Proiezioni di matrici stocastiche su algebre di matrici

C. Di Fiore, F. Tudisco, P. Zellini

Università di Roma, "Tor Vergata"

Genova, 16-17 Febbraio 2012

Algebre di matrici doppiamente stocastiche

- $U \in U_n(\mathbb{C})$ matrice unitaria
- $\boldsymbol{v} \in \mathbb{C}^n$ tale che $(\boldsymbol{v}^T U)_i \neq 0$ per ogni $i = 1, \ldots, n$
- $\mathcal{L} = \{U \operatorname{diag}(\theta) U^* \mid \theta \in \mathbb{C}^n\}$
- $z =$ vettore $\quad z \mapsto \mathcal{L}(z) =$matrice di $\mathcal{L}$ tale che $\boldsymbol{v}^T \mathcal{L}(z) = z^T$
- $A =$ matrice $\quad A \mapsto \mathcal{L}_A =$proiezione di A su $\mathcal{L}$
  $\mathcal{L}_A = \arg\min_{X \in \mathcal{L}} \|A - X\|_F = U \operatorname{diag}(U^* A U) U^*$

A è (doppiamente) p-stocastica se $(A\mathbf{e} =)\, A^T\mathbf{e} = p\mathbf{e}$
A è stocastica se $p = 1$.

Teorema

$\zeta_n \in \mathbb{C}$ tale che $|\zeta_n| = n^{-\frac{1}{2}}$
Se esiste j tale che $U\mathbf{e}_j = \zeta_n \mathbf{e}$ (U ha una colonna costante), allora le seguenti affermazioni sono equivalenti

1. $\mathbf{e}\mathbf{e}^T \in \mathcal{L}, \quad \mathbf{e} = (1, \dots, 1)^T$
2. $\mathcal{L}(z)$ è doppiamente $\left(\frac{\mathbf{z}^T\mathbf{e}}{\mathbf{v}^T\mathbf{e}}\right)$-stocastica
3. $\mathcal{L}_A$ è doppiamente $\left(\frac{1}{n}\mathbf{e}^T A\mathbf{e}\right)$-stocastica

In particolare **se** A **oppure** A^T **è** p**-stocastica**
$\implies \mathcal{L}_A$ **è doppiamente** p**-stocastica**

Circolanti

$U = F =$matrice di Fourier $\qquad \mathcal{L} = \mathcal{C}$
$Fe_1 = n^{-\frac{1}{2}} \boldsymbol{e}$ ovvero la prima colonna di F è costante.

$\mathcal{C} = \{f(\Pi) \mid f \textit{ polinomio}\}$

$$\Pi = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ 1 & & & \end{pmatrix}$$

$\{J_1, \ldots, J_n\}$, $J_k = \Pi^{k-1}$, è una base ortogonale fatta di zeri e uni

Calcolare $\mathcal{C}(A)$ richiede solo operazioni additive fra gli elementi di A

η

$\mathcal{L} = \eta = \mathcal{C}^s + J\mathcal{C}^s \quad \mathcal{C}^s =$ circolanti simmetriche

$U\boldsymbol{e}_1 = U_\eta \boldsymbol{e}_1 = n^{-\frac{1}{2}}\boldsymbol{e}$ ovvero la prima colonna di U_η è costante

$$J_1 = \begin{pmatrix} 1 & 1 & & \\ 1 & 1 & & \\ & & 1 & 1 \\ & & 1 & 1 \end{pmatrix}, \; J_2 = \begin{pmatrix} & & 1 & 1 \\ & & 1 & 1 \\ 1 & 1 & & \\ 1 & 1 & & \end{pmatrix}$$

$$J_3 = \begin{pmatrix} 1 & -1 & & \\ -1 & 1 & & \\ & & 1 & -1 \\ & & -1 & 1 \end{pmatrix}, \; J_4 = \begin{pmatrix} & & 1 & -1 \\ & & -1 & 1 \\ 1 & -1 & & \\ -1 & 1 & & \end{pmatrix}$$

$J_i \in M_n(\{-1, 0, 1\})$ è una base ortogonale per η

Calcolare $\eta(A)$ richiede solo operazioni additive fra gli elementi di A

Haar

U =trasformata di Haar $\quad n = 2^m$

$$U_8 = \left(\begin{array}{cccc|cccc}
1 & & & & 1 & & 1 & 1 \\
1 & & & & 1 & & 1 & -1 \\
1 & & & & 1 & 1 & -1 & \\
1 & & & & 1 & -1 & -1 & \\
\hline
1 & & 1 & 1 & -1 & & & \\
1 & & 1 & -1 & -1 & & & \\
1 & 1 & -1 & & -1 & & & \\
1 & -1 & -1 & & -1 & & &
\end{array}\right) D_8$$

La prima colonna di U_n è costante

Formalmente...

$$Q_1 = 1,\ Q_n = \begin{pmatrix} \mathbf{e}\mathbf{e}_1^T & Q_{\frac{n}{2}} \\ Q_{\frac{n}{2}} & -\mathbf{e}\mathbf{e}_1^T \end{pmatrix} \qquad \mathcal{L} = \mathcal{H}$$

$$D_1 = 1,\ D_n = \begin{pmatrix} D_{\frac{n}{2}} S_{\frac{n}{2}} & O \\ O & D_{\frac{n}{2}} S_{\frac{n}{2}} \end{pmatrix} \qquad S_n = \mathrm{diag}\left(\frac{1}{\sqrt{2}}, 1, \dots, 1\right)$$

$$U_n = Q_n D_n = \begin{pmatrix} \frac{1}{\sqrt{n}}\mathbf{e}\mathbf{e}_1^T & U_{\frac{n}{2}} S_{\frac{n}{2}} \\ U_{\frac{n}{2}} S_{\frac{n}{2}} & -\frac{1}{\sqrt{n}}\mathbf{e}\mathbf{e}_1^T \end{pmatrix} \qquad n = 2, 4, \dots, 2^m$$

Le matrici di rango uno

$$J_k^{(n)} = Q_n e_k e_k^T Q_n^T \quad k = 1, 2, \ldots, n$$

formano una base ortogonale per $\mathcal{H}$ di elementi in $M_n(\{-1, 0, 1\})$

$$J_1^{(4)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \; J_2^{(4)} = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix}$$

$$J_3^{(4)} = \begin{pmatrix} 1 & -1 & & \\ -1 & 1 & & \\ & & & \\ & & & \end{pmatrix}, \; J_4^{(4)} = \begin{pmatrix} & & & \\ & & & \\ & & 1 & -1 \\ & & -1 & 1 \end{pmatrix}$$

Calcolare $\mathcal{H}(A)$ richiede solo operazioni additive fra gli elementi di A

Ulteriori esempi

- U =matrice di Hadamard $\quad n = 2^{2m}$

$$U_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 & -1 \end{pmatrix} D_8$$

- $\mathcal{L}$ = Jacobi $= \{p(T) \mid p \text{ polinomio}\}$ $\qquad T = \begin{pmatrix} 1 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 1 \end{pmatrix}$

- $\mathcal{L}$ = Hartley $= \mathcal{C}^s + J\Pi\mathcal{C}^{sk}$, $\quad \mathcal{C}^{sk}$ =circolanti antisimmetriche

.....

Spazi $\mathcal{L}$ più generali

- $\mu \in M_n(\{0,1\})$ matrice di zeri e uni
- $\circ =$ prodotto di Hadamard
- $U \in U_n(\mathbb{C})$ matrice unitaria
- $\mathcal{L} = \{U(\mu \circ Z)U^* \mid Z \in M_n(\mathbb{C})\}$
- $A =$ matrice $\quad A \mapsto \mathcal{L}_A =$ proiezione di A su $\mathcal{L}$
  $\mathcal{L}_A = \arg\min_{X \in \mathcal{L}} \|A - X\|_F = U(\mu \circ U^*AU)U^*$
- Se $\mu = I$ allora le $\mathcal{L}$ sono le algebre simultaneamente diagonali viste prima

$\mu \circ A + \mu \circ B = \mu \circ (A + B)$, $\mu \circ \alpha A = \alpha\mu \circ A$, allora
$\mathcal{L}$ è spazio vettoriale per ogni scelta di μ

Teorema

Se esiste j tale che
$U\mathbf{e}_j = \zeta_n \mathbf{e}$ (U ha una colonna costante) e $\mu_{jj} = 1$, allora

$$A \quad p\text{-stocastica} \Longrightarrow \mathcal{L}_A \quad p\text{-stocastica}$$

Nota: basta che $\mu_{kj} \neq 0$ per qualche $k \neq j$ per assicurare che $\mathcal{L}_A$ non sia doppiamente p-stocastica, come invece succede quando $\mu = I$

In altre parole, proiezioni di matrici stocastiche su $\mathcal{L}$ rimangono stocastiche.

Dim. *Il fatto che $\mathbf{e}$ è punto fisso di A^T implica che $\mathbf{e}_j$ è punto fisso di $\mu \circ U^* A^T U$ e dunque $(\mathcal{L}_A)^T \mathbf{e}$ è costante.*

Osservazione

Se esiste $\alpha \geq 0$ tale che $\mu^2 \leq \alpha\mu$, allora $\mathcal{L}$ è un'algebra di matrici

Esempi:

$$\mu = I,\ \mu = I + J,\ \mu = \begin{pmatrix} 1 & 1 & \\ & 1 & \\ & 1 & 1 \end{pmatrix}$$

$$\mu = \begin{pmatrix} & & 1 \\ 1 & 1 & 1 \\ & & 1 \end{pmatrix},\ \mu = \begin{pmatrix} 1 & 1 & 1 \\ & & 1 \\ & & 1 \end{pmatrix},\ \mu = \begin{pmatrix} 1 & & \\ 1 & 1 & \\ 1 & & 1 \end{pmatrix}$$

Nota Se $I \leq \mu$, allora $I \in \mathcal{L}$

Questione Investigare sotto quali ipotesi si ha $\mathcal{L}^{-1} \subset \mathcal{L}$

Stima di $\lambda_2(A)$

Data $A \in M_n(\mathbb{C})$ si considerano i suoi n autovalori:

$$\rho(A) = |\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)|$$

Teorema

A è stocastica e $\exists\ k \in \mathbb{N}$ tale che $A^k \geq O$ allora $1 = \lambda_1(A) \in \sigma(A)$

Se inoltre A^k è irriducibile ($\Rightarrow A$ è irriducibile), allora
1 è autovalore semplice di A ed $\exists!\ \boldsymbol{x} > \boldsymbol{o}$ tale che $A\boldsymbol{x} = \boldsymbol{x}$

Dim. *Si applica il teorema di Perron-Frobenius ad A^k per derivare le proprietà su A*

Osservazione

Se $\{J_1, \ldots, J_d\}$ è una base ortogonale per $\mathcal{L}$ tale che $J_i \geq O$, allora

$$A \geq O \Longrightarrow \mathcal{L}_A \geq O$$

Inoltre se A è stocastica, $U\mathbf{e}_j = \zeta_n \mathbf{e}$ e $\mu_{jj} = 1$, allora

$$1 = \rho(A) = \rho(\mathcal{L}_A) \tag{1}$$

Esempio: $\mu = I$, $d = n$, $\mathcal{L} = \mathcal{C} =$ matrici circolanti.

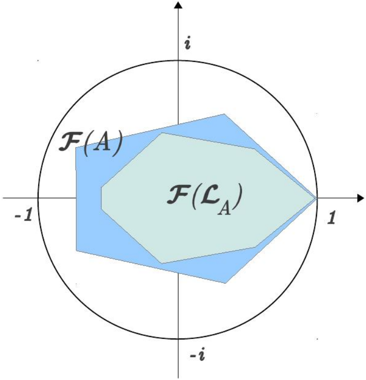
Questione: La (1) si ha più in generale quando $A^k \geq O$ implica $(\mathcal{L}_A)^m \geq O$. Investigare per quali $\mathcal{L}$ vale questa proprietà...

Quando $\mu = I$, indipendentemente dalla scelta di $U \in U_n(\mathbb{C})$, si ha

$$\{\mathbf{x}^* \mathcal{L}_A \mathbf{x} \mid \|\mathbf{x}\| = 1\} = \mathcal{F}(\mathcal{L}_A) \subset \mathcal{F}(A) = \{\mathbf{x}^* A \mathbf{x} \mid \|\mathbf{x}\| = 1\}$$

dunque, se U ha una colonna costante

$A \geq O$, stocastica, normale $\Longrightarrow$ $\mathcal{F}(\mathcal{L}_A) \subset \mathcal{F}(A) \subset \{|\zeta| \leq 1\}$, ovvero l'inviluppo convesso di $\sigma(\mathcal{L}_A)$ è contenuto in quello di $\sigma(A)$.



Questione È possibile mettere in relazione $|\lambda_2(A)|$ e $|\lambda_2(\mathcal{L}_A)|$?

$\gamma(A)$ ed il calcolo dell'autovettore di Perron

Sia $N = \{1, \dots, n\}$, $A \geq O$ stocastica, e

$$\gamma(A) = \min_{V \subset N, |V| = n-1} \left(\max_{j \in N} \sum_{i \in V} a_{ij} - \sum_{i \in V} \min_{j \in N} a_{ij} \right)$$

Osservazione

$0 \leq \gamma(A) \leq 1$

$\gamma(A) = 0 \iff A = \frac{1}{n} \boldsymbol{e}\boldsymbol{e}^T$

Dim.

$\sum_i \min_j a_{ij} \leq \sum_i a_{ij} \implies \gamma(A) \geq 0$,
$\sum_{i \in V} a_{ij} \leq 1 \implies \gamma(A) \leq 1$

$\max_j \sum_i a_{ij} = \sum_i a_{ij_0}$, $\gamma(A) = 0 \implies$
$\sum_i \min_j a_{ij} = \sum_i a_{ij_0} \implies a_{ij_0} = \min_j a_{ij} \ \forall i$

Teorema

Sia $A \geq O$ stocastica, allora $|\lambda_2(A)| \leq \gamma(A)$

Congettura: $\gamma(A) \leq \tau_1(A)$ (lo abbiamo osservato per $n = 2, 3$)

$\tau_1 = \tau_{\|\cdot\|_1}$ = coefficiente di ergodicità [Seneta 1979, Rothblum 1985]

Teorema (?)

Sia $A \geq O$ stocastica, allora $|\lambda_2(A)| = \lim_{k\to\infty} \left(\gamma(A^k)\right)^{\frac{1}{k}}$
*Lo stiamo verificando....*verificato l'altro ieri!

Corollario

Se esiste k tale che $\gamma(A^k) < 1$, allora $1 = \lambda_1(A)$ è semplice e tutti gli altri autovalori hanno modulo strettamente minore di 1.

Teorema

Sia $A \geq O$, stocastica tale che $\gamma(A) < 1$.
L'autovettore $\boldsymbol{x} \geq \boldsymbol{o}$, $\|\boldsymbol{x}\| = 1$ di A relativo a 1 è soluzione del sistema lineare

$$B(A)\,\boldsymbol{x} = (I - \gamma(A)B)\boldsymbol{x} = \boldsymbol{y} \tag{LS}$$

dove B e $\boldsymbol{y}$ si esprimono esplicitamente in termini di a_{ij} ed inoltre $B \geq O$, stocastica

Il calcolo dell'autovettore dominante di A si può ricondurre alla risoluzione di un sistema lineare

Precondizionare con matrici di $\mathcal{L}$

Metodo di Eulero-Richardson precondizionato applicato ad (LS)

$$\mathbf{x}_{k+1} = \mathbf{x}_k + Q^{-1}(\mathbf{y} - B(A)\mathbf{x}_k), \quad k = 0, 1, 2, \dots$$

Matrice di iterazione: $H_Q = I - Q^{-1}B(A) = I - Q^{-1}(I - \gamma(A)B)$

Osservazione

Sia $U\mathbf{e}_1 = \zeta_n \mathbf{e}$ e $\mu_{11} = 1$. Se $\mathcal{L}_{B(A)}$ è invertibile, la scelta $Q = \mathcal{L}_{B(A)}$ ha l'effetto di muovere un autovalore di $H_{\mathcal{L}_{B(A)}}$ in zero.

Questione: quale scelta di μ ($\mu_{11} = 1$) rende $\rho\left(H_{\mathcal{L}_{B(A)}}\right) < 1$?
quale rende $\rho\left(H_{\mathcal{L}_{B(A)}}\right)$ minimo?

Grazie per $\mathcal{L}$'attenzione!