

Indice

1 Stabilità nei sistemi dinamici	2
1.1 Sistema dinamico e punti di equilibrio	2
1.2 Stabilità e Convergenza	2
1.3 Stabilità nei sistemi lineari	3
1.4 Sistema nel dominio di Laplace	4
1.5 Stabilità tramite l'approccio alla Lyapunov	5
1.6 Lyapunov per Sistemi LTI	6
1.7 Stabilità Robusta	10
2 Sistemi MIMO nel dominio di Laplace	12
2.1 Rappresentazione di un sistema MIMO nel dominio di Laplace	12
2.2 Sistema con retroazione dell'uscita MIMO	12
2.3 Guadagno	14
2.4 Poli e zeri per sistemi MIMO	14
3 Assegnamento degli autovalori con feedback di stato	15
3.1 Caso MIMO	15
3.2 Metodo della forma canonica di controllabilità	15
3.3 Metodo dell'equazione di Sylvester	17
3.4 Metodo di Kautsky-Nichols-Van Dooren	19
4 Controllo di sistemi MIMO	21
4.1 Schema di retroazione dell'uscita con integratori	21
4.2 Problema di Mixed Sensitivity Design	30
4.3 Problema H_∞ in forma standard	33
4.4 Problema di Mixed Sensitivity come problema H_∞ in forma standard	34
4.4.1 Costruzione $W_S(s)$	36
5 Appendici	39
5.1 SVD (Singular Value Decomposition)	39
5.2 Diagonalizzare	40
5.3 Uguaglianza matriciale per dimostrare la relazione di $T(s)$	41

Capitolo 1

Stabilità nei sistemi dinamici

1.1 Sistema dinamico e punti di equilibrio

Un sistema dinamico è composta da uno stato $x \in \mathbb{R}^n$ e da una legge di evoluzione

$$\dot{x}(t) = f(x(t), u(t), t) \quad (1.1)$$

Definizione 1.1.1. Un sistema è detto **tempo invariante** se la sua evoluzione non dipende esplicitamente dal tempo t , cioè:

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

Definizione 1.1.2. Un sistema è detto **autonomo** se la sua evoluzione non dipende esplicitamente dall'ingresso $u(t)$, cioè:

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

Se il sistema è sia autonomo che tempo invariante allora si ha:

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

Definizione 1.1.3. Dato un sistema dinamico **tempo invariante** nella forma:

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

un punto di equilibrio (\bar{x}, \bar{u}) è una coppia stato-ingresso tale che:

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

1.2 Stabilità e Convergenza

Definizione 1.2.1. Un punto di equilibrio \bar{x} di un sistema dinamico tempo invariante è detto stabile se:

$$\forall \varepsilon > 0, \exists \delta > 0 : x_0 \in B_\delta(\bar{x}) \Rightarrow |x_{x_0}(t) - \bar{x}| \leq \varepsilon, \forall t \geq 0 \quad (1.2)$$

Dove con $x_{x_0}(t)$ si intende la traiettoria del sistema che parte da x_0 .

Ora dimostriamo che $\delta \leq \varepsilon, \forall \varepsilon > 0$.

Dimostrazione. Supponiamo per assurdo che esista un $\varepsilon > 0$ per cui si ha che la δ per cui è rispettata la definizione di stabilità sia tale che $\delta > \varepsilon$. Allora si ha che per le $x_0 \in B_\delta(\bar{x}) \setminus B_\varepsilon(\bar{x})$ vale:

$$|x_{x_0}(0) - \bar{x}| = |x_0 - \bar{x}| > \varepsilon$$

Ma questo è assurdo perché $x_0 \in B_\delta(\bar{x})$ dunque rispetta la definizione di stabilità per cui si dovrebbe avere :

$$|x_{x_0}(0) - \bar{x}| = |x_0 - \bar{x}| \leq \varepsilon$$

Dunque siamo arrivati ad un assurdo.

□

Definizione 1.2.2. Dato un sistema dinamico un punto di equilibrio \bar{x} è detto **convergente** se esiste un $\delta > 0$ tale che per ogni condizione iniziale $x_0 \in B_\delta(\bar{x})$ si ha che:

$$\lim_{t \rightarrow \infty} \|x(t) - \bar{x}\| = 0$$

Definizione 1.2.3. Un punto di equilibrio \bar{x} è detto **isolato** se esiste un intorno di \bar{x} che non contiene altri punti di equilibrio.

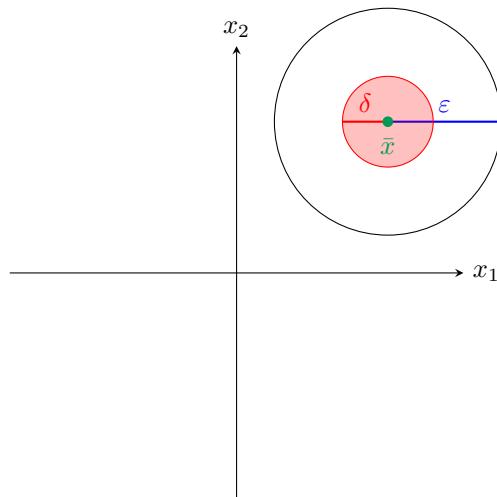


Figura 1.1: In verde si ha il punto di equilibrio \bar{x} . In rosso si ha l'insieme delle condizioni iniziali per cui le traiettorie rimangono confinate all'interno del cerchio di raggio ε .

1.3 Stabilità nei sistemi lineari

Consideriamo un sistema lineare tempo invariante:

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

Dove $A \in \mathbb{R}^{n \times n}$ e $B \in \mathbb{R}^{n \times m}$. Per trovare i punti di equilibrio dobbiamo imporre:

$$\dot{x}(t) = 0 \iff Ax(t) + Bu(t) = 0$$

Dunque i punti di equilibrio sono le coppie (\bar{x}, \bar{u}) che soddisfano:

$$A\bar{x} + B\bar{u} = 0$$

Notiamo subito che il punto $(\bar{x}, \bar{u}) = (0, 0)$ è sicuro un punto di equilibrio. Mentre gli altri punti si trovano risolvendo il sistema lineare omogeneo considerando \bar{x} e \bar{u} come incognite. Se invece di considerare le coppie (\bar{x}, \bar{u}) di equilibrio consideriamo solo gli stati di equilibrio \bar{x} con un fissato \bar{u} , allora l'unica incognita è \bar{x} mentre $B\bar{u}$ è un termine noto. In questo caso il sistema lineare da risolvere è:

$$A\bar{x} = -B\bar{u}$$

Questo sistema al variare di \bar{u} (che funge da parametro) ammette soluzioni differenti, però la matrice A ci dice quante sono queste soluzioni. Se A è invertibile (ovvero $\det(A) \neq 0$) allora esiste un'unica soluzione per ogni \bar{u} :

$$\bar{x} = -A^{-1}B\bar{u}$$

Se invece A non è invertibile (ovvero $\det(A) = 0$) allora il sistema può ammettere infinite soluzioni oppure nessuna soluzione (ricordiamo che il sistema si risolve al variare del parametro \bar{u}).

1.4 Sistema nel dominio di Laplace

Consideriamo il sistema lineare tempo invariante:

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

Applichiamo la trasformata di Laplace ambo i membri otteniamo:

$$sX(s) - x(0) = AX(s) + BU(s)$$

Portando $x(0)$ a destra e portando $AX(s)$ a sinistra otteniamo:

$$sX(s) - AX(s) = BU(s) + x(0)$$

Ricordando che $sX(s) = sIX(s)$ e mettendo in evidenza $X(s)$ otteniamo:

$$(sI - A)X(s) = BU(s) + x(0)$$

Ora premoltiplichiamo ambo i membri per $(sI - A)^{-1}$ otteniamo:

$$X(s) = \color{red}{(sI - A)^{-1}BU(s)} + \color{blue}{(sI - A)^{-1}x(0)}$$

Dove in **rosso** abbiamo la parte dovuta all'ingresso che prende il nome di **evoluzione forzata** dello stato, mentre in **blu** abbiamo la parte dovuta alla **evoluzione libera** dello stato (che ricordiamo essere nulla se $x(0) = 0$). Andando a fare la stessa cosa per l'equazione di uscita:

$$y(t) = Cx(t) + Du(t)$$

Applichiamo la trasformata di Laplace otteniamo:

$$Y(s) = CX(s) + DU(s)$$

Sostituendo $X(s)$ otteniamo:

$$Y(s) = C \left[(sI - A)^{-1} BU(s) + (sI - A)^{-1} x(0) \right] + DU(s)$$

Dove riusciamo un'altra volta a distinguere una parte in **evoluzione forzata** ed un'una in **evoluzione libera**:

$$Y(s) = C \left[(sI - A)^{-1} + D \right] BU(s) + (sI - A)^{-1} x(0)$$

1.5 Stabilità tramite l'approccio alla Lyapunov

Ora ci interessa introdurre un metodo per studiare la stabilità dei sistemi dinamici che prende il nome di **approccio di Lyapunov**, che ci permette di studiare la stabilità dei sistemi andandoci a trovare delle specifiche funzioni scalari $V(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ dette **funzioni di Lyapunov**. Per prima cosa ricordiamo che una funzione $V(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ e un suo punto di equilibrio possono essere classificati come:

- **definita positiva** in \bar{x} se $V(\bar{x}) = 0$ ed esiste un $\delta > 0$ tale che $V(x) > 0$ (notiamo il maggiore stretto) per ogni $x \in B_\delta(\bar{x})$
- **semidefinita positiva** in \bar{x} se $V(\bar{x}) = 0$ ed esiste un $\delta > 0$ tale che $V(x) \geq 0$ per ogni $x \in B_\delta(\bar{x})$
- **definita negativa** in \bar{x} se $V(\bar{x}) = 0$ ed esiste un $\delta > 0$ tale che $V(x) < 0$ per ogni $x \in B_\delta(\bar{x})$
- **semidefinita negativa** in \bar{x} se $V(\bar{x}) = 0$ ed esiste un $\delta > 0$ tale che $V(x) \leq 0$ per ogni $x \in B_\delta(\bar{x})$

Queste proprietà diventano **globali** quando sono vere per ogni $\bar{x} \in \mathbb{R}^n$. La potenza del metodo che stiamo per introdurre sta nel fatto che è valida anche per sistemi non lineari, nel momento in cui aggiungeremo l'ipotesi di linearità andremo ad ottenere anche altri risultati.

Teorema 1.5.1. *Teorema di Lyapunov. Consideriamo il sistema dinamico:*

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

Consideriamo il punto di equilibrio \bar{x} per f . Se f è definita, è continua ed anche la sua derivata è continua in un intorno D di \bar{x} , cioè $f \in C^1(D)$ (f è un vettore di funzioni quindi quando diciamo che è continua intendiamo che ogni sua componente lo è). Se esiste una funzione $V(x) : \mathbb{R}^n \rightarrow \mathbb{R}$, continua e derivabile in D , tale che:

- $V(x)$ è definita positiva in \bar{x} , cioè:

$$\begin{cases} V(\bar{x}) = 0 \\ \exists \delta : V(x) > 0, \quad \forall x \in B_\delta(\bar{x}) \setminus \{\bar{x}\} \end{cases}$$

- $\dot{V}(x)$ è semidefinita negativa in \bar{x} , cioè:

$$\begin{cases} V(\bar{x}) = 0 \\ \exists \delta > 0 : \dot{V}(x) = \nabla V(x) \cdot \frac{dx}{dt} = \nabla V(x(t)) f(x(t)) \leq 0, \quad \forall x \in B_\delta(\bar{x}), \forall t > t_0 \end{cases}$$

Allora si ha che il punto di equilibrio \bar{x} è stabile.

Dimostrazione. Noi vogliamo dimostrare che \bar{x} è un punto di equilibrio stabile, cioè che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che se $x_0 \in B_\delta(\bar{x})$ allora la traiettoria $x(t)$ che parte da x_0 rimane confinata in $B_\varepsilon(\bar{x})$. Dunque partiamo con il fissare un generico $\varepsilon > 0$. Prendiamo la curva di livello (della funzione $V(x)$) a valore maggiore e completamente contenuta in $\bar{B}_\varepsilon(\bar{x})$ (con il trattino sopra intendiamo la chiusura), come in Figura ???. Ora chiamiamo δ la distanza minima tra \bar{x} e la curva di livello di valore \bar{V} . Ora consideriamo una $x(t)$ che parte da un punto $x_0 \in B_\delta(\bar{x})$ ($x(t_0) = x_0$). A noi interessa dimostrare che \bar{x} è stabile, cioè che $\forall x_0 \in B_\delta(\bar{x})$ la traiettoria $x(t)$ rimane confinata all'interno di $B_\varepsilon(\bar{x}) \forall t \geq 0$. Per dimostrare che $x(t)$ rimane confinata in $B_\varepsilon(\bar{x})$ distinguiamo due casi:

- Nel caso in cui si ha $\dot{V}(x) = 0$ la traiettoria $x(t)$ nel peggior caso rimane confinata sulla curva di livello a valore \bar{V} , o su un'altra curva di livello a valore minore di \bar{V} (dunque rimane confinata in $B_\varepsilon(\bar{x})$).
- Se $\dot{V}(x) < 0$ allora la traiettoria $x(t)$ si muove in modo tale da far diminuire il valore di $V(x(t))$, quindi si muove verso l'interno (essendo \bar{V} la curva di livello a valore massimo) della curva di livello a valore \bar{V} (dunque rimane confinata in $B_\varepsilon(\bar{x})$)

Quindi in entrambi i casi la traiettoria $x(t)$ rimane confinata in $B_\varepsilon(\bar{x})$. Dunque abbiamo dimostrato che per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che se $x_0 \in B_\delta(\bar{x})$ allora la traiettoria $x(t)$ rimane confinata in $B_\varepsilon(\bar{x}) \forall t \geq t_0$, ma questa è proprio la definizione di stabilità.

□

Se $f \in C^1(D)$ allora è anche sicuramente localmente Lipschitziana in D (poichè se una funzione è C^1 allora il suo gradiente è limitato) cioè:

$$\exists L > 0 : \|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in I$$

1.6 Lyapunov per Sistemi LTI

Per i sistemi lineari tempo invarianti possiamo utilizzare il metodo di Lyapunov per studiare l'asintotica stabilità del sistema. Consideriamo il sistema LTI:

$$\dot{x} = Ax + Bu$$

Ricordiamo che per tutti i sistemi linearari l'asintotica stabilità è una proprietà del sistema, cioè se un punto di equilibrio è asintoticamente stabile allora lo sono tutti. Tutti i sistemi LTI ammettono come punto di equilibrio l'origine $(\bar{x}, \bar{u}) = (0, 0)$, e quindi possiamo anche andare a studiare direttamente la stabilità di questo punto di equilibrio (che ci permette di

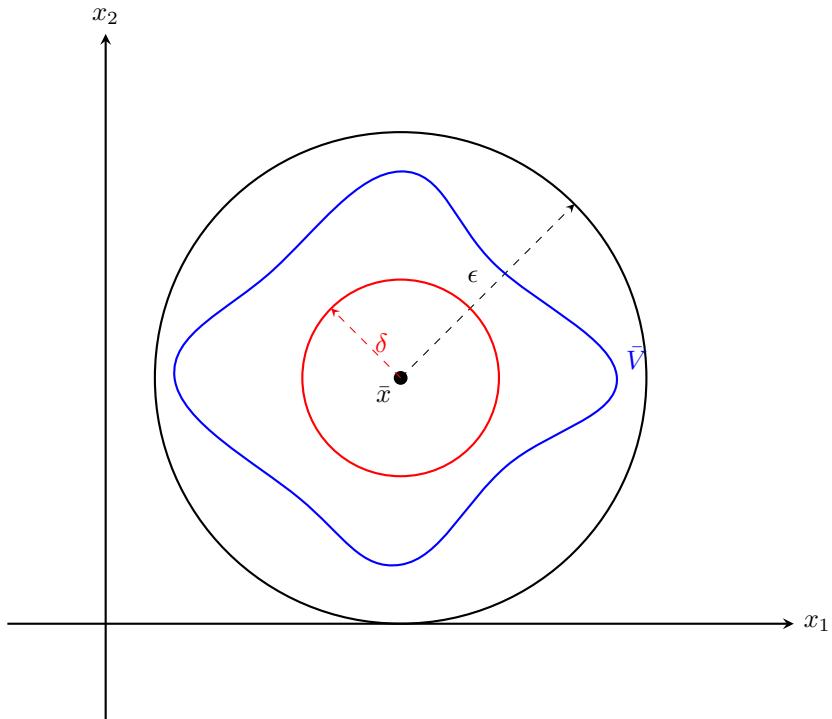


Figura 1.2: Visualizzazione con una curva di livello \bar{V} (blu) di forma generica non ellittica, strettamente contenuta all'interno della circonferenza di raggio ϵ (nera).

non considerare gli ingressi). Dunque consideriamo direttamente il sistema autonomo (che si evolve senza ingressi):

$$\dot{x} = Ax$$

Studiamo con il metodo di Lyapunov la stabilità dell'origine (nello spazio dello stato e degli ingressi) come punto di equilibrio. Per i sistemi LTI vale il seguente teorema:

Teorema 1.6.1. *Teorema di Lyapunov per sistemi LTI. Il sistema LTI:*

$$\dot{x} = Ax$$

è stabile se e solo se per ogni matrice Q simmetrica definita positiva ($Q = Q^T > 0$), esiste una matrice simmetrica definita positiva P ($P = P^T > 0$) tale che:

$$A^T P + PA = -Q \quad (1.3)$$

*quest'ultima equazione prende il nome di **equazione di Lyapunov**.*

Dimostrazione. Partiamo con il mostrare che l'esistenza di P che soddisfa l'equazione di Lyapunov è una **condizione necessaria** alla asintotica stabilità. Prendiamo una generica $Q = Q^T > 0$, vogliamo dimostrare che supposto il sistema sia asintoticamente stabile, allora deve esistere una $P = P^T > 0$ che soddisfa l'equazione di Lyapunov:

$$A^T P + PA = -Q$$

Consideriamo la P definita come:

$$P = \int_0^{+\infty} e^{A^T t} Q e^{At} dt$$

Per prima cosa diciamo che questa P è ben definita (l'integrale converge $\forall Q$) poichè il sistema è asintoticamente stabile, dunque gli esponenziali che usciranno nei calcoli avranno tutti esponenti negativi, dunque si avrà $e^{\lambda t} \rightarrow 0$ per $t \rightarrow +\infty$ (poichè $\lambda < 0$).

Per costruzione questa P è simmetrica. Dimostriamo ora che questa P è definita positiva. Preso un generico $x \in \mathbb{R}^n - \{0\}$ vogliamo dimostrare che $x^T P x > 0$. Calcoliamo:

$$x^T P x = x^T \left(\int_0^{+\infty} e^{A^T t} Q e^{At} dt \right) x$$

Portiamo le x dentro l'integrale:

$$x^T P x = \int_0^{+\infty} x^T e^{A^T t} Q e^{At} x dt$$

Notiamo ora che a primo membro abbiamo $x^T e^{A^T t} = (e^{At} x)^T$:

$$x^T P x = \int_0^{+\infty} (e^{At} x)^T Q (e^{At} x) dt$$

Ora poniamo $y = e^{At} x$, ci interessa dimostrare che:

$$\int_0^{+\infty} y^T Q y dt > 0$$

Ora noi sappiamo che Q è definita positiva, dunque per definizione di matrice definita positiva si ha che:

$$y^T Q y > 0, \forall y \neq 0$$

Il nostro problema è diventato dunque dimostrare che $y \neq 0$ per ogni $x \neq 0$. Ricordiamo che l'esponenziale di matrice è sempre invertibile, dunque la dimensione del $\ker(e^{At})$ è nulla, quindi $y = e^{At} x \neq 0 \forall x \neq 0$. Dunque abbiamo dimostrato che:

$$y^T Q y > 0, \forall x \neq 0$$

Quindi abbiamo dimostrato che l'integrando è sempre positivo per ogni $t \geq 0$. Ma se l'integrandino è sempre positivo anche l'intergrale è positivo dunque la matrice P è definita positiva.

Ora dimostriamo che questa $P > 0$ soddisfa l'equazione di Lyapunov. Calcoliamo:

$$A^T P + P A = A^T \int_0^{+\infty} e^{A^T t} Q e^{At} dt + \int_0^{+\infty} e^{A^T t} Q e^{At} dt A$$

Portando A^T e A dentro gli integrali, ed unendo gli integrali otteniamo:

$$A^T P + P A = \int_0^{+\infty} A^T e^{A^T t} Q e^{At} dt + e^{A^T t} Q e^{At} dt A$$

Ora notiamo che il termine all'interno dell'integrale è la derivata di $e^{A^T t} Q e^{At}$ rispetto a t , dunque possiamo riscrivere:

$$A^T P + P A = \int_0^{+\infty} \frac{d}{dt} (e^{A^T t} Q e^{At}) dt$$

Ma l'integrale della derivata è proprio la funzione valutata agli estremi:

$$A^T P + P A = \left[e^{A^T t} Q e^{At} \right]_0^{+\infty}$$

Ma

$$e^{A^T t} Q e^{At}$$

valutata in 0 vale Q , mentre valutata in $+\infty$ vale 0 (poichè il sistema è asintoticamente stabile), dunque si ha:

$$A^T P + PA = 0 - Q = -Q$$

Dunque abbiamo dimostrato che se il sistema è asintoticamente stabile allora per ogni $Q = Q^T > 0$ esiste una $P = P^T > 0$ che soddisfa l'equazione di Lyapunov.

Ora dimostriamo la **condizione sufficiente**, cioè se $\forall Q = Q^T > 0$ esiste una $P = P^T > 0$ che soddisfa l'equazione di Lyapunov allora il sistema è asintoticamente stabile. Per ipotesi dunque sappiamo che fissata una $Q = Q^T > 0$ esiste almeno una $P = P^T > 0$ che soddisfa l'equazione di Lyapunov, cioè:

$$A^T P + PA = -Q$$

Il nostro obiettivo è dimostrare che il sistema è asintoticamente stabile, dunque andiamo a considerare la funzione di Lyapunov:

$$V(x) = x^T Px$$

Noi sappiamo che P è definita positiva, dunque $V(x)$ è definita positiva. Calcoliamo ora la derivata di $V(x)$ nel tempo:

$$\dot{V}(x) = \dot{x}^T Px + x^T P \dot{x}$$

Ricordiamo che $\dot{x} = Ax$, dunque sostituiamo:

$$\dot{V}(x) = (Ax)^T Px + x^T P(Ax)$$

Facendo il trasposto a primo membro:

$$\dot{V}(x) = x^T A^T Px + x^T PAx$$

Mettiamo in evidenza x^T e x :

$$\dot{V}(x) = x^T (A^T P + PA)x$$

Ora ricordiamo che $A^T P + PA = -Q$, dunque sostituiamo:

$$\dot{V}(x) = x^T (-Q)x$$

Ma Q è definita positiva, dunque $-Q$ è definita negativa, dunque anche $\dot{V}(x)$ è definita negativa. Quindi il sistema è asintoticamente stabile per il teorema di Lyapunov, con funzione di Lyapunov $V(x) = x^T Px$. \square

Notiamo che in realtà non è necessario che la matrice P esista per ogni matrice Q definita positiva, ma basta che esista per una sola matrice Q definita positiva. Notiamo anche che la matrice P che soddisfa l'equazione di Lyapunov è unica per ogni Q .

1.7 Stabilità Robusta

Supponiamo ora che il nostro sistema dinamico sia affetto da incertezze parametriche, cioè che il modello matematico (dunque la matrice delle dinamica A) abbia delle variazioni intorno al suo valore nominale. In questo caso il sistema dinamico si può scrivere come

$$\dot{x}(t) = (A(p(t)))x(t) \quad (1.4)$$

dove $p(t)$ è un parametro incerto che varia nel tempo ed appartiene ad un insieme \mathcal{P} . In generale si studiano sistemi di 3 tipologie:

- **Sistemi Stazionari** : in cui il parametro $p(t) = p \in \mathcal{P}$ non è funzione del tempo (ma rimane incerto).
- **Sistemi Tempo-Varianti** : in cui il parametro $p(t) \in \mathcal{P}$ varia nel tempo in modo arbitrario.
- **Sistemi Quasi Stazionari** : in cui il parametro $p(t) \in \mathcal{P}$ varia nel tempo in modo "lento" (caso per caso si definisce cosa significa lento).

Noi studieremo solo i sistemi stazionari e tempo-varianti.

Partiamo con l'analizzare i **sistemi stazionari**. Dunque il nostro sistema dinamico si può scrivere come:

$$\dot{x} = A(p)x$$

con $p \in \mathcal{P}$. Noi vogliamo studiare la stabilità del sistema per ogni valore del parametro incerto p , ma per fare questo ci basta studiare gli autovalori della matrice A al variare di p . Dunque il sistema è asintoticamente stabile se e solo se per ogni $p \in \mathcal{P}$ gli autovalori di $A(p)$ hanno parte reale negativa. Potrei affrontare anche questo problema con un approccio alla Lyapunov, ricordiamo che un sistema LTI è asintoticamente stabile se e solo se $\forall Q = Q^T > 0$ esiste una $P = P^T > 0$ tale che:

$$A^T(p)P + PA(p) = -Q$$

Quindi dovremmo andare a cercare una P che soddisfa questa equazione per ogni $p \in \mathcal{P}$. Sia con l'approccio agli autovalori che con l'approccio di Lyapunov il problema è che dobbiamo studiare un'infinità continua di casi, su questo problema ritorneremo a breve, diamo ora una veloce occhiata ai sistemi in cui il parametro varia con il tempo.

Andiamo ora a considerare i **sistemi tempo-varianti**, dunque il nostro sistema dinamico si può scrivere come:

$$\dot{x} = A(p(t))x$$

con $p(t) \in \mathcal{P}$. Visto che il parametro varia nel tempo, non possiamo più studiare la stabilità del sistema andando a studiare gli autovalori di $A(p(t))$, poichè si dimostra che anche se la matrice $A(p(t))$ ha tutti gli autovalori con parte reale negativa per ogni t , il sistema potrebbe essere instabile. Proviamo ad affrontare il problema con il teorema di Lyapunov, ricordando che questo vale anche per sistemi tempo-varianti. Il nostro è un sistema lineare tempo-variante, ci manca solo la tempo-invarianza per usare il teorema di Lyapunov per sistemi LTI. Ma se

trovassimo una matrice P che soddisfa l'equazione di Lyapunov per tutti i possibili valori di $p \in \mathcal{P}$, cioè una P tale che:

$$A^T(p)P + PA(p) = -Q, \quad \forall p \in \mathcal{P}$$

allora potremmo usare questa P per costruire una funzione di Lyapunov per il sistema tempo-variante, che è identica a quella che useremmo per il sistema LTI, cioè:

$$V(x) = x^T Px$$

Infatti se la P soddisfa l'equazione di Lyapunov per ogni $p \in \mathcal{P}$, allora la derivata di $V(x)$ nel tempo sarà:

$$\dot{V}(x) < 0, \quad \forall p \in \mathcal{P} \implies \dot{V}(x) < 0 \quad \forall p \in \mathcal{P}, \quad \forall t$$

dunque il sistema sarà asintoticamente stabile per il teorema di Lyapunov (non quello per i sistemi LTI, con i sistemi LTI abbiamo in comune solo la forma della $V(x)$). Ma anche in questo caso dovremmo scrivere un infinità continua di equazioni di Lyapunov, una per ogni $p \in \mathcal{P}$.

Dunque in entrambi i casi (sistemi stazionari e tempo-varianti) dobbiamo risolvere un'infinità continua di equazioni di Lyapunov, una per ogni $p \in \mathcal{P}$, e questo è un problema. Per risolvere questo problema notiamo per prima cosa che per come è costruita la funzione di Lyapunov $V(x) = x^T Px$, non c'è bisogno che fissata una Q esista una P che soddisfa l'equazione di Lyapunov per ogni $p \in \mathcal{P}$, ma in realtà ci basta trovare una P che soddisfa la seguente diseguaglianza per ogni $p \in \mathcal{P}$:

$$A^T(p)P + PA(p) < 0, \quad \forall p \in \mathcal{P}$$

Infatti se questa diseguaglianza è soddisfatta, allora la derivata di $V(x)$ nel tempo sarà:

$$\dot{V}(x) < 0, \quad \forall p \in \mathcal{P}$$

Dunque ora invece di avere un numero infinito di equazioni, abbiamo un numero infinito di diseguaglianze, una per ogni $p \in \mathcal{P}$. Ora per portarci il problema ad un numero finito di diseguaglianze, dobbiamo fare delle ipotesi sull'insieme \mathcal{P} , cioè che l'insieme sia un iperrettangolare, cioè:

$$\mathcal{P} = [p_1^{\min}, p_1^{\max}] \times \dots \times [p_m^{\min}, p_m^{\max}]$$

con m il numero di parametri incerti, e $p \in \mathbb{R}^{m \times 1}$. Sotto questa ipotesi è possibile dimostrare che condizione sufficiente, affinchè la diseguaglianza è soddisfatta per ogni $p \in \mathcal{P}$, è che sia soddisfatta sui vertici dell'iperrettangolo, cioè ci basta trovare una P che soddisfa le diseguaglianze solo sui vertici dell'iperrettangolo:

$$A^T(p_i)P + PA(p_i) < 0, \quad \forall p_i \in \text{vertici}(\mathcal{P}) \implies A^T(p)P + PA(p) < 0, \quad \forall p \in \mathcal{P}$$

Capitolo 2

Sistemi MIMO nel dominio di Laplace

2.1 Rappresentazione di un sistema MIMO nel dominio di Laplace

Consideriamo il seguente sistema LTI a più ingressi e più uscite (MIMO) descritto dalle seguenti equazioni:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad (2.1)$$

Il sistema è MIMO dunque $u \in \mathbb{R}^{m \times 1}$, $y \in \mathbb{R}^{p \times 1}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ e $D \in \mathbb{R}^{p \times m}$. Il sistema può essere rappresentato nel dominio di Laplace come:

$$Y(s) = G(s)U(s) \quad (2.2)$$

Con $G(s) \in \mathbb{R}^{p \times m}$ matrice di trasferimento del sistema definita come:

$$G(s) = C(sI - A)^{-1}B + D \quad (2.3)$$

La matrice $G(s)$ è composta da $p \times m$ funzioni di trasferimento singole $G_{ij}(s)$:

$$G(s) = \begin{pmatrix} G_{11}(s) & \dots & G_{1m}(s) \\ \vdots & \vdots & \vdots \\ G_{p1}(s) & \dots & G_{pm}(s) \end{pmatrix}$$

dove ogni elemento $G_{ij}(s)$ è una funzione polinomiale fratta che rappresenta la funzione di trasferimento tra l'ingresso u_j e l'uscita y_i .

2.2 Sistema con retroazione dell'uscita MIMO

Consideriamo lo schema di controllo a retroazione dell'uscita per un sistema MIMO mostrato in Figura 2.1. Andiamo ora a ricavarci le funzioni di trasferimento del sistema considerando come uscite i due vettori $Y(s) \in \mathbb{R}^p$ e $U(s) \in \mathbb{R}^m$, e come ingresso i quattro vettori $R(s), D_u(s), D_y(s), N(s)$. Per prima cosa definiamo la funzione d'anello aperto $L(s)$ è definita come:

$$L(s) := G(s)K(s)$$

Notiamo che $K(s)G(s)$ a priori è diverso poichè stiamo parlando di matrici ed il prodotto tra matrici non è commutativo. Definiamo anche la **funzione di sensitività**:

$$S(s) := (I + L(s))^{-1}$$

e la **funzione di sensitività complementare**:

$$T(s) := (I + L(s))^{-1} L(s)$$

Andiamo a trovarci le funzioni di trasferimento tra le uscite e gli ingressi, nello specifico saranno otto funzioni di trasferimento:

1. Funzione di trasferimento tra $R(s)$ a $Y(s)$, che vale:

$$K_{YR} = T(s) = (I + L(s))^{-1} L(s)$$

2. Funzione di trasferimento tra $Y(s)$ e $D_u(s)$, che vale:

$$K_{YD_u}(s) := (I + L(s))^{-1} G(s) = S(s)G(s)$$

3. Funzione di trasferimento tra $Y(s)$ e $D_y(s)$, che vale:

$$K_{YD_y}(s) := (I + L(s))^{-1} = S(s)$$

4. Funzione di trasferimento tra $Y(s)$ e $N(s)$, che vale:

$$K_{YN}(s) := - (I + L(s))^{-1} L(s) = -T(s)$$

5. Funzione di trasferimento tra $U(s)$ e $R(s)$, che vale:

$$K_{UR}(s) (I + K(s)G(s))^{-1} K(s)$$

6. Funzione di trasferimento tra $U(s)$ e $D_u(s)$, che vale:

$$(I + K(s)G(s))^{-1}$$

7. Funzione di trasferimento tra $U(s)$ e $D_y(s)$, che vale:

$$- (I + K(s)G(s))^{-1} K(s)$$

8. Funzione di trasferimento tra $U(s)$ e $N(s)$, che vale:

$$- (I + K(s)G(s))^{-1} K(s)$$

Dunque il legame tra $Y(s)$ e $U(s)$ con gli ingressi $R(s), D_u(s), D_y(s), N(s)$ è dato da:

$$Y(s) = S(s)G(s)D_u(s) + S(s)D_y(s) - T(s)N(s) + T(s)R(s)$$

$$U(s) = (I + L_u)^{-1} K(s)(R(s) - D_y(s) - N(s)) + (I + L_u(s))^{-1} D_u(s)$$

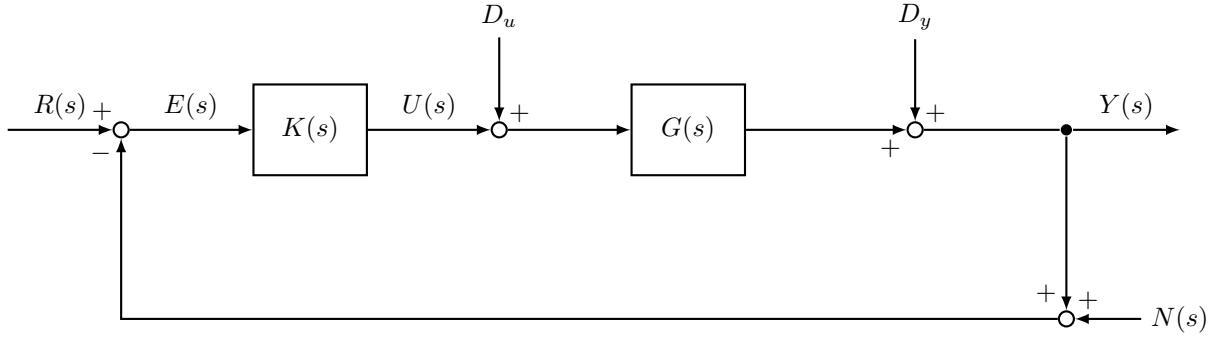


Figura 2.1: Schema a blocchi di un sistema MIMO con retroazione dell'uscita.

2.3 Guadagno

Definiamo gli spazi L^p come gli spazi delle funzioni p -sommabili, cioè in cui esiste la norma p -esima finita:

$$L^p(\mathbb{R}^\times) = \left\{ f : \mathbb{R} \rightarrow \mathbb{R}^n : \int_{-\infty}^{+\infty} \|f(\tau)\|_p^p d\tau < \infty \right\}$$

Dove essendo f una funzione vettoriale a variabile reale la norma p -esima è definita come:

$$\|f(\tau)\|_p = \left(\sum_{i=1}^n |f_i(\tau)|^p \right)^{\frac{1}{p}}$$

Dove con $f_i(\tau)$ si intende la i -esima componente del vettore $f(\tau)$ (che è un vettore di n componenti in cui ognuna è una funzione di τ)

2.4 Poli e zeri per sistemi MIMO

Ora ci poniamo il problema di definire poli e zeri per sistemi MIMO. Per i sistemi SISO i poli sono definiti come gli zeri del denominatore della funzione di trasferimento, mentre gli zeri sono definiti come gli zeri del numeratore della funzione di trasferimento. Il problema è che per i sistemi MIMO ho una matrice di trasferimento $G(s)$ composta da pm funzioni di trasferimento singole $G_{ij}(s)$, quindi bisogna dare una nuova definizione di poli e zeri per sistemi MIMO. Per prima cosa diamo la definizione di polo per sistemi MIMO.

Definizione 2.4.1. Un polo di un sistema MIMO è un valore di s tale che la matrice $G(s)$ diventa singolare, cioè il suo determinante è nullo, dunque per quel valore di s la matrice $G(s)$ non è invertibile.

Definizione 2.4.2. Si definisce rango nominale di una matrice di polinomi $G(s)$, il rango della matrice che si ha per tutti i valori di s eccetto un numero finito di valori.

Definizione 2.4.3. Uno zero di un sistema MIMO è un valore di s tale che la matrice $G(s)$ perde rango, cioè il rango della matrice $G(s)$ diminuisce rispetto al suo rango nominale (quello che ha per tutti gli altri valori di s).

Capitolo 3

Assegnamento degli autovalori con feedback di stato

3.1 Caso MIMO

Consideriamo un sistema LTI descritto dalle equazioni di stato:

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

con $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $x \in \mathbb{R}^n$ e $u \in \mathbb{R}^m$. Il nostro obiettivo è attraverso un feedback di stato del tipo:

$$u(t) = Kx(t) \quad (3.1)$$

con $K \in \mathbb{R}^{m \times n}$ (dunque $Kx \in \mathbb{R}^{m \times 1}$) riuscire ad assegnare gli autovalori del sistema con matrice della dinamica:

$$A + BK \quad (3.2)$$

indicheremo l'insieme degli autovalori di una matrice A con $\sigma(A)$, dunque noi vogliamo scegliere K in modo tale da poter assegnare l'insieme $\sigma(A + BK)$ a piacere:

$$\sigma(A + BK) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

Per i sistemi MIMO esistono infinite soluzioni al problema di assegnamento degli autovalori, cioè esistono infinite matrici K tali che l'insieme degli autovalori di $A + BK$ sia uguale all'insieme desiderato. Noi andremo a vedere delle soluzioni operative che ci permettono di trovare alcune di queste infinite matrici K soluzioni del problema, nello specifico vedremo i seguenti 3 metodi:

- Metodo della forma canonica di controllabilità
- Metodo dell'equazione di Sylvester
- Metodo di Kautsky-Nichols-Van Dooren

3.2 Metodo della forma canonica di controllabilità

Per sistemi MIMO se la coppia (A, B) è completamente controllabile allora è possibile trovare una matrice di trasformazione T che porta il sistema in una forma canonica di controllabilità

a blocchi, detta forma di Brunovsky:

$$z = T^{-1}x \iff x = Tz$$

Con le matrici del sistema nella nuova base che valgono:

$$A_c = T^{-1}AT = \begin{pmatrix} A_{c1} & 0 & \cdots & 0 \\ 0 & A_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{cm} \end{pmatrix}, \quad B_c = T^{-1}B = \begin{pmatrix} B_{c1} & 0 & \cdots & 0 \\ 0 & B_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{cm} \end{pmatrix} \quad (3.3)$$

Con $A_{ci} \in \mathbb{R}^{n_i \times n_i}$ e $B_{ci} \in \mathbb{R}^{n_i \times 1}$ che valgono:

$$A_{ci} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_{i1} & -a_{i2} & -a_{i3} & \cdots & -a_{in_i} \end{pmatrix}, \quad B_{ci} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Inoltre si ha che $n_1 + n_2 + \cdots + n_m = n$. Notiamo che ogni coppia (A_{ci}, B_{ci}) rappresenta una forma canonica di controllabilità rispetto all'ingresso u_i . Notiamo che per come è scritta la matrice A_c con una retroazione di stato del tipo:

$$u = K_c z$$

con $K_c \in \mathbb{R}^{m \times n}$:

$$K_c = \begin{pmatrix} k_{c1} & 0 & \cdots & 0 \\ 0 & k_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_{cm} \end{pmatrix}$$

con $k_{ci} \in \mathbb{R}^{1 \times n_i}$. Con K_c scritto in questo modo si ha che la matrice della dinamica del sistema in retroazione di stato è:

$$A_c + B_c K_c = \begin{pmatrix} A_{c1} & 0 & \cdots & 0 \\ 0 & A_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{cm} \end{pmatrix} + \begin{pmatrix} B_{c1} & 0 & \cdots & 0 \\ 0 & B_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{cm} \end{pmatrix} \begin{pmatrix} k_{c1} & 0 & \cdots & 0 \\ 0 & k_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_{cm} \end{pmatrix}$$

Notando che $B_{ci} k_{ci} \in \mathbb{R}^{n_i \times n_i}$ si ha che:

$$A_c + B_c K_c = \begin{pmatrix} A_{c1} + B_{c1} k_{c1} & 0 & \cdots & 0 \\ 0 & A_{c2} + B_{c2} k_{c2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{cm} + B_{cm} k_{cm} \end{pmatrix}$$

Ma la matrice $A_c + B_c + K_c$ è diagonale a blocchi dunque si ha che gli autovalori della matrice sono gli autovalori dei blocchi diagonali:

$$\sigma(A_c + B_c K_c) = \bigcup_{i=1}^m \sigma(A_{ci} + B_{ci} k_{ci})$$

Ma ogni singolo blocco diagonale $A_{ci} + B_{ci}k_{ci}$ rappresenta un sistema SISO in forma canonica di controllabilità, e la coppia (A_{ci}, B_{ci}) è completamente controllabile per ipotesi (essendo controllabile il sistema originale (A, B)), dunque per il teorema di assegnamento degli autovalori applicato ai sistemi SISO esiste k_{ci} tale che possiamo assegnare gli autovalori di ogni blocco diagonale $A_{ci} + B_{ci}k_{ci}$. Questo k_{ci} si trova imponendo che il polinomio caratteristico del blocco $A_{ci} + B_{ci}k_{ci}$ sia uguale al polinomio caratteristico desiderato:

$$\det(sI - (A_{ci} + B_{ci}k_{ci})) = s^{n_i} + \alpha_{i1}s^{n_i-1} + \alpha_{i2}s^{n_i-2} + \dots + \alpha_{in_i}$$

Ovviamente questo va fatto $\forall i = 1, \dots, m$. Una volta trovati tutti i k_{ci} possiamo costruire la matrice K_c , per riportarla nella base originale del sistema basta ricordare:

$$u = K_c z = K_c T^{-1} x \implies K = K_c T^{-1}$$

3.3 Metodo dell'equazione di Sylvester

Un altro metodo per l'assegnamento degli autovalori per sistemi MIMO è il metodo dell'equazione di Sylvester. L'equazione di Sylvester è la seguente equazione matriciale:

$$AX + XB = C \tag{3.4}$$

in cui $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$ e $C \in \mathbb{R}^{n \times m}$ sono matrici note, mentre $X \in \mathbb{R}^{n \times m}$ è la matrice incognita da trovare. Andiamo ad analizzare questa equazione:

- **Unicità della soluzione:** L'equazione di Sylvester ammette una soluzione unica se e solo se gli spettri (cioè gli insiemi degli autovalori) delle matrici A e $-B$ sono disgiunti, cioè hanno intersezione uguale all'insieme vuoto:

$$\sigma(A) \cap \sigma(-B) = \emptyset \iff \lambda_i(A) + \lambda_j(B) \neq 0 \quad \forall i, j$$

in parole povere le matrici A e B non devono avere autovalori con segni opposti, cioè A e $-B$ non devono avere autovalori in comune. Ad esempio se A ha un autovalore $\lambda_A = 2$ e B ha un autovalore $\lambda_B = -2$ allora l'equazione di Sylvester non ammette una soluzione unica.

- **Esistenza della soluzione:** L'equazione di Sylvester ammette almeno una soluzione se e solo se:

$$\text{vec}(C) \in \text{Im} (I_m \otimes A + B^T \otimes I_n)$$

cioè se la matrice C appartiene all'immagine dell'operatore $L(X)$:

$$L(X) = AX + XB$$

cioè esiste una matrice $X \in \mathbb{R}^{n \times m}$ tale che $L(X) = C$.

Ora torniamo al nostro problema di assegnamento degli autovalori per sistemi MIMO. Ricordiamo che il nostro obiettivo è trovare la matrice K tale che:

$$\sigma(A + BK) = \{\lambda_1, \lambda_2, \dots, \lambda_n\}$$

Ma se esiste questa K allora costruita una matrice T che ha sulle colonne gli autovettori associati agli autovalori desiderati si ha che per la definizione di autovettore e autovalore (poichè il prodotto $(A+BK)T$ equivale a moltiplicare la matrice $A+BK$ per ogni autovettore sulle colonne di T):

$$(A+BK)T = T\Lambda$$

con Λ matrice diagonale(o a blocchi) con gli autovalori desiderati sulla diagonale. Andando a sviluppare i calcoli si hanno

$$AT + BKT = T\Lambda$$

Portando il termine $T\Lambda$ a sinistra e BKT a destra si ha:

$$AT - T\Lambda = -BKT$$

In cui ponendo $F = -KT$ si ha:

$$AT - T\Lambda = BF$$

Questa è un'equazione di Sylvester (a meno del segno del secondo termine a primo membro) in cui $X = T$. Dunque se riusciamo a risolvere l'equazione di Sylvester troviamo una X che ci permette di ricavare una matrice K tale che la matrice $A+BK$ abbia gli autovalori desiderati, la matrice K si ricava come:

$$F = -KT = -KX \implies K = -FX^{-1} = -FT^{-1}$$

Nella pratica per assegnare gli autovalori tramite l'equazione di Sylvester si seguono i seguenti passi:

1. Si scelgono gli autovalori desiderati $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e si costruisce la matrice Λ con questi autovalori sulla diagonale (o a blocchi).
2. Si costruisce una matrice $F \in \mathbb{R}^{m \times n}$ a piacere (ad esempio con elementi casuali).
3. Si risolve l'equazione di Sylvester:

$$AX - X\Lambda = BF$$

la soluzione X trovata è proprio la matrice T degli autovettori associati agli autovalori desiderati.

4. Si calcola la matrice K come:

$$K = -FX^{-1} = -FT^{-1}$$

Ovviamente al variare della matrice F si ottengono differenti matrici K che risolvono il problema, dunque è sulla matrice F che possiamo agire per ottenere differenti soluzioni del problema di assegnamento degli autovalori. Andiamo ora a definire il condizionamento di una matrice. A noi interessa anche che la matrice K non vari troppo al variare delle matrici A , B e Λ , dunque ci interessa che trovare un parametro per valutare la sensibilità numerica della matrice K al variare delle matrici A , B e Λ . Ovviamente la sensibilità numerica di K dipende dalla sensibilità numerica della matrice X .

Definizione 3.3.1. Questo parametro prende il nome di condizionamento di una matrice ed è definito come:

$$cond(X) = \|X\| \|X^{-1}\| \quad (3.5)$$

Maggiore è il condizionamento di una matrice, maggiore è la sensibilità numerica della matrice, dunque il nostro obiettivo è trovare una matrice X con condizionamento il più basso possibile.

Andiamo ora trasformare la condizione di esistenza della soluzione dell'equazione di Sylvester in una condizione sulla coppia (A, B) . Nello specifico la soluzione dell'equazione di Sylvester esiste ed è unica se e solo se la coppia (A, B) è completamente raggiungibile. Infatti, se la coppia (A, B) non è completamente raggiungibile, il sistema può essere portato nella forma di Kalman di controllabilità, in cui \bar{A} e \bar{B} valgono:

$$\bar{A} = \begin{pmatrix} A_C & A_{12} \\ 0 & A_{NC} \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} B_C \\ 0 \end{pmatrix}$$

dove:

- La coppia (A_C, B_C) è completamente controllabile
- A_{NC} rappresenta la parte non raggiungibile del sistema.

Quindi se andiamo a scrivere l'equazione di Sylvester per \bar{A} e \bar{B} si ha:

$$\bar{A}X - X\Lambda = \bar{B}F$$

Dove le equazioni nelle righe di \bar{A} e \bar{B} in cui il sistema non è controllabile si riducono a:

$$A_{NC}X_{NC} - X_{NC}\Lambda = 0$$

Ma l'ultima equazione è soddisfatta solo se gli autovalori di A_{NC} e Λ sono uguali, cioè se non cambiamo gli autovalori della parte non controllabile del sistema. Facciamo ora una considerazione sulla scelta degli autovalori. Se siamo obbligati a scegliere autovalori di Λ uguali ad autovalori di A , per soddisfare delle specifiche di progetto, allora l'equazione di Sylvester non ammette più **soluzione unica**, e i casi possibili sono due:

- L'equazione di Sylvester non ammette soluzioni (il sistema non è raggiungibile)
- L'equazione di Sylvester ammette infinite soluzioni (il sistema è raggiungibile)

Nel primo caso abbiamo che non esiste una X che risolve l'equazione di Sylvester, dunque non possiamo seguire questa strada per trovare K . Nel secondo caso invece esistono infinite matrici X , dunque infinite matrici K che risolvono il problema di assegnamento degli autovalori. Per garantire un condizionamento migliore del problema invece di scegliere autovalori di Λ uguali ad autovalori di A è meglio scegliere autovalori di Λ vicini, ma non proprio uguali agli autovalori di A .

3.4 Metodo di Kautsky-Nichols-Van Dooren

Il metodo di Kautsky-Nichols-Van Dooren è un metodo iterativo per ottenere una matrice K , che risolve il problema di assegnamento degli autovalori per sistemi MIMO, con un buon

condizionamento. Ma il condizionamento di K ricordiamo dipendere dal condizionamento della matrice X (matrice di autovettori) che risolve l'equazione di Sylvester. Il metodo di Kautsky-Nichols-Van Dooren consiste dei seguenti passaggi:

1. Si scelgono gli autovalori desiderati $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e si costruisce la matrice Λ con questi autovalori sulla diagonale (o a blocchi).
2. Si scelgono i vettori f_i in modo casuale (uno per ogni autovalore)
3. Per ogni i si risolve il sistema di equazioni:

$$(A - \lambda_i I) x_i = -B f_i$$

Ottendendo così per ogni i un vettore x_i .

4. Si costruisce la matrice X con i vettori x_i come colonne:

$$X = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}$$

5. Si valuta il condizionamento di X :

$$\text{cond}(X) = \|X\| \|X^{-1}\|$$

6. Se il condizionamento di X non è accettabile si ritorna al passo 2 e si scelgono nuovi vettori f_i . Mentre se il condizionamento è accettabile si procede con il passo successivo.
7. Si calcola la matrice F come:

$$F = \begin{pmatrix} f_1 & f_2 & \dots & f_n \end{pmatrix}$$

8. Si calcola la matrice K come:

$$K = -FX^{-1}$$

Capitolo 4

Controllo di sistemi MIMO

4.1 Schema di retroazione dell'uscita con integratori

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Md(t) \\ y(t) = Cx(t) + Nd(t) \end{cases} \quad (4.1)$$

L'integratore è necessario per eliminare l'errore a regime, ora daremo una giustificazione logica di questa affermazione. Consideriamo che al posto dell'integratore ci sia un blocco $Z(s)$ generico come in figura ??, seguiamo il seguente ragionamento:

1. Noi vogliamo che l'uscita $y(t)$ segua un riferimento costante $r(t) = \bar{r}$, quindi vogliamo che a regime valga $y(t) \rightarrow \bar{y} = \bar{r}$, cioè:

$$\lim_{t \rightarrow +\infty} y(t) = \bar{y} = \bar{r}$$

2. Ma se l'uscita a regime è costante, allora per il teorema del valore finale
3. per $t \rightarrow +\infty$ che $y(t) \rightarrow \bar{y} = \bar{r}$.
4. Ma se $y(t)$ a regime è costante, allora per il teorema del valore finale anche le $u(t)$ e le $v(t)$ a regime devono essere costanti.
5. Inoltre se la $\bar{y} = \bar{r}$ si ha che a regime l'errore è nullo:

$$\bar{e} = \lim_{t \rightarrow +\infty} e(t) = \lim_{t \rightarrow +\infty} r(t) - y(t) = \bar{r} - \bar{y} = 0$$

6. Ma se l'errore è nullo a regime, allora si ha che in ingresso al blocco $Z(s)$ si ha un ingresso nullo, quindi il blocco $Z(s)$ a regime vede un ingresso nullo. Quindi il blocco $Z(s)$ deve essere un integratore, poiché è l'unico operatore che ad un ingresso nullo fornisce in uscita un segnale costante (diverso da zero).

Quindi abbiamo dimostrato che per avere errore nullo a regime il blocco $Z(s)$ deve essere proprio un integratore, cioè:

$$Z(s) = \frac{1}{s}$$

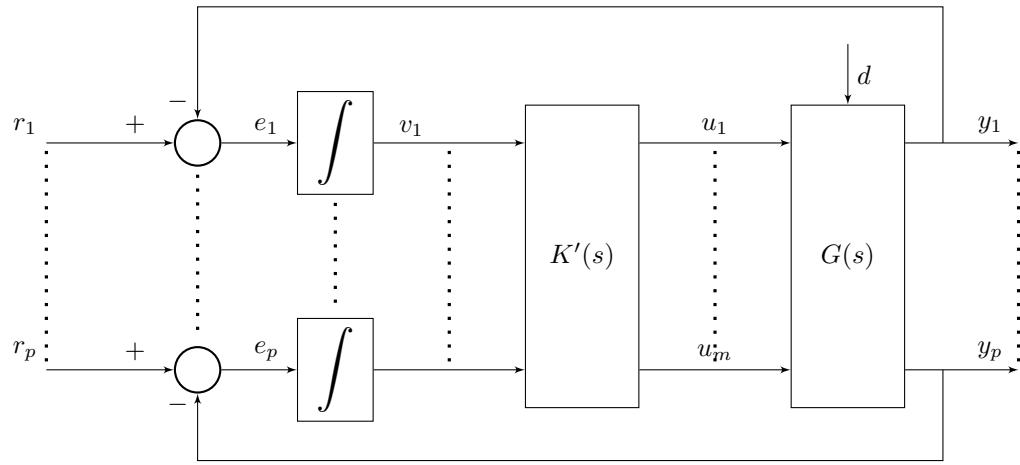


Figura 4.1: Schema di retroazione dell'uscita con integratori

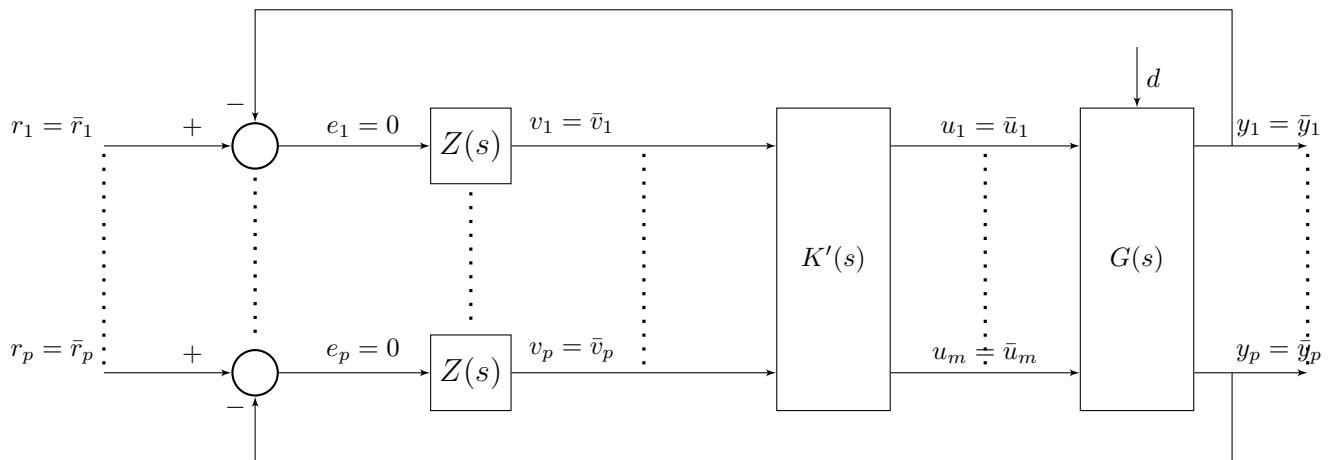


Figura 4.2: Schema per la dimostrazione che l'unico blocco che garantisce errore nullo a regime è l'integratore

Supponiamo ora che il nostro controllore $K'(s)$ riesca a stabilizzare il sistema, cioè che riesca a portare tutti a posizionare tutti poli del sistema aumentato in retroazione dell'uscita con integratori a sinistra dell'asse immaginario (cioè con parte reale negativa). Supposto questo vogliamo andare a studiarci le condizioni necessarie affinchè il sistema riesca a posizionare l'uscita $y(t)$ al valore di riferimento \bar{r} . In termine formali vogliamo studiare le condizioni necessarie affinché a regime si abbia per ogni \bar{y} e \bar{d} (che rappresentano rispettivamente il valore di uscita a regime e il disturbo a regime)

$$\begin{cases} 0 = A\bar{x} + B\bar{u} + M\bar{d} \\ \bar{y} = C\bar{x} + N\bar{d} \end{cases} \quad \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{u} \end{pmatrix} = \begin{pmatrix} -M & 0 \\ -N & I \end{pmatrix} \begin{pmatrix} \bar{d} \\ \bar{y} \end{pmatrix} \quad (4.2)$$

Questo è un sistema di $n+p$ equazioni in $n+m$ (\bar{x} e \bar{u} sono incognite). Definiamo la matrice Σ come:

$$\Sigma := \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathbb{R}^{(n+p) \times (n+m)} \quad (4.3)$$

Quello che andremo a fare ora è determinare delle **condizioni necessarie** sulla matrice Σ affinchè il sistema (4.2) ammetta soluzione $\forall N \in \mathbb{R}^{p \times r}, M \in \mathbb{R}^{n \times r}, \bar{d} \in \mathbb{R}^{r \times 1}, \bar{y} \in \mathbb{R}^{p \times 1}$. Per determinare queste condizioni necessarie andremo a trovare dei casi particolari di Σ in cui il sistema non ammette soluzione, dunque condizione necessaria alla soluzione del sistema è che la matrice Σ non cada in questi casi particolari.

Il **primo caso** particolare che andremo a considerare è quando il numero di incognite ($n+m$) del sistema è minore del numero di equazioni ($n+p$) del sistema, cioè:

$$n+m < n+p \quad (4.4)$$

Questo equivale a dire che il numero di colonne della matrice Σ è minore del numero di righe della matrice Σ . Ma questo significa che la matrice Σ non può essere di rango pieno per righe (poichè ha più righe che colonne, ed il rango ha come massimo valore dunque il numero di colonne). Se la matrice Σ non è di rango pieno per righe significa che esistono delle righe della matrice Σ che sono combinazioni lineari delle altre righe, dunque facendo delle operazioni elementari sulle righe (ricordiamo che queste operazioni non cambiano lo spazio delle soluzioni del sistema) possiamo arrivare rendere nulle queste righe linearmente dipendenti. Ma se una riga della matrice Σ è nulla significa che se a destra il termine noto assume un valore diverso da zero il sistema non ammette soluzione. Ma noi vogliamo che il sistema ammetta soluzione $\forall N \in \mathbb{R}^{p \times r}, M \in \mathbb{R}^{n \times r}, \bar{d} \in \mathbb{R}^{r \times 1}, \bar{y} \in \mathbb{R}^{p \times 1}$, questo significa che i termini noti possono assumere qualunque valore, quindi anche un valore diverso da zero. Quindi se sceglieremo le matrici M ed N e i vettori \bar{d} e \bar{y} in modo tale che il termine noto a destra della riga nulla sia diverso da zero abbiamo che il sistema non ammette soluzione. Dunque abbiamo trovato che nel caso particolare in cui:

$$n+m < n+p$$

il sistema non ammette soluzione $\forall N \in \mathbb{R}^{p \times r}, M \in \mathbb{R}^{n \times r}, \bar{d} \in \mathbb{R}^{r \times 1}, \bar{y} \in \mathbb{R}^{p \times 1}$. Quindi una condizione necessaria affinchè il sistema ammetta soluzione è che la condizione (4.4) non sia verificata, cioè si deve avere:

$$n + m \geq n + p \iff m \geq p$$

cioè il numero di ingressi deve essere maggiore o uguale al numero di uscite, cosa che era già intuibile, infatti se voglio controllare p uscite dovrò avere un numero di ingressi almeno maggiore o uguale a p . Ad esempio se io ho 5 uscite da controllare almeno dovrò avere 5 ingressi per poter controllare tutte le uscite.

Supponiamo ora di aver soddisfatto questa prima condizione necessaria, cioè di esserci messi nella condizione $m \geq p$. Andiamo ora a cercare un **secondo caso particolare** in cui il sistema non ammette soluzione. Un'altro caso particolare in cui il sistema non ammette soluzione è quando il rango della matrice Σ non è massimo, cioè (poichè siamo nella condizione $m \geq p$):

$$\text{rank}(\Sigma) < \min\{n + m, n + p\} = n + p$$

Infatti essendo il numero di righe della matrice Σ uguale a $n + p$, se il rango della matrice Σ è minore di $n + p$ significa che esistono delle righe della matrice Σ che sono combinazioni lineari delle altre righe, dunque come prima facendo delle operazioni elementari sulle righe (possibili perchè non cambiano lo spazio delle soluzioni del sistema) possiamo arrivare rendere nelle queste righe linearmente dipendenti, e mettendo a destra del sistema il termine noto diverso da zero arriviamo ad un sistema che non ammette soluzione. Dunque abbiamo trovato un'altro caso particolare in cui il sistema non ammette soluzione $\forall N \in \mathbb{R}^{p \times r}, M \in \mathbb{R}^{n \times r}, \bar{d} \in \mathbb{R}^{r \times 1}, \bar{y} \in \mathbb{R}^{p \times 1}$. Dunque affinche il sistema ammetta soluzione sicuramente dobbiamo evitare questo caso particolare, cioè dobbiamo fare in modo che il rango di Σ sia massimo, cioè dobbiamo avere:

$$\text{rank}(\Sigma) = n + p$$

Quindi le **condizioni necessarie** affinche il sistema ammetta soluzione $\forall N \in \mathbb{R}^{p \times r}, M \in \mathbb{R}^{n \times r}, \bar{d} \in \mathbb{R}^{r \times 1}, \bar{y} \in \mathbb{R}^{p \times 1}$

1) Il numero di ingressi deve essere maggiore o uguale al numero di uscite:

$$m \geq p \tag{4.5}$$

2) La matrice Σ deve essere di rango pieno:

$$\text{rank}(\Sigma) = \min\{n + m, n + p\} = n + p \tag{4.6}$$

La condizione (4.6) si può determinare in un modo più formale andando anche a considerare la stabilità del sistema in retroazione con integratori (cosa che fino ad ora abbiamo supposto). Questa condizione si può determinare andando a studiare il sistema aumentato in retroazione con integratori. Ricordiamo il vettore di stato del sistema aumentato è $(x(t) \ v(t))^T \in \mathbb{R}^{(n+p) \times 1}$, dunque le derivate degli stati valgono:

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) + Md(t) \\ \dot{v}(t) = e(t) = r(t) - y(t) = r(t) - Cx(t) - Nd(t) \end{cases}$$

Prendiamo come uscita del sistema aumentato $\tilde{y}(t) = v(t)$, dunque si ha che la rappresentazione in forma matriciale del sistema aumentato è:

$$\begin{cases} \begin{pmatrix} \dot{x}(t) \\ \dot{v}(t) \end{pmatrix} = \begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} + \begin{pmatrix} B \\ 0 \end{pmatrix} u(t) + \begin{pmatrix} M \\ -N \end{pmatrix} d(t) \\ \tilde{y}(t) = \begin{pmatrix} 0 & I \end{pmatrix} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \end{cases}$$

Definiamo \tilde{n} :

$$\tilde{n} := n + p$$

Dunque definiamo le seguenti matrici:

$$\tilde{A} := \begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}, \quad \tilde{B} := \begin{pmatrix} B \\ 0 \end{pmatrix} \in \mathbb{R}^{\tilde{n} \times m}, \quad \tilde{C} := \begin{pmatrix} 0 & I \end{pmatrix} \in \mathbb{R}^{p \times \tilde{n}}$$

Che hanno gli ovvi significati nel sistema aumentato. Il nostro obiettivo iniziale era trovare un controllore $K'(s)$ che stabilizzasse il sistema in retroazione con integratori, cioè che rendesse asintoticamente stabile il sistema aumentato. Condizione necessaria affinchè esista un controllore $K'(s)$ che stabilizzi il sistema aumentato è che il sistema aumentato sia:

- 1) Completamente raggiungibile
- 2) Completamente osservabile

Queste due condizioni equivalgono a dire che:

- 1) La coppia (\tilde{A}, \tilde{B}) è completamente raggiungibile
- 2) La coppia (\tilde{A}, \tilde{C}) è completamente osservabile

Quindi dobbiamo ora studiarci la raggiungibilità e l'osservabilità del sistema aumentato. Andiamo prima a studiare l'**osservabilità**, cioè il rango della matrice di osservabilità (ricordiamo che affinchè il sistema sia completamente osservabile il rango deve essere massimo cioè uguale a \tilde{n}):

$$M_O(\tilde{A}, \tilde{C}) = \underbrace{\begin{pmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \vdots \\ \tilde{C}\tilde{A}^{\tilde{n}-1} \end{pmatrix}}_{np \text{ righe}} \quad \text{con } n+p \text{ colonne}$$

Ricordando che $\tilde{C} = \begin{pmatrix} 0 & I \end{pmatrix}$ e $\tilde{C}\tilde{A} = \begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix}$ si ha :

$$\tilde{C}\tilde{A}^2 = \begin{pmatrix} 0 & I \end{pmatrix} \begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix} \begin{pmatrix} A & 0 \\ -C & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \end{pmatrix} \begin{pmatrix} A^2 & 0 \\ -CA & 0 \end{pmatrix} = \begin{pmatrix} -CA & 0 \end{pmatrix}$$

L'espressione generale $\forall k \in \mathbb{N}$ è:

$$\tilde{C}\tilde{A}^k = \begin{pmatrix} 0 & I \end{pmatrix} \begin{pmatrix} A^k & 0 \\ -CA^{k-1} & 0 \end{pmatrix} = \begin{pmatrix} -CA^{k-1} & 0 \end{pmatrix}$$

Quindi $M_O(\tilde{A}, \tilde{C})$ si può scrivere come:

$$M_O(\tilde{A}, \tilde{C}) = \left(\begin{array}{cc|c} 0 & I & p \text{ righe} \\ -C & 0 & \\ -CA & 0 & \\ \vdots & \vdots & (n-1)p \text{ righe} \\ -CA^{n-1} & 0 & \\ \vdots & \vdots & \\ -CA^{\tilde{n}-2} & 0 & \end{array} \right)$$

n colonne p colonne

Come si può vedere la matrice $M_O(\tilde{A}, \tilde{C})$ ha rango pieno $\text{rank}(M_O(\tilde{A}, \tilde{C})) = n + p$ se e solo se la sottomatrice composta dalle seguenti righe:

$$\begin{pmatrix} -C \\ -CA \\ \vdots \\ -CA^{n-1} \\ \vdots \\ -CA^{\tilde{n}-2} \end{pmatrix} \tag{4.7}$$

ha rango pieno, cioè ha rango n . Ricordiamo che il rango è limitato superiormente dal minimo tra il numero di righe (np) e il numero di colonne (n), quindi in questo caso il limite superiore è dato dal numero di colonne. Ora notiamo che le righe aggiunte dopo la riga n (quella corrispondente a $-CA^{n-1}$) sono linearmente dipendenti dalle prime n righe per il teorema di Cayley-Hamilton, quindi per andare a studiare il rango della matrice possiamo limitarci a studiare il rango delle righe:

$$\begin{pmatrix} -C \\ -CA \\ \vdots \\ -CA^{n-1} \end{pmatrix} \tag{4.8}$$

Affinché la matrice (4.7) abbiano rango massimo (n) dobbiamo avere che le righe in (4.8) siano linearmente indipendenti, cioè dobbiamo avere che la matrice (4.8) abbia rango n . Ma questa condizione è proprio la condizione di completa osservabilità della coppia (A, C) del sistema

originale. Quindi il sistema aumentato è completamente osservabile se e solo se il sistema originale è completamente osservabile.

Andiamo ora a studiarci la **raggiungibilità** del sistema aumentato, cioè il rango della matrice di raggiungibilità (ricordiamo che affinchè il sistema sia completamente raggiungibile il rango deve essere massimo cioè uguale a \tilde{n}):

$$M_R(\tilde{A}, \tilde{B}) = \underbrace{\begin{pmatrix} \tilde{B} & \tilde{A}\tilde{B} & \dots & \tilde{A}^{\tilde{n}-1}\tilde{B} \end{pmatrix}}_{\tilde{nm} \text{ colonne}} \quad \left. \right\} n+p \text{ righe}$$

L'espressione generale $\forall k \in \mathbb{N}$ è:

$$\tilde{A}^k \tilde{B} = \begin{pmatrix} A^k & 0 \\ -CA^{k-1} & 0 \end{pmatrix} \begin{pmatrix} B \\ 0 \end{pmatrix} = \begin{pmatrix} A^k B \\ -CA^{k-1} B \end{pmatrix}$$

Quindi la matrice di raggiungibilità del sistema aumentato si può scrivere come:

$$M_R(\tilde{A}, \tilde{C}) = \underbrace{\begin{pmatrix} B & AB & A^2B & \dots & A^{\tilde{n}-1}B \\ 0 & -CB & -CAB & \dots & -CA^{\tilde{n}-2}B \end{pmatrix}}_{\tilde{nm}} \quad \left. \right\} \begin{matrix} n \\ p \end{matrix} \quad (4.9)$$

Andiamo ora ad enunciare il seguente lemma che lega la raggiungibilità del sistema aumentato con la raggiungibilità del sistema originale.

Lemma 4.1.1. *Se il sistema aumentato è completamente raggiungibile allora il sistema originale è completamente raggiungibile.*

Dimostrazione. Noi abbiamo supposto che il sistema aumentato sia completamente raggiungibile, cioè che la matrice di raggiungibilità del sistema aumentato abbia rango pieno. Notiamo che la matrice di raggiungibilità del sistema aumentato ha $n + p$ righe dunque per essere di rango pieno deve avere rango $n + p$:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) = n + p$$

Questo significa che tutte le righe della matrice di raggiungibilità devono essere linearmente indipendenti (infatti la matrice ha più colonne che righe, dunque il rango massimo è dato dal numero di righe). In particolare, anche le prime n righe della matrice di raggiungibilità devono essere linearmente indipendenti, ma questo significa che la seguente matrice, costituita dalle prime n righe della matrice di raggiungibilità, è di rango pieno (cioè ha rango n):

$$\begin{pmatrix} B & AB & A^2B & \dots & A^{n-1}B & \dots & A^{\tilde{n}-1}B \end{pmatrix} \quad (4.10)$$

Ma le colonne dopo la colonna n -esima (cioè le colonne da $n + 1$ a $n + p$) sono linearmente dipendenti dalle prime n colonne per il teorema di Cayley-Hamilton. Quindi se la matrice

(4.10) è di rango pieno n significa che le prime n colonne sono linearmente indipendenti, ma questo significa che la seguente matrice ha rango pieno n :

$$\begin{pmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{pmatrix} \quad (4.11)$$

Ma la matrice (4.11) è proprio la matrice di raggiungibilità del sistema originale (A, B) . Dunque abbia dimostrato che se il sistema aumentato è completamente raggiungibile allora la matrice di raggiungibilità del sistema originale è di rango pieno, cioè che il sistema originale è completamente raggiungibile. \square

Dunque condizione necessaria affinche il sistema aumentato sia completamente raggiungibile è che il sistema originale (A, B) sia completamente raggiungibile.

Andiamo ora a ricavare di nuovo la condizione necessaria (4.6) come abbiamo già detto all'inizio. Per prima cosa notiamo che la matrice di raggiungibilità del sistema aumentato (4.9) si può riscrivere come prodotto di matrici:

$$M_R(\tilde{A}, \tilde{B}) = \begin{pmatrix} A & B \\ -C & 0 \end{pmatrix} \begin{pmatrix} 0 & B & AB & \dots & A^{\tilde{n}-2}B \\ I & 0 & 0 & \dots & 0 \end{pmatrix} \quad (4.12)$$

Definiamo la matrice $\tilde{\Sigma} \in \mathbb{R}^{(n+p) \times (n+m)}$ come:

$$\tilde{\Sigma} := \begin{pmatrix} A & B \\ -C & 0 \end{pmatrix} \quad (4.13)$$

Notiamo che questa matrice $\tilde{\Sigma}$ ha lo stesso rango della matrice Σ definita in (4.3). Mentre chiamiamo la matrice a destra del prodotto in (4.12) come $Q \in \mathbb{R}^{(n+m) \times \tilde{n}m}$:

$$Q := \begin{pmatrix} 0 & B & AB & \dots & A^{\tilde{n}-2}B \\ I & 0 & 0 & \dots & 0 \end{pmatrix} \quad (4.14)$$

Quindi la $M_R(\tilde{A}, \tilde{B})$ si può scrivere come:

$$M_R(\tilde{A}, \tilde{B}) = \tilde{\Sigma}Q \quad (4.15)$$

Dimostriamo ora per assurdo che se la matrice $\tilde{\Sigma}$ non è di rango pieno allora la matrice di raggiungibilità del sistema aumentato non è di rango pieno, vogliamo dimostrare che:

Lemma 4.1.2. *Se la matrice di raggiungibilità del sistema aumentato è di rango pieno allora la matrice $\tilde{\Sigma}$ è di rango pieno.*

Ogliamo quindi dimostrare che, affinché la matrice di raggiungibilità del sistema aumentato sia di rango pieno, è necessario che $\tilde{\Sigma}$ sia di rango pieno. Ora dimostreremo il teorema 4.1.2 per assurdo, cioè dimostreremo che il negato del teorema 4.1.2 implica una contraddizione. Notiamo che il negato del teorema 4.1.2 consiste nel provare a dimostrare che se la matrice $M_R(\tilde{A}, \tilde{B})$ è di rango pieno allora la matrice $\tilde{\Sigma}$ non è di rango pieno.

Dimostrazione. La nostra ipotesi è che la matrice di raggiungibilità del sistema aumentato sia di rango pieno, cioè che:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) = n + p$$

Vogliamo dimostrare, come tesi, che questa ipotesi implica che la matrice $\tilde{\Sigma}$ non sia di rango pieno:

$$\text{rank}(\tilde{\Sigma}) < n + p$$

Ma per la proprietà del prodotto delle matrici si ha che:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) \leq \min\{\text{rank}(\tilde{\Sigma}), \text{rank}(Q)\}$$

Cioè si ha che:

$$n + p \leq \min\{\text{rank}(\tilde{\Sigma}), \text{rank}(Q)\}$$

Ma notiamo che a destra il minimo tra i due ranghi è sicuramente minore o uguale al rango di $\tilde{\Sigma}$, dunque si ha che:

$$n + p \leq \text{rank}(\tilde{\Sigma})$$

Ma noi avevamo supposto all'inizio che la matrice $\tilde{\Sigma}$ non fosse di rango pieno, cioè si ha:

$$n + p < n + p$$

Ma questo è un assurdo. Dunque la nostra tesi iniziale deve essere falsa, cioè se la matrice di raggiungibilità del sistema aumentato è di rango pieno allora la matrice $\tilde{\Sigma}$ è di rango pieno. \square

Ora abbiamo dimostrato che affinche il sistema aumentato sia completamente raggiungibile (che ricordiamo essere condizione necessaria e sufficiente, insieme alla completa osservabilità, per la risoluzione del problema di controllo) è necessario che la matrice $\tilde{\Sigma}$ sia di rango pieno, ma questa matrice come abbiamo detto prima ha lo stesso rango della matrice Σ definita in (4.3), dunque abbiamo ridimostrato la condizione necessaria (4.6), cioè che la matrice Σ deve essere di rango pieno affinche il problema di controllo sia risolvibile. Ora dimostriamo che se la matrice Σ è di rango pieno e il sistema originale è completamente raggiungibile allora il sistema aumentato è completamente raggiungibile.

Lemma 4.1.3. *Se la matrice Σ è di rango pieno e il sistema originale è completamente raggiungibile allora il sistema aumentato è completamente raggiungibile.*

Dimostrazione. Vogliamo dimostrare che sotto le ipotesi del teorema la matrice di raggiungibilità del sistema aumentato è di rango pieno, cioè che:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) = n + p$$

Per la proprietà del prodotto delle matrici si ha che:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) \leq \min\{\text{rank}(\tilde{\Sigma}), \text{rank}(Q)\}$$

Inoltre una nota disegualanza sui ranghi delle matrici dice che ($M_R(\tilde{A}, \tilde{B}) = \tilde{\Sigma}Q$):

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) \geq \text{rank}(\tilde{\Sigma}) + \text{rank}(Q) - (n + m)$$

Unendo le due diseguaglianze si ha che:

$$\text{rank}(\tilde{\Sigma}) + \text{rank}(Q) - (n + m) \leq M_R(\tilde{A}, \tilde{B}) \leq \min\{\text{rank}(\tilde{\Sigma}), \text{rank}(Q)\} \quad (4.16)$$

Ma per ipotesi la matrice Σ è di rango pieno:

$$\text{rank}(\tilde{\Sigma}) = n + p$$

Il sistema originale è completamente raggiungibile, dunque la matrice Q anche essa è di rango pieno:

$$\text{rank}(Q) = n + m$$

Dunque riscrivendo la diseguaglianza (4.16) si ha:

$$(n + p) + (n + m) - (n + m) \leq \text{rank}(M_R(\tilde{A}, \tilde{B})) \leq \min\{n + p, n + m\}$$

Da cui si ricava (dato che $m \geq p$):

$$n + p \leq \text{rank}(M_R(\tilde{A}, \tilde{B})) \leq n + p$$

Dunque si deve per forza avere:

$$\text{rank}(M_R(\tilde{A}, \tilde{B})) = n + p$$

Quindi la matrice di raggiungibilità del sistema aumentato è di rango pieno, cioè il sistema aumentato è completamente raggiungibile. \square

Uniamo tutti i risultati ottenuti nei lemmi 4.1.1, 4.1.2, 4.1.3 in un unico teorema che consiste in una condizione necessaria e sufficiente.

Teorema 4.1.4. *Il sistema aumentato è completamente raggiungibile se e solo se il sistema originale è completamente raggiungibile e la matrice Σ è di rango pieno.*

4.2 Problema di Mixed Sensitivity Design

Il nostro obiettivo ora è andare a formulare le seguenti specifiche di progetto:

- Stabilità
- Stabilità robusta
- Performance inseguimento/disturbo
- Limitazione controllo

sulle funzione di sensitività $S(s)$ e sensitività complementare $S'(s)$.

$$E(s) = R(s) - Y(s) = S(s)(R(s) - D(s)) + T(s)N(s)$$

$$U(s)K(s)S(s)(R(s) - D(s) - N(s))$$

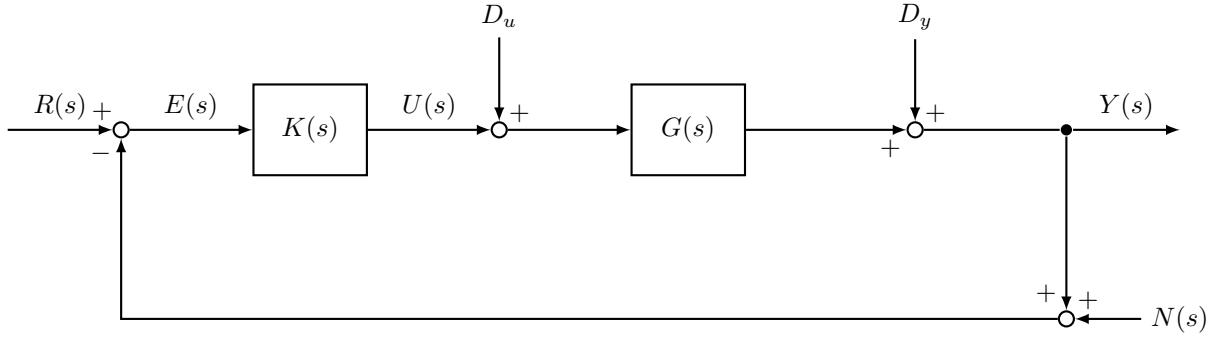


Figura 4.3: Sistema di controllo MIMO con retroazione negativa e disturbi

$$\begin{cases} \|S(s)\|_\infty \rightarrow 0 \\ \|T(s)\|_\infty \rightarrow 0 \end{cases} \quad (4.17)$$

E' impossibile realizzare la condizione precedente poichè vale la relazione

$$S(s) + T(s) = I$$

Questo problema si può risolvere ricordando che gli ingressi del sistema sono di solito segnali presenti a frequenze diverse:

- Il riferimento $R(s)$ è a bassa frequenza
- Il disturbo $D(s)$ è a bassa frequenza
- Il rumore $N(s)$ è ad alta frequenza

Con queste considerazioni possiamo andare a formulare le specifiche di progetto in modo diverso, infatti le specifiche 4.17 sono potenti, ma non necessarie, possiamo accontentarci delle seguenti specifiche, più deboli ma sufficienti a soddisfare le nostre esigenze (e più che altro realizzabili):

$$\begin{cases} \|S(s)(R(s) - D(s))\|_2 \rightarrow 0, & \forall R, D \text{ reali (cioè stiamo considerando il comportamento frequenziale di } R(s) \text{ e } D(s)) \\ \|T(s)N(s)\|_2 \rightarrow 0, & \forall N \text{ reale (cioè stiamo considerando il comportamento frequenziale di } N(s)) \end{cases} \quad (4.18)$$

dove notiamo che $S(s)(R(s) - D(s))$ e $T(s)N(s)$ sono vettori, quindi la norma 2 si intende come norma vettoriale.

Infatti se sono vere le (4.18) si ha che per la diseguaglianza triangolare:

$$\|E(s)\|_2 \leq \|S(s)(R(s) - D(s))\|_2 + \|T(s)N(s)\|_2 \rightarrow 0 \implies \|E(s)\|_2 \rightarrow 0$$

Ma cosa hanno di differente le specifiche (4.18) rispetto alle specifiche (4.17), la differenza principale sta nel fatto che ora possiamo andare ad esplicitare le caratteristiche spettrali dei segnali $R(s)$, $D(s)$ e $N(s)$, infatti possiamo dire che:

- Visto che $R(s)$ e $D(s)$ sono segnali a bassa frequenza, per soddisfare la prima specifica di (4.18) ci basta che $S(s)$ tenda a zero in bassa frequenza, mentre in alta frequenza ci

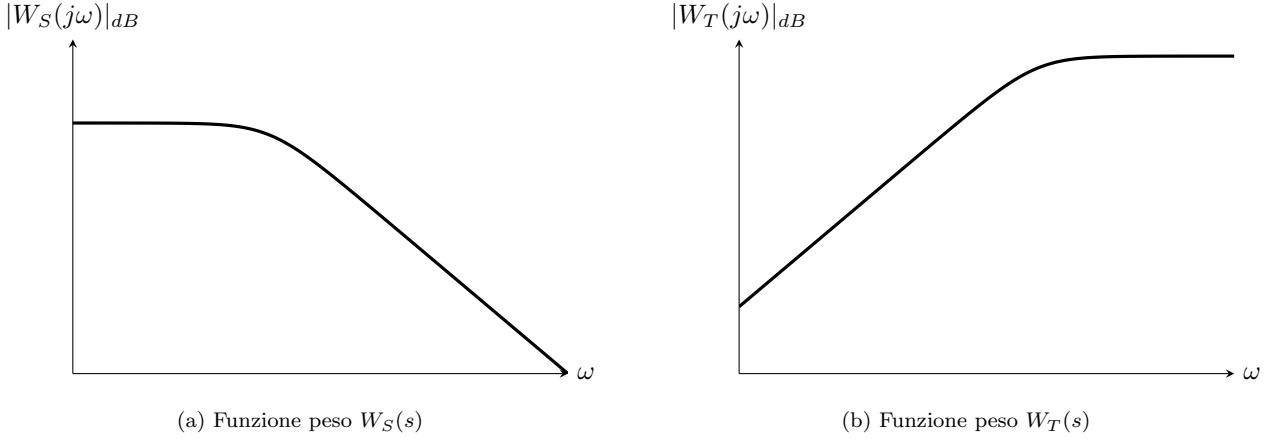


Figura 4.4: Andamenti tipici delle funzioni peso per sistemi SISO

penseranno i segnali $R(s)$ e $D(s)$ a far sì che la norma vada a zero. Questo si traduce in una specifica sulla norma infinito non più di $S(s)$ ma di una funzione $W_S(s)S(s)$:

$$\|W_S(s)S(s)\|_\infty \rightarrow 0$$

dove $W_S(s)$ è una funzione di peso che ha un valore alto in bassa frequenza e un valore basso in alta frequenza, cosicché per soddisfare la specifica la funzione $S(s)$ deve essere piccola in bassa frequenza, come richiesto. L'introduzione di questa funzione di peso ci permette di non imporre vincoli in alta frequenza su $S(s)$, andando quindi a permetterci di imporre vincoli sulla $T(s)$, continuando però a rispettare l'identità $S(s) + T(s) = I$. Una possibile scelta per $W_S(s)$ per un sistema SISO è in Figura 4.4a.

- Visto che $N(s)$ è un segnale ad alta frequenza, per soddisfare la seconda specifica di (4.18) ci basta che $T(s)$ tenda a zero in alta frequenza, mentre in bassa frequenza ci penserà il segnale $N(s)$ a far sì che la norma vada a zero. Questo si traduce in una specifica sulla norma infinito non più di $T(s)$ ma di una funzione $W_T(s)T(s)$:

$$\|W_T(s)T(s)\|_\infty \rightarrow 0$$

dove $W_T(s)$ è una funzione di peso che ha un valore basso in bassa frequenza e un valore alto in alta frequenza, cosicché per soddisfare la specifica la funzione $T(s)$ deve essere piccola in alta frequenza, mentre in bassa frequenza non ci sono vincoli su $T(s)$, poiché la funzione di peso $W_T(s)$ si occuperà di far andare a zero la norma infinito. Una possibile scelta per $W_T(s)$ per un sistema SISO è in Figura 4.4b.

- E' possibile definire anche una funzione di peso $W_K(s)$ per limitare l'azione di controllo $U(s)$. Infatti andando a gestire la norma infinito di $W_K(s)K(s)S(s)$:

$$\|W_K(s)K(s)S(s)\|_\infty \rightarrow 0$$

possiamo limitare l'azione di controllo andando a scegliere opportunamente la funzione di peso $W_K(s)$. Per un sistema SISO una possibile $W_K(s)$ è costante in un largo intervallo di frequenze (quelle possibili per il nostro ingresso di controllo $U(s)$), e va a zero in alta frequenza ed in bassa frequenza. Modificando opportunamente il valore costante possiamo andare a limitare l'azione di controllo. Ad esempio aumentando il valore costante andremo

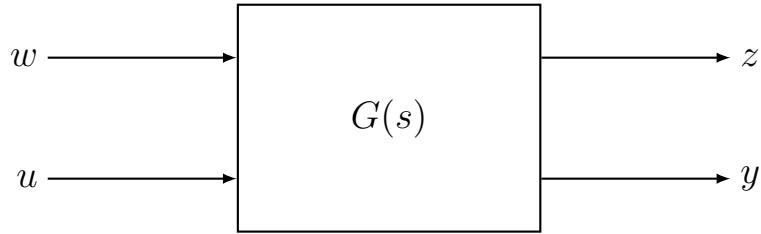


Figura 4.5: Sistema generale per il controllo H_∞

ad abbassare il valore di $|K(s)S(s)|$ (per un sistema SISO $K(s)S(s)$ è uno scalare e quindi si può parlare di modulo), e di conseguenza andremo ad abbassare il valore di $|U(s)|$.

Il problema che si occupa di trovare le espressioni delle funzioni di peso $W_S(s)$, $W_T(s)$ e $W_K(s)$ prende il nome di **Mixed Sensitivity Design**.

Un modo per risolvere questo problema è risolvere il problema H_∞

4.3 Problema H_∞ in forma standard

Consideriamo il sistema in Figura 4.5, questo sistema rappresenta il sistema di base per andare a formulare il **problema di controllo H_∞ in forma standard**. Notiamo che w, u, z, y possono essere vettori. Andiamo ad analizzare questo sistema, nello specifico notiamo che la relazione ingresso uscita del sistema è data da:

$$\begin{pmatrix} Z(s) \\ Y(s) \end{pmatrix} = G(s) \begin{pmatrix} W(s) \\ U(s) \end{pmatrix}$$

La matrice di trasferimento $G(s)$ si può scomporre in quattro sottomatrici:

$$G(s) = \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix}$$

Con questa scomposizione si ha che la relazione ingresso uscita del sistema si può scrivere come:

$$\begin{cases} Z(s) = G_{11}(s)W(s) + G_{12}(s)U(s) \\ Y(s) = G_{21}(s)W(s) + G_{22}(s)U(s) \end{cases} \quad (4.19)$$

Ora consideriamo il sistema in Figura 4.6, dove abbiamo inserito un controllore $K(s)$ in feedback, si ha:

$$U(s) = K(s)Y(s)$$

Andando a sostituire nel sistema (4.19) si ottiene:

$$\begin{cases} Z(s) = G_{11}(s)W(s) + G_{12}(s)Y(s) \\ Y(s) = G_{21}(s)W(s) + G_{22}(s)Y(s) \end{cases}$$

Quella che vogliamo andarci a ricavare noi ora è la funzione di trasferimento ad anello chiuso tra l'ingresso $W(s)$ e l'uscita $Z(s)$, andiamo quindi a ricavarci $Y(s)$ in funzione di $W(s)$ nella seconda equazione:

$$\begin{cases} Z(s) = G_{11}(s)W(s) + G_{12}(s)Y(s) \\ Y(s) - G_{22}(s)Y(s) = G_{21}(s)W(s) \end{cases}$$

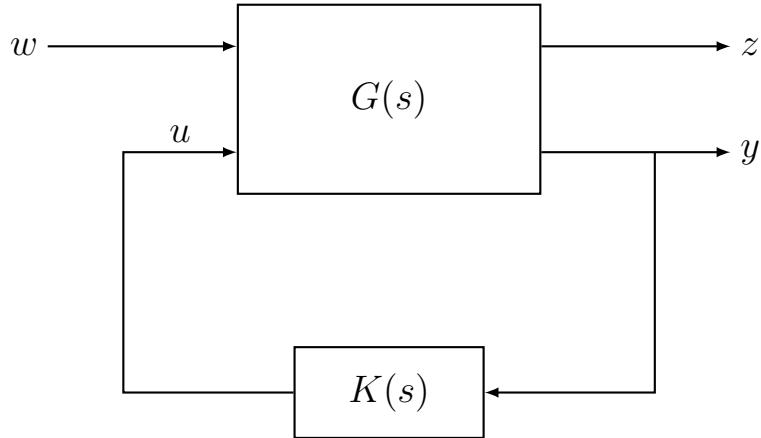


Figura 4.6: Problema di controllo H_∞ in forma standard

Mettendo in evidenza $Y(s)$ nella seconda equazione si ha:

$$\begin{cases} Z(s) = G_{11}(s)W(s) + G_{12}(s)Y(s) \\ (1 - G_{22}(s))Y(s) = G_{21}(s)W(s) \end{cases}$$

Moltiplicando ambo i membri della seconda equazione per $(1 - G_{22}(s))^{-1}$ si ottiene:

$$\begin{cases} Z(s) = G_{11}(s)W(s) + G_{12}(s)Y(s) \\ Y(s) = (1 - G_{22}(s))^{-1}G_{21}(s)W(s) \end{cases}$$

Sostituendo l'espressione di $Y(s)$ nella prima equazione si ottiene:

$$Z(s) = G_{11}(s)W(s) + G_{12}(s)(1 - G_{22}(s))^{-1}G_{21}(s)W(s)$$

Mettendo in evidenza $W(s)$ si ha:

$$Z(s) = (G_{11}(s) + G_{12}(s)(1 - G_{22}(s))^{-1}G_{21}(s))W(s)$$

Quindi la funzione di trasferimento ad anello chiuso tra l'ingresso $W(s)$ e l'uscita $Z(s)$ è data da:

$$T_{zw}(s) = G_{11}(s) + G_{12}(s)(1 - G_{22}(s))^{-1}G_{21}(s) \quad (4.20)$$

Ora siamo in grado di formulare il problema di controllo H_∞ su un sistema in forma standard come quello in Figura 4.5, che però ha una forma diversa dal nostro sistema in retroazione di uscita. Nella prossima sezione andremo a vedere come utilizzare i risultati ottenuti per il problema di controllo H_∞ in forma standard per risolvere il problema di mixed sensitivity design.

4.4 Problema di Mixed Sensitivity come problema H_∞ in forma standard

Ricordiamo che il nostro schema di retroazione dell'uscita è quello in Figura ??, e la funzione di cui vogliamo minimizzare la norma infinito è la seguente:

$$\min_{K(s)} \left\| \begin{pmatrix} W_S(s)S(s) \\ W_T(s)T(s) \\ W_K(s)K(s)S(s) \end{pmatrix} \right\|_\infty$$

Ma se riuscissimo a portare il sistema in Figura ?? nella forma del sistema in Figura 4.5 in modo tale che la funzione di trasferimento $T_{zw}(s)$ sia uguale a:

$$T_{zw}(s) = \begin{pmatrix} W_S(s)S(s) \\ W_T(s)T(s) \\ W_K(s)K(s)S(s) \end{pmatrix}$$

allora potremmo tradurre la risoluzione del problema di mixed sensitivity design in un problema di controllo H_∞ in forma standard. Ora andremo a costruirsi un sistema fittizio a partire dal sistema in Figura ?? in modo tale che sia equivalente al sistema in Figura 4.5 e con una funzione di trasferimento $T_{zw}(s)$ uguale a quella desiderata. Il sistema fittizio è mostrato in Figura ??, questo sistema ha come ingressi i segnali w e u e come uscite i segnali z_1 , z_2 , z_3 e y . Poniamo z come:

$$z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

Ora quello che noi vogliamo andare a trovare è la funzione di trasferimento $\bar{G}(s)$ che lega:

$$\begin{pmatrix} Z(s) \\ Y(s) \end{pmatrix} = \bar{G}(s) \begin{pmatrix} W(s) \\ U(s) \end{pmatrix}$$

$$\bar{G}(s) = \begin{pmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{pmatrix}$$

$$Z_1(s) = W_S(s)Y(s) = W_S(s)(W(s) - G(s)U(s))$$

$$Z_2(s) = W_K(s)U(s)$$

$$Z_3(s) = W_T(s)G(s)U(s)$$

$$\bar{G}(s) = \begin{pmatrix} W_S(s) & -W_S(s)G(s) \\ 0 & W_K(s) \\ 0 & W_T(s)G(s) \\ I & -G(s) \end{pmatrix}$$

Ora una volta ricavata la funzione di trasferimento $\bar{G}(s)$, quello che vogliamo andare a dimostrare è che il nostro problema H_∞ standard è equivalente al problema di mixed sensitivity design, cioè che la funzione di trasferimento ad anello chiuso tra l'ingresso $W(s)$ e l'uscita $Z(s)$ sia uguale a:

$$T_{zw}(s) = \begin{pmatrix} W_S(s)S(s) \\ W_T(s)T(s) \\ W_K(s)K(s)S(s) \end{pmatrix}$$

La sequenza di passi per andare a formulare e risolvere un problema di H_∞ standard è la seguente:

- Definizione dell'impianto $G(s)$

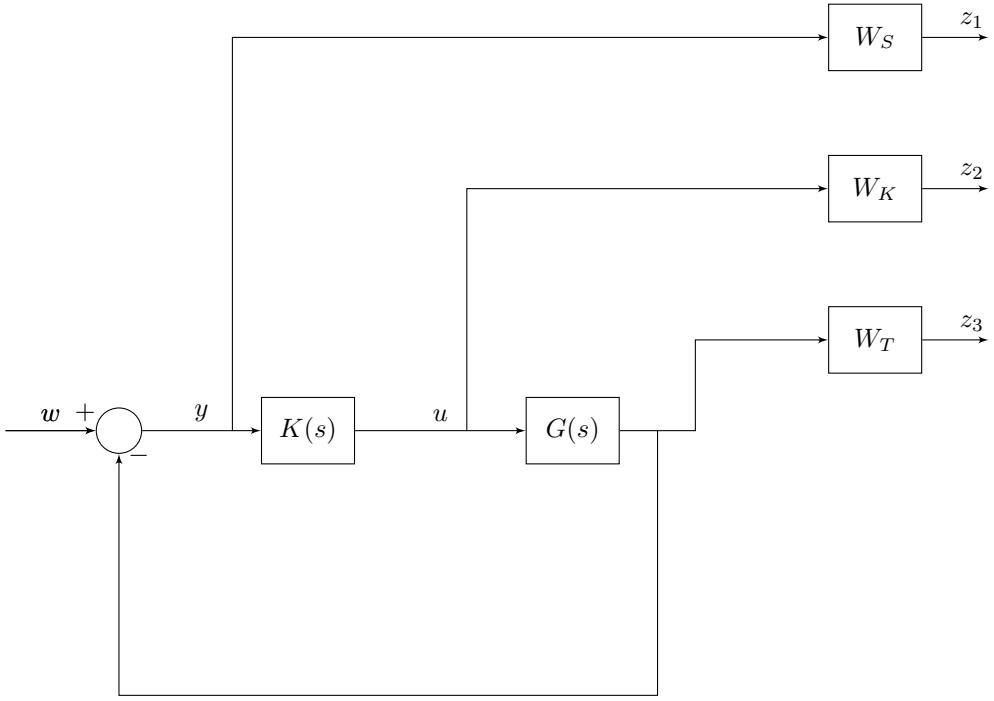


Figura 4.7: Sistema fittizio per trasformare il problema di mixed sensitivity in un problema H_∞ standard. Notiamo che la y non è l'uscita del sistema in retroazione, ma è l'errore $y(t) = e(t)$

- Traduzione delle specifiche di progetto in termini di funzioni di peso $W_S(s), W_T(s), W_K(s)$
- Definire $\bar{G}(s)$ del sistema in forma standard per il problema H_∞
- Risolvere il problema di controllo H_∞

Risolvere il problema di H_∞ ci da un uscita:

- Il controllo $K(s)$
- Un valore $\gamma > 0$ tale che:

$$\|T_{zw}(s)\|_\infty < \gamma \implies \left\| \begin{pmatrix} W_S(s)S(s) \\ W_T(s)T(s) \\ W_K(s)K(s)S(s) \end{pmatrix} \right\|_\infty < \gamma$$

Andiamo a ricordare che noi vogliamo usare il problema H_∞ per risolvere il problema di mixed sensitivity. Per noi il problema di H_∞ associato al problema di mixed sensitivity è risolto se e solo se $\gamma < 1$, quindi si ha:

$$\|T_{zw}(s)\|_\infty < 1 \implies \left\| \begin{pmatrix} W_S(s)S(s) \\ W_T(s)T(s) \\ W_K(s)K(s)S(s) \end{pmatrix} \right\|_\infty < 1$$

4.4.1 Costruzione $W_S(s)$

Notiamo che le progettazioni sulle funzioni di peso che effettueremo ora valgono per sistemi SISO, dunque le funzioni di trasferimento che andremo a considerare sono tutte funzioni scalari. Ricordiamo che l'errore è dato da:

$$E(s) = S(s)(R(s) - D(s)) + T(s)N(s)$$

Per quanto riguarda le specifiche sull'inseguimento del riferimento $R(s)$ le specifiche possibili sono le seguenti:

- Errore a regime $\leq \bar{e}_\infty$. Ricordiamo che l'errore a regime vale:

$$e_\infty = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} sE(s)$$

se il riferimento è un gradino $R(s) = \frac{1}{s}$ allora si ha (ricordiamo che stiamo considerando i disturbi nulli $D(s) = 0$ e $N(s) = 0$):

$$e_\infty = \lim_{t \rightarrow \infty} e(t) = \lim_{s \rightarrow 0} sE(s) = S(0)$$

Quindi la specifica sull'errore a regime si può riscrivere come:

$$S(0) \leq \bar{e}_\infty$$

Ora il nostro obiettivo è convertire questa specifica su $S(0)$ in una specifica su $W_S(s)$, in modo tale che se la specifica su $W_S(s)$ è soddisfatta allora anche la specifica su $S(0)$ è soddisfatta. Partiamo con il ricordare che per noi il H_∞ associato al problema di mixed sensitivity è risolto se e solo se:

$$\|W_S(s)S(s)\|_\infty < 1$$

Ma per definizione della norma H_∞ si ha che la disegualanza è vera $\forall s$, quindi in particolare per $s = 0$:

$$|W_S(0)| |S(0)| < 1 \iff |S(0)| < \frac{1}{|W_S(0)|}$$

Ora se imponiamo la seguente condizione su $W_S(0)$:

$$\frac{1}{|W_S(0)|} \leq \bar{e}_\infty$$

si il risultato voluto, cioè una specifica su $W_S(s)$ che implica la specifica su $S(0)$:

$$|S(0)| < \frac{1}{|W_S(0)|} \leq \bar{e}_\infty \implies |S(0)| < \bar{e}_\infty$$

Quindi la specifica sull'errore a regime si può tradurre in una specifica su $W_S(0)$:

$$- + \frac{1}{|W_S(0)|} \leq \bar{e}_\infty \implies |W_S(0)| \geq \frac{1}{\bar{e}_\infty} \quad (4.21)$$

- Tempo di assestamento, abbiamo delle specifiche ad esempio su t_{a5} o t_{a1} , ricordiamo:

$$t_{a5} = 3\tau \frac{3}{\omega_c}$$

$$t_{a1} = 4.6\tau \frac{3}{\omega_c}$$

In letteratura si trova spesso la specifica sul tempo di assestamento espressa come:

$$t_a \approx \frac{4}{\omega_c}$$

Ricordiamo che la frequenza ω_c è il valore di frequenza in cui la $L(j\omega_c)$ in dB ha modulo 0dB, cioè:

$$|L(j\omega_c)|_{dB} = 0dB \iff |L(j\omega_c)| = 1$$

Questo implica che la funzione di sensitività deve avere:

$$|L(j\omega_c)| = 1 \implies |S(j\omega_c)| = \frac{1}{2}$$

- Sovraelongazione, che ricordiamo essere legata al modulo di picco della $S(j\omega)$:

$$M_p(S(j\omega)) = \max_{\omega} |S(j\omega)|$$

$$|W_S(j\omega)| \text{ in alta frequenza}$$

$$\|W_S(s)S(s)\|_{\infty} < 1 \text{ Disturbi in ingresso}$$

$$\|W_T(s)T(s)\|_{\infty} < 1 \text{ Riferimenti in ingresso}$$

Teorema del piccolo guadagno.

$$\|W_K(s)K(s)S(s)\|_{\infty} < 1$$

$$W_S = \frac{\frac{s}{M_P} + \omega_c}{s + A\omega_c}$$

- Fattore di riduzione M_R
- ω di attenuazione

$$W_T = \frac{\frac{s}{M_P} + \omega_c}{s + A\omega_c}$$

Capitolo 5

Appendici

5.1 SVD (Singular Value Decomposition)

Consideriamo una matrice $A \in \mathbb{R}^{m \times n}$. Inoltre dati due vettori $y \in \mathbb{R}^m$ e $x \in \mathbb{R}^n$ tali che vale la relazione:

$$y = Ax$$

$$A = U\Sigma V^T$$

- La matrice $U \in \mathbb{R}^{m \times m}$ (matrice di rotazione delle uscite) è una matrice ortonormale composta dai vettori colonna u_1, u_2, \dots, u_m , si ha:

$$U = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix}$$

- La matrice $V \in \mathbb{R}^{n \times n}$ (matrice di rotazione degli ingressi) è una matrice ortonormale composta dai vettori colonna v_1, v_2, \dots, v_n , si ha:

$$V = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$$

- La matrice $\Sigma \in \mathbb{R}^{m \times n}$ (matrice dei valori singolari)

$$p = \min \{m, n\}$$

Dove si ha che:

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \end{pmatrix}, \quad \text{se } p = m$$

$$\Sigma = \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix}, \quad \text{se } p = n$$

$$y = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \begin{pmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_m \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix} \cdot x$$

Se io prendo $x = v_i$ si ha:

$$\begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix} \cdot x = \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{pmatrix} \cdot v_i$$

Ricordando che V è ortonormale, si che le sue colonne sono tra loro ortogonal e di norma 1, dunque si ha:

$$v_j^T \cdot v_i = 0, \quad \text{se } i \neq j$$

$$v_j^T \cdot v_i = 1, \quad \text{se } i = j$$

Dunque si ha:

$$\begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_i^T \\ \vdots \\ v_n^T \end{pmatrix} \cdot v_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

Dove l'uno è nella i -esima riga. Andiamo a calcolarci la norma 2 di A :

$$\|A\|_2 = \sqrt{\lambda_{MAX}(A^T A)}$$

Notiamo che per le proprietà della trasposta $A^T = (U\Sigma V)^T = V\Sigma^T U^T$

$$\|A\|_2 = \sqrt{\lambda_{MAX}(V\Sigma^T U^T U \Sigma V^T)}$$

Ora ricordando che U è ortonormale si ha $U^T U = I$ ($U^T = U^{-1}$):

$$\|A\|_2 = \sqrt{\lambda_{MAX}(V\Sigma^T \Sigma V^T)}$$

Ora ricordando che Σ e Σ^T sono diagonali si ha che vale la proprietà commutativa tra le matrici, dunque possiamo scrivere:

$$\|A\|_2 = \sqrt{\lambda_{MAX}(\Sigma^T \Sigma V V^T)}$$

Ora ricordando che Σ è diagonale vale $\Sigma^T = \Sigma$, dunque si ha:

$$\|A\|_2 = \sqrt{\lambda_{MAX}\Sigma^T \Sigma} = \sqrt{\lambda_{MAX}\Sigma \Sigma} = \sqrt{\lambda_{MAX}\Sigma^2}$$

Come si vede la norma 2 di A è legata al valore singolare maggiore di A .

5.2 Diagonalizzare

Sia dato un sistema LTI descritto dalle equazioni di stato:

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

quando noi diamo un sistema in questa forma diamo per assodato che il vettore x sia scritto nella base canonica di \mathbb{R}^n . Se la matrice A è diagonalizzabile, ci poniamo il problema di portare

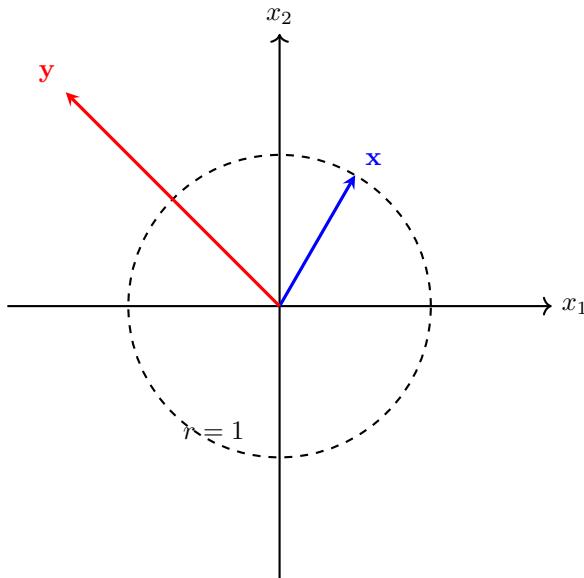


Figura 5.1: Esempio di rotazione del vettore di ingresso $x = v_1$ nel vettore di uscita amplificato $y = \sigma_1 u_1$. In questo caso abbiamo preso $m = n = 2$, dunque la matrice A è quadrata, mentre Σ è diagonale. Notiamo che abbiamo potuto rappresentare tutto sullo stesso grafico solo perché $m = n$.

x in un nuovo stato \hat{x} in cui A è diagonale. Dunque cerchiamo una matrice di trasformazione T tale che:

$$x = T\hat{x} \implies \hat{A} = T^{-1}AT \text{ è diagonale}$$

Ma come si può immediatamente notare la matrice T è la matrice del cambiamento di base dalla base canonica alla base che diagonalizza A . Quindi a noi basta trovare la matrice che permette di portare il vettore \hat{x} , scritto nella base che diagonalizza A , nel vettore x scritto nella base canonica. Ma la base che diagonalizza A è formata dagli autovettori di A , dunque scegliamo un insieme di autovettori linearmente indipendenti di A :

$$\hat{B} = \{v_1, v_2, \dots, v_n\}$$

Per calcolare T dobbiamo ricordare che la matrice del cambiamento di base si calcola attraverso l'applicazione identità (che è un applicazione lineare), dunque applichiamo l'identità agli autovettori scelti:

$$id(\hat{B}) = \{v_1, v_2, \dots, v_n\}$$

Ora scomponiamo i vettori rispetto alla base canonica ed incolonniamoli come colonne della matrice T (la base canonica è molto comoda perchè ci permette di scrivere i vettori proprio come sono), dunque la matrice T vale:

$$T = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$$

In cui i vettori $v_1, v_2, \dots, v_n \in \mathbb{R}^{n \times 1}$ sono vettori colonna $n \times 1$, dunque la matrice $T \in \mathbb{R}^{n \times n}$.

5.3 Uguaglianza matriciale per dimostrare la relazione di $T(s)$

Per dimostrare che per la funzione di sensitività complementare $T(s)$ vale la relazione:

$$T(s) = (I + L(s))^{-1} L(s) = L(s) (I + L(s))^{-1}$$

cioè che le matrici $(I + L(s))^{-1}$ e $L(s)$ commutano, abbiamo usato la seguente identità matriciale:

$$(I + AB)^{-1} A = A (I + BA)^{-1}$$

Dimostrazione. Noi vogliamo dimostrare che:

$$(I + AB)^{-1} A = A (I + BA)^{-1}$$

Studiamoci la matrice differenza:

$$D := (I + AB)^{-1} A - A (I + BA)^{-1}$$

Se riusciamo a dimostrare che $D = 0$ allora abbiamo dimostrato l'uguaglianza. Moltiplichiamo D a sinistra per $(I + AB)$ e a destra per $(I + BA)$, si ha:

$$(I + AB)D(I + BA) = (I + AB)(I + AB)^{-1} A(I + BA) - (I + AB)A(I + BA)^{-1}(I + BA)$$

Notiamo che $(I + AB)^{-1} A(I + BA) = I$ e $(I + BA)^{-1}(I + BA) = I$, quindi si ha:

$$(I + AB)D(I + BA) = A(I + BA) - (I + AB)A = A + ABA - A - ABA = 0$$

Abbiamo dimostrato $D = 0$, quindi anche l'uguaglianza iniziale. \square