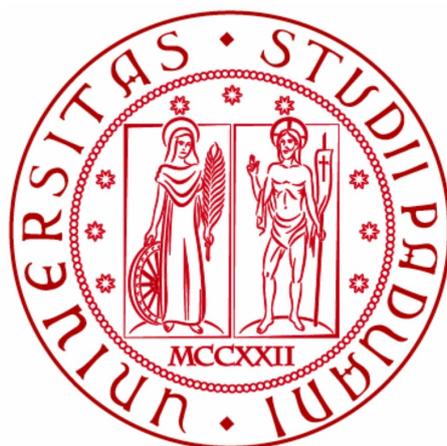


UNIVERSITA' DEGLI STUDI DI PADOVA

Tesi di Laurea Magistrale in Ingegneria dell'Automazione

Controllabilità di sistemi complessi:
energia per il controllo



Laureando: Fassina Riccardo – 1077960

Relatore: Zampieri Sandro

Anno Accademico 2015-2016

Indice

1	Introduzione	1
1.1	Il problema della controllabilità	1
1.2	Struttura della tesi e contributi	2
1.3	Notazione	3
2	Energia per il controllo: il gramiano di controllabilità	5
2.1	Modello per lo studio della rete	6
2.2	Upper bound sull'autovalore minimo del gramiano a tempo discreto	8
2.3	Upper bound sull'autovalore minimo del gramiano a tempo continuo	11
2.4	Caso simmetrico e risultati secondari	20
2.5	Minimizzazione dell'upper bound a tempo continuo	24
3	Simmetrizzazione e “quasi simmetrizzazione” di grafi	32
3.1	Definizioni e risultati introduttivi	33
3.2	Simmetrizzazione e potenziali dei nodi	34
3.3	Procedura di “quasi-simmetrizzazione”	36
3.4	Upper bound sull'autovalore minimo del gramiano di matrici “quasi-simmetriche”	38
3.5	Il ciclo semplice	41
3.5.1	Matrice d'errore di simmetrizzazione	41
3.5.2	Norma dell'errore di simmetrizzazione	47
3.5.3	Errore di simmetrizzazione con pesi i.i.d. gaussiani	50
3.5.4	Controllabilità del ciclo semplice	52
3.5.5	Espansione a grafi con un solo ciclo	55
3.6	Grafi cactus	56
3.7	Il grafo completo	62
4	Conclusioni	73
A	Limite superiore sulla norma del prodotto di Hadamard	75
B	Altre misure di controllabilità del gramiano	78
	Bibliografia	80

Capitolo 1

Introduzione

1.1 Il problema della controllabilità

La controllabilità è una proprietà fondamentale nello studio dei sistemi, ed è stata ampiamente studiata in letteratura. Si tratta infatti di una caratteristica essenziale nell'ambito della modellizzazione che ha rilevanza in varie applicazioni, dall'ingegneria, alla statistica, fino alla biologia.

Se un sistema è *controllabile*, può essere “guidato” opportunamente verso ogni configurazione; in altre parole, definite le variabili che entrano in gioco nella dinamica, e controllando dall'esterno opportunamente alcune di esse, è possibile obbligarle, tramite la loro mutua interazione, ad assumere valori desiderati, il cui significato fisico è stabilito dalla modellizzazione stessa. Tuttavia in molte applicazioni reali la completa controllabilità può “costare” un'elevata quantità d'energia, intesa come difficoltà nel guidare il sistema verso lo stato desiderato. Nella pratica, infatti, molto spesso l'energia messa a disposizione non può essere arbitrariamente grande (sia per limitazione fisiche, ma anche di costo) e quindi alcune configurazioni, di fatto, non sono raggiungibili.

Quindi, la controllabilità di per sè in molte situazioni non è sufficiente, o meglio, non fornisce abbastanza informazione circa la possibilità di modificare liberamente lo stato di un sistema. Proprio in tale contesto, uno strumento essenziale per determinare l'effettiva controllabilità di un sistema, dati alcuni vincoli “energetici” che possono essere di varia natura, è il cosiddetto *gramiano di controllabilità*, cioè la matrice che contiene l'informazione necessaria a discriminare fra le configurazioni del sistema in base al dispendio energetico richiesto per ottenerle.

E' possibile scegliere diverse “funzioni di costo” del gramiano, utilizzabili allo scopo di minimizzare l'energia spesa per la controllabilità, quali ad esempio la traccia o il determinante (per un approfondimento si veda *appendice B*): una delle possibili scelte è l'autovalore minimo del gramiano stesso. Questa può essere considerata un'ottima indicazione circa la controllabilità del sistema, in quanto è strettamente legata allo stato più difficile da raggiungere dal punto di vista energetico, tra gli stati a norma unitaria. In particolare quindi fornisce un limite sull'energia minima necessaria per controllare arbitrariamente il sistema, usando un approccio di tipo *worst case*.

Una volta stabilita questa limitazione è possibile determinare se un sistema è “difficile” oppure “facile” da controllare, con riferimento al comportamento dell’autovalore minimo del gramiano rispetto alla dimensione del sistema, ovvero il numero di variabili o di “agenti” che interagiscono fra loro.

Sfortunatamente, spesso l’autovalore minimo del gramiano non è facilmente calcolabile, soprattutto nei sistemi che hanno una dimensione molto elevata: per ovviare a tale problema, è necessario trovare opportuni limiti che ne approssimino il comportamento, e che permettano di stabilire, senza il calcolo diretto, la difficoltà di controllo.

1.2 Struttura della tesi e contributi

Lo scopo generale della tesi è quello di studiare delle limitazioni sull’autovalore minimo del gramiano di raggiungibilità, le quali consentiranno di determinare se i sistemi in esame sono difficili da controllare o meno, in assenza del calcolo diretto dell’autovalore stesso.

Più precisamente nel *capitolo 2* ricorderemo i risultati già noti precedentemente riportati in [1] sull’upper bound dell’autovalore minimo del gramiano nel caso di sistemi a tempo discreto, e successivamente cercheremo di espandere la trattazione ai sistemi a tempo continuo. In particolare vedremo che, nel caso continuo, è presente un ulteriore grado di libertà sull’insieme degli upper bound dell’autovalore minimo, il cui effetto verrà studiato nel dettaglio, ed inoltre sarà possibile ricavarne un valore ottimale (affinchè il bound risulti più stringente) nel caso di sistemi simmetrici (relativi quindi a grafi non orientati).

Motivati dal fatto che i sistemi simmetrici sono, in generale, più semplici da studiare e si possono ricavare limitazioni migliori dell’autovalore minimo del gramiano, nel *capitolo 3* saranno analizzati i sistemi simmetrizzabili tramite un opportuno cambio di base (più precisamente tramite matrici diagonali), e sarà introdotto un metodo per rendere “quasi-simmetrici” i grafi che non rispettano le condizioni di simmetrizzabilità. Questa tecnica associa ad ogni nodo un potenziale, in analogia a quanto accade nelle reti elettriche. Più nel dettaglio sarà analizzato l’errore di simmetrizzazione (inteso come grado di dissimilarità da un sistema simmetrico) e come esso si ripercuote sull’upper bound dell’autovalore minimo del gramiano.

Sarà studiato approfonditamente il caso di un sistema, con pesi scelti casualmente, costituito da un ciclo semplice, il quale ci consentirà di espandere i risultati ad una classe più ampia di reti, ovvero quella dei grafi *cactus*. In particolare sarà possibile affermare che tali sistemi sono difficili da controllare.

Infine, verranno esaminate anche le limitazioni relative al metodo di quasi-simmetrizzazione, portando come esempio il grafo completo.

1.3 Notazione

Di seguito è riportata la notazione adottata per la trattazione:

- Dato un vettore $v \in \mathbb{R}^n$, la norma euclidea di v sarà indicata da $\|v\|_2$ ed è definita come:

$$\|v\|_2 = \sqrt{v^T v}$$

- Data la matrice M , l'elemento in posizione (i, j) sarà indicato con $M_{i,j}$
- Data la matrice $M \in \mathbb{R}^{n \times n}$, lo spettro di M , ovvero l'insieme dei suoi autovalori, sarà indicato con $spec(M)$; in particolare la seguente notazione:

$$|\lambda_{min}(M)| = \min\{|\lambda| : \lambda \in spec(M)\}$$

$$|\lambda_{max}(M)| = \max\{|\lambda| : \lambda \in spec(M)\}$$

indica il valore assoluto dell'autovalore (o dell'insieme di autovalori) rispettivamente di modulo massimo e minimo della matrice. Il modulo si riferisce al campo dei numeri complessi, ovvero:

$$x = \Re(x) + j\Im(x) \implies |x| = \sqrt{\Re(x)^2 + \Im(x)^2}$$

- L'insieme dei valori singolari della matrice M sarà indicato con:

$$\sigma(M) = \{\lambda^{\frac{1}{2}} : \lambda \in spec(M^T M)\}$$

In particolare:

$$\sigma_{max}(M) = \max\{\lambda : \lambda \in \sigma(M)\}$$

$$\sigma_{min}(M) = \min\{\lambda : \lambda \in \sigma(M)\}$$

- La norma spettrale della matrice M è definita come:

$$\|M\|_2 = \sigma_{max}(M)$$

- La norma di Frobenius della matrice M è definita come:

$$\|M\|_F = \left(\sum_{i,j=1}^n |M_{i,j}|^2 \right)^{1/2}$$

- Il condition number di una matrice M è indicato con

$$cond(M) = \frac{\sigma_{max}(M)}{\sigma_{min}(M)}$$

- Date due matrici di eguale dimensione M_1 e M_2 , il prodotto di Hadamard (o elemento per elemento) fra matrici sarà indicato con:

$$M_h = M_1 \odot M_2$$

ed è tale che:

$$M_{h_{i,j}} = M_{1_{i,j}} M_{2_{i,j}}$$

- Data la variabile aleatoria X , la media e la varianza di X saranno indicate rispettivamente con $E(X)$ e $Var(X)$
- Una variabile aleatoria con distribuzione normale di media μ e varianza θ^2 sarà indicata con:

$$X \sim N(\mu, \theta^2)$$

- Date le funzioni $f(t)$ e $g(t)$, la notazione:

$$f(t) = O(g(t)), \quad t \rightarrow +\infty$$

sta ad indicare che esistono delle costanti positive t_0 e k tali che:

$$|f(t)| \leq k|g(t)|, \quad \forall t > t_0$$

Capitolo 2

Energia per il controllo: il gramiano di controllabilità

In questo capitolo ci occuperemo dello studio del *gramiano di controllabilità* di sistemi, che, come è già stato affermato, gioca un ruolo fondamentale per la controllabilità. In particolare, esso fornisce indicazioni sull'energia necessaria per far sì che il sistema raggiunga un determinato stato.

Inizieremo con l'introduzione del modello generale per lo studio della controllabilità di una rete, basato sulla teoria dei grafi e dei sistemi, ricordando le definizioni e gli strumenti fondamentali che verranno utilizzati in seguito. Quindi, presenteremo i risultati già noti sul gramiano di raggiungibilità di sistemi a tempo discreto, con riferimento al lavoro riportato in [1].

In seguito cercheremo di espandere tali risultati relativi all'autovalore minimo del gramiano per i sistemi a tempo continuo, il cui gramiano di controllabilità assume una forma più complessa rispetto al tempo discreto. Tuttavia è possibile trovare una stretta correlazione fra i due grazie alla teoria di Lyapunov, a discapito dell'introduzione di un'ulteriore variabile, che allargherà l'insieme delle stime sul valore dell'autovalore minimo. Vedremo inoltre che la soluzione risulta notevolmente semplificata nel caso in cui la matrice di adiacenza del grafo sia simmetrica, e tale proprietà consente di ricavare maggiori informazioni circa la controllabilità del sistema.

Infine, ricaveremo il valore ottimale per la variabile libera nel caso simmetrico, ovvero il valore tale da rendere il limite sull'autovalore minimo del gramiano di raggiungibilità più stringente, e quindi più efficace. Nel caso generale non simmetrico, il problema di minimizzazione della stima rispetto alla variabile libera è più complesso e non di facile soluzione: resta quindi un quesito aperto, anche se, dal punto di vista sperimentale, il valore ottimale del caso simmetrico sembra funzionare anche per quello non simmetrico qualora gli autovalori siano reali (come accade per esempio in sistemi di tipo tridiagonale Toeplitz).

2.1 Modello per lo studio della rete

Iniziamo la trattazione introducendo il modello per lo studio della rete.

Una qualsiasi rete può essere rappresentata da un grafo orientato $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, caratterizzato dall'insieme dei suoi nodi $\mathcal{V} = \{1, \dots, n\}$, dove n è la dimensione della rete, e dall'insieme dei vertici $\mathcal{E} = \mathcal{V} \times \mathcal{V}$. Se supponiamo di associare ad ogni vertice $(i, j) \in \mathcal{E}$ un peso a_{ij} , che risulta nullo qualora il vertice non esista, è possibile definire la *matrice delle adiacenze* A con elementi a_{ij} .

D'ora in avanti supporremo altresì che la matrice A sia sempre *diagonalizzabile*, ovvero ammetta una base di autovettori. Possiamo associare ad ogni nodo del grafo uno stato, ovvero un valore reale, e indicare l'insieme degli stati della rete con il vettore $x \in \mathbb{R}^n$: in particolare si può descrivere l'evoluzione degli stati tramite la mappa $x : \mathcal{T}_{\geq 0} \rightarrow \mathbb{R}^n$ (dove $\mathcal{T} = \mathbb{N}$ nel caso a tempo discreto e $\mathcal{T} = \mathbb{R}$ nel caso a tempo continuo). In maniera compatta si indica l'evoluzione dei nodi della rete con i sistemi:

$$x(t+1) = Ax(t)$$

nel caso discreto, e

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

nel caso continuo. Assumiamo che alcuni nodi siano *controllabili*, ovvero sia possibile scegliere il valore da associare al loro stato con un determinato input, e definiamo l'insieme dei nodi di controllo con:

$$\mathcal{K} = \{k_1, \dots, k_m\} \subseteq \mathcal{V}$$

Al sistema definito in precedenza si introduce quindi la *matrice di controllo* $B_{\mathcal{K}}$, la quale, per semplicità, ha la seguente forma:

$$B_{\mathcal{K}} = [e_{k_1} \dots e_{k_m}]$$

dove e_i rappresenta l' i -esimo vettore della base canonica di dimensione n , ovvero un vettore di zeri con un'unica componente, l' i -esima, pari a 1.

Di conseguenza, l'evoluzione temporale del grafo, nel caso continuo, è ora descritta dal sistema:

$$\dot{x}(t) = Ax(t) + B_{\mathcal{K}}u(t)$$

dove $u(t) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ è il segnale di controllo della rete. Nel caso discreto invece:

$$x(t+1) = Ax(t) + B_{\mathcal{K}}u(t)$$

con $u(t) : \mathbb{N}_{\geq 0} \rightarrow \mathbb{R}$.

Introduciamo ora la nozione di *controllabilità* di un sistema:

Definizione 2.1 (Controllabilità): Un sistema (A, B) si dice *controllabile* all'istante $T > 0$ se e solo se, per qualsiasi stato $x_f \in \mathbb{R}^n$, esiste un ingresso u tale da guidare il sistema allo stato $x(T) = x_f$ a partire da $x(0) = 0$.

Un risultato fondamentale e ampiamente noto per la controllabilità di sistemi sia continui che discreti è il seguente (si veda [2] per un riferimento) :

Proposizione 2.1: Ogni stato della rete in \mathbb{R}^n è controllabile in tempo finito se e solo se la matrice di controllabilità (o di *raggiungibilità*), definita come:

$$R_n = [B_{\mathcal{K}} \quad AB_{\mathcal{K}} \quad \dots \quad A^{n-1}B_{\mathcal{K}}]$$

ha rango pieno.

La controllabilità ci permette quindi di dire quali stati possiamo raggiungere con un ingresso arbitrario, ma non dà informazione sull'energia necessaria al raggiungimento dello stato desiderato. A tale scopo si ricorre al cosiddetto *gramiano di raggiungibilità* (al tempo T), definito come:

$$W_{d,T} = \sum_{i=0}^{T-1} A^i B B^T (A^T)^i$$

per i sistemi a tempo discreto, e:

$$W_{c,T} = \int_0^T e^{At} B B^T e^{A^T t} dt$$

per quelli a tempo continuo. Il gramiano è una matrice simmetrica e, se il sistema è controllabile, anche definita positiva.

Supponiamo che lo stato finale che ci proponiamo di raggiungere al tempo T abbia norma unitaria, ovvero $\|x_f\|_2 = 1$: possiamo definire l'energia dell'ingresso di controllo a tempo continuo come:

$$E(u, T) := \|u\|_{2,T}^2 = \int_0^T \|u(t)\|_2^2 dt$$

nel caso continuo, mentre:

$$E(u, T) := \|u\|_{2,T}^2 = \sum_0^{T-1} \|u(t)\|_2^2$$

nel caso discreto. L'ingresso ad energia minima che porta il sistema da $x(0) = 0$ a $x(T) = x_f$ ha la seguente forma nel caso continuo:

$$u^*(t) := B_{\mathcal{K}}^T e^{A^T(T-t)} W_{c,T}^{-1} x_f, \quad t \in [0, T]$$

Nel caso discreto invece si scrive:

$$u^*(t) := B_{\mathcal{K}}^T (A^T)^{T-t-1} W_{d,T}^{-1} x_f, \quad t \in [0, T]$$

A questo punto, vale che:

$$E(u^*, T) = \int_0^T \|u^*(t)\|_2^2 dt = x_f^T W_{c,T}^{-1} x_f \leq \lambda_{\min}^{-1}(W_{c,T})$$

dove l'uguaglianza è verificata quando x_f è autovettore del gramiano associato al suo autovalore minimo (nel caso discreto basta sostituire l'integrale con la sommatoria).

Il risultato ottenuto ci dice che, nella pratica, essendo l'energia a disposizione finita, sistemi il cui gramiano ha autovalori molto piccoli presentano stati che non possono essere raggiunti, sebbene controllabili. L'autovalore minimo del gramiano di raggiungibilità equivale quindi al dispendio minimo di energia per guidare il sistema verso gli stati più difficilmente raggiungibili. In particolare, se il sistema è internamente stabile, ovvero le matrici A sono Hurwitz (nel caso continuo) o Schur (nel caso discreto), possiamo far tendere il limite di integrale e sommatoria all'infinito e studiare l'autovalore minimo del gramiano di raggiungibilità su *orizzonte infinito*: infatti è noto che la minima energia per raggiungere un determinato stato tende ad un limite finito quando l'intervallo temporale di controllo tende all'infinito (per un riferimento si veda ancora una volta [2]).

L'obiettivo ora è di determinare un upper bound sull'autovalore minimo del gramiano, in modo da ottenere un lower bound sull'energia minima per raggiungere lo stato più difficilmente controllabile (in altre parole un limite inferiore sul caso peggiore dal punto di vista energetico).

2.2 Upper bound sull'autovalore minimo del gramiano a tempo discreto

Nel caso a tempo discreto, la forma che assume il gramiano di controllabilità consente di trovare un'upper bound sull'autovalore minimo. In particolare, si ricorda che trovare un limite superiore all'autovalore minimo del gramiano è equivalente a trovare un limite inferiore all'energia necessaria per guidare il sistema verso la "direzione" più difficile da controllare. Il risultato principale è riportato di seguito:

Teorema 2.1 (upper bound sull'autovalore minimo del gramiano a tempo discreto): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con $|\mathcal{V}| = n$, matrice di adiacenza A diagonalizzabile dalla matrice di autovettori V , insieme di nodi di controllo \mathcal{K} e

$$n_\mu = |\{\lambda : \lambda \in \text{spec}(A), \quad |\lambda| \leq \mu\}|$$

si ha che $\forall \mu \in [|\lambda_{\min}(A)|, 1)$ vale:

$$\lambda_{\min}(W_d) \leq \text{cond}^2(V) \frac{\mu^{2 \frac{n_\mu}{|\mathcal{K}|}}}{\mu^2(1 - \mu^2)}$$

dove W_d è il gramiano di controllo su orizzonte infinito del sistema discreto $(A, B_{\mathcal{K}})$

Dimostrazione. Se V è la base di autovettori di A , supponiamo che le matrici $A, B_{\mathcal{K}}$ siano, senza perdita di generalità, tali che:

$$V^{-1}AV = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \quad V^{-1}B_{\mathcal{K}} = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

dove $A_1 \in \mathbb{R}^{n_{\mu} \times n_{\mu}}$ e A_2 sono matrici diagonali e A_1 in particolare contiene gli autovalori di A in modulo minori di μ , cioè:

$$\text{spec}(A_1) = \{\lambda : \lambda \in \text{spec}(A), \quad |\lambda| \leq \mu\}$$

Ora, si può scrivere:

$$\begin{aligned} W_d &= \sum_{i=0}^{\infty} A^i B_{\mathcal{K}} B_{\mathcal{K}}^T (A^T)^i = V \sum_{i=0}^{\infty} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}^i \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} B_1^T \\ B_2^T \end{bmatrix}^T \begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix}^i V^T \\ &= V \tilde{W}_d V^T \end{aligned}$$

dove si è sfruttata la simmetria di A_1, A_2 , essendo matrici diagonali.

Dalla simmetria di W_d si ricava invece:

$$\begin{aligned} \lambda_{\min}(W_d) &= \min_{\|x\|=1} x^T W_d x = \min_{\|x\|=1} x^T V \tilde{W}_d V^T x \leq \|V\|_2^2 \min_{\|y\|=1} y^T \tilde{W}_d y \\ &\leq \|V\|_2^2 \min_{\|y_1\|=1} \begin{bmatrix} y_1^T & 0 \end{bmatrix} \tilde{W}_d \begin{bmatrix} y_1^T \\ 0 \end{bmatrix} = \|V\|_2^2 \lambda_{\min} \left(\sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right) \end{aligned}$$

Consideriamo la matrice di raggiungibilità del sistema (A_1, B_1) al tempo T definita come:

$$R_{k,T} = \begin{bmatrix} B_1 & A_1 B_1 & \dots & A_1^{T-1} B_1 \end{bmatrix}, \quad R_{k,T} \in \mathbb{R}^{n_{\mu} \times T|\mathcal{K}|}$$

Se vale la condizione $T|\mathcal{K}| < n_{\mu}$, ovvero $T \leq \lceil \frac{n_{\mu}}{|\mathcal{K}|} \rceil - 1$, la matrice non ha rango pieno e quindi esiste l'insieme S definito come:

$$S = \{x : x^T R_{k,T} = 0\}$$

Se restringiamo la ricerca del minimo a tale insieme, otteniamo:

$$\begin{aligned}
\lambda_{\min}(W_d) &\leq \|V\|_2^2 \lambda_{\min} \left(\sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right) \\
&\leq \|V\|_2^2 \min_{\|x\|=1, x \in S} x^T \sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\
&\leq \|V\|_2^2 \min_{\|x\|=1} x^T \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\
&\leq \|V\|_2^2 \max_{\|x\|=1} x^T \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\
&= \|V\|_2^2 \left\| \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right\|_2 \\
&\leq \|V\|_2^2 \sum_{i=T}^{\infty} \|A_1^i\|_2 \|B_1\|_2 \|B_1^T\|_2 \|(A_1^T)^i\|_2 \\
&\leq \|V\|_2^2 \|B_1\|_2^2 \sum_{i=T}^{\infty} \|A_1^i\|_2^2
\end{aligned}$$

Per quanto riguarda la norma di B_1 abbiamo che:

$$\|B_1\|_2 \leq \|V^{-1} B_{\mathcal{K}}\|_2 \leq \|V^{-1}\|_2 \|B_{\mathcal{K}}\|_2 \leq \|V^{-1}\|_2$$

perchè la matrice di controllo ha norma unitaria, mentre per la sommatoria vale:

$$\begin{aligned}
\|A_1^i\|_2 &= \max_{\|x\|=1} x^T \begin{bmatrix} \xi^{2i} & & \\ & \ddots & \\ & & \xi^{2i} \end{bmatrix} x = \max_{\|x\|=1} \sqrt{\sum_{k=1}^{n_\mu} |\xi_k^i x_k|^2} \\
&\leq \max_{\|x\|=1} \sqrt{\max_k |\xi^i|^2 \sum_{k=1}^{n_\mu} |x_k|^2} \\
&= \max_k |\xi^i| \leq \mu^i
\end{aligned}$$

Quindi infine:

$$\lambda_{\min}(W_d) \leq \|V\|_2^2 \|V^{-1}\|_2^2 \sum_{i=T}^{\infty} \mu^{2i} = \text{cond}^2(V) \frac{\mu^{2T}}{(1 - \mu^2)}$$

dove si può scegliere $T = \frac{n_\mu}{|\mathcal{K}|} - 1$ (semplificando così la forma del bound), in quanto $\mu < 1$, e di conseguenza $\mu^{\frac{n_\mu}{|\mathcal{K}|} - 1} > \mu^{\lceil \frac{n_\mu}{|\mathcal{K}|} \rceil - 1}$. \square

La dimostrazione appena riportata si basa su due aspetti fondamentali: il primo riguarda il cambio di base tramite V , il quale ci consente di restringere il campo su cui si cerca l'autovalore minimo ad una sottomatrice stabile della matrice A , e quindi la convergenza della serie è assicurata. Il secondo aspetto invece riguarda l'introduzione del sottoinsieme S , che consente di eliminare le componenti della serie che hanno norma "grande", ovvero le prime, e di trovare di conseguenza un upper bound che non sia eccessivamente elevato (quindi inefficace).

A partire da tale risultato è possibile trovare un limite superiore sull'autovalore minimo del gramiano di raggiungibilità a tempo continuo, ma è necessario innanzitutto trovare una relazione che leghi in qualche modo i gramiani di sistemi discreti e continui.

2.3 Upper bound sull'autovalore minimo del gramiano a tempo continuo

Per trasportare i risultati già noti sull'autovalore minimo del gramiano di raggiungibilità dal tempo discreto al continuo sono necessari i seguenti risultati preliminari:

Proposizione 2.2: Data una qualsiasi matrice A quadrata e p costante, se λ è autovalore di A , allora:

- $\lambda + p$ è autovalore di $(A + pI)$
- la matrice $(A - pI)$ ammette inversa se e solo se p non è autovalore di A

Dimostrazione. Il primo risultato è di facile verifica: infatti, se prendiamo λ autovalore di A e v autovettore di A associato a λ , vale che:

$$(A + pI)v = Av + pv = \lambda v + pv = (\lambda + p)v$$

In particolare se v è autovettore di A , lo è anche di $(A + pI)$.

Il secondo risultato deriva dal fatto che, in modo analogo, $\lambda - p$ è autovalore di $(A - pI)$, e se $\lambda - p \neq 0$, ovvero $\lambda \neq p$, la matrice $(A - pI)$ risulta invertibile, non avendo autovalori nulli. In particolare, se $\lambda - p$ è autovalore di $(A - pI)$, allora $\frac{1}{\lambda - p}$ è autovalore di $(A - pI)^{-1}$, poichè l'insieme degli autovalori di una matrice inversa corrisponde all'insieme degli autovalori inversi della matrice, ovvero:

$$\lambda_i \in \text{spec}(A) \implies \frac{1}{\lambda_i} \in \text{spec}(A^{-1}), \quad i = 1, \dots, n$$

□

Nota: D'ora in avanti supporremo che p non sia autovalore di A , e che quindi la matrice $(A - pI)$ sia invertibile.

Proposizione 3: Data una qualsiasi matrice A quadrata e p costante, con $p \notin \text{spec}(A)$ i membri della matrice $(A + pI)(A - pI)^{-1}$ commutano.

Dimostrazione. Dobbiamo dimostrare che

$$(A - pI)^{-1}(A + pI) = (A + pI)(A - pI)^{-1}$$

Moltiplicando ambo i membri a destra e a sinistra per $(A - pI)$, l'equazione si riduce a:

$$(A + pI)(A - pI) = (A - pI)(A + pI)$$

A questo punto e' immediato verificare l'identità, sviluppando i termini. □

Proposizione 2.4: La matrice A è Hurwitz (ovvero ha tutti gli autovalori a parte reale strettamente negativa), se e solo se la matrice $(A + pI)(A - pI)^{-1}$, con $p \in \mathbb{R}_{>0}$, è Schur (ovvero ha tutti gli autovalori in modulo strettamente minore di 1)

Dimostrazione. Per la dimostrazione si sfrutta il fatto che, se λ è autovalore di A , allora $\frac{\lambda+p}{\lambda-p}$ è autovalore di $(A + pI)(A - pI)^{-1}$ (si noti come $(A - pI)^{-1}$ sia sempre invertibile qualora si scelga p positivo in quanto gli autovalori sono tutti a parte reale strettamente negativa). Tale affermazione deriva dalla *proposizione 2.2* e dalla *proposizione 2.3*:

$$(A + pI)(A - pI)^{-1}v = (A - pI)^{-1}(A + pI)v = (A - pI)^{-1}v(\lambda + p) = \frac{\lambda + p}{\lambda - p}$$

A questo punto dobbiamo verificare che $\lambda = \alpha + j\beta$, con $\alpha < 0$, se e solo se:

$$\left| \frac{\lambda + p}{\lambda - p} \right| < 1$$

Dalla definizione di modulo nel campo complesso e sapendo che:

$$\sqrt{x} < 1 \iff x < 1, \quad x > 0$$

otteniamo la serie di disuguaglianze:

$$\begin{aligned} \frac{\sqrt{(\alpha + p)^2 + \beta^2}}{\sqrt{(\alpha - p)^2 + \beta^2}} < 1 &\iff \frac{(\alpha + p)^2 + \beta^2}{(\alpha - p)^2 + \beta^2} < 1 \\ \frac{(\alpha + p)^2 + \beta^2}{(\alpha - p)^2 + \beta^2} < 1 &\iff (\alpha + p)^2 + \beta^2 < (\alpha - p)^2 + \beta^2 \\ (\alpha + p)^2 + \beta^2 < (\alpha - p)^2 + \beta^2 &\iff 4\alpha p < 0 \end{aligned}$$

e l'ultima è verificata. □

Grazie a tali risultati è possibile ora enunciare il seguente teorema, che ci fornisce una relazione fra il gramiano di raggiungibilità a tempo discreto e quello a tempo continuo, e deriva dalla teoria di Lyapunov (riportata più ampiamente in [3]):

Teorema 2.2: Il gramiano su orizzonte infinito del sistema continuo

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

con A matrice Hurwitz, coincide con il gramiano su orizzonte infinito del sistema discreto

$$x(t+1) = A(p)x(t) + B(p)u(t)$$

$\forall p \in \mathbb{R}_{>0}$, dove

$$A(p) := (A + pI)(A - pI)^{-1}, \quad B(p) = \sqrt{(2p)}(A - pI)^{-1}B$$

Dimostrazione. E' noto che il gramiano di raggiungibilità del sistema continuo (A, B) , definito come:

$$W_c = \int_0^\infty e^{At}BB^Te^{A^T t} dt$$

è soluzione, per qualsiasi $p > 0$, della seguente equazione (*ADI Iteration*) ([3]):

$$X_{i+1} = (A + pI)(A - pI)^{-1}X_i(A^T + pI)(A^T - pI)^{-1} + 2p(A - pI)^{-1}BB^T(A^T - pI)^{-1}$$

una volta fissato $X_{i+1} = X_i = X$.

Infatti, data:

$$X = (A + pI)(A - pI)^{-1}X(A^T + pI)(A^T - pI)^{-1} + 2p(A - pI)^{-1}BB^T(A^T - pI)^{-1}$$

moltiplicando ambo i membri a destra per $(A - pI)$ e a sinistra per $(A^T - pI)$ e ricordando che i due membri della matrice $A(p)$ commutano (si veda *proposizione 2.3*) otteniamo:

$$(A - pI)X(A^T - pI) = (A + pI)X(A^T + pI) + 2pBB^T$$

la quale è equivalente, sviluppando i prodotti, a:

$$-pXA^T - pAX = pAX + pXA^T + 2pBB^T$$

ovvero all'equazione di Lyapunov a tempo continuo (CALE)

$$XA^T + XA + BB^T = 0$$

di cui il gramiano a tempo continuo su orizzonte infinito è soluzione. Scegliendo $A(p), B(p)$ come da definizione (in particolare $A(p)$ è Schur come da *proposizione 2.4*) abbiamo che

l'iterazione si può scrivere come:

$$X = A(p)XA(p)^T + B(p)B(p)^T$$

la quale rappresenta l'equazione di Lyapunov a tempo discreto (DALE) e la cui soluzione è il gramiano a tempo discreto su orizzonte infinito

$$W_d = \sum_{i=0}^{\infty} A(p)^i B B^T (A(p)^T)^i$$

del sistema:

$$x(t+1) = A(p)x(t) + B(p)u(t)$$

□

Possiamo sfruttare ora la relazione appena ricavata per trovare un'upper bound sull'autovalore minimo del gramiano di raggiungibilità a tempo continuo, sfruttando i risultati già noti nel caso discreto, in quanto il gramiano del sistema a tempo continuo coincide con un gramiano a tempo discreto con determinate matrici legate a quelle del sistema continuo di partenza. Prima di procedere sono tuttavia necessari i seguenti fatti, che torneranno utili nella dimostrazione del teorema fondamentale nel caso continuo:

Proposizione 2.5: Se la matrice A è diagonalizzabile dalla base di autovettori V , allora anche la matrice $A(p)$, definita in precedenza, è diagonalizzabile dalla stessa base V .

Dimostrazione. Dimostriamo innanzitutto che se V diagonalizza A , allora diagonalizza anche $(A + pI)$; infatti:

$$\begin{aligned} V(A + pI)V^{-1} &= VAV^{-1} + pVV^{-1} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} + pI \\ &= \begin{bmatrix} \lambda_1 + p & & \\ & \ddots & \\ & & \lambda_n + p \end{bmatrix} \end{aligned}$$

Analogamente V diagonalizza anche $A - pI$. Quindi vale che:

$$V(A + pI)V^{-1}(V(A - pI)V^{-1})^{-1} = \begin{bmatrix} \frac{\lambda_1 + p}{\lambda_1 - p} & & \\ & \ddots & \\ & & \frac{\lambda_n + p}{\lambda_n - p} \end{bmatrix}$$

ma sviluppando il primo membro si ottiene:

$$\begin{aligned} V(A + pI)V^{-1}(V(A - pI)V^{-1})^{-1} &= V(A + pI)V^{-1}V(A - pI)^{-1}V^{-1} \\ &= V(A + pI)(A - pI)^{-1}V^{-1} \end{aligned}$$

Quindi V diagonalizza $A(p)$. □

Proposizione 2.6: L'insieme degli autovalori ξ di $A(p)$, $p \in \mathbb{R}_{>0}$, definito come:

$$\Xi_{\mu,p} = \{\xi \in \text{spec}(A(p)) : |\xi| \leq \mu, \mu \in (0, 1)\}$$

coincide con l'insieme:

$$\Lambda_{\mu,p} = \text{spec}(A) \cap I_{\mu,p}$$

dove

$$I_{\mu,p} = \{x \in \mathbb{C} : \Re(x)^2 + \Im(x)^2 + 2p \frac{1 + \mu^2}{1 - \mu^2} \Re(x) + p^2 \leq 0, \mu \in (0, 1)\}$$

Dimostrazione. Ricordando che la relazione che lega gli autovalori di A con gli autovalori di $A(p)$ è la seguente:

$$\xi = \frac{\lambda + p}{\lambda - p}$$

si tratta di dimostrare che:

$$\left| \frac{\lambda + p}{\lambda - p} \right| \leq \mu \iff \lambda \in \Lambda_{\mu,p}$$

Sostituendo nella disuguaglianza $\lambda = \Re(\lambda) + j\Im(\lambda)$ e sfruttando il fatto che $\mu \in (0, 1)$, dopo alcuni passaggi, si ottiene:

$$\Re(\lambda)^2 + \Im(\lambda)^2 + 2p \frac{1 + \mu^2}{1 - \mu^2} \Re(\lambda) + p^2 \leq 0$$

ovvero l'insieme dei punti interni ad una circonferenza sul piano complesso di centro e raggio pari a:

$$(x_c, y_c) = \left(-p \frac{1 + \mu^2}{1 - \mu^2}, 0 \right), \quad r = p \frac{2\mu}{1 - \mu^2}$$

In particolare, se A è una matrice simmetrica, ovvero tutti i suoi autovalori sono reali, l'insieme si riduce a:

$$\Lambda_{\mu,p} = \{\lambda \in \text{spec}(A) : -p \frac{1 + \mu}{1 - \mu} \leq \lambda \leq -p \frac{1 - \mu}{1 + \mu}, \mu \in (0, 1)\}$$

□

Per maggiore chiarezza sono riportati i grafici dell'insieme $I_{\mu,p}$ al variare di μ e p :

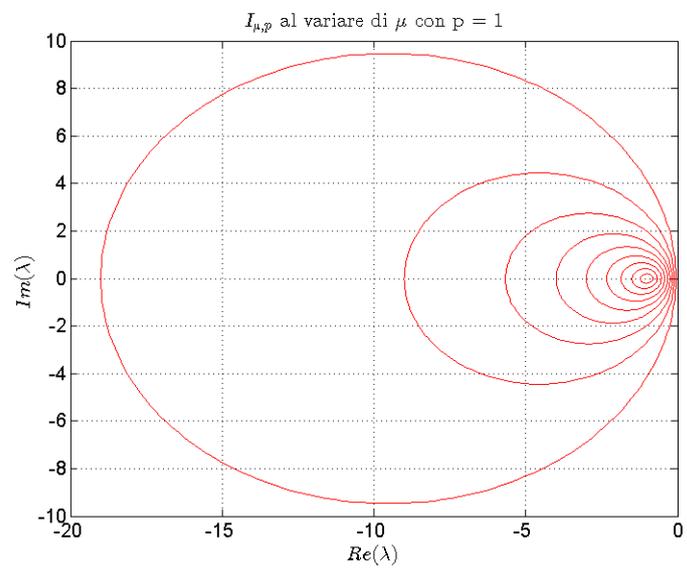


Figura 2.1: variazione dell'insieme $I_{\mu,p}$ rispetto al parametro μ

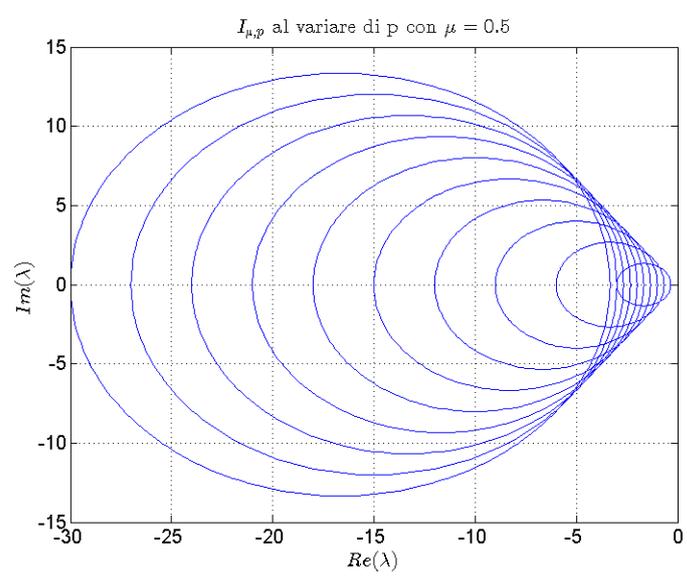


Figura 2.2: variazione dell'insieme $I_{\mu,p}$ rispetto al parametro p

In particolare si può notare che:

- Fissato p , nel caso in cui μ tenda a 1, il raggio del cerchio tende all'infinito, fino a comprendere l'intero semipiano sinistro del piano complesso, in accordo con la *proposizione 2.4*; se invece μ tende a 0, il cerchio si riduce ad un intorno infinitesimale di $-p$ fino a collassare sul punto stesso, in quanto la matrice $A(p)$ risulta non invertibile (cioè presenta autovalori nulli) anche nel caso in cui $(A + pI)$ non sia invertibile (ovvero un autovalore di A coincide con $-p$) (si veda *figura 2.1*).
- Fissato μ è facile immaginare che il raggio e il centro dei cerchi rispettivamente cresca e decresca in modo lineare rispetto a p (si veda *figura 2.2*).

Ecco quindi il teorema fondamentale che consente di stabilire un limite superiore all'autovalore minimo del gramiano di raggiungibilità a tempo continuo:

Teorema 2.3 (upper bound sull'autovalore minimo del gramiano a tempo continuo): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con $|\mathcal{V}| = n$, matrice di adiacenza A Hurwitz diagonalizzabile dalla matrice di autovettori V , insieme di nodi di controllo \mathcal{K} , $p \in \mathbb{R}_{>0}$ e $n_{\mu,p} = |\Lambda_{\mu,p}|$, si ha che $\forall \mu \in [|\lambda_{\min}(A(p))|, 1)$ vale che:

$$\lambda_{\min}(W_c) \leq \text{cond}^2(V) \frac{\mu^{2\frac{n_{\mu,p}}{|\mathcal{K}|}}}{\mu^2(1-\mu^2)} \frac{2p}{(\sigma_{\min}(A-pI))^2}$$

dove W_c è il gramiano di controllo su orizzonte infinito del sistema continuo $(A, B_{\mathcal{K}})$

Dimostrazione. Dal *teorema 2.2* sappiamo che il gramiano W_c del sistema continuo $(A, B_{\mathcal{K}})$ coincide con il gramiano W_d del sistema discreto $(A(p), B(p))$ definito in precedenza, quindi abbiamo anche che i loro autovalori minimi coincidono, cioè $\lambda_{\min}(W_c) = \lambda_{\min}(W_d)$.

La *proposizione 2.5* afferma che se A è diagonalizzabile da V , anche $A(p)$ lo è, e quindi, senza perdita di generalità, supponiamo che le colonne di $A(p)$ siano ordinate in modo tale che

$$V^{-1}A(p)V = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad V^{-1}B(p) = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

con $A_1 \in \mathbb{R}^{n_{\mu,p} \times n_{\mu,p}}$ e A_2 matrici diagonali e A_1 matrice di autovalori tali che:

$$\text{spec}(A_1) = \{\xi : \xi \in \text{spec}(A(p)), |\xi| \leq \mu\}$$

ovvero, per la *proposizione 2.6*, gli autovalori λ appartenenti all'insieme $\Lambda_{\mu,p}$. Osserviamo che, similmente a quanto già visto nella dimostrazione del *teorema 2.1*:

$$\begin{aligned} W_d &= \sum_{i=0}^{\infty} A(p)^i B B^T (A(p)^T)^i = V \sum_{i=0}^{\infty} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}^i \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} B_1^T \\ B_2^T \end{bmatrix}^T \begin{bmatrix} A_1^T & 0 \\ 0 & A_2^T \end{bmatrix}^i V^T \\ &= V \tilde{W}_d V^T \end{aligned}$$

dove si è sfruttato il fatto che le matrici A_1, A_2 sono simmetriche. Siccome anche W_d è simmetrica:

$$\begin{aligned}\lambda_{\min}(W_d) &= \min_{\|x\|=1} x^T W_d x = \min_{\|x\|=1} x^T V \tilde{W}_d V^T x \leq \|V\|_2^2 \min_{\|y\|=1} y^T \tilde{W}_d y \\ &\leq \|V\|_2^2 \min_{\|y_1\|=1} [y_1^T \quad 0] \tilde{W}_d \begin{bmatrix} y_1^T \\ 0 \end{bmatrix} = \|V\|_2^2 \lambda_{\min} \left(\sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right)\end{aligned}$$

Consideriamo ancora una volta la matrice di raggiungibilità del sistema (A_1, B_1) al tempo T definita come:

$$R_{k,T} = [B_1 \quad A_1 B_1 \quad \dots \quad A_1^{T-1} B_1], \quad R_{k,T} \in \mathbb{R}^{n_\mu \times T|\mathcal{K}|}$$

Se vale la condizione $T|\mathcal{K}| < n_{\mu,p}$, ovvero $T \leq \lceil \frac{n_{\mu,p}}{|\mathcal{K}|} \rceil - 1$, la matrice non ha rango pieno e quindi esiste l'insieme S definito come:

$$S = \{x : x^T R_{k,T} = 0\}$$

A questo punto vale la serie di disuguaglianze:

$$\begin{aligned}\lambda_{\min}(W_d) &\leq \|V\|_2^2 \lambda_{\min} \left(\sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right) \\ &\leq \|V\|_2^2 \min_{\|x\|=1, x \in S} x^T \sum_{i=0}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\ &\leq \|V\|_2^2 \min_{\|x\|=1} x^T \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\ &\leq \|V\|_2^2 \max_{\|x\|=1} x^T \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i x \\ &= \|V\|_2^2 \left\| \sum_{i=T}^{\infty} A_1^i B_1 B_1^T (A_1^T)^i \right\|_2 \\ &\leq \|V\|_2^2 \sum_{i=T}^{\infty} \|A_1^i\|_2 \|B_1\|_2 \|B_1^T\|_2 \|(A_1^T)^i\|_2 \\ &\leq \|V\|_2^2 \|B_1\|_2^2 \sum_{i=T}^{\infty} \|A_1^i\|_2^2\end{aligned}$$

Per quanto riguarda la norma di B_1 , stavolta abbiamo che:

$$\begin{aligned}\|B_1\|_2 &\leq \|V^{-1}B(p)\|_2 \leq \|V^{-1}\|_2\|B(p)\|_2 \leq \|V^{-1}\|_2\sqrt{2p}\|(A-pI)^{-1}\|_2\|B_{\mathcal{K}}\|_2 \\ &= \|V^{-1}\|_2\sqrt{2p}\|(A-pI)^{-1}\|_2\end{aligned}$$

dove si è sfruttato il fatto che la norma della matrice $B_{\mathcal{K}}$ è unitaria. In particolare:

$$\|(A-pI)^{-1}\|_2 = \max_{\|x\|_2=1} x^T(A^T - pI)^{-1}(A-pI)^{-1}x$$

Se definiamo $y = (A-pI)^{-1}x$, possiamo riscrivere la norma come:

$$\|(A-pI)^{-1}\|_2 = \frac{1}{\min_{\|x\|_2=1} y^T(A^T - pI)(A-pI)y} = \frac{1}{\sigma_{\min}(A-pI)}$$

Per la sommatoria invece vale:

$$\begin{aligned}\|A_1^i\|_2 &= \max_{\|x\|=1} x^T \begin{bmatrix} \xi_1^{2i} & & \\ & \ddots & \\ & & \xi_{n_\mu}^{2i} \end{bmatrix} x = \max_{\|x\|=1} \sqrt{\sum_{k=1}^{n_{\mu,p}} |\xi_k^i x_k|^2} \\ &\leq \max_{\|x\|=1} \sqrt{\max_k |\xi^i|^2 \sum_{k=1}^{n_{\mu,p}} |x_k|^2} \\ &= \max_k |\xi^i| \leq \mu^i\end{aligned}$$

Quindi:

$$\begin{aligned}\lambda_{\min}(W_d) &\leq \|V\|_2^2 \|V^{-1}\|_2^2 \sum_{i=T}^{\infty} \mu^{2i} \frac{2p}{(\sigma_{\min}(A-pI))^2} \\ &= \text{cond}^2(V) \frac{\mu^{2T}}{(1-\mu^2)} \frac{2p}{(\sigma_{\min}(A-pI))^2}\end{aligned}$$

Essendo $\mu < 1$ possiamo scegliere $T = \frac{n_\mu}{|\mathcal{K}|} - 1 \leq \lceil \frac{n_\mu}{|\mathcal{K}|} \rceil - 1$, da cui la tesi. \square

Il risultato appena provato ci consente di trovare un'upper bound sull'autovalore minimo del gramiano di raggiungibilità a tempo discreto. Tuttavia, è facile notare che l'introduzione della matrice $B(p)$, e in particolare il limite superiore sulla sua norma, complica l'analisi, in quanto ora compare anche il valore singolare minimo della matrice $A-pI$, che in generale non è semplice da trattare. Inoltre, la stessa dipendenza dal parametro p ci fornisce un ulteriore grado di libertà che allarga la famiglia di upper bound, già dipendenti da μ . Malgrado ciò, vedremo che è possibile ovviare a tali problematiche nel caso in cui la matrice del sistema sia simmetrica.

2.4 Caso simmetrico e risultati secondari

Il *teorema 2.3* appena ottenuto può essere semplificato nel caso in cui la matrice A sia simmetrica, oltre che Hurwitz; sotto questa ulteriore ipotesi infatti vale il seguente corollario:

Corollario 2.1: Se la matrice A è Hurwitz e simmetrica, $p \in \mathbb{R}_{>0}$ e valgono le stesse ipotesi del *teorema 2.2*, si ha che:

$$\lambda_{\min}(W_c) \leq \frac{\mu^{2\frac{n\mu,p}{|K|}}}{\mu^2(1-\mu^2)} \frac{2p}{(|\lambda_{\min}(A)| + p)^2}$$

Dimostrazione. Se A è simmetrica, ovviamente $\text{cond}(V) = 1$, perchè la base che la diagonalizza risulta ortonormale.

Inoltre, siccome anche $A - pI$ è simmetrica, i suoi valori singolari risultano essere i moduli dei suoi autovalori, ovvero $|\lambda - p|$, con $\lambda \in \text{spec}(A)$. Quindi:

$$\sigma_{\min}(A - pI) = \min_{\lambda \in \text{spec}(A)} |\lambda - p| = \min_{\lambda \in \text{spec}(A)} |\lambda| + p = |\lambda_{\min}(A) - p| = |\lambda_{\min}(A)| + p$$

dove si è sfruttato il fatto che tutti gli autovalori λ sono reali (essendo A simmetrica) e negativi, di conseguenza l'autovalore corrispondente al valore singolare minimo di $A - pI$ è proprio l'autovalore in modulo più piccolo di A . \square

Nota: Per matrici simmetriche il bound risulta più stringente, perchè vale:

$$\|A^t\|_2 = \|A\|_2^t, \quad \forall t > 0, \quad A = A^T$$

Infatti, se la matrice A è simmetrica, abbiamo che:

$$\|A^t\|_2 = |\lambda_{\max}(A^t)| = |\lambda_{\max}(A)^t| = |\lambda_{\max}(A)|^t = \|A\|_2^t$$

in quanto, se λ_i è autovalore di A , allora λ_i^t è autovalore di A^t . La procedura di maggiorazione dell'autovalore minimo del gramiano di raggiungibilità è quindi più efficace per matrici simmetriche.

Sempre supponendo A simmetrica e Hurwitz è possibile ottenere un ulteriore bound sul minimo autovalore del gramiano di raggiungibilità, ma è necessario il seguente risultato:

Proposizione 2.7: Se la matrice A Hurwitz del sistema continuo è simmetrica allora anche la matrice $A(p)$ definita in precedenza è simmetrica

Dimostrazione. Supposto che A sia simmetrica, dobbiamo dimostrare che lo è anche la matrice

$$A(p) := (A + pI)(A - pI)^{-1}$$

Ovviamente i due membri di $A(p)$ sono simmetrici perche derivati da somme, sottrazioni e inversioni di matrici simmetriche. Per verificare che $A(p)$ è simmetrica sfruttiamo la proprietà che afferma che il prodotto fra due matrici simmetriche è simmetrico se e solo se tale prodotto è commutativo, ovvero che:

$$(A - pI)^{-1}(A + pI) = (A + pI)(A - pI)^{-1}$$

identità già verificata in *proposizione 2.3* □

Corollario 2.2: Se valgono le stesse ipotesi del *teorema 2.3*, con A matrice simmetrica e Hurwitz, allora:

$$\lambda_{\min}(W_c) \leq \min \left\{ \frac{1 - \lambda_{\min}(A(p))^4}{1 - \lambda_{\min}(A(p))^2} \frac{2p}{(\sigma_{\min}(A + pI))^2}, \frac{|\lambda_{\max}(A(p))|^{2(T_{\max}-1)}}{1 - |\lambda_{\max}(A(p))|^2} \frac{2p}{(|\lambda_{\min}(A)| + p)^2} \right\}$$

dove $T_{\max} = \frac{n}{|\mathcal{K}|}$, $p \in \mathbb{R}_{>0}$. Se inoltre $p > \|A\|_2$ valgono le relazioni:

$$|\lambda_{\min}(A(p))| = \left| \frac{\lambda_{\max}(A) + p}{\lambda_{\max}(A) - p} \right|, \quad |\lambda_{\max}(A(p))| = \left| \frac{\lambda_{\min}(A) + p}{\lambda_{\min}(A) - p} \right|$$

Dimostrazione. La prima parte della disuguaglianza deriva dal fatto che, se scegliamo come matrice di controllo B del sistema continuo

$$B_{\mathcal{V}} = [B_{\mathcal{K}} \quad B_{\mathcal{V}/\mathcal{K}}]$$

ovvero controlliamo ogni singolo nodo della rete aggiungendo i vettori della base canonica mancanti alla matrice $B_{\mathcal{K}}$, ne deriva che per il gramiano di controllabilità a tempo discreto corrispondente vale:

$$W_{\mathcal{V}} = W_{\mathcal{K}} + W_{\mathcal{V}/\mathcal{K}}$$

Essendo entrambi i membri matrici definite positive:

$$\lambda_{\min}(W_{\mathcal{K}}) \leq \lambda_{\min}(W_{\mathcal{V}})$$

dove $W_{\mathcal{K}}$ è W_d della precedente notazione.

A questo punto, sapendo che $B_{\mathcal{V}}B_{\mathcal{V}}^T = I$, vale la serie di disuguaglianze:

$$\begin{aligned}
\lambda_{\min}(W_c) &\leq \min_{\|x\|=1} x^T 2p \sum_{t=0}^{\infty} A(p)^t (A - pI)^{-1} B_V B_V^T (A - pI)^{-1} (A(p)^T)^t x \\
&= \min_{\|x\|=1} x^T 2p \sum_{t=0}^{\infty} A(p)^t (A - pI)^{-1} (A - pI)^{-1} (A(p))^t x \\
&= \min_{\|x\|=1} x^T 2p (A + pI)^{-1} \sum_{t=0}^{\infty} (A + pI) ((A - pI)^{-1} (A + pI))^t (A - pI)^{-1} \\
&\quad (A - pI)^{-1} ((A + pI) (A - pI)^{-1})^t (A + pI) (A + pI)^{-1} x \\
&= \min_{\|x\|=1} x^T 2p (A + pI)^{-1} \sum_{t=1}^{\infty} A(p)^{2t} (A + pI)^{-1} x \\
&\leq 2p \|(A + pI)^{-1}\|_2^2 \min_{\|y\|=1} y^T \sum_{t=1}^{\infty} A(p)^{2t} y \\
&\leq \frac{2p}{(\sigma_{\min}(A + pI))^2} \lambda_{\min} \left(\sum_{t=1}^{\infty} A(p)^{2t} \right) \\
&\leq \frac{2p}{(\sigma_{\min}(A + pI))^2} \frac{1 - \lambda_{\min}(A(p))^4}{1 - \lambda_{\min}(A(p))^2}
\end{aligned}$$

dove si è sfruttata la simmetria di $A(p)$ e la commutatività delle sue componenti. La seconda parte della disuguaglianza deriva invece dal risultato ottenuto dal *teorema 2.3*, ricordando che, essendo $A(p)$ simmetrica, $\text{cond}(V) = 1$, e sostituendo $\mu = |\lambda_{\max}(A(p))| < 1$ (in modo tale che $n_{\mu,p} = n$, ovvero T_{\max} perda la dipendenza da p). Per dimostrare che:

$$|\lambda_{\min}(A(p))| = \left| \frac{\lambda_{\max}(A) + p}{\lambda_{\max}(A) - p} \right|, \quad |\lambda_{\max}(A(p))| = \left| \frac{\lambda_{\min}(A) + p}{\lambda_{\min}(A) - p} \right|$$

qualora $p > \|A\|_2$, ricordiamo innanzitutto che, essendo A simmetrica:

$$\|A\|_2 = |\lambda_{\max}(A)| \implies -p < \lambda \quad \forall \lambda \in A$$

Ora, prendiamo due autovalori di A , λ_1 e λ_2 , tali che $\lambda_2 < \lambda_1$. Essendo reali e negativi:

$$|\lambda_2| > |\lambda_1|, \quad -p < \lambda_2 < \lambda_1$$

A questo punto è facile verificare che:

$$\left| \frac{\lambda_1 + p}{\lambda_1 - p} \right| > \left| \frac{\lambda_2 + p}{\lambda_2 - p} \right|$$

Infatti, essendo i due termini all'interno dei moduli strettamente negativi, la disequazione si riduce a:

$$\frac{\lambda_1 + p}{p - \lambda_1} > \frac{\lambda_2 + p}{p - \lambda_2}$$

che risulta valida se e solo se $\lambda_2 < \lambda_1$. Quindi, $\forall \lambda \in A$, abbiamo che:

$$\left| \frac{\lambda_{\min}(A) + p}{\lambda_{\min}(A) - p} \right| \geq \left| \frac{\lambda + p}{\lambda - p} \right| \geq \left| \frac{\lambda_{\max}(A) + p}{\lambda_{\max}(A) - p} \right|$$

□

In sintesi nel caso simmetrico, e sotto opportune condizioni su p , è possibile stabilire anche una relazione diretta fra gli autovalori della matrice A e quelli della matrice $A(p)$. Il seguente corollario invece permette di determinare il numero minimo di nodi da controllare nel grafo in modo tale da soddisfare un limite imposto sull'energia disponibile, ovvero un lower bound all'autovalore minimo del gramiano:

Corollario 2.3: Supponendo le stesse condizioni del *teorema 2.3*, se

$$\lambda_{\min}(W_c) \geq \varepsilon$$

con $\varepsilon \in \mathcal{R}_{>0}$ allora:

$$|\mathcal{K}| \geq R_{\varepsilon,p} n_{\mu,p}$$

dove

$$R_{\varepsilon,p} = \frac{2 \log(\mu)}{\log(\varepsilon) - \log(2p) + \log(\mu^2(1 - \mu^2)) - 2 \log(\text{cond}(V)) + 2 \log(\sigma_{\min}(A - pI))}$$

Dimostrazione. Dal *teorema 2.2* sappiamo che

$$\lambda_{\min}(W_c) \geq \varepsilon \iff \varepsilon \leq \text{cond}^2(V) \frac{\mu^{2(\frac{n_{\mu,p}}{|\mathcal{K}}| - 1)}}{1 - \mu^2} \frac{2p}{(\sigma_{\min}(A - pI))^2}$$

ovvero, sfruttando le proprietà dei logaritmi:

$$\frac{\log(\varepsilon) - \log(2p) + \log(\mu^2(1 - \mu^2)) - 2 \log(\text{cond}(V)) + 2 \log(\sigma_{\min}(A - pI))}{\log(\mu^2)} \geq \frac{n_{\mu,p}}{|\mathcal{K}|}$$

che corrisponde alla tesi del corollario. □

2.5 Minimizzazione dell'upper bound a tempo continuo

I risultati finora enunciati ci forniscono un limite superiore all'autovalore minimo del gramiano di raggiungibilità a tempo continuo; tuttavia sorge spontaneo chiedersi quanto restrittivi siano effettivamente questi bound, ovvero quanto si avvicinano all'autovalore minimo del gramiano. In altri termini, esiste un valore p "ottimale" che renda il bound il più restrittivo possibile?

Il *teorema 2.3* infatti ci fornisce una famiglia di limiti, che assumono valori diversi al variare sia di μ che di p . Ora, fissato un certo μ , cercheremo di capire per quali p è possibile ottenere i risultati migliori. In particolare, una scelta "comoda" è $\mu = |\lambda_{\max}(A(p))|$: in questo caso, la variabile T perde la dipendenza sia da μ che da p , in quanto vengono conteggiati tutti gli autovalori della matrice $A(p)$ (e quindi di A).

Allo scopo di chiarificare il problema, è riportato un esempio:

Esempio 2.1 (sistema tridiagonale Toeplitz): Un sistema tridiagonale Toeplitz è un particolare grafo con matrice di adiacenza:

$$A = \begin{bmatrix} a & b & & & \\ c & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

la quale è una matrice tridiagonale (cioè ha elementi nulli al di fuori della diagonale principale e delle sovra e sotto diagonali), ed inoltre gli elementi appartenenti alla stessa diagonale sono identici fra loro. In questo caso particolare esistono formule chiuse per gli autovalori di tale matrice (si veda [4]), ovvero:

$$\lambda(A) = \begin{cases} a + 2\operatorname{sgn}(b)\sqrt{bc} \cos\left(\frac{k\pi}{n+1}\right) & k = 1, \dots, n \quad \text{se } bc > 0 \\ a + 2j\operatorname{sgn}(b)\sqrt{|bc|} \sin\left(\frac{k\pi}{2h+2}\right) & k = -h, \dots, h \quad \text{se } bc < 0, n = 2h + 1 \\ a + 2j\operatorname{sgn}(b)\sqrt{|bc|} \sin\left(\frac{(2k-1)\pi}{4h+2}\right) & k = -h + 1, \dots, h \quad \text{se } bc < 0, n = 2h \end{cases}$$

Quindi, affinché la matrice sia Hurwitz nel caso in cui b, c siano positivi, è sufficiente scegliere $a < -2\sqrt{bc}$. Inoltre, in questa situazione, tutti gli autovalori risultano reali. Scegliamo quindi $a = -5$, $b = 1$, $c = 2$ e come dimensione del sistema $n = 5$. Infine supponiamo di controllare solo il primo nodo della catena, quindi $|\mathcal{K}| = 1$.

In *figura 2.3* è riportato l'andamento in scala logaritmica dell'upper bound (rosso) sull'autovalore minimo del gramiano di raggiungibilità come espresso nel *teorema 2.3* e delle sue singole componenti rispetto alla variabile p (in blu la parte riguardante l'autovalore massimo di $A(p)$, in fucsia il membro dato dalla norma di $B(p)$, in verde l'autovalore massimo di $A(p)$).

upper bound su $\lambda_{\min}(W_c)$ per A tridiagonale Toeplitz (a = -5, b = 1, c = 2, $|\mathcal{K}| = 1$)

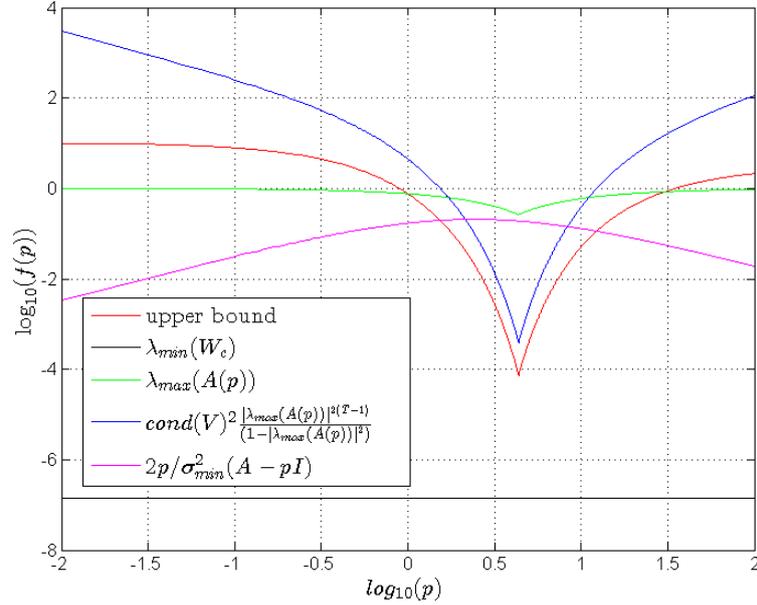


Figura 2.3: variazione dell'upper bound rispetto a p

Si possono formulare alcune osservazioni:

- la componente data da $\frac{2p}{\sigma_{\min}(A-pI)^2}$ presenta un andamento piuttosto “piatto”, con una variazione decisamente trascurabile rispetto alla parte relativa all'autovalore massimo di $A(p)$
- esiste un valore ottimo di p per cui l'upper bound ha un minimo globale e si avvicina maggiormente al valore reale dell'autovalore minimo del gramiano di raggiungibilità
- Tale valore sembra legato direttamente all'autovalore massimo di $A(p)$

Il grafico quindi suggerisce che il valore ottimale di p in modo tale da minimizzare il bound si riferisca all'autovalore massimo di $A(p)$, che a sua volta è legato agli autovalori di A . La verifica che effettivamente esiste questa relazione è evidente dalla figura 2.4 che riporta l'andamento dei moduli degli autovalori di $A(p)$, e quindi di A , rispetto al parametro p .

In particolare λ_1 è l'autovalore di $A(p)$ relativo all'autovalore massimo di A (cioè a parte reale più negativa), mentre λ_4 all'autovalore minimo di A (cioè a parte reale meno negativa). Inoltre, il valore ottimo di p è il punto esatto in cui vi è un cambio di autovalore massimo in $A(p)$, e corrisponde alla media geometrica degli autovalori massimo e minimo di A . La spiegazione di tale comportamento è data dal seguente corollario 2.4

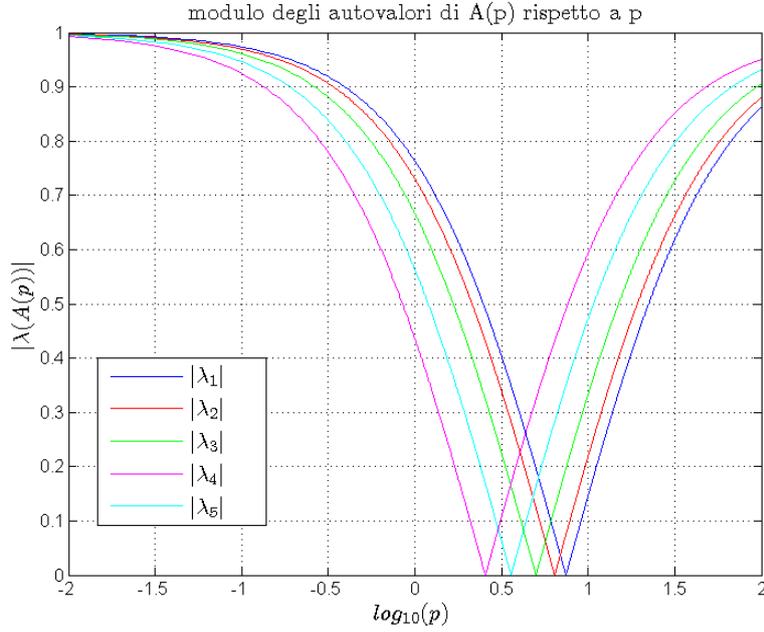


figura 2.4: variazione dei moduli degli autovalori di $A(p)$ rispetto a p

Corollario 2.4: Dato un grafo \mathcal{G} che soddisfa le stesse ipotesi del *teorema 3*, con matrice di adiacenza A Hurwitz tale che tutti i suoi autovalori siano reali, e scelti $\mu = |\lambda_{max}(A(p))|$ e $|\mathcal{K}| \geq 2n$, si ha che la quantità

$$f(p) := \frac{|\lambda_{max}(A(p))|^{2\frac{n}{|\mathcal{K}|}}}{|\lambda_{max}(A(p))|^2(1 - |\lambda_{max}(A(p))|^2)}$$

è minimizzata da $p^* = \sqrt{\lambda_{max}(A)\lambda_{min}(A)}$

Dimostrazione. Innanzitutto dimostriamo che $f(p)$ è una funzione strettamente crescente rispetto a $|\lambda_{max}(A(p))|$ se $T = \frac{n}{|\mathcal{K}|} > 2$ (ricordiamo $n_{\mu,p} = n$ visto che $\mu = |\lambda_{max}(A(p))|$). Infatti:

$$\frac{d}{d|\lambda_{max}(A(p))|} f(p) = \frac{2|\lambda_{max}(A(p))|^{2T+1}(T-1 + (2-T)|\lambda_{max}(A(p))|^2)}{(|\lambda_{max}(A(p))|^2 - |\lambda_{max}(A(p))|^4)^2}$$

risulta positiva se e solo se:

$$\begin{aligned} \frac{d}{d|\lambda_{max}(A(p))|} f(p) > 0 &\iff (T-1 + (2-T)|\lambda_{max}(A(p))|^2) > 0 \\ &\iff |\lambda_{max}(A(p))|^2 < \frac{T-1}{T-2} \end{aligned}$$

e l'ultima disuguaglianza risulta verificata in quanto la funzione $\frac{T-1}{T-2}$ è strettamente maggiore di 1 per $T > 2$, mentre $\lambda_{max}(A(p))^2 < 1$ in quanto A era Hurwitz.

Quindi il minimo di $f(p)$ viene raggiunto quando $|\lambda_{max}(A(p))|$ è minimo. Se $T = 2$ è facile verificare che si ottiene il medesimo risultato.

Ora è necessario capire come si comporta $|\lambda_{max}(A(p))|$ al variare di p . Sappiamo che la relazione che lega gli autovalori di $A(p)$ e A è la seguente:

$$\lambda(A(p)) = \frac{\lambda(A) + p}{\lambda(A) - p}$$

la quale è una funzione strettamente decrescente in p se A è Hurwitz, essendo:

$$\frac{d}{dp} \frac{\lambda(A) + p}{\lambda(A) - p} = \frac{2\lambda(A)}{(\lambda(A) - p)^2} < 0$$

Ma siccome per il calcolo della derivata di $\lambda_{max}(A(p))$ va considerato il modulo di tale funzione, possiamo in alternativa verificare il segno della derivata di $\lambda_{max}(A(p))^2$:

$$\frac{d}{dp} \left(\frac{\lambda(A) + p}{\lambda(A) - p} \right)^2 = \frac{4\lambda(A)(\lambda(A) + p)}{(\lambda(A) - p)^3} < 0$$

la quale risulta:

- decrescente se $p < -\lambda(A)$
- crescente se $p > -\lambda(A)$

Valgono inoltre i seguenti risultati, facilmente dimostrabili:

- se $p^2 \leq \lambda_{min}(A)\lambda_{max}(A)$, si ha che, $\forall \lambda(A) : \lambda_{max}(A) \leq \lambda(A) \leq \lambda_{min}(A)$,

$$|\lambda_{max}(A(p))| = \left| \frac{\lambda_{max}(A) + p}{\lambda_{max}(A) - p} \right| \geq \left| \frac{\lambda(A) + p}{\lambda(A) - p} \right|$$

- se $p^2 \geq \lambda_{min}(A)\lambda_{max}(A)$, si ha che, $\forall \lambda(A) : \lambda_{max}(A) \leq \lambda(A) \leq \lambda_{min}(A)$,

$$|\lambda_{max}(A(p))| = \left| \frac{\lambda_{min}(A) + p}{\lambda_{min}(A) - p} \right| \geq \left| \frac{\lambda(A) + p}{\lambda(A) - p} \right|$$

Quindi, considerando che $\left| \frac{\lambda_{max}(A)+p}{\lambda_{max}(A)-p} \right|$ è una funzione decrescente in $\lambda_{max}(A)$ per $p^2 \leq \lambda_{min}(A)\lambda_{max}(A)$ (essendolo per $-p > \lambda_{max}(A)$), mentre $\left| \frac{\lambda_{min}(A)+p}{\lambda_{min}(A)-p} \right|$ è una funzione crescente in $\lambda_{min}(A)$ per $p^2 \geq \lambda_{min}(A)\lambda_{max}(A)$ (essendolo per $-p < \lambda_{min}(A)$), il minimo di $|\lambda_{max}(A(p))|$ viene raggiunto proprio in $p^2 = \lambda_{min}(A)\lambda_{max}(A)$ \square

Nel caso in cui la matrice A abbia autovalori complessi, studiare il valore ottimo di p risulta decisamente più complicato. Tuttavia, nell'eventualità in cui tutti gli autovalori possano assumere valori complessi ma presentino la stessa parte reale (come accade nei sistemi tridiagonali Toeplitz qualora si scelgano b, c di segno discorde), è possibile ricavare facilmente il p desiderato.

Corollario 2.5: Dato un grafo \mathcal{G} che soddisfa le stesse ipotesi del *teorema 2.3*, con matrice di adiacenza A Hurwitz tale che tutti i suoi autovalori abbiano la medesima parte reale, e scelti $\mu = |\lambda_{max}(A(p))|$ e $|\mathcal{K}| \geq 2n$, si ha che la quantità

$$f(p) := \frac{|\lambda_{max}(A(p))|^{\frac{n}{|\mathcal{K}|}}}{|\lambda_{max}(A(p))|^2(1 - |\lambda_{max}(A(p))|^2)}$$

è minimizzata da $p^* = |\lambda_{max}(A)|$

Dimostrazione. Abbiamo già dimostrato in *corollario 2.4* che la funzione $f(p)$ è crescente rispetto al modulo dell'autovalore massimo di $A(p)$. Inoltre, dati due autovalori di A λ_1, λ_2 tali da avere la stessa parte reale α e parti immaginarie rispettivamente β_1, β_2 con $|\beta_1| > |\beta_2|$, si ha che:

$$\left| \frac{\lambda_1 + p}{\lambda_1 - p} \right| \geq \left| \frac{\lambda_2 + p}{\lambda_2 - p} \right|$$

Tale risultato è facilmente dimostrabile sostituendo agli autovalori la loro forma complessa e sviluppando. Quindi:

$$|\lambda_{max}(A(p))| = \left| \frac{\lambda_{max}(A) + p}{\lambda_{max}(A) - p} \right|$$

Se indichiamo $\lambda_{max}(A) = \alpha + i\beta$, il valore ottimo di p si può ricavare calcolando la derivata rispetto a p di $|\lambda_{max}(A(p))|$ e ponendola uguale a 0:

$$\frac{d}{dp}(|\lambda_{max}(A(p))|) = \frac{(\alpha + p)((\alpha - p)^2 + \beta^2) + (\alpha - p)((\alpha + p)^2 + \beta^2)}{((\alpha - p)^2 + \beta^2)\sqrt{(\alpha + p)^2 + \beta^2}\sqrt{(\alpha - p)^2 + \beta^2}} = 0$$

$$\frac{d}{dp}(|\lambda_{max}(A(p))|) = 0 \iff p = \sqrt{\alpha^2 + \beta^2} = |\lambda_{max}(A)|$$

\square

Quindi, in generale, è possibile trovare un valore ottimale di p che minimizzi la componente dell'upper bound sull'autovalore minimo del gramiano relativa a $|\lambda_{max}(A(p))|$. Tuttavia, non è detto che il p trovato sia un minimo globale per l'intero limite (anche se l'andamento della componente relativa al valore singolare minimo di $A - pI$ nelle simulazioni suggerisce che funzioni ugualmente).

Nel caso in cui A sia una matrice simmetrica, oltre che Hurwitz, è possibile stabilire un p ottimo per l'intero bound sull'autovalore minimo: questo è reso possibile dal fatto che il valore singolare minimo della matrice $A - pI$ ammette una forma chiusa legata all'autovalore minimo di A , come già visto in *corollario 2.1*:

Teorema 2.4: Data una matrice $A \in \mathbb{R}^{n \times n}$ Hurwitz e simmetrica, $p \in \mathbb{R}_{>0}$ e $T = \frac{n}{|K|} \geq 2$, l'upper bound su $\lambda_{min}(W_c)$:

$$g(p) := \frac{|\lambda_{max}(A(p))|^{2T}}{|\lambda_{max}(A(p))|^2(1 - |\lambda_{max}(A(p))|^2)} \frac{2p}{(|\lambda_{min}(A)| + p)^2}$$

è minimizzato da $p^* = \sqrt{\lambda_{max}(A)\lambda_{min}(A)}$, dove W_c è il gramiano di controllo su orizzonte infinito del sistema continuo (A, B_K)

Dimostrazione. Se scriviamo la funzione di $g(p)$ come prodotto dei singoli termini:

$$f(p) = \frac{|\lambda_{max}(A(p))|^{2T}}{|\lambda_{max}(A(p))|^2(1 - |\lambda_{max}(A(p))|^2)}, \quad h(p) = \frac{2p}{(|\lambda_{min}(A)| + p)^2}$$

si tratta di dimostrare che $g(p)$ è una funzione:

- decrescente in p per $p < \sqrt{(\lambda_{max}(A)\lambda_{min}(A))}$
- crescente in p per $p > \sqrt{(\lambda_{max}(A)\lambda_{min}(A))}$

Caso 1: $p < \sqrt{(\lambda_{max}(A)\lambda_{min}(A))}$

La derivata rispetto a p di $g(p)$ ha la seguente forma:

$$\frac{d}{dp}(g(p)) = \frac{d}{dp}(f(p))h(p) + f(p)\frac{d}{dp}(h(p))$$

dove:

$$\begin{aligned} \frac{d}{dp}(f(p)) &= \frac{d}{d\lambda_{max}(A(p))}(f(p)) \frac{d}{dp}(\lambda_{max}(A(p))) \\ &= \frac{2|\lambda_{max}(A(p))|^{2T+1}(T-1 + (2-T)|\lambda_{max}(A(p))|^2)}{(|\lambda_{max}(A(p))|^2 - |\lambda_{max}(A(p))|^4)^2} \frac{d}{dp}(\lambda_{max}(A(p))) \\ \frac{d}{dp}(h(p)) &= \frac{2(|\lambda_{min}(A)| + p)}{(|\lambda_{min}(A)| - p)^3} \end{aligned}$$

In particolare, data la condizione su p , si può scrivere $|\lambda_{max}(A(p))|$ come funzione di $\lambda_{max}(A)$ nel seguente modo (vedi dimostrazione del *corollario 2.4*):

$$|\lambda_{max}(A(p))| = \left| \frac{\lambda_{max}(A) + p}{\lambda_{max}(A) - p} \right| = \frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p}$$

e quindi:

$$\frac{d}{dp} |\lambda_{max}(A(p))| = \frac{-2|\lambda_{max}(A)|}{(|\lambda_{max}(A)| + p)^2}$$

Sostituendo nella derivata di $g(p)$ otteniamo:

$$\begin{aligned} \frac{d}{dp}(g(p)) &= \frac{2 \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^{2T+1} \left(T - 1 + (2 - T) \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^2 \right)}{\left(\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^2 - \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^4 \right)^2} \times \\ &\frac{-2|\lambda_{max}(A)|}{(|\lambda_{max}(A)| + p)^2} \frac{2p}{(|\lambda_{min}(A)| + p)^2} + \frac{\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^{2T}}{\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^2 \left(1 - \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^2 \right)} \times \\ &\frac{2(|\lambda_{min}(A)| + p)}{(|\lambda_{min}(A)| - p)^3} \end{aligned} \quad (2.1)$$

Quindi dimostrare che l'upper bound è una funzione decrescente in p equivale a verificare la seguente disequazione (derivata attuando le dovute semplificazioni di termini positivi):

$$\frac{4p|\lambda_{max}(A)| \left(T - 1 + (2 - T) \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^2 \right)}{\left(\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right) - \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^3 \right) (|\lambda_{max}(A)| + p)^2} > \frac{|\lambda_{min}(A)| - p}{|\lambda_{min}(A)| + p}$$

Siccome il fattore $(T - 1 + (2 - T)|\lambda_{max}(A(p))|^2)$ è una funzione crescente in T e maggiore o uguale a 1 per $T \geq 2$, se riusciamo a dimostrare che la disuguaglianza vale per $T = 2$, automaticamente sarà valida anche per $T > 2$. Quindi, sostituendo $T = 2$ si ottiene:

$$\frac{4p|\lambda_{max}(A)|}{\left(\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right) - \left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right)^3 \right) (|\lambda_{max}(A)| + p)^2} > \frac{|\lambda_{min}(A)| - p}{|\lambda_{min}(A)| + p}$$

Sviluppando, si ricava:

$$\left(\frac{|\lambda_{max}(A)| - p}{|\lambda_{max}(A)| + p} \right) \left(\frac{|\lambda_{min}(A)| - p}{|\lambda_{min}(A)| + p} \right) < 1$$

- se $|\lambda_{\min}(A)| < p < \sqrt{(\lambda_{\max}(A)\lambda_{\min}(A))}$, la quantità a primo membro è negativa e quindi la disuguaglianza è banalmente verificata
- se $p \leq |\lambda_{\min}(A)|$ la disuguaglianza si riduce a

$$\left(\frac{|\lambda_{\max}(A)| - p}{|\lambda_{\max}(A)| - p}\right) \left(\frac{|\lambda_{\min}(A)| - p}{|\lambda_{\min}(A)| + p}\right) = |\lambda_{\max}(A(p))||\lambda_{\min}(A(p))| < 1$$

ed è verificata in quanto, se A è Hurwitz, $A(p)$ è Schur

Caso 2: $p > \sqrt{(\lambda_{\max}(A)\lambda_{\min}(A))}$

In questo caso il modulo dell'autovalore massimo di $A(p)$ si può scrivere come (si veda dimostrazione del *corollario 2.4*):

$$|\lambda_{\max}(A(p))| = \left| \frac{\lambda_{\min}(A) + p}{\lambda_{\min}(A) - p} \right| = -\frac{|\lambda_{\min}(A)| - p}{|\lambda_{\min}(A)| + p}$$

Ripercorrendo gli stessi passaggi del caso precedente e imponendo che la funzione $g(p)$ sia crescente rispetto a p , si giunge a:

$$\frac{4p|\lambda_{\min}(A)| \left(T - 1 + (2 - T) \left(\frac{|\lambda_{\min}(A)| - p}{|\lambda_{\min}(A)| + p} \right)^2 \right)}{\left(\left(-\frac{|\lambda_{\max}(A)| - p}{|\lambda_{\max}(A)| + p} \right) + \left(\frac{|\lambda_{\max}(A)| - p}{|\lambda_{\max}(A)| + p} \right)^3 \right) (|\lambda_{\min}(A)| + p)^2} > -(|\lambda_{\min}(A)| - p)$$

Scegliendo, analogamente a quanto fatto in precedenza, $T = 2$ e sviluppando, la disuguaglianza si riduce a:

$$\left(\frac{|\lambda_{\min}(A)| - p}{|\lambda_{\min}(A)| + p} \right)^2 < 1 \iff |\lambda_{\min}(A(p))|^2 < 1$$

che è ovviamente verificata perchè $A(p)$ è Schur.

Essendo $g(p)$ una funzione continua strettamente decrescente per $p < \sqrt{(\lambda_{\max}(A)\lambda_{\min}(A))}$ e strettamente crescente per $p > \sqrt{(\lambda_{\max}(A)\lambda_{\min}(A))}$, avrà necessariamente un punto di minimo globale proprio per $p^* = \sqrt{(\lambda_{\max}(A)\lambda_{\min}(A))}$. \square

Riassumendo, nel caso di A simmetrica è possibile trovare un valore ottimo per il parametro p che minimizzi l'intero bound sull'autovalore minimo del gramiano, e tale valore coincide con la media geometrica degli autovalori massimo e minimo di A . Nel caso non simmetrico ma ad autovalori reali, a causa dell'impossibilità di riscrivere in maniera più semplice il valore singolare minimo di $A - pI$, si può solo minimizzare una parte del bound, che tuttavia, dalle simulazioni, sembra prevalere sul limite superiore della norma di $B(p)$.

Capitolo 3

Simmetrizzazione e “quasi simmetrizzazione” di grafi

In questo capitolo affronteremo il tema della simmetrizzazione di matrici, ed esanderemo alcuni risultati già noti riportati in [5]. L'utilità di simmetrizzare una matrice, in particolare la matrice di adiacenza di un grafo, tramite una matrice diagonale (quindi, in pratica, applicando un cambiamento di base) è data dal fatto che, oltre a semplificare lo studio della controllabilità (si vedano i risultati ottenuti nel capitolo precedente sull'upper bound per il gramiano di matrici simmetriche), tale procedura mette in evidenza la relazione che sussiste fra la controllabilità di un sistema e il diametro dello stesso, ovvero la “distanza” massima che separa due nodi qualsiasi del grafo. Da tale scoperta, è possibile determinare se un sistema è difficile da controllare o meno: in particolare un grafo che rispetta alcune condizione topologiche e il cui diametro è una funzione sublineare della cardinalità del grafo stesso è difficile da controllare, ovvero l'autovalore minimo del gramiano di controllabilità di tale sistema decresce all'aumentare della dimensione del sistema.

Nella prima parte del capitolo daremo alcune definizione chiave per la trattazione, come quella di matrice *DSS*, ovvero *diagonalmente simile ad una matrice simmetrica*, e riporteremo alcuni risultati già noti per tali matrici. Successivamente cercheremo di capire come espandere questi risultati a matrici che non rispettano la condizione di simmetrizzabilità, e vedremo che esiste un metodo attraverso il quale è possibile “quasi-simmetrizzare” una matrice, ovvero ottenere una matrice molto simile ad una matrice simmetrica in seguito all'applicazione di una matrice diagonale: in particolare, il metodo applicato si basa sull'associazione ad ogni nodo del grafo di un “potenziale”, in analogia a quanto accade fisicamente nelle reti elettriche, e alla minimizzazione di un'opportuna funzione relativa alla parte antisimmetrica della matrice.

Quindi otterremo un limite superiore sull'autovalore minimo del gramiano di raggiungibilità per matrici quasi simmetrizzabili che rispettano alcune condizioni, e in particolare si potrà osservare come tale bound dipenda direttamente dall'errore di simmetrizzazione, ovvero da quanto si discosta la matrice “quasi-simmetrizzata” dalla controparte simmetrica.

In seguito ci focalizzeremo sul caso di un grafo costituito da un ciclo semplice, cioè un *anello*, e vedremo che, nel caso in cui si scelgano casualmente i pesi del grafo (ovvero costituiscono variabili aleatorie i.i.d.), è possibile ottenere una stima efficace dell'errore di simmetrizzazione. Questo ci consentirà anche di affermare che un grafo di tale tipologia è difficile da controllare. I risultati ottenuti saranno successivamente estesi anche al caso di sistemi costituiti al più da un ciclo semplice, ovvero non necessariamente strutturati come un anello, ma tali da avere al loro interno un solo ciclo e varie “ramificazione” (in sintesi rappresentano gli alberi a cui vengono aggiunti archi in modo tale che si venga a formare un ciclo), e ai cosiddetti *grafi cactus*.

Infine, cercheremo di applicare la procedura di simmetrizzazione al caso del grafo completo, ovvero un grafo in cui ogni nodo è collegato ad ogni singolo altro nodo del grafo, e vedremo che in questa situazione il metodo non risulta efficace: di conseguenza non ci consente di trarre conclusioni circa la controllabilità del grafo stesso.

3.1 Definizioni e risultati introduttivi

Iniziamo la trattazione dando innanzitutto la definizione di matrice *DSS*:

Definizione 3.1 (matrice DSS): Una matrice $A \in \mathbb{R}^{n \times n}$ si dice *diagonalmente simile ad una matrice simmetrica* (DSS) se esiste una matrice diagonale non singolare $D = \text{diag}\{d_1, \dots, d_n\}$ tale che:

$$D^{-1}AD$$

è una matrice simmetrica.

In pratica una matrice risulta *DSS* se, tramite un cambio di base ottenuto con un'opportuna matrice diagonale, può essere ridotta ad una matrice simmetrica. Come già visto, le matrici simmetriche sono in generale più semplici da studiare, quindi la possibilità di simmetrizzare una matrice costituisce un utile vantaggio per lo studio della stessa. Sfortunatamente, non tutte le matrici sono simmetrizzabili da una matrice diagonale: in particolare devono verificare alcune condizioni, una delle quali particolarmente stringente, che riguarda, qualora la matrice rappresenti la matrice d'adiacenza di un grafo, i cicli della rete. Il seguente teorema ci consente di affermare se una matrice è *DSS*:

Teorema 3.1: Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ connesso e non orientato, con matrice di adiacenza A tale che $\text{sgn}(A_{i,j}) = \text{sgn}(A_{j,i}) \forall (i,j) \in \mathcal{E}$, le seguenti due affermazioni risultano equivalenti:

1. La matrice A è DSS
2. Per tutti i cicli semplici $\{i_1, \dots, i_s, i_1\}$ del grafo, dove $i_k \in \mathcal{V}, k = 1, \dots, s$, vale che:

$$\prod_{k=1}^s A_{i_k, i_{k+1}} = \prod_{k=1}^s A_{i_{k+1}, i_k}$$

Quindi una matrice è simmetrizzabile da una matrice diagonale se e solo se il prodotto dei pesi degli archi dei due cammini di ogni ciclo semplice del grafo è lo stesso. Di conseguenza un grafo privo di cicli, come può esserlo un albero o una linea, è sicuramente simmetrizzabile. Per i grafi *DSS* vale il seguente risultato sulla controllabilità ([5]):

Teorema 3.2: Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ non orientato e connesso, con matrice di adiacenza $A \in \mathbb{R}^{n \times n}$ *DSS* simmetrizzabile dalla matrice diagonale D , se indichiamo con $W_{k,T}$ il gramiano del sistema discreto $(A, B_{\mathcal{K}})$ al tempo $T \in \mathbb{N}_{>0}$, vale, $\forall \mu \in [\lambda_{\min}(A), 1)$:

$$\lambda_{\min}(W_{k,T}) \leq \text{cond}^2(D) \frac{\mu^{2 \frac{n_{\mu}}{|\mathcal{K}|}}}{\mu^2(1 - \mu^2)}$$

dove $n_{\mu} = |\{\lambda : \lambda \in \text{spec}(A), |\lambda| \leq \mu\}|$

Questo teorema consente di mettere in relazione la controllabilità del sistema con il diametro del grafo stesso, definito come:

$$\text{diam}(\mathcal{G}) = \max_{u,v \in \mathcal{V}} \min |dist(u,v)|$$

dove $dist(u,v)$ indica il cammino fra i nodi u, v , e quindi $|dist(u,v)|$ è la cardinalità del cammino (inteso come numero degli archi facenti parte del cammino).

Dal *teorema 3.2* di [5] infatti deriva che:

$$\text{cond}^2(D) \leq \left(\frac{A_{\max}}{A_{\min}} \right)^{2 \cdot \text{diam}(\mathcal{G})}$$

dove

$$A_{\max} := \max_{(i,j) \in \mathcal{E}} |A_{i,j}|, \quad A_{\min} := \min_{(i,j) \in \mathcal{E}} |A_{i,j}|$$

Quindi, il condition number di D , ovvero il rapporto fra il valore diagonale massimo e il valore diagonale minimo, è legato al diametro del grafo stesso e in particolare si può allora affermare che, se il diametro cresce sublinearmente rispetto alla dimensione del grafo, il sistema è difficile da controllare perchè il termine $\mu^{2 \frac{n_{\mu}}{|\mathcal{K}|}}$ prevale su $\text{cond}(D)$ e fa tendere l'upper bound sull'autovalore minimo del gramiano a zero.

3.2 Simmetrizzazione e potenziali dei nodi

Ora che abbiamo ricordato le definizioni e i risultati già noti, introduciamo una trasformazione della matrice A del grafo che ci consentirà di ricavare il metodo per provare a rendere matrici che non sono simmetrizzabili “quasi-simmetriche”.

D'ora in poi supporremo che la matrice A abbia elementi non negativi, in modo tale da soddisfare automaticamente la condizione sul segno degli elementi simmetrici. La

trasformazione in questione sulla matrice A fa in modo tale che ogni suo elemento venga mappato in un elemento della nuova matrice B secondo la seguente legge:

$$B_{i,j} = \begin{cases} \log A_{i,j} & \text{se } A_{i,j} > 0 \\ -\infty & \text{se } A_{i,j} = 0 \end{cases}$$

A questo punto abbiamo che la matrice B può essere scritta come somma della sua parte simmetrica e antisimmetrica, definite di seguito:

$$B = B_s + B_a, \quad B_s = \frac{B + B^T}{2}, \quad B_a = \frac{B - B^T}{2}$$

Nota: Nel caso in cui $B_{i,j} = B_{j,i} = -\infty$, si assume che:

$$B_{a_{i,j}} = \frac{B_{i,j} - B_{j,i}}{2} = 0$$

In particolare vale il seguente risultato:

Teorema 3.2: Una A ad elementi non negativi risulta simmetrizzabile dalla matrice diagonale non singolare $D = \text{diag}\{d_1, \dots, d_n\}$ se e solo se, data la matrice B definita in precedenza, gli elementi non nulli della parte antisimmetrica B_a possono essere espressi come:

$$B_{a_{i,j}} = v_i - v_j$$

dove i v_i , con $i = 1, \dots, n$, sono i *potenziali* associati ad ogni nodo i , e in particolare:

$$d_i = e^{v_i}, \quad i = 1, \dots, n$$

Dimostrazione. Sappiamo che, se $B_{a_{i,j}} = (v_i - v_j)$ (dove $B_{a_{i,j}} \neq 0$), allora:

$$B_{i,j} - (v_i - v_j) = (B_s)_{i,j}$$

ovvero $B - B_a$ è una matrice simmetrica. Ora, ricordando la relazione che sussiste tra A e B , abbiamo che:

$$\begin{aligned} e^{(B_s)_{i,j}} &= e^{B_{i,j} - v_i + v_j} = e^{B_{i,j}} \frac{e^{v_j}}{e^{v_i}} = A_{i,j} \frac{e^{v_j}}{e^{v_i}} \\ &= (D^{-1}AD)_{i,j} = A_{i,j} \frac{d_j}{d_i} \end{aligned}$$

Quindi $D^{-1}AD$ è una matrice simmetrica perchè esponenziale, componente per componente, di una matrice simmetrica. \square

In sintesi, se gli elementi della parte antisimmetrica di B possono essere espressi come differenze fra “potenziali” associati ai singoli nodi della rete, in analogia a quanto accade nelle reti elettriche, la matrice A risulta DSS. In effetti, la condizione sui pesi dei cicli ricorda la legge di Kirchoff sulle tensioni di maglia. Il vantaggio risiede nel fatto che, se ci riferiamo alla simmetrizzazione della matrice B , abbiamo a che fare con somme e differenze, invece che prodotti, e questo ci consente di applicare il metodo riportato nella prossima sezione.

3.3 Procedura di “quasi-simmetrizzazione”

A questo punto, dopo aver constatato che è possibile simmetrizzare alcune matrici tramite una matrice diagonale, ci chiediamo se esista comunque un metodo per trovare una matrice diagonale D non singolare tale da rendere A , qualora non sia simmetrizzabile, “quasi simmetrica”, ovvero

$$(D^{-1}AD)_{i,j} \simeq (D^{-1}AD)_{j,i}$$

anche se non è verificata la condizione sui pesi dei cicli. Il problema quindi si riduce alla determinazione del metodo attraverso il quale è possibile ricavare i potenziali che meglio simmetrizzano la matrice B , e quindi la matrice A . A tale scopo, conviene scrivere la matrice B_a come un vettore colonna b_a contenente gli elementi non nulli di B_a , affinché la relazione che lega la parte antisimmetrica di B ai potenziali dei singoli nodi sia la seguente:

$$b_a = Ev, \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

dove $E \in \mathbb{R}^{|\mathcal{E}| \times n}$ è la matrice di incidenza del grafo, ovvero una matrice che ha per elementi l'insieme $\{0, 1, -1\}$ secondo il seguente criterio: sulla riga associata all'arco (i, j) l'elemento di colonna i presenta come valore 1 (perchè l'arco è entrante in i), l'elemento di colonna j ha come valore -1 (perchè l'arco è uscente da j), mentre tutti gli altri elementi della riga sono nulli. La soluzione cercata è quindi:

$$v^* = \underset{v}{\operatorname{argmin}} \frac{1}{2} \|b_a - Ev\|_2^2$$

Tuttavia vale il seguente teorema relativo al rango della matrice di incidenza, riportato in [6]:

Teorema 3.3 (rango della matrice di incidenza): Dato un grafo orientato e connesso \mathcal{G} di dimensione n , il rango della sua matrice di incidenza E vale $n - 1$

Per applicare il metodo dei minimi quadrati, allora, è necessario ricorrere ad un espediente per aumentare di un'unità il rango della matrice di incidenza. Ciò si può attuare vari modi, ma quello più semplice consiste nell'aggiungere un vincolo sui potenziali, per esempio mettendone uno a zero, oppure imporre che la media aritmetica dei potenziali sia nulla: ciò equivale nei rispettivi casi ad aggiungere una riga al sistema $b_a = Ev$ del tipo:

$$[b_{a_{|\mathcal{E}|+1,1}}] = [0] = [1 \ 0 \ \dots \ 0]v$$

(cioè il potenziale del primo nodo è nullo), oppure

$$[b_{a_{|\mathcal{E}|+1,1}}] = [0] = \left[\frac{1}{n} \ \frac{1}{n} \ \dots \ \frac{1}{n} \right]v$$

$$[b_{a_{|\varepsilon|+1,1}}] = [0] = [1 \ 1 \ \dots \ 1]v$$

(cioè somma dei potenziali, pesata o meno, nulla).

A questo punto la matrice ha rango pieno per righe e possiamo applicare la formula che utilizza la pseudoinversa di E per risolvere il problema:

$$v^* = (E^T E)^{-1} E^T b_a$$

Tuttavia, anche se abbiamo trovato una soluzione ottima al problema posto, resta comunque da capire quanto effettivamente la matrice “quasi-simmetrizzata” si discosta da una matrice simmetrica. A tale scopo conviene introdurre la matrice dei potenziali, che, d’ora in avanti, sarà indicata con V tale che:

$$V_{i,j} = \begin{cases} v_i - v_j & \text{se } B_{a_{i,j}} \neq 0 \\ 0 & \text{se } B_{a_{i,j}} = 0 \end{cases}$$

V è legata strettamente all’errore sulla simmetrizzazione della matrice. In particolare si può prendere come indice dell’errore la norma spettrale della matrice differenza ΔB definita come:

$$\Delta B = B_a - V$$

e quindi l’errore vale:

$$err_{\Delta B} = \|\Delta B\|_2 = \sigma_{max}(\Delta B)$$

È facile notare che, se la matrice A è simmetrica, e quindi lo è anche B , l’errore risulta nullo, in quanto tutti gli elementi di B_a sono nulli, e di conseguenza anche i potenziali. Se introduciamo la notazione e^V ad indicare che:

$$(e^V)_{i,j} = e^{V_{i,j}}$$

l’errore sulla matrice A invece si può quantificare nel seguente modo:

$$\Delta A_{i,j} = (A_a - e^V)_{i,j} = \left(\sqrt{\frac{A_{i,j}}{A_{j,i}}} - \frac{e^{v_i}}{e^{v_j}} \right)_{i,j}$$

$$err_{\Delta A} = \|\Delta A\|_2 = \sigma_{max}(\Delta A)$$

Infatti:

$$B = B_s + B_a \iff A = A_s \odot A_a$$

dove \odot indica il prodotto elemento per elemento delle matrici e A_s, A_a sono rispettivamente la parte simmetrica e antisimmetrica di A . Queste sono legate a B_s, B_a dalle relazioni:

$$(B_s)_{i,j} = \frac{B_{i,j} + B_{i,j}^T}{2} \iff (A_s)_{i,j} = \sqrt{A_{i,j} A_{j,i}}$$

$$(B_a)_{i,j} = \frac{B_{i,j} - B_{i,j}^T}{2} \iff (A_a)_{i,j} = \sqrt{\frac{A_{i,j}}{A_{j,i}}}$$

Se la matrice A è già simmetrica, A_a risulta essere una matrice di tutti uni che indichiamo con

$$\mathbb{1}\mathbb{1}^T$$

dove $\mathbb{1}$ è il vettore colonna di soli uni, ovvero:

$$\mathbb{1}_{i,j} = 1 \quad \forall i = 1, \dots, n, j = 1$$

In questo caso i potenziali sono tutti nulli: di conseguenza anche l'errore risulta nullo. In particolare l'errore ΔB si ripercuote sulla matrice A nel seguente modo:

$$(D^{-1}AD)_{i,j} = (A_s)_{i,j}(A_a)_{i,j}e^{v_j-v_i} = (A_s)_{i,j}e_{i,j}^{B_a}e_{i,j}^{-V} = (A_s \odot e^{\Delta B})_{i,j}$$

Infatti, se $\Delta B \simeq 0$, allora $e^{\Delta B} \simeq \mathbb{1}$, e di conseguenza $D^{-1}AD \simeq A_s$

Nota: Minimizzare la quantità $\|b_a - Ev\|_2$ equivale a minimizzare la norma di Frobenius della matrice $B_a - V$ al quadrato:

$$\|B_a - V\|_F^2$$

Infatti, la norma di Frobenius è definita come:

$$\|X\|_F = \left(\sum_{i,j=1}^n |X_{i,j}|^2 \right)^{1/2}$$

ed è immediato verificare l'uguaglianza. Ci si potrebbe chiedere perchè non minimizzare invece la quantità

$$\|B_a - V\|_2^2$$

cioè la norma spettrale della matrice. Purtroppo, è un problema molto più difficile da risolvere perchè non ammette una formula chiusa per il valore ottimale del vettore dei potenziali.

3.4 Upper bound sull'autovalore minimo del gramiano di matrici “quasi-simmetriche”

A questo punto, è interessante analizzare come l'errore di simmetrizzazione si ripercuote sulla stima dell'autovalore minimo del gramiano di raggiungibilità. A tale scopo utilizzeremo come indice di errore di simmetrizzazione sulla matrice A il seguente:

$$\Delta = e^{\Delta B} - \mathbb{1}\mathbb{1}^T$$

Corollario 3.1 (upper bound su $\lambda_{\min}(W_d)$ di matrici quasi simmetrizzabili): Dato il grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ con matrice di adiacenza $A \in \mathbb{R}^{n \times n}$ Schur ad elementi non negativi e “quasi simmetrizzabile” dalla matrice diagonale D definita come:

$$D = \text{diag}\{e^{v_1}, \dots, e^{v_n}\}$$

con $v_i, i = 1, \dots, n$ potenziali di A e \mathcal{K} insieme di nodi di controllo, se vale la condizione:

$$\|\Delta\|_2 < \frac{1}{\|A_s\|_2} - 1$$

allora:

$$\lambda_{\min}(W_d) \leq \text{cond}^2(D) \frac{(|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^{2T}}{1 - (|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^2}$$

con W_d gramiano di controllo su orizzonte infinito del sistema $(A, B_{\mathcal{K}})$ e $T = \lceil \frac{n}{|\mathcal{K}|} \rceil - 1$

Dimostrazione. Se definiamo $A^* := D^{-1}AD$, abbiamo che $A = DA^*D^{-1}$. Dalla definizione di gramiano di raggiungibilità a tempo discreto a orizzonte infinito ricaviamo:

$$\begin{aligned} \lambda_{\min}(W_d) &= \min_{\|x\|=1} x^T \sum_{t=0}^{\infty} A^t B_{\mathcal{K}} B_{\mathcal{K}}^T (A^T)^t x \\ &= \min_{\|x\|=1} x^T \sum_{t=0}^{\infty} (DA^*D^{-1})^t B_{\mathcal{K}} B_{\mathcal{K}}^T ((DA^*D^{-1})^T)^t x \\ &\leq \|D\|_2^2 \left(\min_{\|x\|=1} x^T \sum_{t=0}^{\infty} (A^*)^t D^{-1} B_{\mathcal{K}} B_{\mathcal{K}}^T D^{-1} (A^{*T})^t x \right) \end{aligned}$$

Se definiamo $D^{-1}B_{\mathcal{K}} = B_1$ e indichiamo con W_d^* il gramiano di raggiungibilità del sistema (A^*, B_1) , otteniamo:

$$\lambda_{\min}(W_d) \leq \|D\|_2^2 \left(\min_{\|x\|=1} x^T \sum_{t=0}^{\infty} (A^*)^t B_1 B_1^T (A^{*T})^t x \right) = \|D\|_2^2 \lambda_{\min}(W_d^*)$$

Similmente a quanto già fatto nei teoremi precedenti, aggiungiamo la condizione secondo cui x debba appartenere all'insieme S definito come:

$$S = \{x \in \mathbb{R}^n : x^T R_{k,T} = 0\}$$

dove $R_{k,T}$ è la matrice di raggiungibilità al tempo $T \leq \lceil \frac{n}{|\mathcal{K}|} \rceil - 1$ del sistema (A^*, B_1) (tale insieme non è vuoto in quanto la matrice di raggiungibilità non ha rango pieno). Vale quindi:

$$\begin{aligned}
\lambda_{\min}(W_d) &\leq \|D\|_2^2 \left(\min_{\|x\|=1, x \in S} x^T \sum_{t=0}^{\infty} (A^*)^t B_1 B_1^T (A^{*T})^t x \right) \\
&= \|D\|_2^2 \left(\min_{\|x\|=1} x^T \sum_{t=T}^{\infty} (A^*)^t B_1 B_1^T (A^{*T})^t x \right) \\
&\leq \|D\|_2^2 \left(\max_{\|x\|=1} x^T \sum_{t=T}^{\infty} (A^*)^t B_1 B_1^T (A^{*T})^t x \right) \\
&= \|D\|_2^2 \left\| \sum_{t=T}^{\infty} (A^*)^t B_1 B_1^T (A^{*T})^t \right\|_2 \\
&\leq \|D\|_2^2 \sum_{t=T}^{\infty} \|A^*\|_2^{2t} \|B_1\|_2^2
\end{aligned}$$

Siccome per la norma di B_1 vale:

$$\|B_1\|_2 = \|D^{-1} B_K\|_2 \leq \|D^{-1}\|_2 \|B_K\|_2 = \|D^{-1}\|_2$$

la disuguaglianza si riduce a:

$$\lambda_{\min}(W_d) \leq \text{cond}^2(D) \sum_{t=T}^{\infty} \|A^*\|_2^{2t}$$

A questo punto vale:

$$\|A^*\|_2 = \|A_s \odot e^{\Delta B}\|_2 = \|A_s \odot (\mathbb{1}\mathbb{1}^T + \Delta)\|_2 = \|A_s + A_s \odot \Delta\|_2$$

dove si è sfruttata la proprietà distributiva del prodotto elemento per elemento. Ora, applicando le proprietà delle norme matriciali e la seguente disuguaglianza riportata in [7] (la cui dimostrazione è presente in *appendice A*):

$$\forall F, G \in \mathbb{R}^{n \times n} \quad \|F \odot G\|_2 \leq \|F\|_2 \|G\|_2$$

otteniamo:

$$\|A_s + A_s \odot \Delta\|_2 \leq \|A_s\|_2 + \|A_s \odot \Delta\|_2 \leq \|A_s\|_2 + \|A_s\|_2 \|\Delta\|_2$$

A questo punto, se vale la condizione:

$$\|\Delta\|_2 < \frac{1}{\|A_s\|_2} - 1$$

la nuova serie

$$\sum_{t=T}^{\infty} (\|A_s\|_2 + \|A_s\|_2 \|\Delta\|_2)^{2t}$$

converge a:

$$\frac{(|\lambda_{max}(A_s)|(1 + \|\Delta\|_2))^{2T}}{1 - (|\lambda_{max}(A_s)|(1 + \|\Delta\|_2))^2}$$

dove T può essere scelto pari a $\frac{n}{|\mathcal{K}|} - 1$. □

Si può notare come, essendo il valore della serie geometrica crescente rispetto all'argomento della serie stessa, $\|\Delta\|_2$ deteriora l'upper bound aumentandone il valore. Tuttavia, se $\|\Delta\|_2$ è molto piccolo, tale effetto è praticamente irrilevante, e, per $\|\Delta\|_2 = 0$, ovvero non vi è errore di simmetrizzazione, l'upper bound assume la forma nota per matrici DSS (*Teorema 3.2*, con $\mu = \lambda_{max}(A)$), in quanto $A_s = A^*$.

Non resta che capire come quantificare effettivamente tale errore $\|\Delta\|_2$. A tale scopo è utile notare che la matrice Δ può essere considerata rappresentativa dell'errore di simmetrizzazione ΔB in quanto ricavabile dallo sviluppo in serie di Taylor elemento per elemento dell'esponenziale di matrice per $\Delta B \rightarrow 0$ (cioè quando i singoli valori di ΔB sono piuttosto piccoli):

$$e^{\Delta B} \simeq \mathbb{1}\mathbb{1}^T + \Delta B \implies \Delta B \simeq e^{\Delta B} - \mathbb{1}\mathbb{1}^T = \Delta$$

Quindi, in questa situazione, le matrici Δ e ΔB sono simili, e di conseguenza utilizzare l'una o l'altra per stimare l'errore è arbitrario.

Un caso interessante e che conduce ad una soluzione piuttosto semplice è quello del grafo costituito da un ciclo semplice.

3.5 Il ciclo semplice

3.5.1 Matrice d'errore di simmetrizzazione

Un grafo costituito da un ciclo semplice presenta una matrice A pari a:

$$A = \begin{bmatrix} e^{a_1} & e^{b_1} & 0 & \dots & e^{c_n} \\ e^{c_1} & e^{a_2} & e^{b_2} & \dots & 0 \\ 0 & e^{c_2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ e^{b_n} & 0 & 0 & \dots & e^{a_n} \end{bmatrix} \quad (*)$$

Di conseguenza la matrice B corrispondente è:

$$B = \begin{bmatrix} a_1 & b_1 & -\infty & \dots & c_n \\ c_1 & a_2 & b_2 & \dots & -\infty \\ -\infty & c_2 & \ddots & \ddots & -\infty \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b_n & -\infty & -\infty & \dots & a_n \end{bmatrix}$$

con relativa parte antisimmetrica uguale a:

$$B_a = \frac{1}{2} \begin{bmatrix} 0 & b_1 - c_1 & 0 & \dots & c_n - b_n \\ c_1 - b_1 & 0 & b_2 - c_2 & \dots & 0 \\ 0 & c_2 - b_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ b_n - c_n & 0 & 0 & \dots & 0 \end{bmatrix}$$

(gli elementi che non corrispondono ad archi del ciclo possono essere, in alternativa, posti uguali a $-\infty$, ma ciò è ininfluente per i risultati che otterremo). Sappiamo che la matrice V che meglio simmetrizza B è ricavabile con il metodo dei minimi quadrati dal vettore colonna v ottimo definito come:

$$v = (E^T E)^{-1} E^T b_a$$

dove b_a è il vettore colonna degli elementi non nulli (o non infiniti) di B_a , mentre E è la matrice di incidenza del grafo. Siccome la matrice E ha rango $n - 1$, quindi non pieno, è necessario aggiungere una condizione affinché la soluzione esista, per esempio porre la prima componente di v uguale a 0. Piuttosto di aggiungere una riga sia alla matrice E che alla matrice b_a , conviene riscrivere il problema in una forma equivalente, scegliendo come nuove matrici:

$$E \in \mathbb{R}^{n \times (n-1)} : E = \begin{bmatrix} -1 & 0 & 0 & \dots \\ 1 & -1 & 0 & \dots \\ 0 & 1 & -1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_a \in \mathbb{R}^{n \times 1} : b_a = \frac{1}{2} \begin{bmatrix} b_1 - c_1 \\ b_2 - c_2 \\ \vdots \\ b_n - c_n \end{bmatrix}$$

$$v \in \mathbb{R}^{(n-1) \times 1} : v = \begin{bmatrix} v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}$$

in quanto abbiamo scelto $v_1 = 0$. Si sono inoltre ignorate le componenti corrispondenti ai $c_i - b_i$ in quanto sono informazione ridondante ai fini della risoluzione. Abbiamo quindi il seguente risultato:

Teorema 3.3 (matrice d'errore di simmetrizzazione per cicli semplici): Dato un grafo

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$ costituito da un ciclo semplice con matrice A come in (*), di dimensione $n > 2$ la corrispondente matrice d'errore di simmetrizzazione ΔB assume la forma:

$$\Delta B = \frac{1}{2} \begin{bmatrix} 0 & d & 0 & \dots & -d \\ -d & 0 & d & \dots & 0 \\ 0 & -d & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ d & 0 & 0 & \dots & 0 \end{bmatrix}$$

dove $d = \frac{\sum_{h=1}^n (b_h - c_h)}{n}$.

Di seguito sono riportate due dimostrazioni, entrambe egualmente valide:

Dimostrazione. Se supponiamo che il primo potenziale $v_1 = 0$, possiamo utilizzare la forma ridotta riportata in precedenza per ricavare il vettore ottimo v . In particolare abbiamo che la matrice $E^T E \in \mathbb{R}^{(n-1) \times (n-1)}$ equivale a una matrice tridiagonale Toeplitz simmetrica pari a:

$$E^T E = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 2 \end{bmatrix}$$

Per ricavare l'inversa di una matrice di tale tipologia esiste un metodo piuttosto semplice riportato in [8], il quale afferma che, se prendiamo come matrice tridiagonale Toeplitz simmetrica di dimensione n la seguente:

$$F = \begin{bmatrix} f & g & 0 & \dots & 0 \\ g & f & g & \dots & 0 \\ 0 & g & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & f \end{bmatrix}$$

e definiamo la funzione ricorsiva $U(x)$ come:

$$U_i(x) = 2xU_{i-1}(x) - U_{i-2}(x), \quad U_0 = 1, \quad U_1 = 2x \quad i = 2, \dots, n$$

abbiamo che:

$$F_{i,j}^{-1} = \begin{cases} (-1)^{i+j} \frac{1}{g} \frac{U_{i-1}(f/2g)U_{n-j}(f/2g)}{U_n(f/2g)} & \text{se } i \leq j \\ (-1)^{i+j} \frac{1}{g} \frac{U_{j-1}(f/2g)U_{n-i}(f/2g)}{U_n(f/2g)} & \text{se } i > j \end{cases}$$

Nel nostro caso $f/2g = -1$ e vale per la funzione U :

$$U_k(-1) = (-1)^k(k+1)$$

Sfruttando tale proprietà, sostituendo i valori specifici di f, g e ricordando che la matrice $E^T E$ è quadrata di dimensione $n-1$, otteniamo:

$$\begin{aligned} (E^T E)_{i,j}^{-1} &= \begin{cases} (-1)^{i+j+1} \frac{U_{i-1}(-1)U_{n-1-j}(-1)}{U_{n-1}(-1)} & \text{se } i \leq j \\ (-1)^{i+j+1} \frac{U_{j-1}(-1)U_{n-1-i}(-1)}{U_{n-1}(-1)} & \text{se } i > j \end{cases} \\ &= \begin{cases} (-1)^{i+j+1} \frac{(-1)^{i-1}(i)(-1)^{n-1-j}(n-j)}{(-1)^{n-1}n} & \text{se } i \leq j \\ (-1)^{i+j+1} \frac{(-1)^{j-1}(j)(-1)^{n-1-i}(n-i)}{(-1)^{n-1}n} & \text{se } i > j \end{cases} \\ &= \begin{cases} (-1)^{2i} i \left(1 - \frac{j}{n}\right) & \text{se } i \leq j \\ (-1)^{2i} j \left(1 - \frac{i}{n}\right) & \text{se } i > j \end{cases} \end{aligned}$$

e quindi infine:

$$(E^T E)_{i,j}^{-1} = \begin{cases} i \left(1 - \frac{j}{n}\right) & \text{se } i \leq j \\ j \left(1 - \frac{i}{n}\right) & \text{se } i > j \end{cases}$$

Ora è facile vedere che per la matrice $(E^T E)^{-1} E^T \in \mathbb{R}^{(n-1) \times n}$ vale:

$$(E^T E)^{-1} E_{i,j}^T = \begin{cases} -(E^T E)_{i,1}^{-1} = \frac{i}{n} - 1 & \text{se } j = 1 \\ (E^T E)_{i,j-1}^{-1} - (E^T E)_{i,j}^{-1} & \text{se } 2 \leq j \leq n-1 \\ (E^T E)_{i,n-1}^{-1} = \frac{i}{n} & \text{se } j = n \end{cases}$$

In particolare:

$$(E^T E)_{i,j-1}^{-1} - (E^T E)_{i,j}^{-1} = \begin{cases} i \left(1 - \frac{j-1}{n}\right) - i \left(1 - \frac{j}{n}\right) = \frac{i}{n} & \text{se } i < j \\ (j-1) \left(1 - \frac{i}{n}\right) - j \left(1 - \frac{i}{n}\right) = \frac{i}{n} - 1 & \text{se } i > j \\ (i-1) \left(1 - \frac{i}{n}\right) - i \left(1 - \frac{i}{n}\right) = \frac{i}{n} - 1 & \text{se } i = j \end{cases}$$

Di conseguenza, possiamo scrivere in forma più compatta:

$$(E^T E)^{-1} E_{i,j}^T = \begin{cases} \frac{i}{n} & \text{se } i < j \\ \frac{i}{n} - 1 & \text{se } i \geq j \end{cases}$$

In sintesi, la generica k -esima riga di $(E^T E)^{-1} E^T$ ha le prime k componenti pari a $\frac{k}{n} - 1$, e le restanti $n - k$ uguali a $\frac{k}{n}$.

La formula per ricavare il vettore dei potenziali è:

$$v = (E^T E)^{-1} E^T b_a$$

Siccome il potenziale v_k è alla $(k-1)$ -esima riga, abbiamo quindi che:

$$\begin{aligned}
v_{k+1} &= \left(\frac{k}{n} - 1\right) \sum_{h=1}^k \frac{b_h - c_h}{2} + \frac{k}{n} \sum_{h=k+1}^n \frac{b_h - c_h}{2} \\
&= \left(\frac{k}{n}\right) \sum_{h=1}^n \frac{b_h - c_h}{2} - \sum_{h=1}^k \frac{b_h - c_h}{2} \\
&= \left(\frac{k-1}{n}\right) \sum_{h=1}^n \frac{b_h - c_h}{2} - \sum_{h=1}^{k-1} \frac{b_h - c_h}{2} + \frac{1}{n} \sum_{h=1}^n \frac{b_h - c_h}{2} - \frac{b_k - c_k}{2} \\
&= v_k + \frac{1}{n} \sum_{h=1}^n \frac{b_h - c_h}{2} - \frac{b_k - c_k}{2}
\end{aligned}$$

Se indichiamo con:

$$d_k := b_k - c_k, \quad d := \frac{1}{n} \sum_{h=1}^n b_h - c_h$$

vale per i potenziali la relazione ricorsiva:

$$v_k - v_{k+1} = \frac{d_k}{2} - \frac{d}{2}, \quad k = 2, \dots, n$$

con $v_1 = 0$

Da tale fatto ricaviamo infine:

$$\begin{aligned}
\Delta B &= B_a - V \\
&= \begin{bmatrix} 0 & \frac{b_1 - c_1}{2} & 0 & \dots & \frac{c_n - b_n}{2} \\ \frac{c_1 - b_1}{2} & 0 & \frac{b_2 - c_2}{2} & \dots & 0 \\ 0 & \frac{c_2 - b_2}{2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{b_n - c_n}{2} & 0 & 0 & \dots & 0 \end{bmatrix} - \begin{bmatrix} 0 & v_1 - v_2 & 0 & \dots & v_1 - v_n \\ v_2 - v_1 & 0 & v_2 - v_3 & \dots & 0 \\ 0 & v_3 - v_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v_n - v_1 & 0 & 0 & \dots & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & \frac{d_1}{2} & 0 & \dots & -\frac{d_n}{2} \\ -\frac{d_1}{2} & 0 & \frac{d_2}{2} & \dots & 0 \\ 0 & -\frac{d_2}{2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{d_n}{2} & 0 & 0 & \dots & 0 \end{bmatrix} - \begin{bmatrix} 0 & v_1 - v_2 & 0 & \dots & v_1 - v_n \\ v_2 - v_1 & 0 & v_2 - v_3 & \dots & 0 \\ 0 & v_3 - v_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v_n - v_1 & 0 & 0 & \dots & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & \frac{d}{2} & 0 & \dots & -\frac{d}{2} \\ -\frac{d}{2} & 0 & \frac{d}{2} & \dots & 0 \\ 0 & -\frac{d}{2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{d}{2} & 0 & 0 & \dots & 0 \end{bmatrix}
\end{aligned}$$

□

La seconda dimostrazione, invece di focalizzarsi sulla forma che assumono le matrici nella formula per il vettore potenziale ottimale, ragiona direttamente sulla procedura dei minimi quadrati:

Dimostrazione. Se chiamiamo la funzione da minimizzare:

$$F(v_1, \dots, v_n) = \frac{1}{2} \|b_a - Ev\|_2^2$$

essa può essere riscritta in forma esplicita come:

$$F(v_2, \dots, v_n) = \frac{1}{2} \left(\frac{d_1}{2} + v_2 \right)^2 + \frac{1}{2} \sum_{i=2}^{n-1} \left(\frac{d_i}{2} - (v_i - v_{i+1}) \right)^2 + \frac{1}{2} \left(\frac{d_n}{2} - v_n \right)^2$$

dove $d_i = (b_i - c_i)$, $i = 1, \dots, n$ e $v_1 = 0$.

Ora, per minimizzare tale funzione rispetto ai potenziali, è necessario porre il sistema delle derivate parziali della funzione F rispetto ai v_i a zero, ottenendo:

$$\begin{aligned} \frac{dF}{dv_2}(v_2, \dots, v_n) &= \left(\frac{d_1}{2} + v_2 \right) - \left(\frac{d_2}{2} - v_2 + v_3 \right) = 0 \\ \frac{dF}{dv_i}(v_2, \dots, v_n) &= \left(\frac{d_{i-1}}{2} - v_{i-1} + v_i \right) - \left(\frac{d_i}{2} - v_i + v_{i+1} \right) = 0 \quad i = 3, \dots, n-1 \\ \frac{dF}{dv_n}(v_2, \dots, v_n) &= \left(\frac{d_{n-1}}{2} - v_{n-1} + v_n \right) - \left(\frac{d_n}{2} - v_n \right) = 0 \end{aligned}$$

A questo punto, manipolando opportunamente le equazioni del sistema, si ricava:

$$\begin{aligned} 2v_2 - v_3 &= \frac{1}{2}(d_2 - d_1) \\ 2v_i - v_{i-1} - v_{i+1} &= \frac{1}{2}(d_i - d_{i-1}) \quad i = 3, \dots, n-1 \\ 2v_n - v_{n-1} &= \frac{1}{2}(d_n - d_{n-1}) \end{aligned}$$

E quindi, riscrivendo in maniera compatta:

$$\frac{d_{i-1}}{2} - v_{i-1} + v_i = \frac{d_i}{2} - v_i + v_{i+1} = \frac{d}{2}, \quad i = 1, \dots, n$$

con $v_1 = v_{n+1} = 0$. Quindi gli elementi della matrice d'errore di simmetrizzazione ΔB non nulli sono posizionati come gli elementi non nulli delle matrici B_a e V , e in particolare sono tutti identici. Per capire qual è effettivamente il valore di $\frac{d}{2}$, è utile notare che, dalla precedente equazione:

$$\frac{d_n}{2} - v_n + v_{n+1} = \frac{d}{2} \implies v_n = \frac{d_n}{2} - \frac{d}{2}$$

perchè $v_{n+1} = 0$, e ricorsivamente:

$$v_i = \frac{d_i}{2} + \frac{d_{i+1}}{2} + \dots + \frac{d_n}{2} - (n - i + 1) \frac{d}{2}, \quad i = 1, \dots, n$$

fino ad arrivare a $v_1 = 0$:

$$v_1 = \frac{d_1}{2} + \frac{d_2}{2} + \dots + \frac{d_n}{2} - n \frac{d}{2} = 0$$

da cui infine si ricava:

$$\frac{d}{2} = \frac{1}{2n} \sum_{i=1}^n d_i$$

ovvero:

$$d = \frac{1}{n} \sum_{i=1}^n b_i - c_i$$

□

Quindi, nel caso in cui il grafo sia un ciclo semplice, la matrice d'errore di simmetrizzazione ΔB è una matrice che ha la stessa "forma" di A , è antisimmetrica e tutte le componenti sono identiche (a meno del segno) e pari a $\frac{d}{2}$; tale valore d rappresenta la media, sul numero di nodi totali, della differenza del peso dei due cammini del ciclo dati da b_1, \dots, b_n e c_1, \dots, c_n . Infatti, d può essere riscritto come:

$$d = \frac{1}{n} \left(\sum_{h=1}^n b_h - \sum_{h=1}^n c_h \right)$$

3.5.2 Norma dell'errore di simmetrizzazione

Ora che abbiamo ottenuto l'espressione della matrice ΔB non resta che calcolare la sua norma. A tale scopo è utile la seguente proposizione:

Proposizione 3.1: Se $F \in \mathbb{R}^{n \times n}$ è una matrice antisimmetrica, allora per i suoi valori singolari vale:

$$\sigma_i(F) = |\lambda_i(F)| \quad i = 1, \dots, n$$

dove $\lambda_i(F)$ sono gli autovalori di F

Dimostrazione. Essendo F antisimmetrica, vale che:

$$F^T = -F$$

Dalla definizione di valore singolare ne consegue direttamente la tesi:

$$\sigma_i(F) = \lambda_i^{1/2}(F^T F) = \lambda_i^{1/2}(-FF) = \lambda_i^{1/2}((jF)^2) = |\lambda_i(jF)| = |\lambda_i(F)|$$

poichè $\lambda_i(F^k) = \lambda_i(F)^k$, $\forall k$, e $\lambda_i(jF) = j\lambda_i(F)$ □

Grazie a tale risultato è possibile dimostrare il seguente teorema:

Teorema 3.4 (norma della matrice d'errore di simmetrizzazione per cicli semplici): Data la matrice d'errore di simmetrizzazione per cicli semplici di dimensione $n > 2$:

$$\Delta B = \begin{bmatrix} 0 & \frac{d}{2} & 0 & \cdots & -\frac{d}{2} \\ -\frac{d}{2} & 0 & \frac{d}{2} & \cdots & 0 \\ 0 & -\frac{d}{2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{d}{2} & 0 & 0 & \cdots & 0 \end{bmatrix}$$

con $d = \frac{\sum_{h=1}^n (b_h - c_h)}{n}$, la corrispondente norma vale:

$$\|\Delta B\|_2 = |d| \max_k \left| \sin\left(\frac{2\pi}{n}k\right) \right|, \quad k = 0, \dots, n-1$$

Dimostrazione. Grazie alla *proposizione 3.1*, sappiamo che:

$$\|\Delta B\|_2 = \sigma_{max}(\Delta B) = |\lambda_{max}(\Delta B)|$$

poichè ΔB è una matrice antisimmetrica. ΔB è inoltre una matrice *circolante*, cioè è una matrice nella forma:

$$F = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & \ddots & \ddots & c_{n-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}$$

Tali matrici hanno come autovalori (si veda [9]):

$$\lambda_k(F) = c_0 + c_1 w_k + c_2 w_k^2 + \dots + c_{n-1} w_k^{n-1}, \quad k = 0, \dots, n-1$$

dove

$$w_k = e^{\frac{2\pi j}{n}k}$$

con j unità immaginaria.

Quindi gli autovalori di ΔB assumono la forma:

$$\lambda_k(\Delta B) = \frac{d}{2} e^{\frac{2\pi j}{n}k} - \frac{d}{2} e^{\frac{2\pi(n-1)j}{n}k}, \quad k = 0, \dots, n-1$$

poichè:

$$c_1 = \frac{d}{2}, \quad c_{n-1} = -\frac{d}{2}, \quad c_i = 0 \quad \forall i \neq 1, n-1$$

Sviluppando gli esponenziali si ottiene:

$$\lambda_k(\Delta B) = \frac{d}{2} \left(\cos\left(\frac{2\pi}{n}k\right) + j \sin\left(\frac{2\pi}{n}k\right) \right) - \frac{d}{2} \left(\cos\left(\frac{2\pi(n-1)}{n}k\right) + j \sin\left(\frac{2\pi(n-1)}{n}k\right) \right)$$

Ma è facile vedere che:

$$\begin{aligned} \cos\left(\frac{2\pi}{n}k\right) &= \cos\left(\frac{2\pi(n-1)}{n}k\right) \quad \forall k \in \mathbb{N} \\ \sin\left(\frac{2\pi}{n}k\right) &= -\sin\left(\frac{2\pi(n-1)}{n}k\right) \quad \forall k \in \mathbb{N} \end{aligned}$$

Quindi:

$$\lambda_k(\Delta B) = \frac{d}{2} 2j \left(\sin\left(\frac{2\pi}{n}k\right) \right) = jd \sin\left(\frac{2\pi}{n}k\right) \quad k = 0, \dots, n-1$$

In effetti, gli autovalori delle matrici antisimmetriche compaiono o a coppie coniugate puramente immaginarie oppure sono nulli. Dall'ultimo risultato deriva direttamente la tesi:

$$|\lambda_{max}(\Delta B)| = \max_k \left| jd \left(\sin\left(\frac{2\pi}{n}k\right) \right) \right| = |d| \max_k \left| \sin\left(\frac{2\pi}{n}k\right) \right| \quad k = 0, \dots, n-1$$

□

Quindi abbiamo che la norma della matrice errore di simmetrizzazione è direttamente proporzionale al valore d e ad una funzione che d'ora in avanti indicheremo con $s_{max}(n)$ data da:

$$s_{max}(n) = \max_k \left| \sin\left(\frac{2\pi}{n}k\right) \right| \quad k = 0, \dots, n-1$$

In particolare, vale che:

$$0 < s_{max}(n) \leq 1 \quad \forall n \in \mathbb{N}$$

e

$$\lim_{n \rightarrow \infty} s_{max}(n) = 1$$

in quanto gli autovalori di ΔB si equidistribuiscono lungo la circonferenza complessa (cioè si posizionano in modo tale da avere la stessa distanza angolare pari a $\frac{2\pi}{n}$): via via che la dimensione cresce, aumentano anche gli autovalori, fino a coprire l'intera circonferenza.

3.5.3 Errore di simmetrizzazione con pesi i.i.d. gaussiani

Ora che abbiamo trovato un'espressione per la norma dell'errore di simmetrizzazione, risulta interessante capire il suo comportamento qualora i pesi degli archi del grafo in un ciclo semplice siano scelti in modo casuale.

Ci chiediamo quindi: se le variabili $b_i, c_i, i = 1, \dots, n$ sono i.i.d. (ovvero indipendenti ed identicamente distribuite), che forma avrà la variabile $|d|$? In questo contesto probabilistico, la funzione $s_{max}(n)$ non costituisce un variabile aleatoria ma una costante determinata unicamente dalla dimensione del sistema.

Supponiamo quindi che le variabili b_i, c_i abbiano una distribuzione gaussiana di media μ e varianza θ^2 , e siano indipendenti fra loro, ovvero:

$$b_i, c_i \sim N(\mu, \theta^2), \quad i = 1, \dots, n$$

Ricordiamo un risultato che ci sarà utile per capire la distribuzione di $|d|$, riportato in [10]:

Proposizione 3.2: Se la variabile f è distribuita come una gaussiana di media nulla e varianza θ_f^2 :

$$f \sim N(0, \theta_f^2)$$

allora la variabile $|f|$ è distribuita come una *folded normal variable* di media $\sqrt{\frac{2}{\pi}}\theta_f$ e varianza $\theta_f^2(1 - \frac{2}{\pi})$:

$$|f| \sim NF\left(\sqrt{\frac{2}{\pi}}\theta_f, \theta_f^2\left(1 - \frac{2}{\pi}\right)\right)$$

In particolare la densità di probabilità di $|f|$ è la seguente:

$$p(|f(x)|) = \begin{cases} \frac{2}{\sqrt{2\pi\theta_f^2}} e^{-\frac{x^2}{2\theta_f^2}} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Qualora la variabile f della proposizione non avesse media nulla, l'espressione della densità di probabilità, della media e della varianza di $|f|$ risulterebbero leggermente più complesse, ma come vedremo nel seguente teorema, tale situazione non interessa la nostra trattazione:

Teorema 3.5 (Distribuzione di $|d|$): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ costituito da un ciclo semplice con matrice A come in (*), di dimensione $n > 2$, se i pesi b_i, c_i sono indipendenti e identicamente distribuiti come gaussiane di media μ e varianza θ^2 :

$$b_i, c_i \sim N(\mu, \theta^2) \quad i = 1, \dots, n$$

allora la variabile $|d|$ definita come:

$$|d| = \left| \frac{\sum_{i=1}^n (b_i - c_i)}{n} \right|$$

è distribuita come una *folded normal variable* di media e varianza pari a:

$$|d| \sim NF\left(\frac{2}{\sqrt{\pi n}}\theta, \frac{2\theta^2}{n}\left(1 - \frac{2}{\pi}\right)\right)$$

Dimostrazione. Abbiamo che:

$$b_i, c_i \sim N(\mu, \theta^2) \quad i = 1, \dots, n$$

Di conseguenza:

$$d_i = b_i - c_i \sim N(0, 2\theta^2) \quad i = 1, \dots, n$$

poichè differenza di variabili gaussiane indipendenti. Sfruttando ancora la proprietà di combinazione lineare di gaussiane otteniamo:

$$\sum_{i=1}^n d_i \sim N(0, 2n\theta^2) \quad i = 1, \dots, n$$

e quindi:

$$d = \frac{1}{n} \sum_{i=1}^n d_i \sim N\left(0, \frac{2\theta^2}{n}\right) \quad i = 1, \dots, n$$

Infine, dalla *proposizione 3.2* deriva la tesi:

$$|d| \sim NF\left(\sqrt{\frac{2}{\pi}} \frac{\sqrt{2}\theta}{\sqrt{n}}, \frac{2\theta^2}{n}\left(1 - \frac{2}{\pi}\right)\right) = NF\left(\frac{2\theta}{\sqrt{\pi n}}, \frac{2\theta^2}{n}\left(1 - \frac{2}{\pi}\right)\right)$$

□

Conseguenza fondamentale di questo fatto è che, all'aumentare della dimensione del sistema, il valore atteso della norma dell'errore di simmetrizzazione tende a 0. Infatti:

$$E(\|\Delta B\|_2) = E(|d|s_{max}(n)) = s_{max}(n)E(|d|) = s_{max}(n)\frac{\sqrt{2}\theta}{\sqrt{n\pi}}$$

E quindi:

$$\lim_{n \rightarrow \infty} E(\|\Delta B\|_2) = \lim_{n \rightarrow \infty} s_{max}(n)\frac{\sqrt{2}\theta}{\sqrt{n\pi}} = \lim_{n \rightarrow \infty} \frac{\sqrt{2}\theta}{\sqrt{n\pi}} = 0$$

3.5.4 Controllabilità del ciclo semplice

Il fatto che il valore atteso dell'errore di simmetrizzazione tenda a zero è importante in quanto non solo prova che la procedura di simmetrizzazione tramite una matrice diagonale funziona bene per un grafo con molti nodi costituito da un ciclo semplice, i cui pesi sono variabili gaussiane i.i.d. (cioè la simmetrizzazione è quasi perfetta), ma afferma anche che tale tipologia di rete (qualora la parte simmetrica abbia norma minore di 1) è *difficile da controllare*, ovvero l'autovalore minimo del gramiano di raggiungibilità del sistema tende a zero all'aumentare della dimensione del sistema.

Infatti, dal *corollario 3.1*, abbiamo che:

$$\lambda_{\min}(W_{d,T}) \leq \text{cond}^2(D) \frac{(|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^{2T}}{1 - (|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^2}$$

Se passiamo al logaritmo di entrambi i membri otteniamo:

$$\log(\lambda_{\min}(W_d)) \leq 2 \log(\text{cond}(D)) + 2T \log(|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2)) - \log(1 - (|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^2)$$

Siccome abbiamo supposto che $|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2) < 1$, conviene riscrivere:

$$\log(\lambda_{\min}(W_d)) \leq 2 \log(\text{cond}(D)) - 2T \log((|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^{-1}) - \log(1 - (|\lambda_{\max}(A_s)|(1 + \|\Delta\|_2))^2)$$

Ora, visto che $E(\|\Delta B\|_2) \rightarrow 0$ e $\text{Var}(\|\Delta B\|_2) \rightarrow 0$ per $n \rightarrow \infty$, necessariamente anche $\|\Delta B\|_2 \rightarrow 0$, e quindi $\|\Delta\|_2 \rightarrow 0$, in quanto strettamente legati dalla relazione precedentemente riportata. Quindi, passando al valore atteso, per $n \rightarrow \infty$:

$$E(\log(\lambda_{\min}(W_d))) \leq 2E(\log(\text{cond}(D))) - 2TE(\log((|\lambda_{\max}(A_s)|))) - E(\log(1 - (|\lambda_{\max}(A_s)|)^2))$$

Per dimostrare che il sistema è difficile da controllare bisogna verificare che $E(\log(\lambda_{\min}(W_d))) \rightarrow -\infty$. Notiamo che la disuguaglianza può essere riscritta come:

$$E(\log(\lambda_{\min}(W_d))) \leq 2E(\log(\text{cond}(D))) - 2Tf_1 - f_2$$

dove f_1, f_2 sono costanti indipendenti dalla dimensione del sistema. Per definizione, $T = \frac{n}{|\mathcal{K}|} - 1$, e quindi il termine $-2Tf_1$ decresce linearmente rispetto alla dimensione del sistema. Per quanto riguarda $\log(\text{cond}(D))$, può essere espresso come:

$$\log(\text{cond}(D)) = \log(e^{v_{\max} - v_{\min}}) = v_{\max} - v_{\min}$$

dove, per il ciclo semplice, si ha che:

$$v_{max} = \max_{k=2, \dots, n} \left\{ 0, k \frac{d}{2} - \frac{d_k}{2} \right\}$$

$$v_{min} = \min_{k=2, \dots, n} \left\{ 0, k \frac{d}{2} - \frac{d_k}{2} \right\}$$

Ora, si possono distinguere due casi:

- se $\frac{d}{2} > 0$ allora, per n sufficientemente grande, v_{max} è:

$$v_{max} = n \frac{|d|}{2} - \frac{d_n}{2}$$

e quindi:

$$\begin{aligned} E(\log(\text{cond}(D))) &= E(v_{max}) - E(v_{min}) \simeq E(v_{max}) \\ &= E\left(n \frac{|d|}{2}\right) - E\left(\frac{d_n}{2}\right) = \frac{n}{2} \frac{2\theta}{\sqrt{\pi n}} - 0 \\ &= \frac{\theta\sqrt{n}}{\sqrt{\pi}} = O(\sqrt{n}) \end{aligned}$$

- se $\frac{d}{2} < 0$, allora, per n sufficientemente grande, v_{min} è:

$$v_{min} = -n \frac{|d|}{2} - \frac{d_n}{2}$$

e quindi:

$$\begin{aligned} E(\log(\text{cond}(D))) &= E(v_{max}) - E(v_{min}) \simeq -E(v_{min}) \\ &= -E\left(-n \frac{|d|}{2}\right) + E\left(\frac{d_n}{2}\right) = \frac{n}{2} \frac{2\theta}{\sqrt{\pi n}} + 0 \\ &= \frac{\theta\sqrt{n}}{\sqrt{\pi}} = O(\sqrt{n}) \end{aligned}$$

In entrambe le situazione, il $\log(\text{cond}(D))$ cresce come \sqrt{n} . Ma di conseguenza il termine dominante in $E(\log(\lambda_{min}(W_d)))$ resta $-2Tf_1$, il quale, per $n \rightarrow \infty$, tende a $-\infty$.

A conferma dei risultati appena ottenuti, sono riportate alcune simulazioni numeriche, dove i pesi del ciclo sono stati scelti i.i.d. con distribuzione gaussiana normalizzata, ovvero $b_i, c_i \sim N(0, 1)$:

Alcune osservazioni:

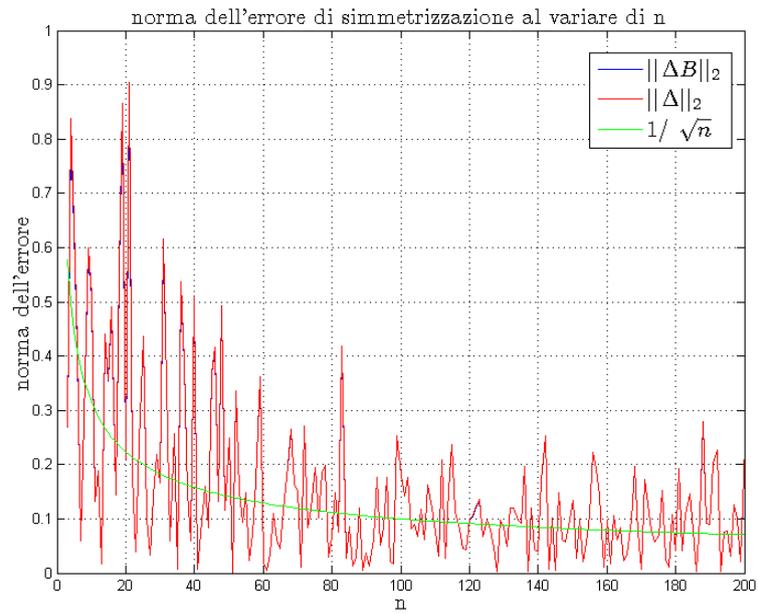


Figura 3.1: variazione dell'errore di simmetrizzazione rispetto alla dimensione del ciclo

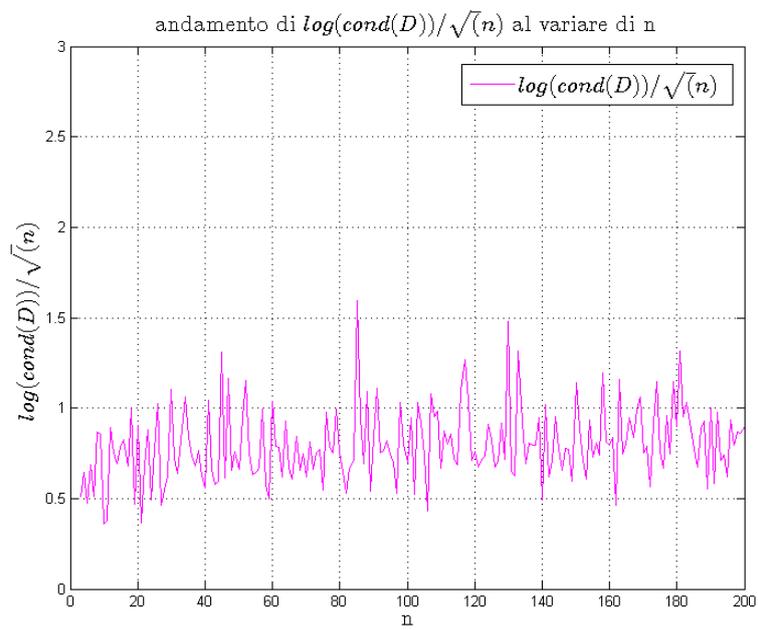


Figura 3.2: variazione del condition number di D rispetto alla dimensione del ciclo

- Si può notare come $\|\Delta B\|_2$ (riportata in *figura 3.1* in blu) e $\|\Delta\|_2$ (riportata invece in rosso) sono praticamente identiche in quanto, nel caso in questione, l'errore di simmetrizzazione è piuttosto piccolo, e di conseguenza l'approssimazione in serie di Taylor funziona molto bene: inoltre, all'aumentare della dimensione, la norma dell'errore di simmetrizzazione tende a zero, e quindi $\|\Delta B\|_2 - \|\Delta\|_2 \rightarrow 0$.
- L'errore di simmetrizzazione tende a zero all'aumentare del sistema all'incirca come $\frac{1}{\sqrt{n}}$ (riportata in verde sempre in *figura 3.1*), in quanto, come precedentemente provato, il valore atteso di ΔB è $\frac{\sqrt{2}\theta}{\sqrt{n\pi}}$ (nel particolare caso in esame $\theta^2 = 1$).
- Per quanto riguarda invece $\log(\text{cond}(D))$, riportato in *figura 3.2*, si può notare come il suo andamento sia molto simile a \sqrt{n} , in quanto $\frac{\log(\text{cond}(D))}{\sqrt{n}}$ è all'incirca costante, fatto precedentemente ricavato dal punto di vista teorico.

3.5.5 Espansione a grafi con un solo ciclo

E' possibile estendere i risultati relativi al ciclo semplice a un qualsiasi grafo che presenti al suo interno un solo ciclo e varie "ramificazioni". Infatti vale la seguente generalizzazione del *teorema* sulla matrice d'errore di simmetrizzazione per cicli semplici:

Teorema 3.6 (matrice d'errore di simmetrizzazione per grafi con un solo ciclo): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ di dimensione n avente all'interno un solo ciclo di dimensione h , se supponiamo che i nodi siano ordinati in modo tale che la matrice A possa essere espressa come:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

dove $A_{11} \in \mathbb{R}^{h \times h}$ è la sottomatrice relativa al ciclo, $A_{12} \in \mathbb{R}^{h \times (n-h)}$, $A_{21} \in \mathbb{R}^{(n-h) \times h}$, $A_{22} \in \mathbb{R}^{(n-h) \times (n-h)}$, allora la corrispondente matrice d'errore di simmetrizzazione ΔB assume la forma:

$$\Delta B = \begin{bmatrix} \Delta B_1 & 0 \\ 0 & 0 \end{bmatrix}$$

dove

$$\Delta B_1 = \frac{1}{2} \begin{bmatrix} 0 & d & 0 & \dots & -d \\ -d & 0 & d & \dots & 0 \\ 0 & -d & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ d & 0 & 0 & \dots & 0 \end{bmatrix}, \quad \Delta B_1 \in \mathbb{R}^{h \times h}$$

$$\text{e } d = \frac{\sum_{i=1}^h (b_i - c_i)}{h}$$

Dimostrazione. Per la dimostrazione, definiamo un nodo del grafo come *esterno* se è connesso solamente ad un altro nodo del grafo e non fa parte del ciclo. Se indichiamo la

funzione che il vettore dei potenziali deve minimizzare come:

$$F(v_1, \dots, v_n) = \frac{1}{2} \|b_a - Ev\|_2^2$$

e supponiamo che $v_1 = 0$, allora per qualsiasi nodo esterno v_e vale:

$$\frac{dF}{dv_e}(v_2, \dots, v_n) = \frac{d}{dv_e} \frac{1}{2} \left(\frac{d_{e,f}}{2} - (v_e - v_f) \right)^2 = - \left(\frac{d_{e,f}}{2} - (v_e - v_f) \right)$$

dove $d_{e,f}$ è la differenza fra il peso dell'unico arco entrante nel nodo e e il peso dell'unico arco uscente dal nodo e , mentre f è l'unico nodo confinante di e . Se annulliamo la derivata otteniamo quindi:

$$\frac{dF}{dv_e}(v_2, \dots, v_n) = 0 \iff v_e - v_f = \frac{d_{e,f}}{2}$$

e di conseguenza la componente $\Delta B_{e,f} = 0$. Inoltre, la derivata della funzione $F(v_1, \dots, v_n)$ rispetto a v_f è:

$$\frac{dF}{dv_f}(v_2, \dots, v_n) = F_1(v_2, \dots, v_n) + \left(\frac{d_{e,f}}{2} - v_e + v_f \right) = F_1(v_2, \dots, v_n)$$

dove la componente $F_1(v_2, \dots, v_n)$ non dipende da v_e . Se il nodo f confina con più nodi esterni, il ragionamento non cambia: una volta annullata la derivata della funzione F rispetto ai nodi esterni che confinano con f , la derivata della funzione F rispetto a v_f perde la dipendenza da quei nodi.

Ora, se il nodo f non fa parte del ciclo, si può considerare come nuovo nodo esterno e applicare il metodo in maniera ricorsiva, finchè non si raggiunge un nodo del ciclo.

Al termine della procedura, i potenziali da minimizzare rimasti sono quelli relativi al ciclo ed è possibile ricavare la tesi direttamente dai risultati già noti per il ciclo semplice. \square

Quindi, per gli archi che non fanno parte del ciclo, la procedura di simmetrizzazione risulta perfetta. I risultati successivi sulla norma della matrice ΔB qualora si scelgano casualmente i pesi degli archi restano egualmente validi, con l'unica differenza che non dipendono più dalla dimensione totale del sistema ma esclusivamente dalla dimensione del ciclo. In particolare un grafo con un solo ciclo stabile è ancora difficile da controllare, perché stavolta $\text{cond}(D)$ cresce come \sqrt{h} , con $h \leq n$, per $h \rightarrow \infty$.

3.6 Grafi cactus

Una classe più generale di grafi a cui possono essere applicati gli stessi ragionamenti riguardanti il ciclo semplice sono i cosiddetti *grafi cactus*. La definizione è la seguente:

Definizione 3.2 (grafo cactus): Un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ si dice *grafo cactus* (o *albero cactus*) se è connesso e, presi due qualsiasi cicli del grafo, essi hanno al più un vertice in comune.

In alternativa, un grafo cactus è un grafo connesso in cui un qualsiasi arco appartiene al più ad un solo ciclo semplice. Un esempio di grafo cactus è riportato di seguito:

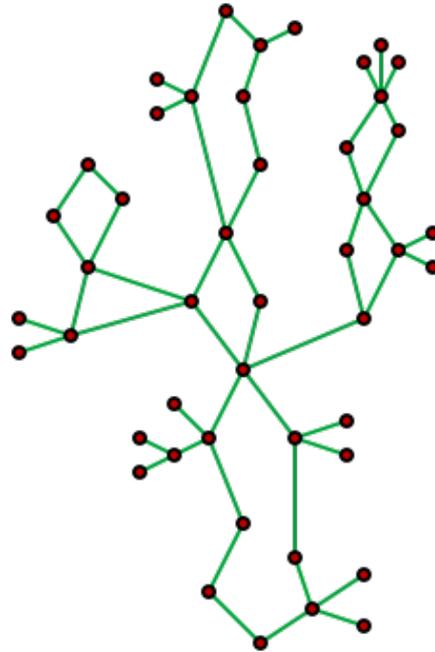


Figura 3.3: esempio di grafo cactus

I grafi cactus hanno numerose applicazioni pratiche, che spaziano dall'elettronica, alla genomica comparativa ([11]), fino alle comunicazioni radio ([12]). Per dimostrare che è possibile giungere alle stesse conclusioni del caso di grafi con un solo ciclo circa la controllabilità, distinguiamo innanzitutto due casi:

- Un grafo cactus costituito da due cicli che non hanno in comune nodi ma sono congiunti da un arco (o più in generale da una sottografo di tipo linea, ovvero con sottomatrice tridiagonale)
- Un grafo cactus costituito da due cicli che hanno in comune un nodo

Un grafo cactus è infatti costituito da tali unità fondamentali a cui si sommano eventuali ramificazioni (la cui simmetrizzazione è perfetta, come già dimostrato), e quindi i risultati che otterremo potranno essere naturalmente applicati anche a configurazioni più generiche.

Caso 1: Nel primo caso i cicli sono separati da uno o più archi (per semplicità supporremo che vi sia solo un arco di separazione). In particolare i nodi del primo ciclo appartengono all'insieme:

$$C_1 = \{\nu_i \in \mathcal{V} : i = 1, \dots, m\}$$

e sono ordinati in sequenza, mentre i nodi del secondo ciclo appartengono all'insieme:

$$C_2 = \{\nu_i \in \mathcal{V} : i = m + 1, \dots, n\}$$

e sono anch'essi numerati sequenzialmente, come riportato in *figura 3.4*:

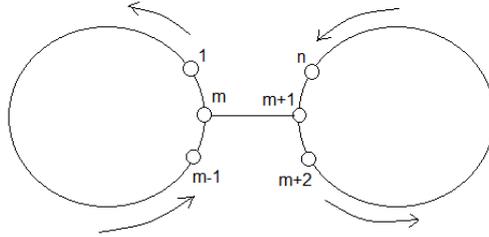


Figura 3.4: caso 1

Se richiamiamo la funzione di minimizzazione (supponendo $v_1 = 0$):

$$F(v_2, \dots, v_n) = \frac{1}{2} \left((B_{a_{1,2}} + v_2)^2 + (B_{a_{m,1}} - v_m)^2 + \sum_{i=2}^{m-1} (B_{a_{i,i+1}} - v_i + v_{i+1})^2 + \right. \\ \left. (B_{a_{m,m+1}} - v_m + v_{m+1})^2 + \sum_{i=m+1}^{n-1} (B_{a_{i,i+1}} - v_i + v_{i+1})^2 + \right. \\ \left. (B_{a_{n,m+1}} - v_n + v_{m+1})^2 \right)$$

e poniamo le derivate rispetto ai potenziali a zero, otteniamo:

$$\frac{dF}{dv_i}(v_2, \dots, v_n) = -(B_{a_{i,i+1}} - v_i + v_{i+1}) + (B_{a_{i-1,i}} - v_{i-1} + v_i) = 0, \quad i = m + 2, \dots, n - 1 \\ \frac{dF}{dv_n}(v_2, \dots, v_n) = -(B_{a_{n,m+1}} - v_n + v_{m+1}) + (B_{a_{n-1,n}} - v_{n-1} + v_n) = 0$$

e quindi vale l'uguaglianza:

$$B_{a_{n,m+1}} - v_n + v_{m+1} = B_{a_{n-1,n}} - v_{n-1} + v_n = \dots = B_{a_{m+1,m+2}} - v_{m+1} + v_{m+2}$$

che implica:

$$\frac{dF}{dv_{m+1}}(v_2, \dots, v_n) = B_{a_{m,m+1}} - v_m + v_{m+1} = 0$$

Quindi i nodi m -esimo e $(m+1)$ -esimo si occupano della simmetrizzazione dell'arco di congiunzione fra i cicli (la quale risulta perfetta), mentre le equazioni risolutive legate al primo e al secondo ciclo formano sottosistemi disgiunti fra loro che non hanno variabili in comune. A tali sottosistemi possono essere applicati gli stessi ragionamenti relativi al ciclo semplice.

Si può generalizzare la soluzione ottenuta a una catena di p cicli legati fra loro da uno o più archi: in tal caso la matrice d'errore di simmetrizzazione ΔB assume la forma diagonale a blocchi, dopo essere stata "ripulita" dalla righe e colonne nulle:

$$\Delta B = \begin{bmatrix} \Delta B_1 & 0 & \dots & 0 \\ 0 & \Delta B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Delta B_p \end{bmatrix}$$

Ora, sappiamo che:

$$\Delta B_i = \begin{bmatrix} 0 & \frac{d_i}{2} & 0 & \dots & -\frac{d_i}{2} \\ -\frac{d_i}{2} & 0 & \frac{d_i}{2} & \dots & 0 \\ 0 & -\frac{d_i}{2} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{d_i}{2} & 0 & 0 & \dots & 0 \end{bmatrix} \quad i = 1, \dots, p$$

con d_i definito come:

$$d_i = \frac{2}{n_i} \sum_{(j,k) \in \mathcal{E}_i} B_{a_{j,k}}, \quad i = 1, \dots, p$$

dove n_i è la cardinalità (ovvero il numero di nodi) del ciclo i -esimo, mentre \mathcal{E}_i indica l'insieme degli archi appartenenti al ciclo i -esimo.

Da ciò è possibile ricavare la norma dell'errore di simmetrizzazione:

$$\|\Delta B\|_2 = \max_i \left(|d_i| \max_k \left| \sin \left(\frac{2\pi k}{n_i} \right) \right| \right), \quad i = 1, \dots, p, \quad k = 0, \dots, n_i - 1$$

che può essere massimizzata da:

$$\|\Delta B\|_2 \leq \max_i |d_i|, \quad i = 1, \dots, p$$

Se supponiamo che gli archi del ciclo siano scelti casualmente secondo gaussiane i.i.d., e

che i nodi dei cicli presenti nel grafo crescano all'aumentare della dimensione del sistema, chiaramente la norma dell'errore di simmetrizzazione tende ad annullarsi, fatto già dimostrato in precedenza (si ricorda in particolare che sia il valore atteso che la varianza delle variabili d_i tendono a zero al crescere di n).

Caso 2: Nel secondo caso i cicli hanno un nodo in comune, che indicheremo con v_m ; i primi $m - 1$ nodi appartengono al primo ciclo, mentre i rimanenti, successivi a m , al secondo. Inoltre i nodi sono ordinati numericamente come riportato in *figura 3.5*:

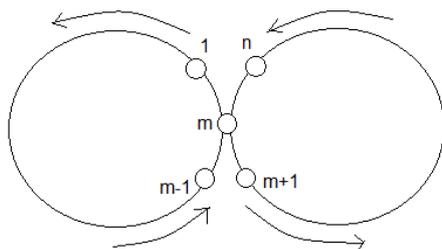


Figura 3.5: caso 2

La funzione di minimizzazione è data da:

$$F(v_2, \dots, v_n) = \frac{1}{2} \left((B_{a_{1,2}} + v_2)^2 + (B_{a_{m,1}} - v_m)^2 + \sum_{i=2}^{m-1} (B_{a_{i,i+1}} - v_i + v_{i+1})^2 + \sum_{i=m}^{n-1} (B_{a_{i,i+1}} - v_i + v_{i+1})^2 + (B_{a_{n,m}} - v_n + v_m)^2 \right)$$

dove ancora una volta si è supposto $v_1 = 0$. Ripercorrendo gli stessi passaggi del caso precedente si ottiene:

$$\frac{dF}{dv_i}(v_2, \dots, v_n) = -(B_{a_{i,i+1}} - v_i + v_{i+1}) + (B_{a_{i-1,i}} - v_{i-1} + v_i) = 0, \quad i = m + 1, \dots, n - 1$$

$$\frac{dF}{dv_n}(v_2, \dots, v_n) = -(B_{a_{n,m}} - v_n + v_m) + (B_{a_{n-1,n}} - v_{n-1} + v_n) = 0$$

$$\begin{aligned} \frac{dF}{dv_m}(v_2, \dots, v_n) &= -(B_{a_{m,1}} - v_m) + (B_{a_{m-1,m}} - v_{m-1} + v_m) - \\ &- (B_{a_{m,m+1}} - v_m + v_{m+1}) + (B_{a_{n,m}} - v_n + v_m) = 0 \end{aligned}$$

da cui è possibile ricavare ancora una volta la catena di uguaglianze:

$$B_{a_{n,m}} - v_n + v_m = B_{a_{n-1,n}} - v_{n-1} + v_n = \dots = B_{a_{m,m+1}} - v_m + v_{m+1}$$

dove i d_i sono definiti come nel caso precedente. Se supponiamo che i pesi del grafo siano i.i.d. gaussiani e che la cardinalità dei cicli presenti cresca all'aumentare della dimensione del sistema, abbiamo ancora una volta che l'errore di simmetrizzazione tende ad annullarsi, poichè si annullano i d_i .

Quindi, anche i grafi cactus con pesi scelti casualmente e i.i.d. gaussiani, e che rispettano le condizioni imposte dal *corollario 3.1* sull'autovalore del gramiano nel caso di matrici "quasi-simmetriche", sono difficili da controllare, in quanto anche stavolta il condition number della matrice diagonale D cresce sublinearmente con la dimensione del grafo (più precisamente con la cardinalità del ciclo che "pesa" maggiormente nella simmetrizzazione).

3.7 Il grafo completo

Finora abbiamo esaminato nel dettaglio il caso di un grafo costituito al più da un ciclo semplice e siamo giunti alla conclusione che l'applicazione di un'apposita matrice diagonale riesce a "quasi simmetrizzare" la matrice della rete, e anzi la gravità dell'errore di simmetrizzazione diminuisce al crescere della dimensione. In particolare i risultati ottenuti valgono anche per i grafi cactus. Ora invece ci concentriamo sul caso "estremo" di un grafo completo, ovvero con matrice A tale che:

$$A_{i,j} \neq 0 \quad \forall i, j = 1, \dots, n$$

In particolare la matrice V risulta anch'essa "piena" (cioè nessun elemento è nullo, escludendo ovviamente la diagonale) se e solo se:

$$B_{a,i,j} \neq 0 \quad \forall i, j = 1, \dots, n, i \neq j$$

A questo punto cerchiamo di verificare se la procedura di simmetrizzazione tramite matrice diagonale risulta ancora valida. A tale scopo, vale il seguente risultato:

Teorema 3.7 (potenziali ottimali per grafi completi): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ di dimensione n completo, cioè con matrice A tale che:

$$A_{i,j} \neq 0 \quad \forall i, j = 1, \dots, n$$

matrice B data da:

$$B_{i,j} = \log(A_{i,j}) \quad \forall i, j = 1, \dots, n$$

e B_a componente antisimmetrica di B , allora il vettore dei potenziali che minimizza la quantità:

$$v^* = \operatorname{argmin}_{v \in K} \frac{1}{2} \|b_a - Ev\|_2^2, \quad K = \left\{ v \in \mathbb{R}^n : \sum_{i=1}^n v_i = 0 \right\}$$

dove b_a è il vettore colonna degli elementi non nulli di B_a e E matrice di incidenza del grafo, è dato da:

$$v^* = \frac{1}{n} B_a \mathbb{1}$$

dove $\mathbb{1}$ è il vettore colonna di dimensione n tale che:

$$\mathbb{1}_{i,j} = 1 \quad i = 1, \dots, n, j = 1$$

Dimostrazione. Per dimostrare la tesi ricordiamo innanzitutto che:

$$v^* = \operatorname{argmin}_{v \in K} \frac{1}{2} \|b_a - Ev\|_2^2 = \operatorname{argmin}_{v \in K} \frac{1}{2} \|B_a - V\|_F^2$$

dove $\|\cdot\|_F$ indica la norma di Frobenius della matrice. Se supponiamo che anche la matrice B_a sia piena (eccettuata la diagonale), possiamo riscrivere la norma come:

$$v^* = \operatorname{argmin}_{v \in K} \frac{1}{2} \|B_a - V\|_F^2 = \operatorname{argmin}_{v \in K} \frac{1}{2} \|B_a - (v\mathbb{1}^T - \mathbb{1}v^T)\|_F^2$$

dove $\mathbb{1}$ è il vettore colonna di soli uni. Abbiamo quindi che:

$$\frac{1}{2} \|B_a - V\|_F^2 = \frac{1}{2} \left(\sum_{i,j=1}^n (B_{a_{i,j}} - (v_i - v_j))^2 \right) = F(v_1, \dots, v_n)$$

La derivata della funzione F rispetto ad un qualsiasi potenziale v_h è:

$$\frac{dF}{dv_h}(v_1, \dots, v_n) = - \sum_{j \neq h} (B_{a_{h,j}} - (v_h - v_j)) + \sum_{i \neq h} (B_{a_{i,h}} - (v_i - v_h))$$

dove il primo membro rappresenta gli elementi della riga h -esima della matrice ΔB , mentre il secondo membro rappresenta gli elementi della colonna h -esima di ΔB . Ora, ricordando che $B_{a_{i,j}} = -B_{a_{j,i}}$, si trova:

$$\begin{aligned} & - \sum_{j \neq h} (B_{a_{h,j}} - (v_h - v_j)) + \sum_{i \neq h} (B_{a_{i,h}} - (v_i - v_h)) \\ &= \sum_{j \neq h} (B_{a_{j,h}} - (v_j - v_h)) + \sum_{i \neq h} (B_{a_{i,h}} - (v_i - v_h)) \\ &= 2 \sum_{i \neq h} (B_{a_{i,h}} - (v_i - v_h)) \end{aligned}$$

Eguagliando a zero la derivata otteniamo:

$$\begin{aligned}
2 \sum_{i \neq h} (B_{a_i, h} - (v_i - v_h)) = 0 &\iff \sum_i (B_{a_i, h} - (v_i - v_h)) = 0 \\
&\iff nv_h = - \sum_i B_{a_i, h} + \sum_i v_i \\
&\iff nv_h = \sum_i B_{a_h, i} \\
&\iff v_h = \frac{1}{n} \sum_i B_{a_h, i}
\end{aligned}$$

dove si è sfruttato il fatto che $v \in K = \left\{ v \in \mathbb{R}^n : \sum_{i=1}^n v_i = 0 \right\}$. Concludendo:

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \frac{1}{n} \begin{bmatrix} \sum_i B_{a_1, i} \\ \vdots \\ \sum_i B_{a_n, i} \end{bmatrix} = \frac{1}{n} B_a \mathbf{1}$$

□

Quindi, nel caso di un grafo completo, il valore ottimo del potenziale di un nodo è paragonabile alla media della differenza fra i pesi degli archi che escono dal nodo e i pesi di quelli che entrano nel nodo stesso. Infatti:

$$v_h = \frac{1}{n} \sum_i B_{a_h, i} = \frac{1}{2n} \sum_i (B_{h, i} - B_{i, h})$$

Ora, in analogia a quanto già ricavato dal caso del ciclo semplice, cerchiamo di capire il comportamento dell'errore di simmetrizzazione nel caso in cui i pesi del grafo siano scelti casualmente. In particolare supporremo che siano variabili i.i.d. distribuite secondo gaussiane di media μ e varianza θ^2 . In questa situazione vale il seguente risultato:

Teorema 3.8 (distribuzione della matrice d'errore di simmetrizzazione): Dato un grafo $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ di dimensione n completo con matrice A , se i pesi della matrice B definita come

$$B_{i, j} = \log(A_{i, j})$$

sono i.i.d. e distribuiti secondo gaussiane di media μ e varianza θ^2 :

$$B_{i, j} \sim N(\mu, \theta^2), \quad i, j = 1, \dots, n$$

allora gli elementi non diagonali della matrice errore di simmetrizzazione ΔB sono distribuiti secondo gaussiane a media nulla e varianza $\frac{\theta^2(n^2 + 2n - 4)}{2n^2}$:

$$\Delta B_{i, j} \sim N\left(0, \frac{\theta^2(n^2 + 2n - 4)}{2n^2}\right), \quad i, j = 1, \dots, n$$

Dimostrazione. Per definizione, gli elementi della matrice d'errore di simmetrizzazione possono essere espressi come:

$$\Delta B_{i,j} = B_{a_{i,j}} - (v_i - v_j), \quad i, j = 1, \dots, n$$

Se $B_{i,j} \sim N(\mu, \theta^2)$, allora abbiamo che:

$$\begin{aligned} E(B_{a_{i,j}}) &= \frac{1}{2}E(B_{i,j} - B_{j,i}) = \frac{1}{2}(E(B_{i,j}) - E(B_{j,i})) = \frac{1}{2}(\mu - \mu) = 0 \\ \text{Var}(B_{a_{i,j}}) &= \frac{1}{4}\text{Var}(B_{i,j} - B_{j,i}) = \frac{1}{4}(\text{Var}(B_{i,j}) + \text{Var}(B_{j,i})) = \frac{1}{4}(\theta^2 + \theta^2) = \frac{\theta^2}{2} \end{aligned}$$

dove nel calcolo della varianza si è sfruttato il fatto che gli elementi della matrice B sono i.i.d.. Quindi gli elementi della parte antisimmetrica di B sono gaussiane a media nulla e varianza pari a $\theta^2/2$:

$$B_{a_{i,j}} \sim N\left(0, \frac{\theta^2}{2}\right), \quad i, j = 1, \dots, n, \quad i \neq j$$

Per quanto riguarda i potenziali invece:

$$\begin{aligned} E(v_k) &= \frac{1}{n}E\left(\sum_{h=1}^n B_{a_{i,j}}\right) = \frac{1}{n}\sum_{h=1}^n E(B_{a_{i,j}}) = 0 \\ \text{Var}(v_k) &= \frac{1}{n^2}\text{Var}\left(\sum_{h=1}^n B_{a_{k,h}}\right) = \frac{1}{n^2}\text{Var}\left(\sum_{\substack{h=1 \\ h \neq k}}^n B_{a_{k,h}}\right) = \frac{n-1}{n^2} \frac{\theta^2}{2} = \frac{\theta^2(n-1)}{2n^2} \end{aligned}$$

Quindi:

$$v_k \sim N\left(0, \frac{\theta^2(n-1)}{2n^2}\right), \quad k = 1, \dots, n$$

Di conseguenza per il valore atteso degli elementi dell'errore di simmetrizzazione vale:

$$E(\Delta B_{i,j}) = E(B_{a_{i,j}} - (v_i - v_j)) = 0$$

mentre per la varianza:

$$\begin{aligned}
\text{Var}(\Delta B_{i,j}) &= \text{Var}\left(\frac{B_{i,j} - B_{j,i}}{2} - \left(\frac{1}{n} \sum_{\substack{h=1 \\ h \neq i}}^n B_{a_{i,h}} - \frac{1}{n} \sum_{\substack{h=1 \\ h \neq j}}^n B_{a_{j,h}}\right)\right) \\
&= \text{Var}\left(\frac{B_{i,j} - B_{j,i}}{2} - \left(\frac{1}{n} \sum_{\substack{h=1 \\ h \neq i,j}}^n B_{a_{i,h}} - \frac{1}{n} \sum_{\substack{h=1 \\ h \neq j,i}}^n B_{a_{j,h}}\right)\right) \\
&= \text{Var}\left(\frac{B_{i,j} - B_{j,i}}{2}\right) + \text{Var}\left(\frac{1}{n} \sum_{\substack{h=1 \\ h \neq i,j}}^n B_{a_{i,h}}\right) + \text{Var}\left(\frac{1}{n} \sum_{\substack{h=1 \\ h \neq j,i}}^n B_{a_{j,h}}\right) \\
&= \frac{1}{4}2\theta^2 + \frac{1}{n^2} \sum_{\substack{h=1 \\ h \neq i,j}}^n \text{Var}(B_{a_{i,h}}) + \frac{1}{n^2} \sum_{\substack{h=1 \\ h \neq j,i}}^n \text{Var}(B_{a_{j,h}}) \\
&= \frac{1}{4}2\theta^2 + \frac{n-2}{n^2} \frac{\theta^2}{2} + \frac{n-2}{n^2} \frac{\theta^2}{2} = \frac{\theta^2(n^2 + 2n - 4)}{2n^2}
\end{aligned}$$

dove si è sfruttata ancora una volta l'indipendenza delle variabili in gioco. Concludendo:

$$\Delta B_{i,j} \sim N\left(0, \frac{\theta^2(n^2 + 2n - 4)}{2n^2}\right), \quad i, j = 1, \dots, n$$

□

In particolare, dal teorema appena enunciato, si deduce che, al crescere della dimensione del sistema, gli elementi della matrice d'errore di simmetrizzazione tendono ad assumere la stessa distribuzione degli elementi della parte antisimmetrica di B . Infatti per $n \rightarrow \infty$:

$$\Delta B_{i,j} \sim N\left(0, \frac{\theta^2(n^2 + 2n - 4)}{2n^2}\right) \rightarrow N\left(0, \frac{\theta^2}{2}\right)$$

Quindi, l'azione della simmetrizzazione diventa via via sempre più ininfluenza, e di conseguenza meno efficace. A questo punto, che cosa si può dire sulla norma dell'errore di simmetrizzazione? Per rispondere alla domanda è utile il seguente risultato riportato in [13] (sotto il nome di *Corollario 2.3.6*):

Teorema 3.9: Data una matrice M quadrata di dimensione n , con $M_{i,j}$ indipendenti fra loro tali che:

$$M_{i,j} \sim N(0, 1), \quad \forall j \geq i$$

e $M_{i,j} = M_{j,i}$, allora esistono delle costanti f_1, f_2 tali che:

$$P(\|M\|_2 > f_3\sqrt{n}) \leq f_1 e^{-f_2 f_3 n}$$

per qualsiasi $f_3 \geq f_1$. In particolare $\|M\|_2 = O(\sqrt{n})$ quasi certamente.

Il teorema appena riportato ci dice quindi che l'andamento della norma spettrale di una matrice simmetrica con elementi distribuiti secondo gaussiane normalizzate indipendenti fra loro è proporzionale alla radice della dimensione della matrice stessa. Tale risultato può essere espanso al caso in cui la matrice sia antisimmetrica, come accade per ΔB . Vale infatti:

Corollario 3.2 (norma dell'errore di simmetrizzazione ΔB per grafi completi): Nel caso di grafo completo di dimensione n , la norma dell'errore di simmetrizzazione ΔB cresce all'aumentare della dimensione del grafo secondo \sqrt{n} per n sufficientemente grande, cioè:

$$\|\Delta B\|_2 = O(\sqrt{n})$$

Dimostrazione. Innanzitutto è fondamentale notare che gli elementi della matrice d'errore di simmetrizzazione ΔB non sono indipendenti fra loro perchè non lo sono gli elementi di V . Tuttavia, per n sufficientemente grande, l'influenza di tale matrice diventa irrilevante e quindi possiamo approssimare ΔB con B_a (i cui elementi invece sono i.i.d.). Ora è possibile normalizzare l'approssimazione della matrice ΔB (ovvero rendere i suoi elementi distribuiti secondo gaussiane normalizzate i.i.d.), premoltiplicandola per la radice della varianza dei suoi elementi, ovvero:

$$\Delta B_{i,j} \sim N\left(0, \frac{\theta^2}{2}\right) \implies \frac{\theta}{\sqrt{2}}\Delta B_{i,j} \sim N(0, 1), \quad n \longrightarrow \infty$$

Inoltre, per fare in modo che la matrice non abbia la diagonale nulla, si può definire:

$$\frac{\theta}{\sqrt{2}}\Delta B = \Delta B_p - D_p$$

dove ΔB_p è una matrice antisimmetrica "piena" con elementi distribuiti secondo gaussiane normalizzate i.i.d., mentre D_p è una matrice diagonale con elementi anch'essi distribuiti secondo gaussiane normalizzate i.i.d.. A questo punto vale che:

$$\frac{\theta}{\sqrt{2}}\|\Delta B\|_2 = \|\Delta B_p - D_p\|_2 \leq \|\Delta B_p\|_2 + \|D_p\|_2$$

e

$$\frac{\theta}{\sqrt{2}}\|\Delta B\|_2 = \|\Delta B_p - D_p\|_2 \geq |\|\Delta B_p\|_2 - \|D_p\|_2|$$

Alla norma di ΔB_p è possibile applicare il *teorema 3.9* (anche se la matrice è antisimmetrica invece che simmetrica, dal punto di vista della distribuzione delle variabili non cambia nulla, poichè se una variabile è distribuita secondo una gaussiana normalizzata, lo è anche l'opposto della variabile) e quindi:

$$\|\Delta B_p\|_2 = O(\sqrt{n})$$

mentre la norma di D_p corrisponde semplicemente al massimo in modulo dei valori diagonali della matrice stessa, che risulta una costante indipendente dalla dimensione del sistema. Di conseguenza, la norma di ΔB (ovvero la norma di B_a per n sufficientemente grande) è compresa fra due quantità che crescono come \sqrt{n} . Quindi:

$$\frac{\theta}{\sqrt{2}} \|\Delta B\|_2 = O(\sqrt{n}) \implies \|\Delta B\|_2 = O(\sqrt{n})$$

□

La norma della matrice d'errore di simmetrizzazione cresce quindi, a meno di una costante moltiplicativa, come la radice della dimensione del sistema: questo fatto rende la stima della matrice Δ tramite la matrice ΔB inutile, in quanto l'approssimazione risultava valida se e solo se la norma di ΔB era sufficientemente piccola. Non resta quindi che stimare la norma della matrice Δ . Come sappiamo gli elementi di Δ sono legati a quelli di ΔB dalla relazione:

$$\Delta_{i,j} = e^{\Delta B_{i,j}} - 1, \quad \forall i, j = 1, \dots, n$$

Quindi, per capire la distribuzione di tali elementi è necessario il seguente risultato:

Teorema 3.10 (distribuzione log-normale): Data una variabile aleatoria f con distribuzione gaussiana tale che:

$$f \sim N(\mu, \theta_f^2)$$

allora la variabile aleatoria data dall'esponenziale di f ha una distribuzione log-normale di media $e^{\mu + \frac{\theta_f^2}{2}}$ e varianza $(e^{\theta_f^2} - 1)e^{2\mu + \theta_f^2}$:

$$e^f \sim \ln N\left(e^{\mu + \frac{\theta_f^2}{2}}, (e^{\theta_f^2} - 1)e^{2\mu + \theta_f^2}\right)$$

In particolare e^f ha densità di probabilità pari a:

$$p(e^{f(x)}) = \frac{1}{x\theta_f\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\theta_f^2}}$$

Dimostrazione. Se la variabile f ha distribuzione gaussiana tale che:

$$f \sim N(\mu, \theta_f^2)$$

e indichiamo con g l'esponenziale di f , ovvero $g = e^f$, abbiamo che:

$$P(g < k) = P(e^f < k) = P(f < \ln k) = \int_{-\infty}^{\ln k} \frac{1}{\sqrt{2\pi}\theta_f} e^{-\frac{(x-\mu)^2}{2\theta_f^2}} dx$$

Se applichiamo il cambio di variabile $y = e^x$, ovvero $x = \ln y$, si ha che $dx = \frac{1}{y}dy$ e:

$$P(g < k) = \int_{-\infty}^k \frac{1}{y\sqrt{2\pi\theta_f}} e^{-\frac{(\ln y - \mu)^2}{2\theta_f^2}} dy$$

Quindi la densità di probabilità di g è proprio:

$$p(g(y)) = \frac{1}{y\sqrt{2\pi\theta_f}} e^{-\frac{(\ln y - \mu)^2}{2\theta_f^2}}$$

Ora, per la media di g vale:

$$E(g) = E(e^f) = \int_{-\infty}^{\infty} e^x \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{(x-\mu)^2}{2\theta_f^2}} dy$$

Applicando la trasformazione $y = x - \mu$, otteniamo:

$$\begin{aligned} E(g) &= \int_{-\infty}^{\infty} e^{\mu+y} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{y^2}{2\theta_f^2}} dy = e^{\mu} \int_{-\infty}^{\infty} e^y \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{y^2}{2\theta_f^2}} dy \\ &= e^{\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{\frac{2y\theta_f^2 - y^2}{2\theta_f^2}} dy = e^{\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{\frac{-(y-\theta_f^2)^2 + \theta_f^4}{2\theta_f^2}} dy \\ &= e^{\mu + \frac{1}{2}\theta_f^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{(y-\theta_f^2)^2}{2\theta_f^2}} dy = e^{\mu + \frac{1}{2}\theta_f^2} \end{aligned}$$

perchè l'integrale di una densità di probabilità è 1.

Per quanto riguarda la varianza di g invece:

$$\begin{aligned} E(g^2) &= E(e^{2f}) = \int_{-\infty}^{\infty} e^{2x} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{(x-\mu)^2}{2\theta_f^2}} dy = \int_{-\infty}^{\infty} e^{2(y+\mu)} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{y^2}{2\theta_f^2}} dy \\ &= e^{2\mu} \int_{-\infty}^{\infty} e^{2y} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{y^2}{2\theta_f^2}} dy = e^{2\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{\frac{4y\theta_f^2 - y^2}{2\theta_f^2}} dy \\ &= e^{2\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{(y-2\theta_f^2)^2 + 4\theta_f^4}{2\theta_f^2}} dy \\ &= e^{2\mu + 2\theta_f^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\theta_f}} e^{-\frac{(y-2\theta_f^2)^2}{2\theta_f^2}} dy = e^{2\mu + 2\theta_f^2} \end{aligned}$$

e quindi:

$$Var(g) = E(g^2) - E(g)^2 = e^{2\mu + 2\theta_f^2} - e^{2\mu + \theta_f^2} = e^{2\mu + \theta_f^2} (e^{\theta_f^2} - 1)$$

□

Da tale teorema si ricava immediatamente la distribuzione di $e^{\Delta B_{i,j}}$, ovvero:

$$e^{\Delta B_{i,j}} \sim \ln N\left(e^{\frac{\theta^2}{2} \frac{n^2+2n-4}{2n^2}}, (e^{\frac{\theta^2}{2} \frac{n^2+2n-4}{2n^2}} - 1)e^{\frac{\theta^2}{2} \frac{n^2+2n-4}{2n^2}}\right), \quad i \neq j$$

mentre gli elementi diagonali sono banalmente pari a 1. In particolare, per $n \rightarrow \infty$, abbiamo che:

$$e^{\Delta B_{i,j}} \sim \ln N\left(e^{\frac{\theta^2}{4}}, (e^{\frac{\theta^2}{2}} - 1)e^{\frac{\theta^2}{2}}\right), \quad i \neq j$$

Quindi se n è sufficientemente grande, il valore atteso degli elementi dell'esponenziale di ΔB è una costante indipendente dalla dimensione del sistema. Passando al valore atteso degli elementi di Δ si ricava:

$$E(\Delta_{i,j}) = E(e^{\Delta B_{i,j}} - 1) = E(e^{\Delta B_{i,j}}) - 1 = e^{\frac{\theta^2}{4}} - 1, \quad i \neq j$$

E di conseguenza:

$$E(\Delta) = (e^{\frac{\theta^2}{4}} - 1)(\mathbb{1}\mathbb{1}^T - I), \quad n \rightarrow \infty$$

dove $\mathbb{1}\mathbb{1}^T$ è una matrice quadrata di dimensione n con tutti gli elementi pari a 1, mentre I è la matrice identità. La norma spettrale della matrice $E(\Delta)$ è quindi approssimabile con:

$$\|E(\Delta)\|_2 = (e^{\frac{\theta^2}{4}} - 1)\|\mathbb{1}\mathbb{1}^T - I\|_2, \quad n \rightarrow \infty$$

Dalla *proposizione 2.2* del capitolo precedente è possibile affermare che:

$$\lambda_i(\mathbb{1}\mathbb{1}^T - I) = \lambda_i(\mathbb{1}\mathbb{1}^T) - 1, \quad i = 1, \dots, n$$

ed è facilmente verificabile che la matrice $\mathbb{1}$ ha un solo autovalore pari a n e tutti gli altri nulli, poichè ha rango $n - 1$. Si ricava infine:

$$\lambda_{max}(\mathbb{1}\mathbb{1}^T - I) = \lambda_{max}(\mathbb{1}\mathbb{1}^T) - 1 = n - 1 \implies \|E(\Delta)\|_2 = (e^{\frac{\theta^2}{4}} - 1)(n - 1)$$

Quindi la matrice Δ presenta addirittura una crescita direttamente proporzionale alla dimensione del sistema, a ulteriore conferma del fatto che la procedura di simmetrizzazione nel caso del grafo completo è inefficace.

A conclusione dello studio del grafo completo, sono riportati i grafici dell'andamento della norma dell'errore di simmetrizzazione e del condition number della matrice diagonale D di esponenziali dei potenziali, nel caso in cui gli elementi della matrice B siano variabili aleatorie i.i.d. gaussiane normalizzate, ovvero:

$$B_{i,j} \sim N(0, 1) \quad i, j=1, \dots, n$$

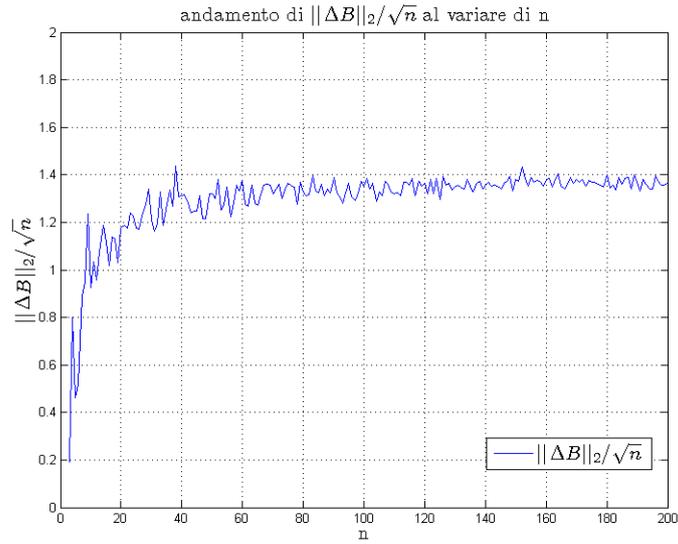


Figura 3.6: variazione dell'errore di simmetrizzazione $\|\Delta B\|_2$ rispetto alla dimensione del ciclo

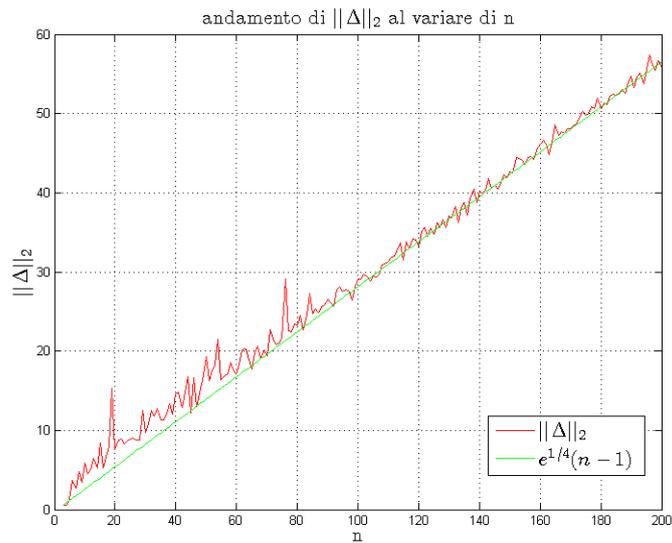


Figura 3.7: variazione dell'errore di simmetrizzazione $\|\Delta\|_2$ rispetto alla dimensione del ciclo

Alcune osservazioni:

- Nei primi due grafici sulla norma degli errori di simmetrizzazione ΔB (in blu) e Δ (in rosso) sono confermati i risultati ottenuti dalla teoria: in particolare l'andamento

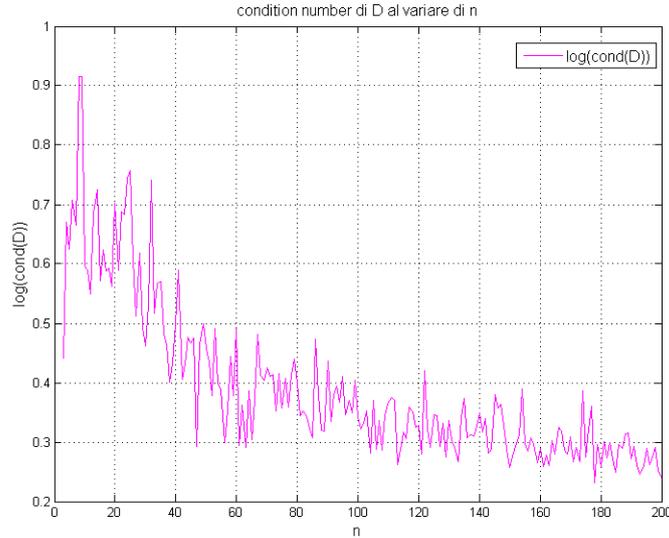


Figura 3.8: variazione del condition number di D rispetto alla dimensione del ciclo

mento della norma di ΔB è dell'ordine di \sqrt{n} a meno di una costante moltiplicativa, mentre l'andamento della norma di Δ è chiaramente lineare rispetto alla dimensione del sistema. Inoltre, è graficata la retta $y = (e^{\frac{1}{4}} - 1)(n - 1)$ (in verde), che, come si evince, approssima in modo ottimale la crescita della norma dell'errore Δ : in effetti, nel caso di variabili gaussiane normalizzate, $\theta^2 = 1$. Non è tuttavia vero che la norma assume all'infinito tale valore, in quanto la varianza degli elementi di Δ tende ad una costante indipendente dalla dimensione del grafo: più precisamente, nel caso in questione, tale valore è $(e^{\frac{1}{2}} - 1)e^{\frac{1}{2}}$.

- Come è facile immaginare, il $\log(\text{cond}(D))$ che è definito come:

$$\log(\text{cond}(D)) = v_{max} - v_{min} = \max_k \frac{1}{n} \sum_{j=1}^n B_{a_{k,j}} - \min_k \frac{1}{n} \sum_{j=1}^n B_{a_{k,j}}$$

tende ad annullarsi al crescere della dimensione del sistema, in quanto i valori dei potenziali stessi tendono ad annullarsi. In particolare quindi $\text{cond}(D) \rightarrow 1$ per $n \rightarrow \infty$

Riassumendo, la procedura di quasi-simmetrizzazione perde la sua efficacia nel caso di grafi che hanno contemporaneamente un grado di “completezza”, inteso come numero di vicini per ogni nodo, molto elevato e cicli innestati, dove la situazione più estrema è rappresentata dal grafo completo.

Capitolo 4

Conclusioni

Scopo della tesi era lo studio della controllabilità di sistemi con particolare attenzione all'energia spesa per il controllo: a tal proposito si è studiato il gramiano di controllabilità a tempo discreto e continuo e l'andamento del suo autovalore minimo rispetto alla dimensione del sistema, ottenendo i seguenti risultati:

- E' possibile espandere l'upper bound sull'autovalore minimo del gramiano a tempo discreto nel caso continuo qualora sia su orizzonte infinito e il sistema risulti stabile, utilizzando la relazione nota come *ADI iteration* e la teoria di Lyapunov.
- Il bound ottenuto a tempo continuo presenta un ulteriore grado di libertà rappresentato da una variabile positiva p , e la minimizzazione del limite superiore rispetto a tale variabile risulta in generale complesso.
- Nel caso di grafi con matrice di adiacenza simmetrica, è possibile trovare un valore "ottimale" per la variabile p , cioè in modo tale da rendere l'upper bound più stringente, e ciò è reso possibile dal fatto che sussiste una relazione diretta fra autovalori e valori singolari relativi alla perturbazione della matrice di adiacenza stessa con la matrice identità.

Motivati dalla maggiore semplicità analitica dei sistemi simmetrici (ovvero dei grafi non orientati), si è successivamente studiata una procedura di quasi-simmetrizzazione di matrici tramite matrici diagonali, traendo le seguenti conclusioni:

- E' possibile "quasi-simmetrizzare" una matrice che non rispetti le condizioni di simmetrizzabilità, tramite un cambio di base dato da una matrice diagonale, grazie ad una procedura volta a minimizzare la norma della differenza fra la parte antisimmetrica della matrice di partenza e una matrice di potenziali associati ai singoli nodi della rete.
- L'upper bound trovato in precedenza sull'autovalore minimo del gramiano a tempo discreto è ancora valido per matrici "quasi-simmetrizzabile" ed è possibile modificarlo opportunamente in modo che dipenda dall'errore di simmetrizzazione: in particolare è richiesto che la parte simmetrica e l'errore di simmetrizzazione soddisfino alcune condizione relative alla loro norma.

- Nel caso del ciclo semplice con pesi i.i.d. gaussiani è possibile affermare che la procedura di simmetrizzazione porta a ottimi risultati al crescere della dimensione del sistema, in quanto l'errore di simmetrizzazione tende ad annullarsi; in particolare, grazie a questo fatto, è possibile affermare che un sistema costituito da un ciclo semplice è difficile da controllare.
- I risultati ottenuti per il ciclo semplice possono essere naturalmente espansi ai grafi cactus, che rappresentano una classe più ampia di sistemi dove, qualora siano presenti cicli, essi hanno in comune al più un nodo.
- La procedura esposta non è efficace nel caso del grafo completo, in quanto i potenziali ottimali tendono ad annullarsi al crescere della cardinalità del sistema, annullando l'utilità del metodo stesso.

Possibili ricerche future relative agli argomenti trattati potrebbero riguardare la dimostrazione che il valore ottimale di p trovato per matrici simmetriche risulta valido anche per la classe di matrici ad autovalori reali, la scoperta di un valore ottimale di p per matrici generali che non abbiano vincoli sugli autovalori, l'espansione della procedura di quasi-simmetrizzazione ad altre tipologie di grafi e il suo grado di efficacia in relazione al numero di cicli presenti nel sistema.

Appendice A

Limite superiore sulla norma del prodotto di Hadamard

Teorema (upper bound sulla norma del prodotto di Hadamard): Date due matrici $G, F \in \mathbb{R}^{n \times n}$, vale che:

$$\|F \odot G\|_2 \leq \|F\|_2 \|G\|_2$$

dove \odot indica il prodotto di Hadamard (o elemento per elemento) di matrici.

Dimostrazione. La dimostrazione del teorema si divide in due parti. Nella prima parte dimostriamo che vale la seguente disuguaglianza

$$\|F \odot G\|_2 \leq r_{max}(F) c_{max}(G)$$

dove $r_{max}(F)$ e $c_{max}(G)$ indicano rispettivamente la riga di norma euclidea massima di F e la colonna di norma euclidea massima di G , ovvero:

$$r_{max}(F) = \max_{i=1, \dots, n} \left(\sum_{j=1}^n |F_{i,j}|^2 \right)^{1/2}$$

$$c_{max}(G) = \max_{j=1, \dots, n} \left(\sum_{i=1}^n |G_{i,j}|^2 \right)^{1/2}$$

E' noto che, dati due vettori $x, y \in \mathbb{R}^n$, la norma spettrale del prodotto di Hadamard di F e G può essere espressa come:

$$\|F \odot G\|_2 = \max_{\|x\|_2=1, \|y\|_2=1} \|x^T (F \odot G) y\|_2$$

Ora abbiamo che:

$$\|x^T(F \odot G)y\|_2 = \left| \sum_{i,j=1}^n x_i F_{i,j} G_{i,j} y_j \right|$$

Se applichiamo la disuguaglianza di Holder nello spazio euclideo otteniamo:

$$\begin{aligned} \left| \sum_{i,j=1}^n x_i F_{i,j} G_{i,j} y_j \right| &\leq \left(\sum_{i,j=1}^n |x_i F_{i,j}|^2 \right)^{1/2} \left(\sum_{i,j=1}^n |G_{i,j} y_j|^2 \right)^{1/2} \\ &= \left(\sum_{i=1}^n |x_i|^2 \sum_{j=1}^n |F_{i,j}|^2 \right)^{1/2} \left(\sum_{j=1}^n |y_j|^2 \sum_{i=1}^n |G_{i,j}|^2 \right)^{1/2} \\ &= \left(\sum_{i=1}^n |x_i|^2 r_i(F)^2 \right)^{1/2} \left(\sum_{j=1}^n |y_j|^2 c_j(G)^2 \right)^{1/2} \\ &\leq \left(\sum_{i=1}^n |x_i|^2 r_{\max}(F)^2 \right)^{1/2} \left(\sum_{j=1}^n |y_j|^2 c_{\max}(G)^2 \right)^{1/2} \\ &= r_{\max}(F) c_{\max}(G) \|x\|_2 \|y\|_2 = r_{\max}(F) c_{\max}(G) \end{aligned}$$

Per la commutatività del prodotto di Hadamard, vale anche:

$$\|F \odot G\|_2 \leq r_{\max}(G) c_{\max}(F)$$

Nella seconda parte invece è necessario dimostrare la seguente disuguaglianza, che conduce direttamente alla tesi:

$$r_{\max}(F) c_{\max}(G) \leq \|F\|_2 \|G\|_2$$

A tale scopo, è utile il seguente fatto facilmente dimostrabile: data una matrice $F \in \mathbb{R}^{n \times n}$, vale che:

$$\begin{aligned} \text{diag}(FF^T) &= \{r_1(F)^2, \dots, r_n(F)^2\} \\ \text{diag}(F^T F) &= \{c_1(F)^2, \dots, c_n(F)^2\} \end{aligned}$$

Inoltre il teorema di Schur-Horn afferma che: dati due vettori $d, \lambda \in \mathbb{R}^n$ i cui valori sono ordinati in modo non crescente, ovvero

$$d_1 \geq d_2 \geq \dots \geq d_n, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$$

esiste una matrice hermitiana $H \in \mathbb{R}^{n \times n}$ tale che:

$$\text{diag}(H) = \{d_1, \dots, d_n\}, \quad \text{spec}(H) = \lambda_1, \dots, \lambda_n$$

se e solo se:

$$\sum_{i=1}^k d_i \leq \sum_{i=1}^k \lambda_i, \quad k = 1, \dots, n-1$$

e

$$\sum_{i=1}^n d_i = \sum_{i=1}^n \lambda_i$$

In particolare, se scegliamo $H = F^T F$ abbiamo quindi che:

$$d_1 = \max \text{diag}(F^T F) = r_{\max}(F)^2 \leq \lambda_1(F^T F) = \lambda_{\max}(F^T F) = \sigma_{\max}^2(F) = \|F\|_2$$

e allo stesso modo, se $H = G G^T$:

$$d_1 = \max \text{diag}(G G^T) = c_{\max}(G)^2 \leq \lambda_1(G G^T) = \lambda_{\max}(G G^T) = \sigma_{\max}^2(G) = \|G\|_2$$

Quindi, concludendo, abbiamo che:

$$\|F \odot G\|_2 \leq r_{\max}(F) c_{\max}(G) \leq \|F\|_2 \|G\|_2$$

da cui la tesi □

Appendice B

Altre misure di controllabilità del gramiano

Nella trattazione, allo scopo di studiare il problema della controllabilità dal punto di vista energetico, si è fatto riferimento comunemente all'autovalore minimo del gramiano di controllabilità, in quanto esso è direttamente legato allo stato più difficile da raggiungere. Tuttavia, al posto di una prospettiva di tipo *worst case*, è possibile adottare altre metriche relative al gramiano (si veda [1], [14]):

- $tr(W)$: la traccia del gramiano è forse la scelta più semplice dal punto di vista computazionale e rappresenta in un certo senso l'energia *media* necessaria per controllare la rete. In particolare essa è legata anche alla norma H_2 del sistema, definita come:

$$\|H\|_2 = tr(CWC^T)$$

dove la matrice C lega gli stati alle uscite (se si suppone che sia possibile misurare direttamente tutte le variabili, la norma H_2 coincide proprio con la traccia del gramiano). Inoltre, come riportato in [1], la scelta dei nodi di controllo relativa alla massimizzazione della traccia del gramiano ammette una forma chiusa, anche se è bene ricordare che tale massimizzazione non assicura la controllabilità del sistema.

- $tr(W^{-1})$: in alternativa è possibile analizzare la traccia dell'inverso del gramiano, che, sebbene risulti più difficile, tiene conto anche dell'effettiva controllabilità del sistema rispetto alla precedente metrica. In particolare, rappresenta ancora l'energia media necessaria per controllare stati scelti in maniera casuale; in effetti vale la relazione ([1]):

$$\frac{tr(W^{-1})}{n} \geq \frac{n}{tr(W)}$$

dove n rappresenta la dimensione del sistema.

- $\log \det(W)$: un'altra possibile scelta per quantificare la controllabilità ricade sul determinante del gramiano (o, più comunemente, sul suo logaritmo). Esso è strettamente legato alla funzione volumetrica:

$$V(\Phi) = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)} \sqrt[n]{\det(W)}$$

dove

$$\Phi = \{x \in \mathbb{R}^n : x^T W^{-1} x \leq 1\}$$

e Γ è la funzione nota in letteratura come *Gamma function* ([15]). In pratica, il determinante del gramiano è proporzionale al volume degli stati del sistema che possono essere raggiunti spendendo una quantità di energia minore o uguale a uno. Di solito, siccome il calcolo del determinante può essere numericamente difficoltoso, soprattutto nei sistemi di dimensioni elevate, se ne studia il logaritmo.

Bibliografia

- [1] F. Pasqualetti, S. Zampieri, F. Bullo, *Controllability Metrics, Limitations and Algorithms for Complex Networks*, Transactions on control of network systems, Vol. 1, no. 1, march 2014
- [2] E. Fornasini, *Appunti di teoria dei sistemi*, ISBN 978-88-96477-32-8
- [3] T. Penzl, *Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case*, Systems & Control Letters 40 (2000) 139–144
- [4] G. Lombardi, R. Rebaudo, *Eigenvalues and eigenvectors of a special class of band matrices*, in: Rendiconti dell'Istituto di Matematica dell'Università di Trieste. An International Journal of Mathematics, 20 (1988), pp. 113-128
- [5] G. Bianchin, F. Pasqualetti, S. Zampieri, *The Role of Diameter in the Controllability of Complex Networks*, 2015 54th IEEE Conference on Decision and Control (CDC), 980-985, ISBN 978-1-4799-7884-7
- [6] R.B. Bapat, *Graphs and Matrices*, Universitext, Springer-Verlag London 2014
- [7] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Pr., 1991.
- [8] C.M. da Fonseca, J. Petronilho, *Explicit inverses of some tridiagonal matrices*, Linear Algebra and its Applications 325 (2001) 7-21
- [9] P.J. Davis, *Circulant Matrices*, AMS Chelsea Publishing, 1994
- [10] Michail Tsagris, Christina Beneki and Hossein Hassani, *On the Folded Normal Distribution*, Mathematics 2014, 2, 12-28; doi:10.3390/math2010012
- [11] Paten B., Diekhans M., Earl D., John J.S., Ma J., Suh B., Haussler D., *Cactus graphs for genome comparisons*, J Comput Biol. 2011 Mar;18(3):469-81. doi: 10.1089/cmb.2010.0252.
- [12] K. Das, *Cactus Graphs and Some Algorithms*, <https://arxiv.org/pdf/1408.4005>
- [13] Terence Tao, *Topics in random matrix theory*, Publication Year: 2012, ISBN-10: 0-8218-7430-6, ISBN-13: 978-0-8218-7430-1, Graduate Studies in Mathematics, vol. 132, American Mathematical Society

- [14] F. L. Cortesi, T. H. Summers, J. Lygeros, *Submodularity of energy related controllability metrics*, 53rd IEEE Conference on Decision and Control, pag. 2883 - 2888, doi: 10.1109/CDC.2014.7039832
- [15] G. E. Andrews, R. Askey, R. Roy, *Special Functions*, Cambridge University Press, 2001. ISBN 978-0-521-78988-2.