



Funzioni di matrici di Toeplitz

Dislocare, Comprimere, Spelare.

Seminario PYSANUM

Andrea Baleani

24 Marzo 2026



SCUOLA
NORMALE
SUPERIORE



Table of Contents

1 Introduzione

► Introduzione

► Struttura di dislocamento e decadimento dei valori singolari

► Matrici HSS e decomposizione telescopica

► Aggiornamenti di basso rango e funzioni di matrici HSS

► Esperimenti

► Conclusioni



Matrici Toeplitz

1 Introduzione

Le matrici Toeplitz dipendono da $2n - 1 = \mathcal{O}(n)$ parametri (*data-sparse*).

$$T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-n+1} \\ t_1 & t_0 & \cdots & t_{-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \cdots & t_0 \end{bmatrix}, \quad t_{i,j} = t_{i-j}.$$



Matrici Toeplitz

1 Introduzione

Le matrici Toeplitz dipendono da $2n - 1 = \mathcal{O}(n)$ parametri (*data-sparse*).

$$T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-n+1} \\ t_1 & t_0 & \cdots & t_{-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \cdots & t_0 \end{bmatrix}, \quad t_{i,j} = t_{i-j}.$$

In generale le funzioni non preservano la struttura Toeplitz (ad esempio T^2).



Matrici Toeplitz

1 Introduzione

Le matrici Toeplitz dipendono da $2n - 1 = \mathcal{O}(n)$ parametri (*data-sparse*).

$$T = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{-n+1} \\ t_1 & t_0 & \cdots & t_{-n+2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & \cdots & t_0 \end{bmatrix}, \quad t_{i,j} = t_{i-j}.$$

In generale le funzioni non preservano la struttura Toeplitz (ad esempio T^2).

Domanda: come possiamo sfruttare la proprietà *data-sparse* per calcolare $f(T)$ o $f(T)v$ in modo efficiente preservando struttura?



Da dove vengono le matrici di Toeplitz?

1 Introduzione

La struttura di Toeplitz nasce spesso da proprietà di invarianza per traslazione.

$$b(s) = \int_{-\infty}^{+\infty} p(s-t) x(t) dt \quad \rightsquigarrow \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} p_0 & p_{-1} & p_{-2} & p_{-3} & p_{-4} \\ p_1 & p_0 & p_{-1} & p_{-2} & p_{-3} \\ p_2 & p_1 & p_0 & p_{-1} & p_{-2} \\ p_3 & p_2 & p_1 & p_0 & p_{-1} \\ p_4 & p_3 & p_2 & p_1 & p_0 \end{bmatrix} \begin{bmatrix} x_{-1} \\ x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$



Da dove vengono le matrici di Toeplitz?

1 Introduzione

La struttura di Toeplitz nasce spesso da proprietà di invarianza per traslazione.

$$b(s) = \int_{-\infty}^{+\infty} p(s-t) x(t) dt \quad \rightsquigarrow \quad \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} p_0 & p_{-1} & p_{-2} & p_{-3} & p_{-4} \\ p_1 & p_0 & p_{-1} & p_{-2} & p_{-3} \\ p_2 & p_1 & p_0 & p_{-1} & p_{-2} \\ p_3 & p_2 & p_1 & p_0 & p_{-1} \\ p_4 & p_3 & p_2 & p_1 & p_0 \end{bmatrix} \begin{bmatrix} x_{-1} \\ x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- Filtri per immagini: Space-invariant blurring models, RBF edge-detection...
- Equazioni integrali: metodo di Nyström o Kansa, equazioni di tipo Fredholm o Wiener-Hopf.



Applicazioni di $f(T)$ per matrici Toeplitz

1 Introduzione

Esponenziale di matrice: $f(T) = e^T$

- Evoluzione temporale in sistemi dinamici lineari: $u(t) = e^{tT}u_0$;
- Integratori esponenziali per PDE.

Funzioni di Markov: $f(T) = \int_a^b (T - tI)^{-1} d\mu(t)$, dove μ è una misura positiva su $[a, b]$.

- Diffusione frazionaria: $T^{-\alpha}$, $\alpha \in (0, 1)$;
- Processi gaussiani: $T^{1/2}$, $\log(T)$;
- Problemi inversi: risolventi $(T + \sigma I)^{-1}$.



Roadmap

1 Introduzione

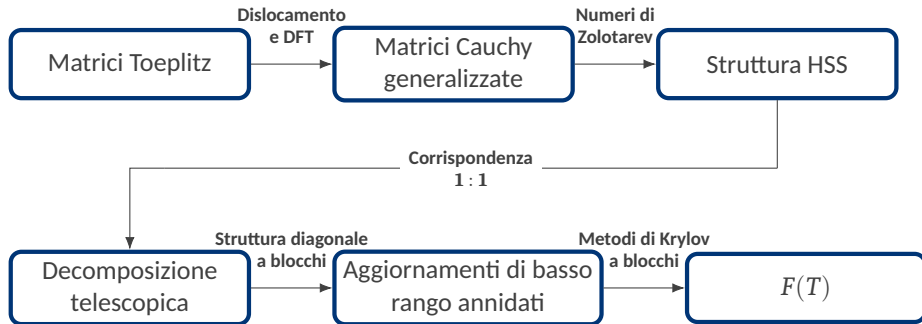




Table of Contents

2 Struttura di dislocamento e decadimento dei valori singolari

- ▶ Introduzione
- ▶ **Struttura di dislocamento e decadimento dei valori singolari**
- ▶ Matrici HSS e decomposizione telescopica
- ▶ Aggiornamenti di basso rango e funzioni di matrici HSS
- ▶ Esperimenti
- ▶ Conclusioni



Basso rango di dislocamento

2 Struttura di dislocamento e decadimento dei valori singolari

Una matrice M soddisfa un'equazione di Sylvester con termine di rango basso:

$$\nabla_{A,B}(M) := AM - MB = R = GH^*, \quad \text{rango}(GH^*) \ll n.$$



Basso rango di dislocamento

2 Struttura di dislocamento e decadimento dei valori singolari

Una matrice M soddisfa un'equazione di Sylvester con termine di rango basso:

$$\nabla_{A,B}(M) := AM - MB = R = GH^*, \quad \text{rango}(GH^*) \ll n.$$

Le matrici Toeplitz hanno rango di dislocamento (Z, Z) pari a 2, dove Z è la matrice di shift:

$$\nabla_{Z,Z}(T) = \begin{bmatrix} y^T & * \\ \mathbf{0}_{n-1} & x \end{bmatrix} = UV^*, \quad Z = \begin{bmatrix} \mathbf{0}^T & 1 \\ I_{n-1} & \mathbf{0} \end{bmatrix}.$$



Basso rango di dislocamento

2 Struttura di dislocamento e decadimento dei valori singolari

Una matrice M soddisfa un'equazione di Sylvester con termine di rango basso:

$$\nabla_{A,B}(M) := AM - MB = R = GH^*, \quad \text{rango}(GH^*) \ll n.$$

Le matrici Toeplitz hanno rango di dislocamento (Z, Z) pari a 2, dove Z è la matrice di shift:

$$\nabla_{Z,Z}(T) = \begin{bmatrix} y^T & * \\ \mathbf{0}_{n-1} & x \end{bmatrix} = UV^*, \quad Z = \begin{bmatrix} \mathbf{0}^T & 1 \\ I_{n-1} & \mathbf{0} \end{bmatrix}.$$

Z è circolante e diagonalizzabile tramite DFT $\Omega_{ij} = \frac{1}{\sqrt{n}} \omega^{(i-1)(j-1)}$:

$$\Omega^* Z \Omega = \text{diag}(\omega^0, \omega^1, \dots, \omega^{n-1}) =: D, \quad \omega = e^{-2\pi i/n}.$$



Equazione di Sylvester singolare

2 Struttura di dislocamento e decadimento dei valori singolari

Da $\nabla_{Z,Z}(T)$ si ottiene un'equazione di Sylvester singolare con una matrice Cauchy generalizzata $C = \Omega^* T \Omega$:

$$DC - CD = (\Omega^* G)(\Omega^* H)^* =: \tilde{G} \tilde{H}^* \quad \Longrightarrow \quad C_{i,j} = \frac{1}{n} \frac{1}{\omega^{i-1} - \omega^{j-1}} (\tilde{G} \tilde{H}^*)_{i,j}, \quad i \neq j.$$



Equazione di Sylvester singolare

2 Struttura di dislocamento e decadimento dei valori singolari

Da $\nabla_{Z,Z}(T)$ si ottiene un'equazione di Sylvester singolare con una matrice Cauchy generalizzata $C = \Omega^* T \Omega$:

$$DC - CD = (\Omega^* G)(\Omega^* H)^* =: \tilde{G} \tilde{H}^* \implies C_{i,j} = \frac{1}{n} \frac{1}{\omega^{i-1} - \omega^{j-1}} (\tilde{G} \tilde{H}^*)_{i,j}, \quad i \neq j.$$

Le sottomatrici $C_{I,J}$ restano di tipo Cauchy generalizzato:

$$D_I C_{I,J} - C_{I,J} D_J = \tilde{G}(I, :) \tilde{H}(J, :)^*, \quad D_I = \text{diag}(\omega^{i-1})_{i \in I}.$$



Equazione di Sylvester singolare

2 Struttura di dislocamento e decadimento dei valori singolari

Da $\nabla_{Z,Z}(T)$ si ottiene un'equazione di Sylvester singolare con una matrice Cauchy generalizzata $C = \Omega^* T \Omega$:

$$DC - CD = (\Omega^* G)(\Omega^* H)^* =: \tilde{G} \tilde{H}^* \implies C_{i,j} = \frac{1}{n} \frac{1}{\omega^{i-1} - \omega^{j-1}} (\tilde{G} \tilde{H}^*)_{i,j}, \quad i \neq j.$$

Le sottomatrici $C_{I,J}$ restano di tipo Cauchy generalizzato:

$$D_I C_{I,J} - C_{I,J} D_J = \tilde{G}(I, :) \tilde{H}(J, :)^*, \quad D_I = \text{diag}(\omega^{i-1})_{i \in I}.$$

Le entrate diagonali non sono un problema

$\ker(\nabla_{Z,Z})$ sono le matrici circolanti quindi $\Omega^* \Pi_F^\perp[T] \Omega = \text{diag}(C)$. Basta calcolare la FFT della prima colonna di $\Pi_F^\perp[T]$, che è la media lungo le diagonali cicliche di T .



Decadimento dei valori singolari

2 Struttura di dislocamento e decadimento dei valori singolari

Teorema

Sia X una matrice con rango di dislocamento (A, B) pari a ν , con A, B normali. Se $\sigma(A) \subseteq E$ e $\sigma(B) \subseteq F$, allora i valori singolari di X soddisfano:

$$\sigma_{j+\nu k}(X) \leq \mathcal{Z}_k(E, F) \sigma_j(X), \quad 1 \leq j + \nu k \leq n.$$



Decadimento dei valori singolari

2 Struttura di dislocamento e decadimento dei valori singolari

Teorema

Sia X una matrice con rango di dislocamento (A, B) pari a ν , con A, B normali. Se $\sigma(A) \subseteq E$ e $\sigma(B) \subseteq F$, allora i valori singolari di X soddisfano:

$$\sigma_{j+\nu k}(X) \leq \mathcal{Z}_k(E, F) \sigma_j(X), \quad 1 \leq j + \nu k \leq n.$$

$$\mathcal{Z}_k(E, F) = \inf_{r \in \mathcal{R}_{k,k}} \frac{\sup_{z \in E} |r(z)|}{\inf_{z \in F} |r(z)|}, \quad (\text{Numero di Zolotarev})$$

dove $\mathcal{R}_{k,k}$ è l'insieme delle funzioni razionali di grado al più (k, k) .



Backstage

2 Struttura di dislocamento e decadimento dei valori singolari

Per alcuni insiemi particolari, $\mathcal{Z}_k$ può essere calcolato "esplicitamente"

$$\mathcal{Z}_k([-b, -a], [a, b]) = 4\rho^{-2k} \prod_{\tau=1}^{\infty} \frac{(1 + \rho^{-8\tau k})^4}{(1 + \rho^{4k} \rho^{-8\tau k})^4} \leq 4\rho^{-2k}, \quad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right),$$

dove $\mu(\lambda) = \frac{\pi}{2} \frac{K(\sqrt{1-\lambda^2})}{K(\lambda)}$ è costruita da $K(\lambda) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-\lambda^2 t^2)}}$.



Backstage

2 Struttura di dislocamento e decadimento dei valori singolari

Per alcuni insiemi particolari, $\mathcal{Z}_k$ può essere calcolato "esplicitamente"

$$\mathcal{Z}_k([-b, -a], [a, b]) = 4\rho^{-2k} \prod_{\tau=1}^{\infty} \frac{(1 + \rho^{-8\tau k})^4}{(1 + \rho^{4k} \rho^{-8\tau k})^4} \leq 4\rho^{-2k}, \quad \rho = \exp\left(\frac{\pi^2}{2\mu(a/b)}\right),$$

dove $\mu(\lambda) = \frac{\pi}{2} \frac{K(\sqrt{1-\lambda^2})}{K(\lambda)}$ è costruita da $K(\lambda) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-\lambda^2 t^2)}}$.

Sia $(C)_{i,j} = \frac{u_i v_j^*}{a_i - b_j}$ reale, sia γ è il modulo del birapporto degli estremi degli intervalli che contengono u, v , allora

$$\sigma_{j+k}(C) \leq 4 \exp\left(\frac{\pi^2}{4\mu(1/\sqrt{\gamma})}\right)^{-2k} \sigma_j(C).$$



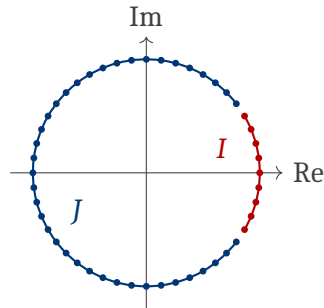
Blocchi fuori diagonale di matrici Cauchy generalizzate

2 Struttura di dislocamento e decadimento dei valori singolari

Teorema

Sia $C_{I,J}$ un blocco colonna di ampiezza $2 \leq m \leq \lceil n/2 \rceil$,
cioè $J = \{j, \dots, j + m - 1\}$ e $I = \{1, \dots, n\} \setminus J$. Allora:

$$\text{rank}_\varepsilon(C_{I,J}) \leq 2 \left\lceil \frac{2}{\pi^2} \log \left(\frac{4}{\varepsilon} \right) \log(4m) \right\rceil.$$





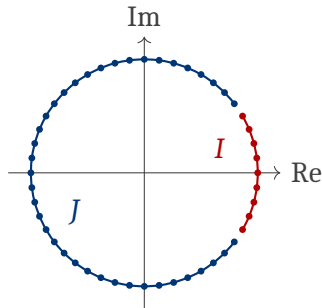
Blocchi fuori diagonale di matrici Cauchy generalizzate

2 Struttura di dislocamento e decadimento dei valori singolari

Teorema

Sia $C_{I,J}$ un blocco colonna di ampiezza $2 \leq m \leq \lceil n/2 \rceil$,
cioè $J = \{j, \dots, j + m - 1\}$ e $I = \{1, \dots, n\} \setminus J$. Allora:

$$\text{rank}_\varepsilon(C_{I,J}) \leq 2 \left\lceil \frac{2}{\pi^2} \log \left(\frac{4}{\varepsilon} \right) \log(4m) \right\rceil.$$



Ed ora? Questa stima ci permette di comprimere la matrice C in formato HSS.



Table of Contents

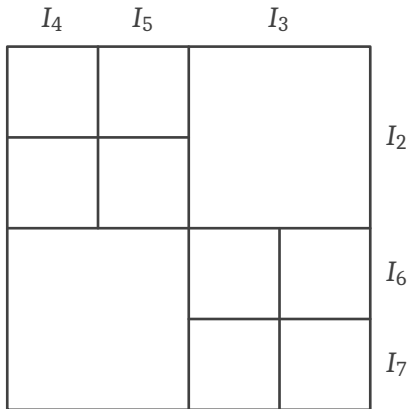
3 Matrici HSS e decomposizione telescopica

- ▶ Introduzione
- ▶ Struttura di dislocamento e decadimento dei valori singolari
- ▶ **Matrici HSS e decomposizione telescopica**
- ▶ Aggiornamenti di basso rango e funzioni di matrici HSS
- ▶ Esperimenti
- ▶ Conclusioni

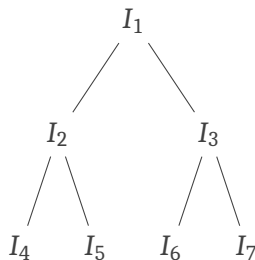


Matrici gerarchicamente semi-separabili (HSS)

3 Matrici HSS e decomposizione telescopica



Partizionamento gerarchico (*cluster tree*) $\mathcal{T}$ di profondità $L = 2$.



$$I_1 = [1, 2, \dots, n]$$

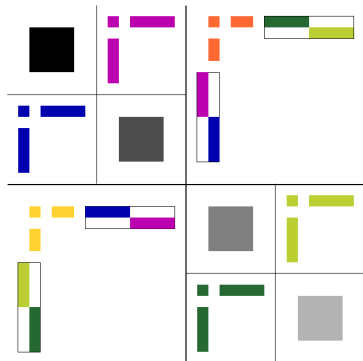
$$I_2 = [1, \dots, \lceil n/2 \rceil], I_3$$

$$I_4 = [1, \dots, \lceil n/4 \rceil], \dots, I_7$$



Matrici gerarchicamente semi-separabili (HSS)

3 Matrici HSS e decomposizione telescopica



1. Blocchi non diagonali di rango basso:

Per $\tau, \tau' \in \mathcal{T}$ fratelli,

$$A_{\tau, \tau'} = U_{\tau}^{\text{big}} \tilde{A}_{\tau, \tau'} (V_{\tau'}^{\text{big}})^*,$$

dove $U_{\tau}^{\text{big}}, V_{\tau'}^{\text{big}}$ hanno CON e $\tilde{A}_{\tau, \tau'} \in \mathbb{C}^{r \times r}$.

2. Struttura delle basi annidate:

Per $\tau \in \mathcal{T}$ non foglia con figli α e β ,

$$U_{\tau}^{\text{big}} = \begin{bmatrix} U_{\alpha}^{\text{big}} \\ U_{\beta}^{\text{big}} \end{bmatrix} U_{\tau}, \quad V_{\tau}^{\text{big}} = \begin{bmatrix} V_{\alpha}^{\text{big}} \\ V_{\beta}^{\text{big}} \end{bmatrix},$$

dove U_{τ}, V_{τ} hanno CON.



Caratteristiche del formato HSS

3 Matrici HSS e decomposizione telescopica

Vantaggiose:

- Complessità di memoria ridotta: $\mathcal{O}(nk)$ contro $\mathcal{O}(n^2)$;
- Prodotto matrice-vettore veloce: $\mathcal{O}(nk)$ contro $\mathcal{O}(n^2)$;
- Chiuse (in modo approssimato) rispetto a somma, prodotto e inversa.
- ULV in $\mathcal{O}(n)$ (SDP) e Cholesky in $\mathcal{O}(n^2)$ (S).



Caratteristiche del formato HSS

3 Matrici HSS e decomposizione telescopica

Vantaggiose:

- Complessità di memoria ridotta: $\mathcal{O}(nk)$ contro $\mathcal{O}(n^2)$;
- Prodotto matrice-vettore veloce: $\mathcal{O}(nk)$ contro $\mathcal{O}(n^2)$;
- Chiuse (in modo approssimato) rispetto a somma, prodotto e inversa.
- ULV in $\mathcal{O}(n)$ (SDP) e Cholesky in $\mathcal{O}(n^2)$ (S).

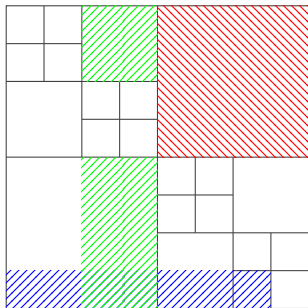
Tecniche: Per usufruirne, la rappresentazione HSS va costruita. Ci sono diversi metodi molto tecnici basati su ID, SVD, RRQR, fADI ecc... Non ne parleremo qui. Esistono compressioni $\mathcal{O}(n)$ che sono black-box!



Rappresentazione HSS approssimata di C

3 Matrici HSS e decomposizione telescopica

A ha rango HSS k sse ogni blocco riga/colonna HSS ha rango al più k .



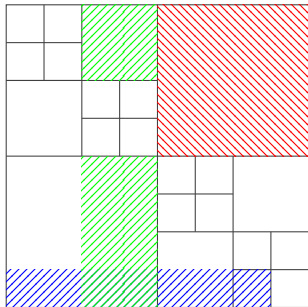


Rappresentazione HSS approssimata di C

3 Matrici HSS e decomposizione telescopica

A ha rango HSS k sse ogni blocco riga/colonna HSS ha rango al più k .

Sappiamo che i blocchi riga e colonna di C hanno basso rango numerico.

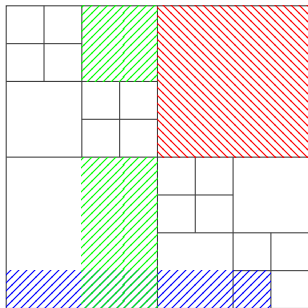




Rappresentazione HSS approssimata di C

3 Matrici HSS e decomposizione telescopica

A ha rango HSS k sse ogni blocco riga/colonna HSS ha rango al più k .



Sappiamo che i blocchi riga e colonna di C hanno basso rango numerico.

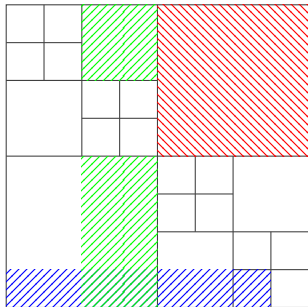
Una ε -HSS è HSS a meno di una perturbazione $\|\delta\|_2 \sim \varepsilon\sqrt{n}$



Rappresentazione HSS approssimata di C

3 Matrici HSS e decomposizione telescopica

A ha rango HSS k sse ogni blocco riga/colonna HSS ha rango al più k .



Sappiamo che i blocchi riga e colonna di C hanno basso rango numerico.

Una ε -HSS è HSS a meno di una perturbazione $\|\delta\|_2 \sim \varepsilon\sqrt{n}$

Corollario

$$\varphi(n, \varepsilon) = 2 \left\lceil \frac{2}{\pi^2} \log \left(\frac{4}{\varepsilon} \right) \log(2n) \right\rceil$$

è un limite superiore per l' ε -rango HSS di C .



Decomposizione telescopica e riduzione della profondità

3 Matrici HSS e decomposizione telescopica

Ogni matrice HSS può essere “pelata”, rimuovendo le basi livello per livello a partire dalle foglie.

$$A = \begin{bmatrix} A_{44} & & & \\ & A_{55} & & \\ & & A_{66} & \\ & & & A_{77} \end{bmatrix} + \begin{bmatrix} U_4 & & & \\ & U_5 & & \\ & & U_6 & \\ & & & U_7 \end{bmatrix} \left[\begin{array}{c|c} \tilde{A}_{45} & U_2 \tilde{A}_{23} V_3^* \\ \hline \tilde{A}_{54} & \tilde{A}_{67} \\ \hline U_3 \tilde{A}_{32} V_2^* & \tilde{A}_{76} \end{array} \right] \begin{bmatrix} V_4 & & & \\ & V_5 & & \\ & & V_6 & \\ & & & V_7 \end{bmatrix}^* .$$



Decomposizione telescopica e riduzione della profondità

3 Matrici HSS e decomposizione telescopica

Ogni matrice HSS può essere “pelata”, rimuovendo le basi livello per livello a partire dalle foglie.

$$A = \begin{bmatrix} A_{44} & & & \\ & A_{55} & & \\ & & A_{66} & \\ & & & A_{77} \end{bmatrix} + \begin{bmatrix} U_4 & & & \\ & U_5 & & \\ & & U_6 & \\ & & & U_7 \end{bmatrix} \left[\begin{array}{c|c} \tilde{A}_{45} & U_2 \tilde{A}_{23} V_3^* \\ \hline \tilde{A}_{54} & \tilde{A}_{67} \\ \hline U_3 \tilde{A}_{32} V_2^* & \tilde{A}_{76} \end{array} \right] \begin{bmatrix} V_4 & & & \\ & V_5 & & \\ & & V_6 & \\ & & & V_7 \end{bmatrix}^* .$$

Il fattore interno è HSS ma con un livello in meno. Si ripete fino alla radice.



Decomposizione telescopica e riduzione della profondità

3 Matrici HSS e decomposizione telescopica

Ogni matrice HSS può essere “pelata”, rimuovendo le basi livello per livello a partire dalle foglie.

$$A = \begin{bmatrix} A_{44} & & & \\ & A_{55} & & \\ & & A_{66} & \\ & & & A_{77} \end{bmatrix} + \begin{bmatrix} U_4 & & & \\ & U_5 & & \\ & & U_6 & \\ & & & U_7 \end{bmatrix} \left[\begin{array}{c|c} \tilde{A}_{45} & U_2 \tilde{A}_{23} V_3^* \\ \hline \tilde{A}_{54} & \tilde{A}_{67} \\ \hline U_3 \tilde{A}_{32} V_2^* & \tilde{A}_{76} \end{array} \right] \begin{bmatrix} V_4 & & & \\ & V_5 & & \\ & & V_6 & \\ & & & V_7 \end{bmatrix}^* .$$

Il fattore interno è HSS ma con un livello in meno. Si ripete fino alla radice.

Ed ora? il fattore interno ha dimensione $r 2^L \ll n$. Possiamo usarlo per calcolare $f(A)$ in modo efficiente.



Table of Contents

4 Aggiornamenti di basso rango e funzioni di matrici HSS

- ▶ Introduzione
- ▶ Struttura di dislocamento e decadimento dei valori singolari
- ▶ Matrici HSS e decomposizione telescopica
- ▶ Aggiornamenti di basso rango e funzioni di matrici HSS
- ▶ Esperimenti
- ▶ Conclusioni



Funzioni di un aggiornamento di basso rango

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Sia $B \in \mathbb{C}^{n \times n}$ e $\mathcal{E}$ con $\text{rango}(\mathcal{E}) \ll n$. Calcoliamo $f(B + \mathcal{E}) - f(B)$.

- Se $f(z) = z^{-1}$ e $\mathcal{E} = EF^*$, formula di Sherman–Morrison–Woodbury:

$$(B + EF^*)^{-1} - B^{-1} = -B^{-1}E(I_k + F^*B^{-1}E)^{-1}F^*B^{-1}.$$



Funzioni di un aggiornamento di basso rango

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Sia $B \in \mathbb{C}^{n \times n}$ e $\mathcal{E}$ con $\text{rango}(\mathcal{E}) \ll n$. Calcoliamo $f(B + \mathcal{E}) - f(B)$.

- Se $f(z) = z^{-1}$ e $\mathcal{E} = EF^*$, formula di Sherman–Morrison–Woodbury:

$$(B + EF^*)^{-1} - B^{-1} = -B^{-1}E(I_k + F^*B^{-1}E)^{-1}F^*B^{-1}.$$

- In generale, dobbiamo costruire una matrice più grande:

$$f\begin{pmatrix} B & \mathcal{E} \\ 0 & B + \mathcal{E} \end{pmatrix} = \begin{pmatrix} f(B) & f(B + \mathcal{E}) - f(B) \\ 0 & f(B + \mathcal{E}) \end{pmatrix}.$$



Funzioni di un aggiornamento di basso rango

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Sia $B \in \mathbb{C}^{n \times n}$ e $\mathcal{E}$ con $\text{rango}(\mathcal{E}) \ll n$. Calcoliamo $f(B + \mathcal{E}) - f(B)$.

- Se $f(z) = z^{-1}$ e $\mathcal{E} = EF^*$, formula di Sherman–Morrison–Woodbury:

$$(B + EF^*)^{-1} - B^{-1} = -B^{-1}E(I_k + F^*B^{-1}E)^{-1}F^*B^{-1}.$$

- In generale, dobbiamo costruire una matrice più grande:

$$f\begin{pmatrix} B & \mathcal{E} \\ 0 & B + \mathcal{E} \end{pmatrix} = \begin{pmatrix} f(B) & f(B + \mathcal{E}) - f(B) \\ 0 & f(B + \mathcal{E}) \end{pmatrix}.$$

La proiettiamo su sottospazi piccoli $\mathcal{W} \oplus \mathcal{Z} \subseteq \mathbb{C}^{2n}$ con basi W, Z :

$$f(B + \mathcal{E}) - f(B) \approx W \left[f \begin{pmatrix} W^*BW & W^*\mathcal{E}Z \\ 0 & Z^*BZ + Z^*\mathcal{E}Z \end{pmatrix} \right]_{12} Z^*.$$



Applicazione alla decomposizione telescopica

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A) = f(D + UA^{(L-1)}V^*)$ usiamo

$$f(A) \approx f(D) + W_U \begin{bmatrix} I & 0 \end{bmatrix} f \left(\begin{bmatrix} W_U^* D W_U & W_U^* U A^{(L-1)} V^* W_V \\ 0 & W_V^* A W_V \end{bmatrix} \right) \begin{bmatrix} 0 \\ I \end{bmatrix} W_V^*,$$

dove W_U, W_V sono basi di spazi di Krylov razionali a blocchi $\mathcal{Q}^\square(D, U, \xi) \oplus \mathcal{Q}^\square(D^*, V, \xi)$.
Sono matrici *diagonali a blocchi*!



Applicazione alla decomposizione telescopica

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A) = f(D + UA^{(L-1)}V^*)$ usiamo

$$f(A) \approx f(D) + W_U \begin{bmatrix} I & 0 \end{bmatrix} f \left(\begin{bmatrix} W_U^* D W_U & W_U^* U A^{(L-1)} V^* W_V \\ 0 & W_V^* A W_V \end{bmatrix} \right) \begin{bmatrix} 0 \\ I \end{bmatrix} W_V^*,$$

dove W_U, W_V sono basi di spazi di Krylov razionali a blocchi $\mathcal{Q}^\square(D, U, \xi) \oplus \mathcal{Q}^\square(D^*, V, \xi)$.
Sono matrici *diagonali a blocchi*!

$$\mathcal{Q}^\square(D, U, \xi) = \left\{ q(D)^{-1} \sum_{j=0}^{k-1} D^j U C_j : C_j \in \mathbb{C}^{b \times b} \right\}, \quad q(x) := \prod_{\substack{\xi \in \xi \\ \xi \neq \infty}} (x - \xi).$$



Ricerca della struttura telescopica

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per impostare una ricorsione, vogliamo che l'argomento di f sia in decomposizione telescopica:

$$\begin{bmatrix} W_U^* & \\ & W_V^* \end{bmatrix} \left(\begin{bmatrix} D & \\ & D \end{bmatrix} + \begin{bmatrix} U \\ U \end{bmatrix} A^{(L-1)} \begin{bmatrix} V \\ V \end{bmatrix}^* \right) \begin{bmatrix} W_U & \\ & W_V \end{bmatrix}.$$



Ricerca della struttura telescopica

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per impostare una ricorsione, vogliamo che l'argomento di f sia in decomposizione telescopica:

$$\begin{bmatrix} W_U^* & \\ & W_V^* \end{bmatrix} \left(\begin{bmatrix} D & \\ & D \end{bmatrix} + \begin{bmatrix} U \\ U \end{bmatrix} A^{(L-1)} \begin{bmatrix} V \\ V \end{bmatrix}^* \right) \begin{bmatrix} W_U & \\ & W_V \end{bmatrix}.$$

Lemma

Questa matrice ha una decomposizione telescopica se:

- il fattore interno è in decomposizione telescopica rispetto a un albero $\mathcal{T}'$;
- i fattori esterni sono diagonali a blocchi e compatibili con $\mathcal{T}'$.

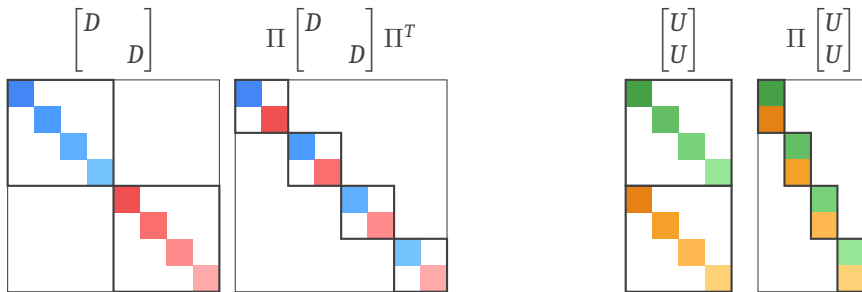
Ciò è possibile *a meno di permutare righe e colonne*.



Interlacciamento dei blocchi

4 Aggiornamenti di basso rango e funzioni di matrici HSS

$$A_{\text{aug}} = \tilde{\Pi} \begin{bmatrix} W_U^* & \\ & W_V^* \end{bmatrix} \Pi^T \left(\Pi \begin{bmatrix} D & \\ & D \end{bmatrix} \Pi^T + \Pi \begin{bmatrix} U \\ U \end{bmatrix} A^{(L-1)} \begin{bmatrix} V \end{bmatrix}^* \Pi^T \right) \Pi \begin{bmatrix} W_U & \\ & W_V \end{bmatrix} \tilde{\Pi}^T.$$





Decomposizione telescopica della matrice aumentata

4 Aggiornamenti di basso rango e funzioni di matrici HSS

$$A_{\text{aug}} = D_{\text{aug}} + U_{\text{aug}} A^{(L-2)} V_{\text{aug}}^*$$

Proposizione

- ha profondità $L - 1$ se A ha profondità L ;
- il livello più esterno è $A^{(L-1)}$ modificato da W_U e W_V ;
- i livelli interni restano invariati (decomposizione telescopica di $A^{(L-2)}$).



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));
2. **Parte diagonale:** calcola $f(D)$ direttamente $(f(D_i))$;



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));
2. **Parte diagonale:** calcola $f(D)$ direttamente $(f(D_i))$;
3. **Matrice aumentata:** costruisci la decomposizione telescopica di A_{aug} ;



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));
2. **Parte diagonale:** calcola $f(D)$ direttamente ($f(D_i)$);
3. **Matrice aumentata:** costruisci la decomposizione telescopica di A_{aug} ;
4. **Discesa di un livello:** itera su A_{aug} fino alla radice;



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));
2. **Parte diagonale:** calcola $f(D)$ direttamente $(f(D_i))$;
3. **Matrice aumentata:** costruisci la decomposizione telescopica di A_{aug} ;
4. **Discesa di un livello:** itera su A_{aug} fino alla radice;
5. **Permutazioni:** “taglia” il livello esterno con $[I \ 0] \tilde{\Pi}^T$ e $\tilde{\Pi} [0 \ I]^T$.



Schema dell'algoritmo

4 Aggiornamenti di basso rango e funzioni di matrici HSS

Per calcolare $f(A)$:

1. **Basi:** calcola W_U, W_V (Krylov razionale su ciascun blocco diagonale (D_i, U_i) (D_i^*, V_i));
2. **Parte diagonale:** calcola $f(D)$ direttamente $(f(D_i))$;
3. **Matrice aumentata:** costruisci la decomposizione telescopica di A_{aug} ;
4. **Discesa di un livello:** itera su A_{aug} fino alla radice;
5. **Permutazioni:** “taglia” il livello esterno con $[I \ 0] \tilde{\Pi}^T$ e $\tilde{\Pi} [0 \ I]^T$.

L'algoritmo fornisce una rappresentazione di $f(A)$ in formato telescopico.



Taglio del livello esterno

4 Aggiornamenti di basso rango e funzioni di matrici HSS

$$f(A) \approx f(D) + [I \ 0] \tilde{\Pi}^T \underbrace{\left(f(D_{\text{aug}}) + W_{U_{\text{aug}}} f(\cdots) W_{V_{\text{aug}}}^* \right)}_{\approx f(A_{\text{aug}})} \tilde{\Pi} [0 \ I]^T.$$



Taglio del livello esterno

4 Aggiornamenti di basso rango e funzioni di matrici HSS

$$f(A) \approx f(D) + [I \ 0] \tilde{\Pi}^T \underbrace{\left(f(D_{\text{aug}}) + W_{U_{\text{aug}}} f(\cdots) W_{V_{\text{aug}}}^* \right)}_{\approx f(A_{\text{aug}})} \tilde{\Pi} [0 \ I]^T.$$

Le matrici $[I \ 0] \tilde{\Pi}^T$ e $\tilde{\Pi} [0 \ I]^T$ sono diagonali a blocchi rispetto a $\mathcal{T}'$ e selezionano righe e colonne appropriate per ripristinare la dimensione originale.

$$[I \ 0] \tilde{\Pi}^T f(D_{\text{aug}}) \tilde{\Pi} [0 \ I]^T \rightsquigarrow \begin{bmatrix} I & 0 \\ & I & 0 \end{bmatrix} [f(D_{\text{aug}})]_{i,i} \begin{bmatrix} 0 \\ I \\ & 0 \\ I \end{bmatrix}$$

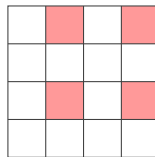




Table of Contents

5 Esperimenti

- ▶ Introduzione
- ▶ Struttura di dislocamento e decadimento dei valori singolari
- ▶ Matrici HSS e decomposizione telescopica
- ▶ Aggiornamenti di basso rango e funzioni di matrici HSS
- ▶ **Esperimenti**
- ▶ Conclusioni



Matrici HSS casuali: errore

5 Esperimenti

N	$\epsilon(\text{rel}, \ \cdot\ _F)$	HSS-rank($\exp(A)$)
128	4.6224×10^{-15}	2
256	6.4653×10^{-15}	4
512	3.0635×10^{-14}	5
1024	1.4482×10^{-13}	5
2048	5.6917×10^{-13}	5
4096	4.8544×10^{-12}	5
8192	5.6909×10^{-12}	5
16 384	1.4305×10^{-11}	5

Table: Errore di approssimazione relativo per il calcolo di $\exp(\text{HSS}(A))$ di A HSS di rango 1.



Matrici HSS casuali: tempi.

5 Esperimenti

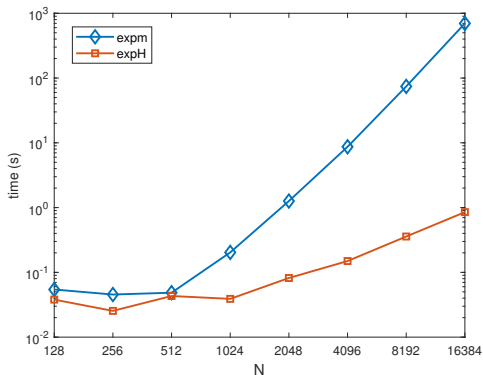


Figure: Tempi per rango 1 ed $n_{\min} = 64$.

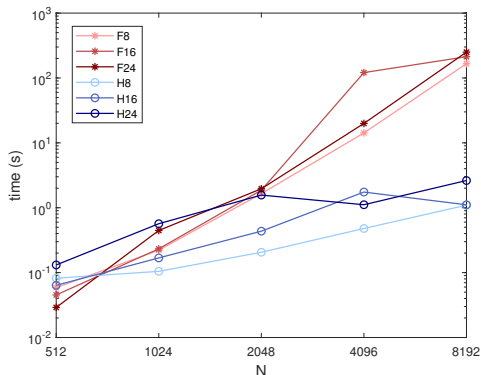


Figure: Tempi per rango variabile ed $n_{\min} = 256$.



Integratori esponenziali

5 Esperimenti

$$\begin{cases} u_t(x, t) = -u_{xx}(x, t) + 10 u_x(x, t), & x \in (0, 1), t > 0, \\ u(0, t) = 0, \quad u(1, t) = 0, & t > 0, \\ u(x, 0) = u_0(x), & x \in [0, 1]. \end{cases}$$



Integratori esponenziali

5 Esperimenti

$$\begin{cases} u_t(x, t) = -u_{xx}(x, t) + 10 u_x(x, t), & x \in (0, 1), t > 0, \\ u(0, t) = 0, \quad u(1, t) = 0, & t > 0, \\ u(x, 0) = u_0(x), & x \in [0, 1]. \end{cases}$$

Discretizzando il dominio spaziale e approssimando le derivate con differenze finite centrali otteniamo il sistema di ODE $U'(t) = A U(t)$ dove

$$A = (N + 1)^2 \text{ tridiag}(-1, 2, -1) + 5(N + 1) \text{ tridiag}(-1, 0, 1).$$



Integratori esponenziali

5 Esperimenti

$$\begin{cases} u_t(x, t) = -u_{xx}(x, t) + 10 u_x(x, t), & x \in (0, 1), t > 0, \\ u(0, t) = 0, \quad u(1, t) = 0, & t > 0, \\ u(x, 0) = u_0(x), & x \in [0, 1]. \end{cases}$$

Discretizzando il dominio spaziale e approssimando le derivate con differenze finite centrali otteniamo il sistema di ODE $U'(t) = A U(t)$ dove

$$A = (N + 1)^2 \text{ tridiag}(-1, 2, -1) + 5(N + 1) \text{ tridiag}(-1, 0, 1).$$

Con una griglia uniforme $t_n = n\Delta t$, $U^{n+1} = e^{\Delta t A} U^n \approx \text{expmv}(\Delta t, A, U^n)$.



Integratori esponenziali

5 Esperimenti

$$\begin{cases} u_t(x, t) = -u_{xx}(x, t) + 10 u_x(x, t), & x \in (0, 1), t > 0, \\ u(0, t) = 0, \quad u(1, t) = 0, & t > 0, \\ u(x, 0) = u_0(x), & x \in [0, 1]. \end{cases}$$

Discretizzando il dominio spaziale e approssimando le derivate con differenze finite centrali otteniamo il sistema di ODE $U'(t) = A U(t)$ dove

$$A = (N + 1)^2 \text{ tridiag}(-1, 2, -1) + 5(N + 1) \text{ tridiag}(-1, 0, 1).$$

Con una griglia uniforme $t_n = n\Delta t$, $U^{n+1} = e^{\Delta t A} U^n \approx \text{expmv}(\Delta t, A, U^n)$.

La nostra alternativa è $U^{n+1} \approx \Omega \exp_{\text{hss}}(\tilde{A}) \Omega^* U^n$ dove $\tilde{A} \approx \Omega^* A \Omega$ approssimazione HSS.

Costo totale: $\mathcal{O}(n \log(n)) + \mathcal{O}(nk) + \mathcal{O}(n \log(n))$ più la compressione T_c (una volta sola).



expmv vs exp(HSS)

5 Esperimenti

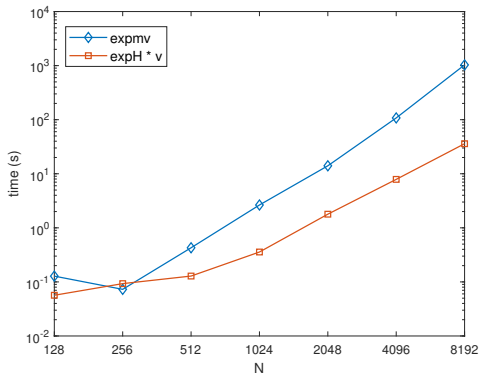


Figure: Tempo per 100 mat-vects $\exp(tA)v$ con $t = 1e - 2$. Consideriamo anche T_c .



Table of Contents

6 Conclusioni

- ▶ Introduzione
- ▶ Struttura di dislocamento e decadimento dei valori singolari
- ▶ Matrici HSS e decomposizione telescopica
- ▶ Aggiornamenti di basso rango e funzioni di matrici HSS
- ▶ Esperimenti
- ▶ Conclusioni



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;
- mostrato come approssimare tale matrice in formato HSS;



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;
- mostrato come approssimare tale matrice in formato HSS;
- descritto un algoritmo per il calcolo di funzioni di matrici telescopiche.



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;
- mostrato come approssimare tale matrice in formato HSS;
- descritto un algoritmo per il calcolo di funzioni di matrici telescopiche.

Per il futuro c'è ancora molto da fare:

- migliorare l'implementazione: scelta dei poli ottimi con AAA e numerical range Toeplitz;



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;
- mostrato come approssimare tale matrice in formato HSS;
- descritto un algoritmo per il calcolo di funzioni di matrici telescopiche.

Per il futuro c'è ancora molto da fare:

- migliorare l'implementazione: scelta dei poli ottimi con AAA e numerical range Toeplitz;
- studiare la stabilità del metodo: stima dell'errore con Crouzeix-Palencia bilineare iterato;



Riepilogo e sviluppi futuri

6 Conclusioni

In questa presentazione abbiamo:

- trasformato una matrice di Toeplitz in una matrice di Cauchy generalizzata;
- mostrato come approssimare tale matrice in formato HSS;
- descritto un algoritmo per il calcolo di funzioni di matrici telescopiche.

Per il futuro c'è ancora molto da fare:

- migliorare l'implementazione: scelta dei poli ottimi con AAA e numerical range Toeplitz;
- studiare la stabilità del metodo: stima dell'errore con Crouzeix-Palencia bilineare iterato;
- estendere questa costruzione: famiglie a basso rango di dislocamento.



Funzioni di matrici di Toeplitz

Grazie dell'attenzione!

Domande?