Matrici di Bassa Complessità Computazionale ed Applicazioni

Lucian Vespan
Relatore: Prof. Carmine Di Fiore
Macroarea di Scienze Matematiche Fisiche e Naturali
Corso di Laurea Specialistica in Matematica

23/07/2025

Matrici strutturate considerate nella tesi

Contesto: La tesi analizza matrici che presentano una struttura interna ricorrente e regolare, che consente una riduzione della complessità computazionale.

- **Matrici di Toeplitz**: definite dalla condizione $T_{ij} = (t_{i-j})_{i,j=1}^{n}$
- **Matrici triangolari di Toeplitz e matrici ϵ-Circolanti** Le matrici Circolanti sono definite cosi: $C_{k,j} = c_{(j-k) mod n}$
- **Matrici τ, matrici di Hartley**
- **Matrici di Hessenberg**: superiore o inferiore è una matrice quadrata "quasi" triangolare: ha nulli tutti gli elementi a_{ij} con rispettivamente $i > j + 1$ oppure $j > i + 1$.
- **Matrici di Hankel e Hilbert**: definite dalla condizione $H(i,j) = c_{i+j-1}$ e $H_{ij} = \frac{1}{i+j-1} \qquad i,j = 1 \cdots n$
- **Matrici di Tartaglia**: triangolari superiori definite cosi: $P_{ij} = \begin{pmatrix} j-1 \\ i-1 \end{pmatrix}$ per $i \leq j$ e $P_{ij} = 0$ altrimenti, per $i,j = 1,\ldots,s$.

STRUTTURA $\rightarrow$ Bassa Complessità Computazionale

CASO TOEPLITZ
Vediamo un modo per calcolare il prodotto di una matrice di Toeplitz T $n \times n$ per un vettore in al più $O(n \log n)$ operazioni aritmetiche.
Si usa il fatto che ogni matrice di Toeplitz può essere immersa in una matrice circolante che è diagonalizzabile dalla Fast Fourier Trasform (FFT).

Esempio per $n = 4$

Si consideri una generica matrice di Toeplitz T 4×4 ed un vettore $\mathbf{v}$ 4×1. Allora T può essere vista come la sottomatrice in alto a sinistra di una matrice circolante C 8×8, e per il vettore $T\mathbf{v}$ vale la seguente rappresentazione:

$$T\mathbf{v} = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & t_{-3} \\ t_1 & t_0 & t_{-1} & t_{-2} \\ t_2 & t_1 & t_0 & t_{-1} \\ t_3 & t_2 & t_1 & t_0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix} = \{C \begin{bmatrix} \mathbf{v} \\ \mathbf{0} \end{bmatrix}\}_4$$

dove

$$C = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & t_{-3} & 0 & t_3 & t_2 & t_1 \\ t_1 & t_0 & t_{-1} & t_{-2} & t_{-3} & 0 & t_3 & t_2 \\ t_2 & t_1 & t_0 & t_{-1} & t_{-2} & t_{-3} & 0 & t_3 \\ t_3 & t_2 & t_1 & t_0 & t_{-1} & t_{-2} & t_{-3} & 0 \\ 0 & t_3 & t_2 & t_1 & t_0 & t_{-1} & t_{-2} & t_{-3} \\ t_{-3} & 0 & t_3 & t_2 & t_1 & t_0 & t_{-1} & t_{-2} \\ t_{-2} & t_{-3} & 0 & t_3 & t_2 & t_1 & t_0 & t_{-1} \\ t_{-1} & t_{-2} & t_{-3} & 0 & t_3 & t_2 & t_1 & t_0 \end{bmatrix}$$

e con il simbolo $\{\mathbf{z}\}_4$ si intende il vettore le cui componenti sono le prime quattro componenti del vettore $\mathbf{z}$.

Per n generico:
Se T è $n \times n$ e $\mathbf{v}$ è $n \times 1$, allora l'osservazione vale ancora e può essere generalizzata, infatti
$T\mathbf{v} = \{C \begin{bmatrix} \mathbf{v} \\ \mathbf{0}_n \end{bmatrix}\}_n$, dove $C = \mathcal{C}(\mathbf{a})$ è la matrice circolante con prima riga
$\mathbf{a}^T = [t_0 \, t_{-1} \, \cdots \, t_{-n+1} \, 0 \, t_{n-1} \, \cdots \, t_1]$
Se n è una potenza di 2, sfruttando la rappresentazione spettrale di $\mathcal{C}(\mathbf{a})$,

$$\mathcal{C}(\mathbf{a}) = \sqrt{2n} \cdot F_{2n} \cdot d(F_{2n}\mathbf{a}) \cdot F_{2n}^H,$$

e le proprietà della matrice di Fourier F_{2n}, si deduce immediatamente una procedura di costo $O(n \log_2 n)$ per il calcolo del prodotto di una matrice di Toeplitz $n \times n$ per un vettore. Tale procedura ha come sotto-procedura l'algoritmo FFT.

APPLICAZIONI

Le matrici di bassa complessità computazionale possono essere utilizzate(in particolare):

- In espressioni dell'inversa di A $n \times n$ di tipo Toeplitz+Hankel, che consentono la risoluzione di $A\mathbf{x} = \mathbf{b}$ con metodi diretti di costo computazionale minore di $O(n^3)$.
- Nel rendere più efficienti metodi iterativi per risolvere gli stessi sistemi con A definita positiva (GC), o per minimizzare funzioni generiche $f : \mathrm{R}^n \to \mathrm{R}$ quando n è grande (BFGS);
- Nel definire algoritmi alternativi per il calcolo dei numeri di Bernoulli;
- **Nell'analizzare come partendo dalla forma canonica di Jordan di A sia possibile costruire la forma canonica di Jordan dell'inversa di A.**
  A tal scopo, come vedremo ora, si usano matrici triangolari di Toeplitz e di Tartaglia.

Forma canonica di Jordan

Definizione: Per una matrice $A \in \mathbb{C}^{n \times n}$ esiste una matrice invertibile χ tale che:

$$\chi^{-1} A \chi = J$$

con J la forma canonica di Jordan di A.

Struttura dei blocchi di Jordan:

$$J_s(\mu) = \mu I_s + Z_s^T$$

dove:

- μ autovalore di A
- I_s matrice identità $s \times s$
- Z_s^T matrice con 1 sulla sovradiagonale e 0 altrove

Forma canonica di Jordan dell'inversa: il problema

Sia $A \in \mathbb{C}^{n \times n}$ una matrice invertibile per cui si conosce una coppia di Jordan (χ, J) tale che $\chi^{-1} A \chi = J$. Vogliamo determinare una coppia di Jordan $(\tilde{\chi}, \tilde{J})$ per A^{-1}:

$$\tilde{\chi}^{-1} A^{-1} \tilde{\chi} = \tilde{J}$$

Osservazioni:

- La struttura di $\tilde{J}$ è immediata: se J contiene blocchi $\mu I_s + Z_s^T$, allora $\tilde{J}$ conterrà blocchi $\frac{1}{\mu} I_s + Z_s^T$.
- La matrice $\tilde{\chi}$ invece non è ovvia, a meno che A sia diagonalizzabile (in tal caso $\tilde{\chi} = \chi$).

Soluzione del problema: $\underline{\tilde{\chi} = \chi \mathcal{M}}$

Dove $\mathcal{M}$ è una matrice diagonale a blocchi (con lo stesso schema di J_A e $J_{A^{-1}}$), i cui blocchi diagonali M sono matrici *triangolari superiori di tipo Tartaglia-Toeplitz*.

Risultati preliminari

Prima identità:

$$(I + Z^T)P(I - Z^T) = P$$

Dove P è la matrice di Tartaglia, triangolare superiore, con elementi binomiali:

$$P_{ij} = \binom{j-1}{i-1} \text{ per } i \leq j, \quad P_{ij} = 0 \text{ altrimenti}$$

Seconda identità:

$$d(\mu^j)(I + Z^T) = (I + \frac{1}{\mu}Z^T)d(\mu^j)$$

Terza identità:

$$(I - Z^T)d((-\mu)^j) = d((-\mu)^j)(I + \mu Z^T)$$

con: $d(\mu^j) = d_s(\mu^j) = diag(\mu^j, j = 0, \ldots, s-1)$ e $\mu \neq 0$

Dalle precedenti tre identità matriciali segue che:

$$d(\mu^j)Pd((-\mu)^j) = d(\mu^j)(I + Z^T)P(I - Z^T)d((-\mu)^j)$$
$$= (I + \tfrac{1}{\mu}Z^T)\, d(\mu^j)Pd((-\mu)^j)\, (I + \mu Z^T)$$

cioè, se $\mu \in \mathbb{C}$, $\mu \neq 0$, ponendo:

$$M = d(\mu^j)Pd((-\mu)^j)$$

si ha:

$$M = (I + \tfrac{1}{\mu}Z^T)M(I + \mu Z^T) \qquad (\star)$$

Una M con la proprietà $(\star)$ è proprio la trasformazione che ci permette di ottenere $\tilde{\chi}$ da χ:

Lo spazio $\mathcal{L}_s^\mu$ associato al blocco di Jordan $J_s(\mu)$

Siano $A \in \mathbb{C}^{n\times n}$, $X \in \mathbb{C}^{n\times s}$, e $\mu \in \mathbb{C}$ con $\mu \neq 0$ tali che: $AX = X(\mu I_s + Z_s^T)$
Allora:

$$A^{-1}X = X(\mu I_s + Z_s^T)^{-1}$$

e se esiste $M \in \mathbb{C}^{s\times s}$ tale che: $(\mu I_s + Z_s^T)^{-1}M = M\left(\frac{1}{\mu}I_s + Z_s^T\right)$
allora:

$$A^{-1}\widetilde{X} = \widetilde{X}\left(\frac{1}{\mu}I_s + Z_s^T\right), \quad \text{dove } \widetilde{X} = XM$$

Si può definire lo spazio vettoriale:

$$\mathcal{L} = \mathcal{L}_s^\mu = \left\{M \in \mathbb{C}^{s\times s} \,:\, (\mu I_s + Z_s^T)^{-1}M = M\left(\frac{1}{\mu}I_s + Z_s^T\right)\right\}$$

in cui M viene scritta in termini della matrice di Tartaglia e delle matrici triangolari di Toeplitz. Vedremo come:

Risultati su $\mathcal{L} = \mathcal{L}_s^\mu$

Si ottengono i seguenti risultati fondamentali:

(a) La matrice $\Gamma = \Gamma_s = d_s(\mu^j)P_s d_s((-\mu)^j)$ appartiene a $\mathcal{L}$ con $e_1^T \Gamma_s = [1 \mid -\mu \mid \mu^2 \mid -\mu^3 \mid \cdots \mid (-\mu)^{s-1}]$

(b) Sia $\Gamma_{s-1} \in \mathcal{L}_{s-1}^\mu$. Allora la matrice a blocchi

$$G = G_s := \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & -\mu^2 \Gamma_{s-1} \end{bmatrix} \in \mathcal{L}_s^\mu$$

(c) Per ogni matrice $T \in \mathbb{C}^{s \times s}$ triangolare superiore di Toeplitz, si ha $GT \in \mathcal{L}_s^\mu$.

(d) La matrice $M \in \mathcal{L}$ è esprimibile come:

$$M = aG + bGZ^T + cG(Z^T)^2 + \cdots + G(Z^T)^{s-1}$$

Caratterizzazione di $\mathcal{L}$ usando le matrici triangolari di Toeplitz e di Tartaglia

Sia $M \in \mathcal{L}$. Se $L \in \mathcal{L}$ è invertibile, allora:

$$L^{-1}M = (I + \mu Z^T)^{-1} L^{-1} M (I + \mu Z^T)$$

cioè $L^{-1}M$ commuta con $I + \mu Z^T$, e pertanto è triangolare superiore di Toeplitz.

Segue che ogni $M \in \mathcal{L}$ appartiene allo spazio:

$$\mathcal{L}_s^\mu = \{LT \mid T \text{ triangolari superiore di Toeplitz, } L \in \mathcal{L} \text{ fissata invertibile}\}$$

Si ottiene che:

$$\mathcal{L} = \mathcal{L}_s^\mu = \text{span}\left\{G, GZ^T, G(Z^T)^2, \dots, G(Z^T)^{s-1}\right\}$$

dove ogni matrice di $\mathcal{L}$ è completamente determinata dalla sua prima riga.

OSSERVAZIONE: Costruzione di Γ in funzione di G

Imponendo la condizione $\Gamma = GT$:

$$[1 \mid -\mu \mid \mu^2 \mid -\mu^3 \mid \cdots \mid (-\mu)^{s-1}] = e_1^T \Gamma = e_1^T GT = e_1^T T$$

si ottiene:

$$T = T((-\mu)^j) = (I + \mu Z^T)^{-1}$$

Dunque si ha:

$$\Gamma = G(I + \mu Z^T)^{-1} \qquad (*)$$

Ricordando che Γ è definita in termini di P_s (matrice triangolare superiore di Tartaglia d'ordine s), mentre G è definita in termini di P_{s-1}, l'uguaglianza $(*)$ stabilisce una relazione tra matrici Tartaglia di ordine consecutivo ottenendo così una analoga relazione anche rispetto alle loro inverse,

Matrice P_k

$$P_k = \begin{bmatrix} 1 & [1 \quad 1 \quad \cdots \quad 1] \\ \underline{0} & P_{k-1} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \end{bmatrix}$$

Inversa P_k^{-1}

$$P_k^{-1} = \begin{bmatrix} 1 & -\,[1 \quad 0 \quad \cdots \quad 0]\; P_{k-1}^{-1} \\ \underline{0} & \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} P_{k-1}^{-1} \end{bmatrix} \qquad (\#)$$

Applicazione: Si hanno vantaggi nella risoluzione di sistemi $P_s\mathbf{x} = \mathbf{b}$ via $\mathbf{x} = P_s^{-1}\mathbf{b}$. Infatti applicando la formula $(\#)$ per $k = s, s-1, \cdots, 2$, si ottiene un algoritmo per il calcolo di $P_s^{-1}\mathbf{b}$ che richiede soltanto operazioni additive. Quindi:

- Nessun calcolo degli elementi di P_s^{-1}.
- Solo $s(s-1)/2$ operazioni additive.

Conclusioni

- Abbiamo descritto come ottenere la forma canonica di Jordan per A^{-1} partendo da quella di A, utilizzando una trasformata M.
- Lo spazio $\mathcal{L}$ delle matrici M è stato completamente caratterizzato .
- L'inversa della matrice di Tartaglia può essere applicata con sole $s(s-1)/2$ operazioni additive, evitando calcoli espliciti dei suoi elementi.