste di due funzioni fondamentali:

- **Funzione di stato (State Function)** — descrive l'evoluzione degli stati del processo nel tempo,
- **Funzione di uscita (Output Function)** — calcola le uscite in funzione degli stati e degli ingressi.

Queste funzioni devono essere implementate e associate alle proprietà dell'oggetto `nlpmpc`.

#### Funzione di stato

La funzione di stato può essere continua o discreta:

- Per un modello continuo, la funzione deve restituire le derivate degli stati ( $\dot{x}$ ),
- Per un modello discreto, la funzione deve restituire lo stato al passo successivo  $x[k+1]$ .

Poiché il controllore NMPC opera in tempo discreto, se la funzione è fornita in forma continua, MATLAB la discretizza internamente tramite un metodo numerico (ad esempio la regola trapezoidale implicita) basato sul tempo di campionamento del controllore.

Le funzioni possono essere specificate in forma di nome di funzione come stringa o char vector, e.g., "stateFCN".

La sintassi per associare la funzione di stato è:

```
1 nlobj.Model.StateFcn = 'stateFCN';
```

### Esempio 11.2.1 Esempio di definizione di funzione di stato

```
1 function dx = stateFCN(x,u)
2     dx = zeros(6,1);
3     dx(1) = x(4);
4     dx(2) = x(5);
5     dx(3) = x(6);
6     dx(4) = (u(1) - u(2) + u(3) - u(4))*cos(x(3));
7     dx(5) = (u(1) - u(2) + u(3) - u(4))*sin(x(3));
8     dx(6) = 0.2*(u(1) - u(2) - u(3) + u(4));
9 end
10
```

### Funzione di uscita

La funzione di uscita calcola le uscite del sistema in base agli stati e agli ingressi. Anche questa può essere specificata come nome di funzione:

```
1 nlobj.Model.OutputFcn = 'outputFCN';
```

In forma discreta la funzione restituisce direttamente le uscite al passo successivo, mentre in forma continua restituisce le uscite istantanee.

### Esempio 11.2.2 Esempio di definizione di funzione di uscita

```
1 function y = outputFCN(x,u)
2     y = zeros(6,1);
3     y(1) = x(1);
4     y(2) = x(2)+0.2*x(3);
5     y(3) = x(3)*x(4);
6 end
7
```

### Parametri opzionali del modello

Se il modello di predizione dipende da parametri esterni (ad esempio costanti fisiche o condizioni operative), è possibile passarli alle funzioni. In questo caso, è necessario specificare anche:

```
1 nlobj.Model.NumberOfParameters = nParams;
```

e includere questi parametri negli argomenti in ingresso delle funzioni di stato, uscita e degli Jacobiani.

*Validazione delle funzioni* Prima di simulare o utilizzare il controllore, è buona pratica verificare la correttezza delle funzioni personalizzate (stato, uscita e degli Jacobiani) tramite il comando `validateFcns`. Questo aiuta a individuare errori di dimensionamento o di definizione delle funzioni.

```
1 validateFcns(nlmpcobj,x,mv)
2 validateFcns(nlmpcobj,x,mv,md)
3 validateFcns(nlmpcobj,x,mv,md,parameters)
4 validateFcns(nlmpcobj,x,mv,md,parameters,ref)
```

### Tempo di campionamento e orizzonti

Analogamente all'MPC lineare, si impostano:

```
1 nlobj.Ts = 0.1;
2 nlobj.PredictionHorizon = 10;
3 nlobj.ControlHorizon = 2;
```

### Jacobiani

Per migliorare l'efficienza numerica e la convergenza dell'ottimizzazione non lineare, è possibile fornire i Jacobiani delle funzioni di stato e di uscita:

```
1 nlobj.Jacobian.StateFcn = 'jacobianStateFcn';
2 nlobj.Jacobian.OutputFcn = 'jacobianOutputFcn';
```

- La funzione `jacobianStateFcn` definisce le derivate parziali di `stateFcn` rispetto allo stato e all'ingresso,
- La funzione `jacobianOutputFcn` definisce le derivate parziali di `outputFcn`.

**Esempio 11.2.3** Esempio di definizione degli Jacobiani per le funzioni di stato e uscita definite in precedenza

```
1 % Funzione di Jacobiano di stato
2 function [A,B] = jacobianStateFcn(x,u)
3     A = zeros(6,6);
4     A(1,4) = 1;
5     A(2,5) = 1;
6     A(3,6) = 1;
7     A(4,3) = -(u(1) - u(2) + u(3) - u(4))*sin(x(3));
8     A(5,3) = (u(1) - u(2) + u(3) - u(4))*cos(x(3));
9     B = zeros(6,4);
10    B(4,:) = cos(x(3))*[1 -1 1 -1];
11    B(5,:) = sin(x(3))*[1 -1 1 -1];
12    B(6,:) = 0.2*[1 -1 -1 1];
13 end
14
15 % Funzione di Jacobiano di uscita
16 function C = jacobianOutputFcn(x,u)
17     C = zeros(3,4);
18     C(1,1) = 1;
19     C(2,2) = 1;
20     C(2,3) = 0.2;
21     C(3,3) = x(4);
22     C(3,4) = x(3);
23 end
```

### Considerazioni pratiche

- L'NMPC risolve ogni volta un problema di ottimizzazione non lineare (NLP), il che può richiedere tempi computazionali maggiori rispetto alla MPC lineare.
- La qualità delle soluzioni dipende fortemente dalla qualità del modello non lineare e dalle condizioni iniziali.
- Specificare funzioni di Jacobiano accurate può migliorare sensibilmente la prestazione e ridurre il tempo di calcolo.

### Comando `nlpmove` — Calcolo dell'azione di controllo per controllori NMPC

Il comando `nlpmove` viene utilizzato per calcolare l'azione di controllo ottimale in un controllore predittivo non lineare (`nlpmpc`). In un controllore NMPC, ad ogni istante di campionamento si risolve un problema di ottimizzazione non lineare sulla base del modello di predizione, delle condizioni correnti del sistema e dei riferimenti desiderati.

All'interno di un loop di controllo, `nlpmove`:

- utilizza lo stato corrente del processo o dello stimatore;
- considera l'ultimo valore di ingresso applicato;
- incorpora i riferimenti desiderati per l'uscita;
- opzionalmente utilizza misurazioni di disturbi o valori nominali;



- risolve un problema di ottimizzazione non lineare per determinare la migliore sequenza futura di ingressi;
- restituisce la prima variabile manipolata ottimale da applicare al processo.

La funzione aggiorna internamente la storia delle variabili manipolate e gli stati previsti necessari per il passo successivo. Il comportamento complessivo è analogo a quello di `mpcmove` per MPC lineare, ma la componente di ottimizzazione è non lineare.

```
1 mv = nlmpcmove(nlobj, xk, lastMV, ref)
2 [mv, ~, info] = nlmpcmove(nlobj, xk, lastMV, ref)
```

Dove:

- `nlobj`: oggetto controllore NMPC creato con `nlmpc`;
- `xk`: vettore degli stati correnti all'istante attuale;
- `lastMV`: ultimo valore di variabile manipolata applicato al sistema;
- `ref`: vettore di riferimenti (setpoint) desiderati;
- `mv`: variabile manipolata calcolata da applicare;
- `info`: struttura opzionale contenente informazioni aggiuntive sulla soluzione di ottimizzazione (ad esempio convergenza, costi, stato predetto).

### 11.3 Interpretazione dei grafici di simulazione MPC

L'analisi dei risultati di una simulazione MPC si basa sull'interpretazione congiunta di diversi grafici, ciascuno dei quali fornisce informazioni complementari sul comportamento del controllore e del sistema in anello chiuso. In particolare, è fondamentale analizzare:

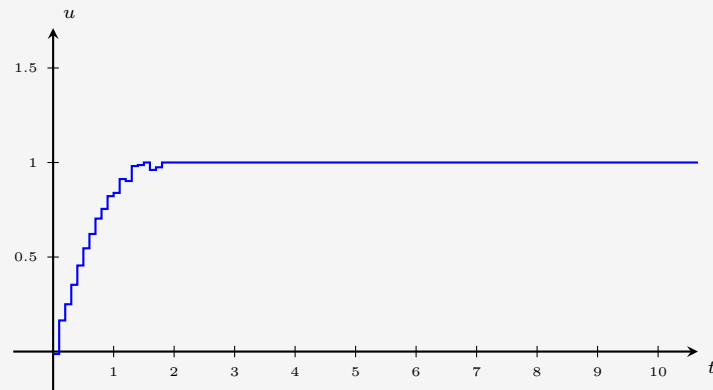
- l'andamento dell'ingresso di controllo (variabile manipolabile o controllabile),
- l'evoluzione delle uscite del sistema,
- i tempi computazionali di risoluzione del problema di ottimizzazione,
- l'evoluzione della funzione di costo.

L'interpretazione corretta di questi grafici consente di valutare le prestazioni del controllore e di guidare il processo di tuning.

#### 11.3.1 Interpretazione dei grafici dell'ingresso di controllo

Il grafico dell'ingresso di controllo  $u(k)$  (o delle variabili manipolate) mostra come il controllore agisce sul processo nel tempo. Da questo grafico è possibile dedurre l'aggressività del controllo, l'attivazione dei vincoli e la qualità del tuning.

##### Esempio 11.3.1 Ingresso saturo per lunghi intervalli



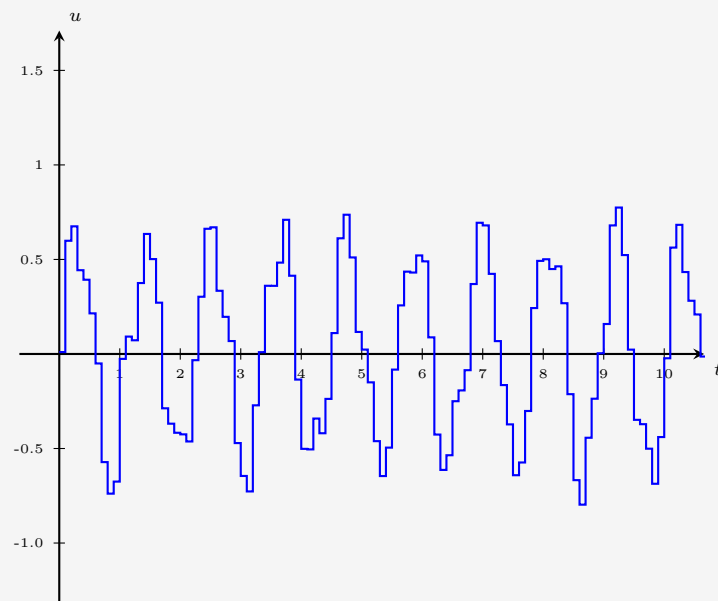
L'ingresso raggiunge rapidamente il valore massimo o minimo consentito e rimane saturo per un tempo significativo. Interpretazione:

- il controllore sta cercando di compensare rapidamente un errore elevato;
- i vincoli sull'ingresso sono attivi e dominano il comportamento del sistema.

Possibili azioni di tuning:

- aumentare l'orizzonte di predizione  $H_p$  per consentire una strategia più anticipativa;
- ridurre il peso sulle uscite per rendere il controllo meno aggressivo;
- verificare se i vincoli sono eccessivamente stringenti rispetto agli obiettivi.

### Esempio 11.3.2 Ingresso oscillante



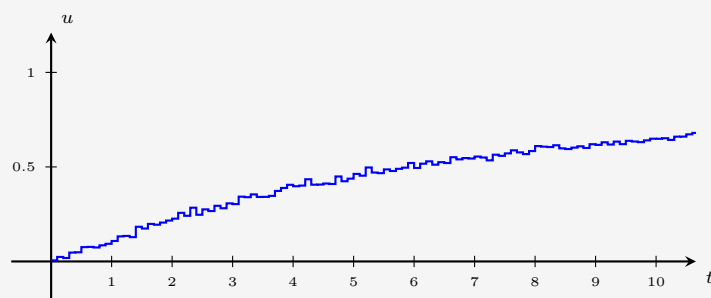
L'ingresso presenta oscillazioni rapide o cambiamenti di segno frequenti.

Interpretazione:

- il controllore reagisce in modo eccessivo a piccoli errori;
- possibile presenza di rumore o di pesi non bilanciati.

Possibili azioni di tuning:

- aumentare il peso sulla variazione dell'ingresso (`ManipulatedVariablesRate`);
- ridurre il tempo di campionamento solo se necessario;
- introdurre o rafforzare un filtro di stima.

**Esempio 11.3.3** Ingresso molto lento

L'ingresso varia lentamente e impiega molto tempo a raggiungere valori significativi.

Interpretazione:

- il controllore è eccessivamente conservativo;
- il peso sulla variazione dell'ingresso è troppo elevato.

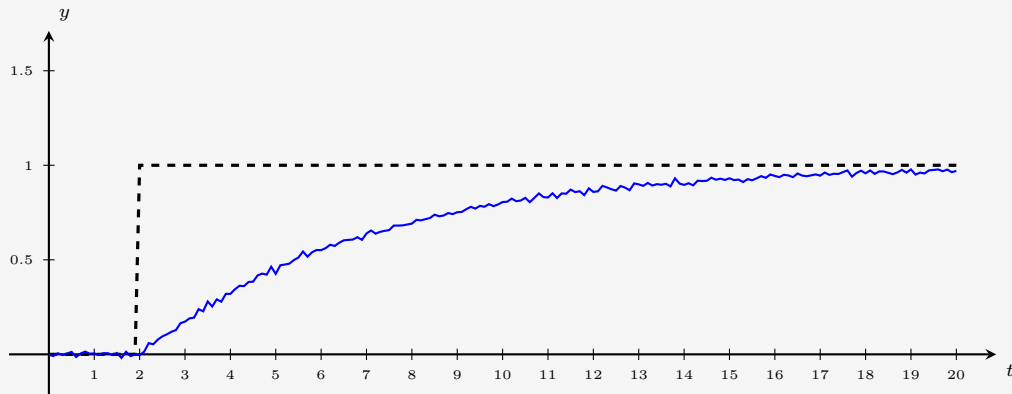
Possibili azioni di tuning:

- ridurre il peso su  $\Delta u$ ;
- aumentare il peso sull'errore di tracking delle uscite.

### 11.3.2 Interpretazione dei grafici delle uscite

I grafici delle uscite  $y(k)$  permettono di valutare direttamente le prestazioni del controllo in termini di tracking, rispetto dei vincoli e dinamica transitoria.

#### Esempio 11.3.4 Tracking lento del riferimento



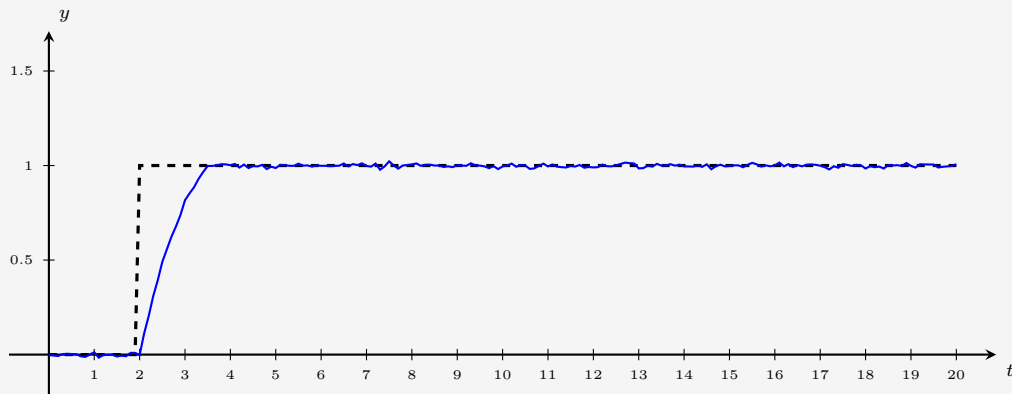
L'uscita raggiunge il riferimento con un tempo di assestamento elevato.

Interpretazione:

- il controllore privilegia il rispetto dei vincoli o la regolarità dell'ingresso;
- l'orizzonte di predizione potrebbe essere troppo corto.

Possibili azioni di tuning:

- aumentare il peso sull'uscita nella funzione di costo;
- aumentare  $H_p$  per includere più dinamica futura.

**Esempio 11.3.5** Superamento dei vincoli di uscita

L'uscita raggiunge o resta a lungo sul limite massimo/minimo consentito.

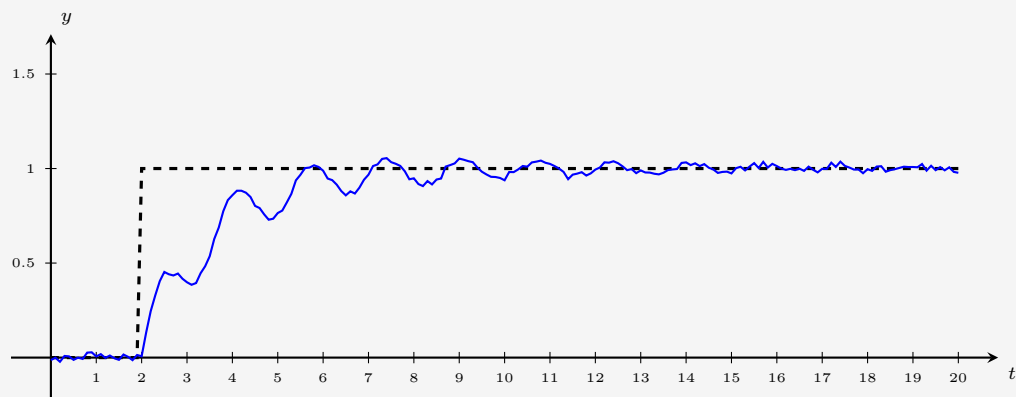
Interpretazione:

- il vincolo di uscita è attivo;
- il riferimento richiesto è difficilmente raggiungibile senza violazioni.

Possibili azioni di tuning:

- verificare la coerenza del riferimento con i vincoli fisici;
- introdurre o regolare le ECR (soft constraints);
- ribilanciare i pesi tra ingresso e uscita.

**Esempio 11.3.6** Oscillazioni sull'uscita



L'uscita oscilla attorno al riferimento. Interpretazione:

- possibile eccessiva aggressività del controllo;
- modello o stima non accurata.

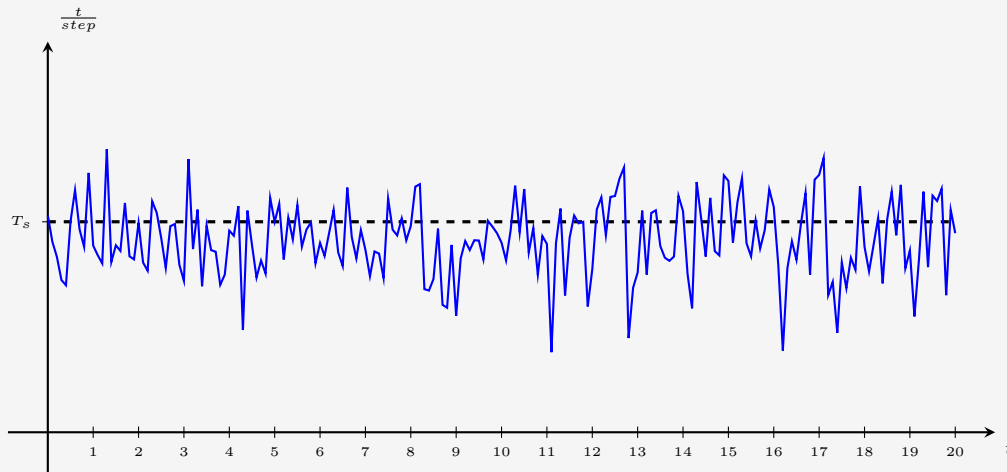
Possibili azioni di tuning:

- aumentare il peso su  $\Delta u$ ;
- migliorare la qualità del modello o della stima.

### 11.3.3 Interpretazione dei tempi computazionali

Il grafico dei tempi computazionali mostra il tempo necessario a risolvere il problema di ottimizzazione ad ogni istante.

**Esempio 11.3.7** Tempo medio vicino a  $T_s$



Il tempo di calcolo è vicino o superiore al tempo di campionamento.

Interpretazione:

- il controllore non è implementabile in tempo reale;
- il problema di ottimizzazione è troppo complesso.

Possibili azioni di tuning:

- ridurre l'orizzonte di controllo  $H_c$ ;
- ridurre il numero di vincoli attivi;
- valutare MPC esplicito o semplificazione del modello.



**Esempio 11.3.8** Picchi occasionali



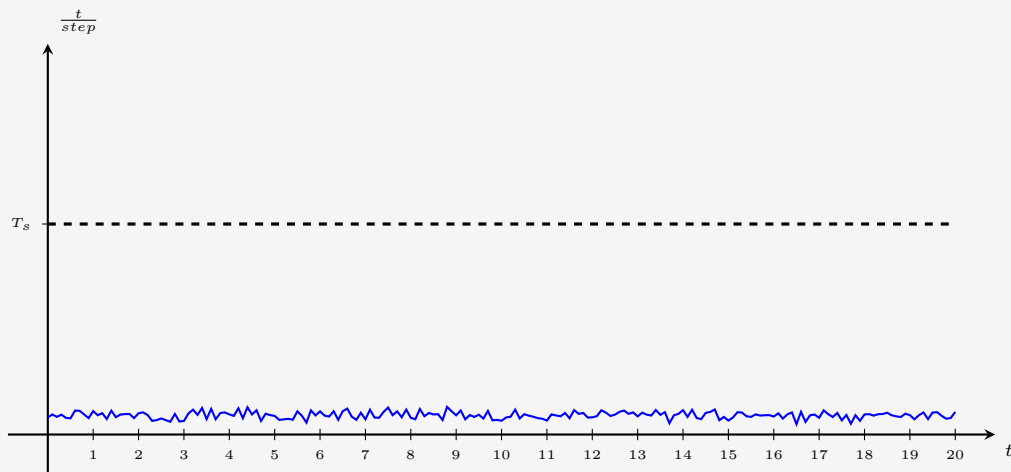
Il tempo computazionale presenta picchi isolati.

Interpretazione:

- cambiamenti improvvisi nel setpoint o nell'attivazione dei vincoli;
- difficoltà di convergenza del solver.

Possibili azioni di tuning:

- migliorare l'inizializzazione;
- rivedere i pesi o i limiti delle variabili.

**Esempio 11.3.9** Tempo stabile e basso

Il tempo computazionale è stabile e molto inferiore a  $T_s$ .

Interpretazione:

- il controllore è ben dimensionato ed eseguibile real-time;
- possibile margine per migliorare le prestazioni.

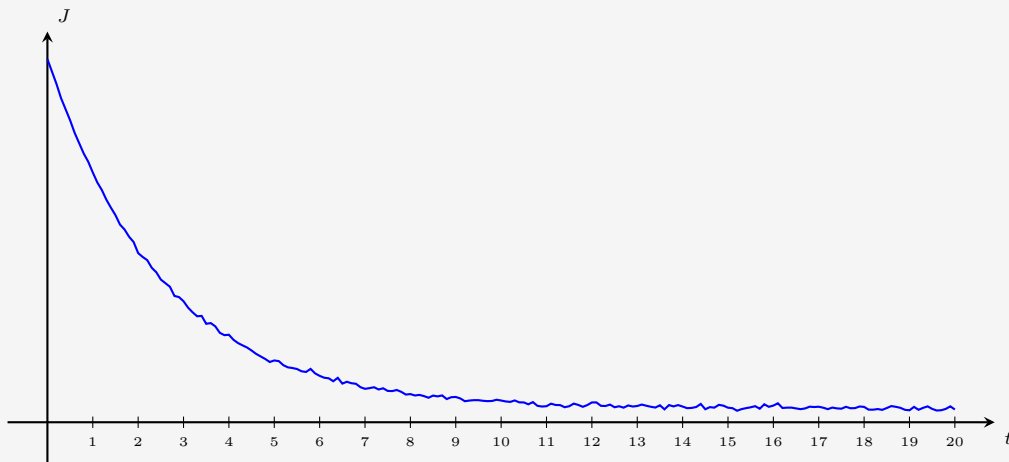
Possibili azioni di tuning:

- aumentare leggermente  $H_p$  per migliorare il tracking;
- introdurre vincoli più stringenti se necessario.

### 11.3.4 Interpretazione dell'evoluzione della funzione di costo

Il grafico del costo  $J(k)$  fornisce una misura sintetica della qualità della soluzione ottimale nel tempo.

#### Esempio 11.3.10 Costo decrescente e convergente



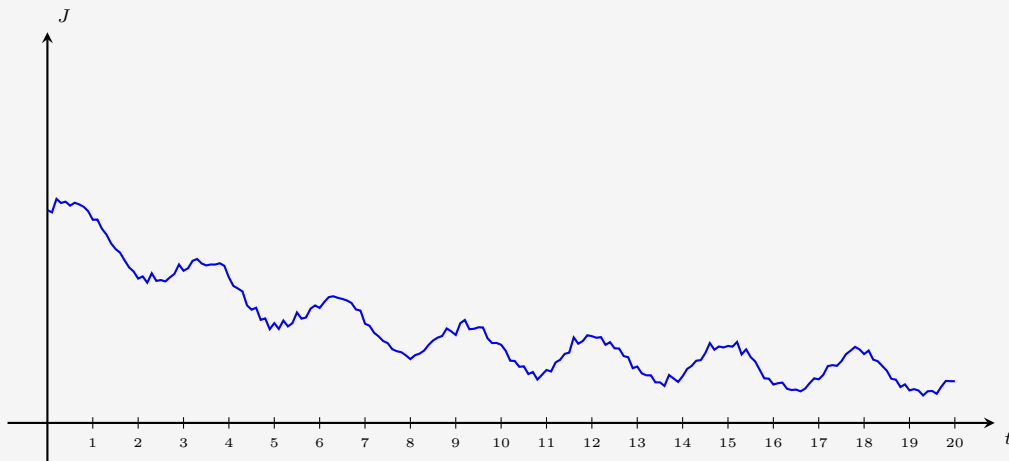
Il costo diminuisce rapidamente e converge a un valore costante.

Interpretazione:

- il sistema si avvicina al punto di equilibrio desiderato;
- il controllore è ben tarato.

Azione:

- nessuna modifica necessaria.

**Esempio 11.3.11** Costo oscillante

Il costo oscilla significativamente nel tempo.

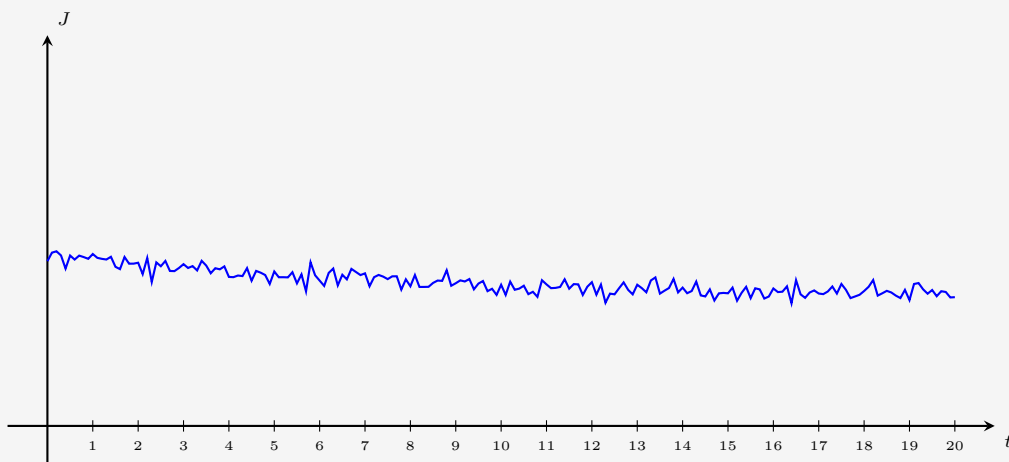
Interpretazione:

- presenza di compromessi instabili tra tracking e sforzo di controllo;
- possibile attivazione intermittente dei vincoli.

Possibili azioni di tuning:

- ribilanciare i pesi tra uscite e variazioni dell'ingresso;
- aumentare  $H_p$  per una visione più ampia del futuro.

**Esempio 11.3.12** Costo elevato persistente



Il costo rimane elevato anche a regime.

Interpretazione:

- il riferimento non è raggiungibile date le limitazioni;
- il modello di predizione non è adeguato.

Possibili azioni di tuning:

- verificare la correttezza del modello;
- rivedere vincoli e obiettivi di controllo;
- introdurre o regolare i soft constraints.