

# Equazioni matriciali che dipendono da un parametro: proprietà di decadimento

Beatrice Meini  
in collaborazione con D. Bini e V. Ramaswami

Dipartimento di Matematica, Università di Pisa

Giornate di algebra lineare numerica, 2007

# Outline

- 1 Il problema
  - Motivazioni
  - Risultati noti
  - Problemi d'interesse
- 2 Nuovi risultati
  - Idea
  - Caso scalare
  - Caso generale
- 3 Ulteriori sviluppi

# Processi di Markov di rinnovo

Definiti dal kernel

$$K(x) = \begin{bmatrix} \tilde{B}_0(x) & \tilde{B}_1(x) & \tilde{B}_2(x) & \tilde{B}_3(x) & \dots \\ \tilde{A}_{-1}(x) & \tilde{A}_0(x) & \tilde{A}_1(x) & \tilde{A}_2(x) & \dots \\ & \tilde{A}_{-1}(x) & \tilde{A}_0(x) & \tilde{A}_1(x) & \ddots \\ & & \tilde{A}_{-1}(x) & \tilde{A}_0(x) & \ddots \\ 0 & & & \ddots & \ddots \end{bmatrix}, \quad x \geq 0,$$

con  $\tilde{A}_i(x)$ ,  $\tilde{B}_i(x)$  distribuzioni di probabilità su  $[0, \infty]$ .

Si definisce

$$\tilde{G}_i(x) = P\{\text{passare dal livello 1 al livello 0 in } i \text{ passi} \\ \text{e in un tempo } \leq x\}$$

# Processi di Markov di rinnovo

Si definiscono

$$A_i(s) = \int_0^\infty e^{-st} \tilde{A}_i(ds), \quad i \geq -1,$$

$$G(w, s) = \sum_{j=1}^{\infty} w^j \int_0^\infty e^{-sx} \tilde{G}_j(dx), \quad 0 \leq w \leq 1, s \geq 0$$

Vale:

$$G(w, s) = w \sum_{i=-1}^{\infty} A_i(s) G(w, s)^{i+1}, \quad |w| \leq 1, s \in \mathbb{C}^+$$

# Processi di Markov di rinnovo

Si definiscono

$$G(w) = G(w, 0), \quad |w| \leq 1,$$
$$A_i = A_i(0), \quad i \geq -1.$$

Proprietà:

- $A_i \geq 0$  per  $i \geq -1$ , e  $\sum_{i=-1}^{\infty} A_i$  è stocastica;
- $G(w)$  è soluzione di:

Equazione matriciale

$$G(w) = w \sum_{i=-1}^{\infty} A_i G(w)^{i+1}$$

# Processi di Markov di rinnovo

Si definiscono

$$G(w) = G(w, 0), \quad |w| \leq 1,$$
$$A_i = A_i(0), \quad i \geq -1.$$

Proprietà:

- $A_i \geq 0$  per  $i \geq -1$ , e  $\sum_{i=-1}^{\infty} A_i$  è stocastica;
- $G(w)$  è soluzione di:

Equazione matriciale

$$G(w) = w \sum_{i=-1}^{\infty} A_i G(w)^{i+1}$$

# Risultati noti

- Per ogni  $0 \leq w \leq 1$  la matrice  $G(w)$  è la minima soluzione non negativa di

$$X = w \sum_{i=-1}^{\infty} A_i X^{i+1}$$

- $G(w) = \sum_{j=0}^{\infty} w^j G_j$
- $G_j \geq 0$  per ogni  $j$
- $G(w)$  è analitica per  $|w| < 1$ , convergente per  $|w| = 1$ .

Interpretazione probabilistica

$G_j$  è la probabilità di passare dal livello 1 al livello 0 in  $j$  passi

# Risultati noti

- Per ogni  $0 \leq w \leq 1$  la matrice  $G(w)$  è la minima soluzione non negativa di

$$X = w \sum_{i=-1}^{\infty} A_i X^{i+1}$$

- $G(w) = \sum_{j=0}^{\infty} w^j G_j$
- $G_j \geq 0$  per ogni  $j$
- $G(w)$  è analitica per  $|w| < 1$ , convergente per  $|w| = 1$ .

Interpretazione probabilistica

$G_j$  è la probabilità di passare dal livello 1 al livello 0 in  $j$  passi

# Risultati noti

- Per ogni  $0 \leq w \leq 1$  la matrice  $G(w)$  è la minima soluzione non negativa di

$$X = w \sum_{i=-1}^{\infty} A_i X^{i+1}$$

- $G(w) = \sum_{j=0}^{\infty} w^j G_j$
- $G_j \geq 0$  per ogni  $j$
- $G(w)$  è analitica per  $|w| < 1$ , convergente per  $|w| = 1$ .

Interpretazione probabilistica

$G_j$  è la probabilità di passare dal livello 1 al livello 0 in  $j$  passi

# Risultati noti

- Per ogni  $0 \leq w \leq 1$  la matrice  $G(w)$  è la minima soluzione non negativa di

$$X = w \sum_{i=-1}^{\infty} A_i X^{i+1}$$

- $G(w) = \sum_{j=0}^{\infty} w^j G_j$
- $G_j \geq 0$  per ogni  $j$
- $G(w)$  è analitica per  $|w| < 1$ , convergente per  $|w| = 1$ .

Interpretazione probabilistica

$G_j$  è la probabilità di passare dal livello 1 al livello 0 in  $j$  passi

# Risultati noti

- Per ogni  $0 \leq w \leq 1$  la matrice  $G(w)$  è la minima soluzione non negativa di

$$X = w \sum_{i=-1}^{\infty} A_i X^{i+1}$$

- $G(w) = \sum_{j=0}^{\infty} w^j G_j$
- $G_j \geq 0$  per ogni  $j$
- $G(w)$  è analitica per  $|w| < 1$ , convergente per  $|w| = 1$ .

## Interpretazione probabilistica

$G_j$  è la probabilità di passare dal livello 1 al livello 0 in  $j$  passi

# Interesse

Studiare proprietà di decadimento dei coefficienti  $G_j$ , ossia:

Determinare  $\theta$  tale che  $\|G_j\| = O(\theta^j)$  per  $j \rightarrow \infty$

Si assume:

- $\sum_{i=-1}^{\infty} A_i$  irriducibile;
- il raggio di convergenza di  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$  sia  $r_a > 1$ .

# Interesse

Studiare proprietà di decadimento dei coefficienti  $G_j$ , ossia:

Determinare  $\theta$  tale che  $\|G_j\| = O(\theta^j)$  per  $j \rightarrow \infty$

Si assume:

- $\sum_{i=-1}^{\infty} A_i$  irriducibile;
- il raggio di convergenza di  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$  sia  $r_a > 1$ .

# Idea

Utilizzare il:

## Teorema (Stima di Cauchy)

Sia  $R$  il raggio di convergenza della serie  $f(w) = \sum_{j=0}^{\infty} w^j f_j$ .  
Allora per ogni  $0 < r < R$  e per ogni  $j$  vale

$$|f_j| \leq \frac{\mu(r)}{r^j}$$

dove  $\mu(r) = \sup_{|w|=r} |f(w)|$ .

Problema: determinare il raggio di convergenza di  $G(w)$

# Idea

Utilizzare il:

## Teorema (Stima di Cauchy)

Sia  $R$  il raggio di convergenza della serie  $f(w) = \sum_{j=0}^{\infty} w^j f_j$ .  
Allora per ogni  $0 < r < R$  e per ogni  $j$  vale

$$|f_j| \leq \frac{\mu(r)}{r^j}$$

dove  $\mu(r) = \sup_{|w|=r} |f(w)|$ .

Problema: determinare il raggio di convergenza di  $G(w)$

# Un esempio scalare

Siano

$$A_i = e^{-\lambda} \frac{\lambda^{i+1}}{(i+1)!}, \quad i \geq -1$$

con  $\lambda > 0$  fissato.

Vale

$$G(w) = \sum_{j=1}^{\infty} w^j e^{-j(\lambda-1)} \lambda^{j-1}$$

Il raggio di convergenza è  $R = (\lambda e^{-(\lambda-1)})^{-1}$

Come possiamo stimare  $R$  senza conoscere  $G(w)$ ?

# Un esempio scalare

Siano

$$A_i = e^{-\lambda} \frac{\lambda^{i+1}}{(i+1)!}, \quad i \geq -1$$

con  $\lambda > 0$  fissato.

Vale

$$G(w) = \sum_{j=1}^{\infty} w^j e^{-j(\lambda-1)} \lambda^{j-1}$$

Il raggio di convergenza è  $R = (\lambda e^{-(\lambda-1)})^{-1}$

Come possiamo stimare  $R$  senza conoscere  $G(w)$ ?

# Caso scalare

Siano  $A_i$  scalari e sia  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$ .

Si cercano condizioni su  $w$  affinché l'equazione

$$z = wA(z) \tag{1}$$

abbia una soluzione in  $(0, r_a)$ .

Valgono le seguenti proprietà:

- La funzione  $A(z)$  è convessa e crescente in  $(0, r_a)$ ;
- la funzione  $f_w(z) = wA(z) - z$  ha un minimo in  $\sigma \in (0, r_a)$ , dove  $\sigma$  risolve l'equazione  $wA'(\sigma) - 1 = 0$ ;
- l'equazione (1) ha una soluzione in  $(0, r_a)$  se e solo se  $f_w(\sigma) \leq 0$ .

# Caso scalare

Siano  $A_i$  scalari e sia  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$ .

Si cercano condizioni su  $w$  affinché l'equazione

$$z = wA(z) \tag{1}$$

abbia una soluzione in  $(0, r_a)$ .

Valgono le seguenti proprietà:

- La funzione  $A(z)$  è convessa e crescente in  $(0, r_a)$ ;
- la funzione  $f_w(z) = wA(z) - z$  ha un minimo in  $\sigma \in (0, r_a)$ , dove  $\sigma$  risolve l'equazione  $wA'(\sigma) - 1 = 0$ ;
- l'equazione (1) ha una soluzione in  $(0, r_a)$  se e solo se  $f_w(\sigma) \leq 0$ .

# Caso scalare

Siano  $A_i$  scalari e sia  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$ .

Si cercano condizioni su  $w$  affinché l'equazione

$$z = wA(z) \tag{1}$$

abbia una soluzione in  $(0, r_a)$ .

Valgono le seguenti proprietà:

- La funzione  $A(z)$  è convessa e crescente in  $(0, r_a)$ ;
- la funzione  $f_w(z) = wA(z) - z$  ha un minimo in  $\sigma \in (0, r_a)$ , dove  $\sigma$  risolve l'equazione  $wA'(\sigma) - 1 = 0$ ;
- l'equazione (1) ha una soluzione in  $(0, r_a)$  se e solo se  $f_w(\sigma) \leq 0$ .

# Caso scalare

Siano  $A_i$  scalari e sia  $A(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i$ .

Si cercano condizioni su  $w$  affinché l'equazione

$$z = wA(z) \tag{1}$$

abbia una soluzione in  $(0, r_a)$ .

Valgono le seguenti proprietà:

- La funzione  $A(z)$  è convessa e crescente in  $(0, r_a)$ ;
- la funzione  $f_w(z) = wA(z) - z$  ha un minimo in  $\sigma \in (0, r_a)$ , dove  $\sigma$  risolve l'equazione  $wA'(\sigma) - 1 = 0$ ;
- l'equazione (1) ha una soluzione in  $(0, r_a)$  se e solo se  $f_w(\sigma) \leq 0$ .

# Raggio di convergenza nel caso scalare

Il raggio di convergenza di  $G(w)$  è

$$R = \frac{1}{A'(\sigma)}$$

dove  $\sigma$  è la soluzione in  $(0, r_a)$  dell'equazione

$$A'(z)z = A(z)$$

Posso estendere queste considerazioni al caso a blocchi?

# Raggio di convergenza nel caso scalare

Il raggio di convergenza di  $G(w)$  è

$$R = \frac{1}{A'(\sigma)}$$

dove  $\sigma$  è la soluzione in  $(0, r_a)$  dell'equazione

$$A'(z)z = A(z)$$

Posso estendere queste considerazioni al caso a blocchi?

# Premesse

Se  $\mu_w$  e  $\mathbf{u}_w$  sono l'autovalore/autovettore di Perron di  $G(w)$ ,  
dalla

$$G(w)\mathbf{u}_w = w \sum_{i=-1}^{\infty} A_i G(w)^{i+1} \mathbf{u}_w$$

segue

$$\mu_w \mathbf{u}_w = w \left( \sum_{i=-1}^{\infty} A_i \mu_w^{i+1} \right) \mathbf{u}_w.$$

Dunque,  $\mu_w$  risolve l'equazione scalare

$$z = w\rho(A(z)). \quad (2)$$

Intuizione: il raggio di convergenza è l'estremo superiore dei  
 $w \in (0, r_a)$  per cui l'equazione (2) ha soluzione

# Premesse

Se  $\mu_w$  e  $\mathbf{u}_w$  sono l'autovalore/autovettore di Perron di  $G(w)$ ,  
dalla

$$G(w)\mathbf{u}_w = w \sum_{i=-1}^{\infty} A_i G(w)^{i+1} \mathbf{u}_w$$

segue

$$\mu_w \mathbf{u}_w = w \left( \sum_{i=-1}^{\infty} A_i \mu_w^{i+1} \right) \mathbf{u}_w.$$

Dunque,  $\mu_w$  risolve l'equazione scalare

$$z = w\rho(A(z)). \quad (2)$$

Intuizione: il raggio di convergenza è l'estremo superiore dei  
 $w \in (0, r_a)$  per cui l'equazione (2) ha soluzione

# Proprietà

Sia  $\theta(z) = \rho(A(z))$ .

È noto che:

- $\theta(z)$  è una funzione reale analitica per  $0 < z < r_a$ ;
- $\theta(z)$  è strettamente crescente per  $0 < z < r_a$ ;
- la funzione  $\log \theta(e^t)$  è convessa e crescente per  $-\infty < t < \log r_a$ .

# Risultato principale

Definiamo  $\phi_w(z) = w\theta(z)$ .

## Teorema

*Valgono i seguenti risultati:*

① *l'equazione*

$$\theta'(z)z = \theta(z) \tag{3}$$

*ha un'unica soluzione  $\sigma$  in  $(0, r_a)$ ;*

② *l'equazione  $\phi_w(z) = z$  ha una soluzione in  $(0, r_a)$  per ogni  $0 < w \leq 1/\theta'(\sigma)$ ;*

③ *il raggio di convergenza di  $G(w)$  è  $R = 1/\theta'(\sigma)$ .*

# Ulteriori sviluppi

- Algoritmi efficienti per il calcolo di  $R$ ;
- algoritmi efficienti per il calcolo di un numero arbitrario di coefficienti  $G_i$