Teorema di Lin

Confronto di alcune dimostrazioni costruttive

Nikita Deniskin

Relatore: prof. Michele Benzi

Scuola Normale Superiore

19 aprile 2021

Teorema (Lin 1995)

Per ogni ε esiste un δ con la seguente proprietà. Se A, B sono matrici Hermitiane in $M(n, \mathbb{C})$ con $||[A, B]|| \leq \delta$, allora esistono A', B' tali che:

- $||A - A'|| \leq \varepsilon$ e $||B - B'|| \leq \varepsilon$
- A' e B' sono Hermitiane
- A' e B' commutano: $[A', B'] = 0$.

Teorema (Lin 1995)

Per ogni ε esiste un δ con la seguente proprietà. Se A, B sono matrici Hermitiane in $M(n, \mathbb{C})$ con $||[A, B]|| \leq \delta$, allora esistono A', B' tali che:

- $||A - A'|| \leq \varepsilon$ e $||B - B'|| \leq \varepsilon$
- A' e B' sono Hermitiane
- A' e B' commutano: $[A', B'] = 0$.

È abbastanza facile ottenere per n fissato una dipendenza ε - δ. La peculiarità di questo teorema è che la stima è indipendente da n.

Teorema (Lin 1995)

Per ogni ε esiste un δ con la seguente proprietà. Se A, B sono matrici Hermitiane in $M(n, \mathbb{C})$ con $||[A, B]|| \leq \delta$, allora esistono A', B' tali che:

- $||A - A'|| \leq \varepsilon$ e $||B - B'|| \leq \varepsilon$
- A' e B' sono Hermitiane
- A' e B' commutano: $[A', B'] = 0$.

È abbastanza facile ottenere per n fissato una dipendenza ε - δ. La peculiarità di questo teorema è che la stima è indipendente da n.

- Come si ottengono A', B' a partire da A, B?
- È possibile ottenere una stima di ε in funzione di δ?

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.
- Davidson (1985) presenta una formulazione equivalente $(\star)$.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.
- Davidson (1985) presenta una formulazione equivalente $(\star)$.
- Szarek (1990), stima dipendente da n, costruttiva, usando $(\star)$.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.
- Davidson (1985) presenta una formulazione equivalente $(\star)$.
- Szarek (1990), stima dipendente da n, costruttiva, usando $(\star)$.
- Lin (1995, pubbl.1997) dimostrazione del teorema, non costruttiva.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.
- Davidson (1985) presenta una formulazione equivalente $(\star)$.
- Szarek (1990), stima dipendente da n, costruttiva, usando $(\star)$.
- Lin (1995, pubbl.1997) dimostrazione del teorema, non costruttiva.
- Friis, Rørdam (1996) semplificano la dimostrazione di Lin.

Storia del problema e della soluzione:

- Formulato da Rosenthal (1969). Elencato da Halmos (1976) tra vari problemi aperti, con particolare interesse nel caso di dimensione infinita.
- Berg, Olsen (1980) trovano un controesempio per dimensione infinita.
- Pearcy, Shields (1979) stima dipendente da n, costruttiva, indebolendo alcune ipotesi.
- Davidson (1985) presenta una formulazione equivalente $(\star)$.
- Szarek (1990), stima dipendente da n, costruttiva, usando $(\star)$.
- Lin (1995, pubbl.1997) dimostrazione del teorema, non costruttiva.
- Friis, Rørdam (1996) semplificano la dimostrazione di Lin.
- Hastings (2009), mostra una stima asintotica per ε-δ, non completamente costruttiva, usando $(\star)$.

Teorema (Pearcy, Shields 1979)

Siano A, B in $M(n, \mathbb{C})$, con B Hermitiana e con al massimo m autovalori distinti. Sia $\delta = ||[A, B]||$. Allora esistono A', B' tali che:

- $||A - A'||, ||B - B'|| \leq \frac{m-1}{\sqrt{2}} \delta^{\frac{1}{2}}$
- B' è Hermitiana
- A' e B' commutano: $[A', B'] = 0$.

Teorema (Pearcy, Shields 1979)

Siano A, B in $M(n, \mathbb{C})$, con B Hermitiana e con al massimo m autovalori distinti. Sia $\delta = ||[A, B]||$. Allora esistono A', B' tali che:

- $||A - A'||, ||B - B'|| \leq \frac{m-1}{\sqrt{2}} \, \delta^{\frac{1}{2}}$
- B' è Hermitiana
- A' e B' commutano: $[A', B'] = 0$.

Confronto con il Teorema di Lin:

- La relazione $\varepsilon = (m-1)(\delta/2)^{\frac{1}{2}}$ è esplicita, però dipende dalla dimensione .
- A', B' si possono costruire esplicitamente.
- Si richiede che solo una delle due matrici sia Hermitiana. Per questo caso, la stima è ottimale.

Dimostrazione: Scriviamo A e B in una base di autovettori di B, ordinati per autovalori crescenti $\lambda_1 < \ldots < \lambda_m$. B è diagonale, multipla dell'identità a blocchi. Per esempio, per $m = 4$:

$$B = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \lambda_3 I & \\ & & & \lambda_4 I \end{pmatrix} \quad A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

Dimostrazione: Scriviamo A e B in una base di autovettori di B, ordinati per autovalori crescenti $\lambda_1 < \ldots < \lambda_m$. B è diagonale, multipla dell'identità a blocchi. Per esempio, per $m = 4$:

$$B = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \lambda_3 I & \\ & & & \lambda_4 I \end{pmatrix} \quad A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

$$[A, B] = \begin{pmatrix} 0 & (\lambda_2 - \lambda_1)A_{12} & (\lambda_3 - \lambda_1)A_{13} & (\lambda_4 - \lambda_1)A_{14} \\ (\lambda_1 - \lambda_2)A_{21} & 0 & (\lambda_3 - \lambda_2)A_{23} & A_{24} \\ (\lambda_1 - \lambda_3)A_{31} & (\lambda_2 - \lambda_3)A_{32} & 0 & (\lambda_4 - \lambda_3)A_{34} \\ (\lambda_1 - \lambda_4)A_{41} & (\lambda_2 - \lambda_4)A_{42} & (\lambda_3 - \lambda_4)A_{43} & 0 \end{pmatrix}$$

Si ha $[A, B]_{rs} = (\lambda_s - \lambda_r)\, A_{rs}$, dunque:

Dimostrazione: Scriviamo A e B in una base di autovettori di B, ordinati per autovalori crescenti $\lambda_1 < \ldots < \lambda_m$. B è diagonale, multipla dell'identità a blocchi. Per esempio, per $m = 4$:

$$B = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \lambda_3 I & \\ & & & \lambda_4 I \end{pmatrix} \quad A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix}$$

$$[A, B] = \begin{pmatrix} 0 & (\lambda_2 - \lambda_1)A_{12} & (\lambda_3 - \lambda_1)A_{13} & (\lambda_4 - \lambda_1)A_{14} \\ (\lambda_1 - \lambda_2)A_{21} & 0 & (\lambda_3 - \lambda_2)A_{23} & A_{24} \\ (\lambda_1 - \lambda_3)A_{31} & (\lambda_2 - \lambda_3)A_{32} & 0 & (\lambda_4 - \lambda_3)A_{34} \\ (\lambda_1 - \lambda_4)A_{41} & (\lambda_2 - \lambda_4)A_{42} & (\lambda_3 - \lambda_4)A_{43} & 0 \end{pmatrix}$$

Si ha $[A, B]_{rs} = (\lambda_s - \lambda_r)\, A_{rs}$, dunque:

$$||(\lambda_s - \lambda_r)\, A_{rs}|| \leq ||\, [A, B]_{rs}|| \leq ||\, [A, B]\, || = \delta,$$

$$||A_{rs}|| \leq \frac{\delta}{|\lambda_r - \lambda_s|}.$$

Si fissi una larghezza h. Suddividiamo lo spettro di B in intervalli $J_1, \ldots, J_k$ in modo che se $\lambda_r \in J_i, \lambda_s \in J_{i+1}$, allora $|\lambda_s - \lambda_r| \geq h$ (tra un intervallo e un altro c'è distanza almeno h). Gli autovalori appartenenti allo stesso J_i vengono uniti in un solo superblocco.

Si fissi una larghezza h. Suddividiamo lo spettro di B in intervalli $J_1, \ldots, J_k$ in modo che se $\lambda_r \in J_i, \lambda_s \in J_{i+1}$, allora $|\lambda_s - \lambda_r| \geq h$ (tra un intervallo e un altro c'è distanza almeno h). Gli autovalori appartenenti allo stesso J_i vengono uniti in un solo superblocco.
Visto che ci sono al massimo m autovalori distinti, la lunghezza di ciascun intervallo è $|J_i| \leq (m-1)h$. Inoltre il numero di blocchi è $k \leq m$.

Si fissi una larghezza h. Suddividiamo lo spettro di B in intervalli $J_1, \dots, J_k$ in modo che se $\lambda_r \in J_i, \lambda_s \in J_{i+1}$, allora $|\lambda_s - \lambda_r| \geq h$ (tra un intervallo e un altro c'è distanza almeno h). Gli autovalori appartenenti allo stesso J_i vengono uniti in un solo superblocco.
Visto che ci sono al massimo m autovalori distinti, la lunghezza di ciascun intervallo è $|J_i| \leq (m-1)h$. Inoltre il numero di blocchi è $k \leq m$.

Definiamo la matrice B': per ogni entrata diagonale $\lambda \in J_i$ di B, la modifichiamo con il punto medio dell'intervallo J_i. Dunque $||B - B'|| \leq \max\limits_{1 \leq i \leq k} \frac{1}{2}|J_i| = \frac{1}{2}(m-1)h$.

$$B = \begin{pmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \lambda_3 I & \\ & & & \lambda_4 I \end{pmatrix} \quad B' = \begin{pmatrix} \mu_1 I & & & \\ & \mu_1 I & & \\ & & \mu_2 I & \\ & & & \mu_2 I \end{pmatrix}$$

La matrice A' è ottenuta da A seguendo la stessa divisione a blocchi in J_i, e mettendo a 0 i blocchi non sulla diagonale.

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ A_{21} & A_{22} & A_{23} & A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{pmatrix} \quad A' = \begin{pmatrix} A_{11} & A_{12} & & \\ A_{21} & A_{22} & & \\ & & A_{33} & A_{34} \\ & & A_{43} & A_{44} \end{pmatrix}$$

A' e B' commutano, perché hanno la stessa struttura diagonale a blocchi, e B' è multipla dell'identità su ciascun blocco.

Stimiamo ora $||A - A'||$. Nella prima suddivisione a blocchi, $A - A'$ ha blocchi non nulli $(A - A')_{rs} = A_{rs}$ quando $|\lambda_r - \lambda_s| > h$. Dunque

$$||(A - A')_{rs}|| = ||A_{rs}|| \leq \frac{\delta}{|\lambda_r - \lambda_s|} \leq \frac{\delta}{h}$$

Stimiamo ora $||A - A'||$. Nella prima suddivisione a blocchi, $A - A'$ ha blocchi non nulli $(A - A')_{rs} = A_{rs}$ quando $|\lambda_r - \lambda_s| > h$. Dunque

$$||(A - A')_{rs}|| = ||A_{rs}|| \leq \frac{\delta}{|\lambda_r - \lambda_s|} \leq \frac{\delta}{h}$$

In $A - A'$ ci sono m blocchi per riga (colonna), e su ciascuna ci sono al massimo $m - 1$ blocchi non nulli.

$$||A - A'|| \leq \max_{r,s} (m-1) \, ||A_{r,s}|| \leq (m-1) \frac{\delta}{h}$$

Stimiamo ora $||A - A'||$. Nella prima suddivisione a blocchi, $A - A'$ ha blocchi non nulli $(A - A')_{rs} = A_{rs}$ quando $|\lambda_r - \lambda_s| > h$. Dunque

$$||(A - A')_{rs}|| = ||A_{rs}|| \leq \frac{\delta}{|\lambda_r - \lambda_s|} \leq \frac{\delta}{h}$$

In $A - A'$ ci sono m blocchi per riga (colonna), e su ciascuna ci sono al massimo $m - 1$ blocchi non nulli.

$$||A - A'|| \leq \max_{r,s} (m-1)\, ||A_{r,s}|| \leq (m-1)\frac{\delta}{h}$$

Abbiamo ottenuta prima $||B - B'|| \leq \frac{1}{2}(m-1)h$.
Scegliendo $h = \left(\frac{\delta}{2}\right)^{\frac{1}{2}}$, si ottiene $||A - A'||, ||B - B'|| \leq \frac{m-1}{\sqrt{2}}\,\delta^{\frac{1}{2}}$ $\square$

L'idea dietro alla costruzione di A', B' è la seguente: per B lasciamo fissi gli autospazi e modifichiamo autovalori vicini per avere un blocco identità. Per A, la rendiamo diagonale a blocchi, e scartiamo i blocchi fuori dalla diagonale.
Useremo lo stesso approccio anche per il teorema di Lin. Tuttavia scartare tutti i blocchi non diagonali di A rende la stima troppo larga.

L'idea dietro alla costruzione di A', B' è la seguente: per B lasciamo fissi gli autospazi e modifichiamo autovalori vicini per avere un blocco identità. Per A, la rendiamo diagonale a blocchi, e scartiamo i blocchi fuori dalla diagonale.
Useremo lo stesso approccio anche per il teorema di Lin. Tuttavia scartare tutti i blocchi non diagonali di A rende la stima troppo larga.

Proposizione 1

Esistono costanti c_0, c_1 per cui date $A, B \in M(n, \mathbb{C})$ Hermitiane e $\Delta > 0$, esiste H Hermitiana tale che

- $||A - H|| \leq \frac{c_0}{\Delta}||[A, B]||$
- $||[H, B]|| \leq c_1||[A, B]||$
- H ha **range finito**: presi due autovettori v_1, v_2 per B con autovalori λ_1, λ_2, se $|\lambda_1 - \lambda_2| \geq \Delta$ allora $\langle v_1, Hv_2 \rangle = 0$.

Si scelga $f(x)$ una bump function continua su $[-1,1]$, $f(0)=1$ e con trasformata di fourier $\hat{f}(y)$, in modo che le costanti $c_0 = \int |y\hat{f}(y)|dy$ e $c_1 = \int |\hat{f}(y)|dy$ siano finite.

$$H = \int_{\mathbb{R}} e^{i\frac{y}{\Delta}B}\, A\, e^{-i\frac{y}{\Delta}B}\, \hat{f}(y)dy$$

Si scelga $f(x)$ una bump function continua su $[-1,1]$, $f(0)=1$ e con trasformata di fourier $\hat{f}(y)$, in modo che le costanti $c_0 = \int |y\hat{f}(y)|dy$ e $c_1 = \int |\hat{f}(y)|dy$ siano finite.

$$H = \int_{\mathbb{R}} e^{i\frac{y}{\Delta}B}\, A\, e^{-i\frac{y}{\Delta}B}\, \hat{f}(y)dy$$

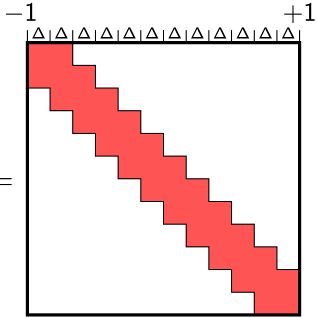
Mostriamo che H ha range finito. Presi v_1, v_2 autovettori per B, si ha

$$\begin{aligned}
\langle v_1, Hv_2\rangle &= \int_{\mathbb{R}} \langle e^{i\frac{y}{\Delta}B} v_1, A\, e^{-i\frac{y}{\Delta}B} v_2\rangle\, \hat{f}(y)dy = \\
&= \int_{\mathbb{R}} \langle e^{i\frac{y\lambda_1}{\Delta}} v_1, A\, e^{-i\frac{y\lambda_2}{\Delta}} v_2\rangle\, \hat{f}(y)dy = \\
&= \langle v_1, Av_2\rangle \int_{\mathbb{R}} e^{i\frac{y(\lambda_1-\lambda_2)}{\Delta}}\, \hat{f}(y)dy = \langle v_1, Av_2\rangle\, f\left(\frac{\lambda_1-\lambda_2}{\Delta}\right)
\end{aligned}$$

f è nulla al di fuori di $[-1,1]$, dunque l'ultimo termine è zero.

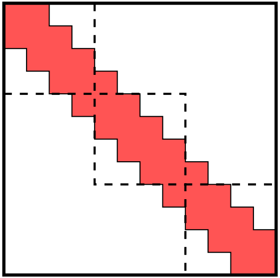
Scriviamo le matrici in una base di autovettori di B. B viene diagonale. Assumiamo che $||A||, ||B|| \leq 1$, dunque $\sigma(A), \sigma(B) \subseteq [-1, 1]$.
Si divida $[-1, 1]$ in intervalli lunghi Δ, e poi si dividano le matrici a blocchi, dove ciascun blocco è il sottospazio generato dagli autovettori di B con autovalore in un dato intervallo. Il numero di blocchi è $n_1 = \lceil \frac{2}{\Delta} \rceil$.

Scriviamo le matrici in una base di autovettori di B. B viene diagonale. Assumiamo che $||A||, ||B|| \leq 1$, dunque $\sigma(A), \sigma(B) \subseteq [-1, 1]$.
Si divida $[-1, 1]$ in intervalli lunghi Δ, e poi si dividano le matrici a blocchi, dove ciascun blocco è il sottospazio generato dagli autovettori di B con autovalore in un dato intervallo. Il numero di blocchi è $n_1 = \lceil \frac{2}{\Delta} \rceil$.

$H =$ 

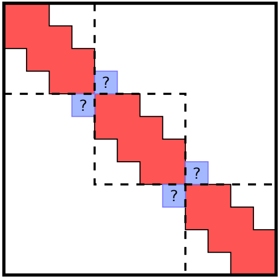
H viene tridiagonale a blocchi, perché presi v, w in blocchi a distanza almeno 2, la proprietà di range finito implica $\langle v, Hw \rangle = 0$.

Raggruppiamo i blocchi in gruppi di L, ottenendo dei blocchi più grandi. Ci saranno quindi $n_2 = \left\lceil \frac{n_1}{L} \right\rceil$ superblocchi.

$H =$ 

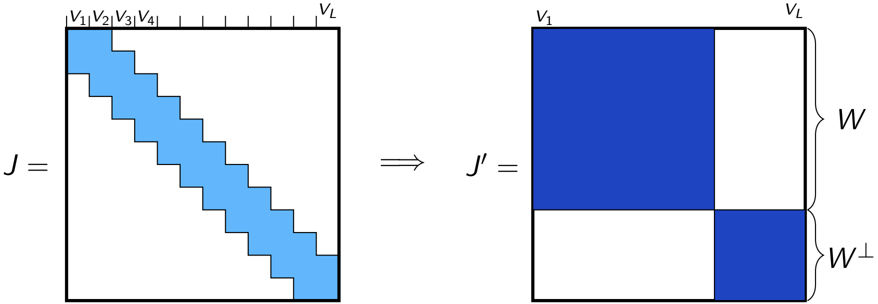
Si potrebbe provare lo stesso approccio del caso precedente: tagliare via tutto quello che non entra nella suddivisione a superblocchi.

Raggruppiamo i blocchi in gruppi di L, ottenendo dei blocchi più grandi. Ci saranno quindi $n_2 = \left\lceil \frac{n_1}{L} \right\rceil$ superblocchi.

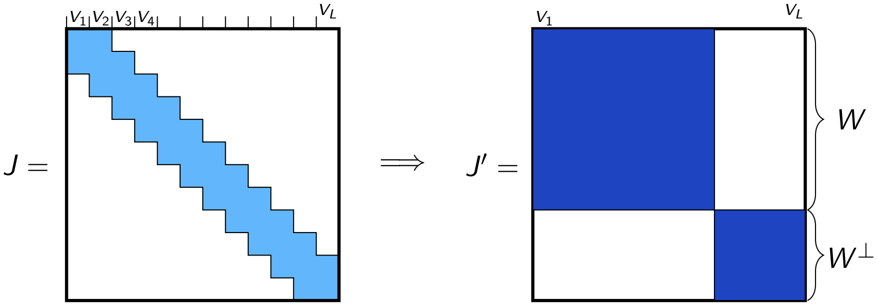


Si potrebbe provare lo stesso approccio del caso precedente: tagliare via tutto quello che non entra nella suddivisione a superblocchi.
Tuttavia, adesso non abbiamo controllo sulle entrate di H: quelle che verrebbero tagliate, potrebbero essere grandi.

Concentriamoci su un solo superblocco. Esso eredita la struttura tridiagonale a blocchi di H, con L blocchi larghi Δ. Chiamiamo J la restrizione di H al superblocco.
Vorremmo trasformare J da una struttura tridiagonale, a una struttura a due blocchi.

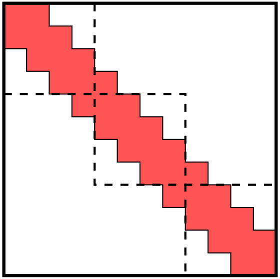


Questo non si può fare per J stessa, ma troveremo J' vicina a J con questa struttura: ha un sottospazio invariante W, tale che $V_1 \subseteq W$ e $V_L^\perp \subseteq W^\perp$.

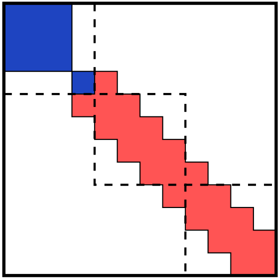


Questo non si può fare per J stessa, ma troveremo J' vicina a J con questa struttura: ha un sottospazio invariante W, tale che $V_1 \subseteq W$ e $V_L^\perp \subseteq W^\perp$. Trovare J' è equivalente a trovare un sottospazio W, in modo che sia W che $W^\perp$ siano quasi invarianti per J.

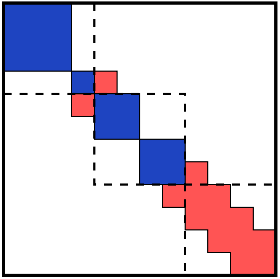
H è la somma diretta dei J, e sia A' ottenuta sostituendo ogni J con il relativo J'. Allora $||A - A'|| \leq \max ||J - J'||$.

$H =$ 

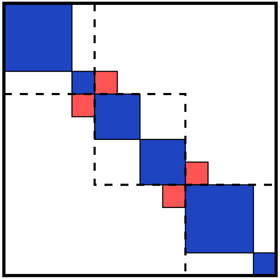
H è la somma diretta dei J, e sia A' ottenuta sostituendo ogni J con il relativo J'. Allora $||A - A'|| \leq \max ||J - J'||$.



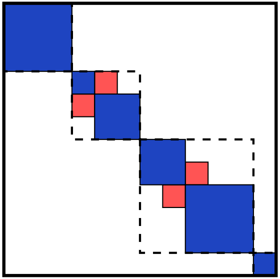
H è la somma diretta dei J, e sia A' ottenuta sostituendo ogni J con il relativo J'. Allora $||A - A'|| \leq \max ||J - J'||$.



H è la somma diretta dei J, e sia A' ottenuta sostituendo ogni J con il relativo J'. Allora $||A - A'|| \leq \max ||J - J'||$.

$A' =$ 

H è la somma diretta dei J, e sia A' ottenuta sostituendo ogni J con il relativo J'. Allora $||A - A'|| \leq \max ||J - J'||$.

$A' =$ 

Cambiamo suddivisione a blocchi. Si prenda B' diagonale a blocchi con la stessa suddivisione. In questo modo A' e B' commutano.

Proposizione ($\star$)

Sia J una matrice Hermitiana, con $||J|| \leq 1$ e tridiagonale a blocchi rispetto a una suddivisione $V_1, \dots, V_L$. Allora esiste un sottospazio W che soddisfa:

1. $V_1 \subseteq W \subseteq V_L^\perp$
2. Per ogni $w \in W$ e $v \in W^\perp$ con $|v| = |w| = 1$, $\langle v, Jw \rangle \leq \xi$

Siano P e Q le proiezioni ortogonali su W e $W^\perp$. Nella base data da W e $W^\perp$:

$$J = \begin{pmatrix} PJP & C \\ C^{\mathrm{H}} & QJQ \end{pmatrix}$$

con $C = PJQ$. La condizione (2) equivale a:

$$||PJQ|| = ||C|| \leq \xi$$

Proposizione ($\star$)

Sia J una matrice Hermitiana, con $||J|| \leq 1$ e tridiagonale a blocchi rispetto a una suddivisione $V_1, \ldots, V_L$. Allora esiste un sottospazio W che soddisfa:

1. $V_1 \subseteq W \subseteq V_L^\perp$
2. Per ogni $w \in W$ e $v \in W^\perp$ con $|v| = |w| = 1$, $\langle v, Jw \rangle \leq \xi$

Definiamo J' ponendo a 0 i blocchi fuori non diagonali:

$$J = \begin{pmatrix} PJP & C \\ C^{\mathrm{H}} & QJQ \end{pmatrix} \qquad J' = \begin{pmatrix} PJP & \\ & QJQ \end{pmatrix}$$

Allora $||J - J'|| = 2||C|| \leq 2\xi$.

Proposizione ($\star$)

Sia J una matrice Hermitiana, con $||J|| \leq 1$ e tridiagonale a blocchi rispetto a una suddivisione $V_1, \ldots, V_L$. Allora esiste un sottospazio W che soddisfa:

1. $V_1 \subseteq W \subseteq V_L^{\perp}$
2. Per ogni $w \in W$ e $v \in W^{\perp}$ con $|v| = |w| = 1$, $\langle v, Jw \rangle \leq \xi$

Quanto vale ξ?

Nella dimostrazione di Hastings, $\xi = \frac{g(L)}{L^{\gamma}}$, dove $g(L)$ cresce più lentamente di ogni potenza di L e possiamo scegliere $\gamma = 1/4$. Tuttavia per questa stima c'è bisogno di usare il teorema di Lin in un caso particolare. Questo non ci permette di ottenere W in maniera costruttiva, e neanche di ottenere una stima di $g(L)$.

Proposizione ($\star$)

Sia J una matrice Hermitiana, con $||J|| \leq 1$ e tridiagonale a blocchi rispetto a una suddivisione $V_1, \dots, V_L$. Allora esiste un sottospazio W che soddisfa:

1. $V_1 \subseteq W \subseteq V_L^{\perp}$
2. Per ogni $w \in W$ e $v \in W^{\perp}$ con $|v| = |w| = 1$, $\langle v, Jw \rangle \leq \xi$

Quanto vale ξ?

Nella dimostrazione di Hastings, $\xi = \frac{g(L)}{L^{\gamma}}$, dove $g(L)$ cresce più lentamente di ogni potenza di L e possiamo scegliere $\gamma = 1/4$. Tuttavia per questa stima c'è bisogno di usare il teorema di Lin in un caso particolare. Questo non ci permette di ottenere W in maniera costruttiva, e neanche di ottenere una stima di $g(L)$.

Nella dimostrazione di Szarek, otteniamo una stima esplicita per ξ, che però dipende dalla dimensione n delle matrici. Si ottiene W in maniera costruttiva.

Indichiamo $||[A, B]|| = \delta$.

- $||A - H|| \leq c_0 \frac{\delta}{\Delta}$
- $||H - A'|| \leq 2\xi = 2\frac{g(L)}{L^\gamma}$
- $||B - B'|| \leq \frac{2}{n_2}$

n_2 indica il numero di superblocchi. $L \cdot n_2 = \frac{2}{\Delta}$

Indichiamo $||[A, B]|| = \delta$.

- $||A - H|| \leq c_0 \frac{\delta}{\Delta}$
- $||H - A'|| \leq 2\xi = 2\frac{g(L)}{L^\gamma}$
- $||B - B'|| \leq \frac{2}{n_2}$

n_2 indica il numero di superblocchi. $L \cdot n_2 = \frac{2}{\Delta}$
Scegliamo $\Delta = \delta^{\alpha_0}$, $n_2 = \Delta^{\alpha_1} = \delta^{\alpha_0\,\alpha_1}$. Si ottiene $L = 2\delta^{\alpha_0(\alpha_1 - 1)}$

Indichiamo $||[A, B]|| = \delta$.

- $||A - H|| \leq c_0 \frac{\delta}{\Delta}$
- $||H - A'|| \leq 2\xi = 2\frac{g(L)}{L^\gamma}$
- $||B - B'|| \leq \frac{2}{n_2}$

n_2 indica il numero di superblocchi. $L \cdot n_2 = \frac{2}{\Delta}$
Scegliamo $\Delta = \delta^{\alpha_0}$, $n_2 = \Delta^{\alpha_1} = \delta^{\alpha_0\,\alpha_1}$. Si ottiene $L = 2\delta^{\alpha_0(\alpha_1 - 1)}$

$$||A - A'|| \leq c_0\,\delta^{1-\alpha_0} + 2\,g(L(\delta))\;\delta^{\alpha_0(1-\alpha_1)\gamma}$$

$$||B - B'|| \leq 2\,\delta^{\alpha_0\alpha_1}$$

Indichiamo $||[A, B]|| = \delta$.

- $||A - H|| \leq c_0 \frac{\delta}{\Delta}$
- $||H - A'|| \leq 2\xi = 2\frac{g(L)}{L^\gamma}$
- $||B - B'|| \leq \frac{2}{n_2}$

n_2 indica il numero di superblocchi. $L \cdot n_2 = \frac{2}{\Delta}$
Scegliamo $\Delta = \delta^{\alpha_0}$, $n_2 = \Delta^{\alpha_1} = \delta^{\alpha_0\,\alpha_1}$. Si ottiene $L = 2\delta^{\alpha_0(\alpha_1 - 1)}$

$$||A - A'|| \leq c_0\, \delta^{1-\alpha_0} + 2\, g(L(\delta))\; \delta^{\alpha_0(1-\alpha_1)\gamma}$$

$$||B - B'|| \leq 2\, \delta^{\alpha_0 \alpha_1}$$

Il miglior andamento asintotico è con i tre esponenti uguali, che con $\gamma = 1/4$ sono pari a $1/6$. Usando $g_1(\delta^{-1}) = \max(2, c_0 + g(L(\delta)))$, otteniamo la seguente relazione per ε.

$$\varepsilon = \; g_1(\delta^{-1})\, \delta^{\frac{1}{6}}$$

Teorema (Hastings 2009)

Preso $\delta > 0$, se A, B sono matrici Hermitiane in $M(n, \mathbb{C})$ con $|||[A, B]||| \leq \delta$, allora esistono A', B' Hermitiane tali che:

- $||A - A'||, ||B - B'|| \leq \varepsilon$
- $\varepsilon(\delta) = g_1(\delta^{-1})\, \delta^{\frac{1}{6}}$
- A' e B' commutano: $[A', B'] = 0$.

La funzione $g_1(x)$ è $O(x^t)$ per ogni $t > 0$.

La stima è indipendente da n, ma non si sa quanto dev'essere grande δ per avere $\varepsilon(\delta)$ piccolo.

Teorema (Szarek 1990)

Esiste una costante c, tale che se A, B sono matrici Hermitiane in $M(n, \mathbb{C})$ di rango minore o uguale a m e $||[A, B]|| = \delta$, allora esistono A', B' Hermitiane tali che:

- $||A - A'||, ||B - B'|| \leq \varepsilon$
- $\varepsilon(\delta, m) \; = \; c \, m^{\frac{1}{1}3} \, \delta^{\frac{2}{13}}$
- A' e B' commutano: $[A', B'] = 0$.

La dimostrazione è completamente costruttiva, ma la stima dipende da m.

Applicazioni del teorema di Lin

- In meccanica quantistica, se due osservabili P, Q commutano, allora il risultato della misura delle due quantità a loro legate è indipendente dall'ordine in cui vengono misurate. Viceversa, se P e Q non commutano, misurare una influisce sul risultato della seconda misura.

Applicazioni del teorema di Lin

- In meccanica quantistica, se due osservabili P, Q commutano, allora il risultato della misura delle due quantità a loro legate è indipendente dall'ordine in cui vengono misurate. Viceversa, se P e Q non commutano, misurare una influisce sul risultato della seconda misura.
- Nel contesto dei computer quantistici, è utile rendere le osservabili il più possibile "indipendenti" tra di loro.

Applicazioni del teorema di Lin

- In meccanica quantistica, se due osservabili P, Q commutano, allora il risultato della misura delle due quantità a loro legate è indipendente dall'ordine in cui vengono misurate. Viceversa, se P e Q non commutano, misurare una influisce sul risultato della seconda misura.
- Nel contesto dei computer quantistici, è utile rendere le osservabili il più possibile "indipendenti" tra di loro.
- Se P, Q commutano, è possibile costruire una misura operatoriale per misurarli contemporaneamente. Se P, Q hanno commutatore piccolo, è possibile misurarli contemporaneamente con un certo errore (dato dal teorema di Lin).

Applicazioni del teorema di Lin

- In meccanica quantistica, se due osservabili P, Q commutano, allora il risultato della misura delle due quantità a loro legate è indipendente dall'ordine in cui vengono misurate. Viceversa, se P e Q non commutano, misurare una influisce sul risultato della seconda misura.
- Nel contesto dei computer quantistici, è utile rendere le osservabili il più possibile "indipendenti" tra di loro.
- Se P, Q commutano, è possibile costruire una misura operatoriale per misurarli contemporaneamente. Se P, Q hanno commutatore piccolo, è possibile misurarli contemporaneamente con un certo errore (dato dal teorema di Lin).
- Nei problemi di Blind Source Separation, un algoritmo molto usato è JADE (Cardoso, Souloumiac) che si basa sul trovare una diagonalizzazione simultanea approssimata di due matrici.

Bibliografia

- D. Herrera, *Constructive Approaches to Lin's Theorem for Almost Commuting Matrices*, `https://arxiv.org/abs/2011.11800`
- H. Lin, *Almost commuting self-adjoint matrices and applications*, Fields. Inst. Commun. 13, (1995), pp. 193
- M. Hastings, *Making Almost Commuting Matrices Commute*, Commun. Math. Phys. 291, (2009) pp. 321–345
- K. Davidson, *Almost Commuting Hermitian Matrices*, Math. Scand. 56 (1985), pp. 222-240
- S. Szarek, *On Almost Commuting Hermitian Operators*, Rocky Mountain Journal of Mathematics. Vol. 20, No. 2, Spring 1990
- C. Pearcy, A. Shields, *Almost Commuting Matrices*, Journal of Functional Analysis. 33, (1979), pp. 332-338