UNIVERSITÀ DI PISA

DIPARTIMENTO DI MATEMATICA

Appunti del corso "Metodi numerici per catene di Markov"

ANNO ACCADEMICO 2017/2018

Indice

Introduzione

Appunti non ufficiali del corso "Metodi numerici per catene di Markov" tenuto dalla prof.ssa Beatrice Meini nell'anno accademico 2017/2018, le note sono state raccolte a lezione da C.Ginevra Biondi.

Lo scopo principale del corso e' studiare metodi numerici per il calcolo della distribuzione stazionaria di una catena di Markov. All'inizio del corso richiamiamo alcune definizioni e proprieta' delle catene di Markov. Successivamente studiamo proprieta' delle matrici nonnegative e metodi numerici per catene di Markov con un numero finito di stati. Nell'ultima parte del corso trattiamo il caso con un numero infinito di stati e con struttura M/G/1.

Capitolo 1

Nozioni preliminari

1.1 Generalità sulle catene di Markov

Definizione 1.1.1. Sia (E, T, P) uno spazio probabilizzato con T un insieme di tempi numerabile, $(X_n)_{n \in T}$ un processo stocastico è detto catena di Markov se è tale che

$$\mathbb{P}(X_{n+1} = j_{n+1} \mid X_0 = j_0, \ldots, X_n = j_n) = \mathbb{P}(X_{n+1} = j_{n+1} \mid X_n = j_n)$$

Notazione: $\mathbb{P}(X_n = j \mid X_{n-1} = i) := P_{ij}$

Notazione: A una matrice, diciamo che $A \geq 0$ se $a_{i,j} \geq 0 \; \forall i, j$

Definizione 1.1.2. Siano P_{ij} definiti come sopra, chiamiamo matrice di transizione associata alla catena di Markov $P := (P_{ij})_{ij}$
Osserviamo che certamente vale $0 \leq P_{ij} \leq 1$ e inoltre

$$\sum_{j \in E} P_{ij} = 1 \; \forall i$$

ovvero

$$P \begin{pmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{pmatrix}$$

Definizione 1.1.3. A una matrice si dice stocastica se

$$\sum_{j} A_{ij} = 1 \; \forall i$$

Proposizione 1.1.4. *Sia P la matrice di transizione associata ad una catena di Markov, allora*

$$\mathbb{P}(X_n = j \mid X_0 = i) = (P_{ij})^n$$

Dimostrazione. Diamo un accenno del perchè vale questo fatto

$$\mathbb{P}(X_2 = j \mid X_0 = i) = \sum_k \mathbb{P}(X_2 = j \mid X_1 = k)\mathbb{P}(X_1 = k \mid X_0 = i) = \sum_k P_{kj}P_{ik} = (P^2)_{ij}$$

□

Definizione 1.1.5. $\pi = (\pi_i)_{i\in E}$ è un vettore di probabilità se $\pi_i \geq 0\ \forall i$ e $\sum \pi_i = 1$

Ad esempio data una catena di Markov è un vettore di probabilità il seguente $\pi_i^{(0)} := \mathbb{P}(X_0 = i)$

Osserviamo inoltre che ponendo $\pi_j^{(n)} := \mathbb{P}(X_n = j \mid X_0)$, $\pi^{(n)}$ continua ad essere un vettore di probabilità a causa della stocasticità di P

Vale $\pi^{(n+1)T} = \pi^{(n)T}P$ ed osserviamo che se esiste il limite

$$\lim_{n\to\infty} \pi^{(n)} = \pi$$

allora

$$\pi^T = \pi^T P \text{ (autovettore sinistro per P)}$$

Definizione 1.1.6. Tale π se esiste si chiama invariante di probabilità

Esempio 1.1.7. *(Random-walk) Consideriamo il caso in cui $E = \{0, 1, .., n\}$ e consideriamo la matrice di transizione*

$$\begin{pmatrix} 0 & 1 & 0 & ... \\ 1-p & 0 & p & ... \\ 0 & 1-p & 0 & p \\ ... & ... & ... & ... \end{pmatrix}$$

Chiamiamo la catena di Markov associata ad una tale matrice di transizione "Random-walk semplice"

Esempio 1.1.8. *(Teoria delle code) Immaginiamo di voler modellizzare una coda di clienti in fila che vengono serviti da un server, poniamo X_n:= numero di clienti in fila al tempo n, $n \to n+1$ quando uno dei clienti viene servito e indichiamo con α il numero di arrivi al tempo $n \to n+1$, chiamiamo inoltre $q_k = \mathbb{P}(\alpha = k)$. Abbiamo che*

$$X_{n+1} = \begin{cases} X_n + \alpha - 1 & \textit{se } X_n + \alpha \geq 1 \\ 0 & \textit{altrimenti} \end{cases}$$

1.2 Classificazine degli stati e grafo associato alla catena di Markov

Definizione 1.2.1. Siano $j \in E$ e $T_j := min\{n \geq 1 \text{ tale che } X_n = j\}$ il tempo della prima visita allo stato j

chiamiamo $f_j := \mathbb{P}(T_j < \infty \mid X_0 = j)$ la probabilità di ritorno in j

Definizione 1.2.2. Lo stato j si dice

- transiente se $f_j < 1$
- positivo ricorrente se $f_j = 1$ e $\mathbb{E}\left[T_j \mid X_0 = j\right] < \infty$
- ricorrente nullo se se $f_j = 1$ e $\mathbb{E}\left[T_j \mid X_0 = j\right] = \infty$

Esempio 1.2.3. $E = \{1, 2, 3\}$, *(ricordiamo che in questo caso* $P_{ij} = \mathbb{P}(X_n = j \mid X_{n-1} = i)$*)*

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

Lo stato 1 è positivo ricorrente mentre 2,3 sono transienti

Ad ogni data catena di Markov possiamo associare un grafo nel seguente modo:

- $\{Nodi\} := E$
- se $P_{ij} > 0$ allora esiste un arco $i \to j$

Definizione 1.2.4. Diciamo che i e j comunicano se $i \to j$ e $j \to i$

Definizione 1.2.5. Diciamo che una catena di Markov è irriducibile se lo è il grafo ad essa associato, riducibile altrimenti

Nota 1.2.6. Assumiamo per convenzione che ogni nodo comunica con se stesso, in questo modo la relazione di comunicazione è una relazione di equivalenza

Nota 1.2.7. Per l'osservazione sopra notiamo che una catena di Markov è irriducibile se e solo se c'è una sola classe di equivalenza sui nodi

Nota 1.2.8. Supponiamo di avere k classi di equivalenza per la relazione di comunicazione, allora esiste una matrice Π di permutazione tale che

$$\Pi P \Pi^T = \begin{pmatrix} P_{11} & 0 & ... & & \\ P_{21} & P_{22} & 0 & ... & \\ P_{31} & P_{32} & P_{33} & 0 & ... \\ ... & & & & \end{pmatrix}$$

Dove i P_{ii} sono quadrati irriducibili e ciascuno corrisponde ad una classe di equivalenza Osserviamo che sia k che i P_{ii} possono essere infiniti

Nota 1.2.9. P stocastico $\Rightarrow \Pi P \Pi^T$ stocastico

1.3 Classi di equivalenza

C una classe irriducibile si dice

- finale se una volta in C non si può più uscire, ovvero se $\forall i \in C \nexists k \notin C$ tale che $\exists i \to k$
- *absorbing* se è finale e $|C| = 1$
- di passaggio se non è finale

Nota 1.3.1. Le proprietà di essere positivo ricorrente, ricorrente nullo o transiente sono proprietà di classe

Nota 1.3.2. Gli elementi di una classe di passaggio sono sempre transienti mentre la natura degli elementi di una classe finale dipende da $|C|$

Nota 1.3.3. Se X irriducibile, $|E| < \infty$ allora X è positiva ricorrente (tutti gli elementi o sono)

Teorema 1.3.4. *X catena di Markov irriducibile, allora X è positivo ricorrente se e solo se $\exists \pi$ vettore invariante di probabilità tale che $\pi_i > 0\ \forall i$. Inoltre se tale π esiste allora è unico*

Definizione 1.3.5. $i \in E$ si dice periodico se tutti i cammini $i \to i$ sono formati da un numero di archi multiplo di un intero $\delta \geq 2$

Esempio 1.3.6. $E = \{1, 2\}$, $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

In questo caso entrambi gli stati hanno periodo 2

Nota 1.3.7. $(P^n)_{ij} =$ "probabilità di trovarsi in j partendo da i dopo n passi"

Nota 1.3.8. Se $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ abbiamo $P^{2k} = Id$, $P^{2k+1} = P$ per cui certamente P^n non è una successione convergente

Teorema 1.3.9. *X irriducibile, positiva ricorrente e aperiodica (ossia nessuno stato è periodico), allora esiste il limite*

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = \pi_j = P^n_{ij}$$

Dove π è il vettore invariante di probabilità per P. In particolare significa che

$$P^n \to \begin{pmatrix} \pi^T \\ \pi^T \\ ... \\ \pi^T \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix} \pi^T$$

Notazione: $\begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix} := (1)$

Nota 1.3.10. Se poniamo $\begin{cases} X^{(0)} \geq 0 \ , \ X^{(0)}(1) = 1 \\ X^{(k+1)T} = X^{(k)T} P \end{cases}$

$X^{(k)T} P = X^{(0)T} P^k \to X^{(0)T}(1)\pi^T$ e ad ogni passo il vettore $X^{(n)}$ rimane stocastico se $X^{(0)}$ lo era.

Nota 1.3.11. Nel caso transiente $\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = 0$

Esempio 1.3.12. *(Teoria delle code)*

Siano $E = \mathbb{N}$, $P = \begin{pmatrix} b_0 & b_1 & 0 & ... \\ a_{-1} & a_0 & a_1 & 0 \\ 0 & a_{-1} & a_0 & a_1 \\ 0 & 0 & a_{-1} & a_0 \end{pmatrix}$

P è irriducibile periodica, vogliamo risolvere il problema $\pi^T(I - P) = 0$, ossia

$$\begin{pmatrix} \pi_0 & \pi_1 & \pi_2 & ... \end{pmatrix} \begin{pmatrix} (1-b_0) & -b_1 & 0 & ... \\ -a_{-1} & (1-a_0) & -a_1 & 0 \\ 0 & -a_{-1} & (1-a_0) & -a_1 \\ 0 & 0 & -a_{-1} & (1-a_0) \end{pmatrix} = 0$$

Osserviamo che la matrice sopra ha struttura di Toeplitz e le equazioni da risolvere diventano

$$\begin{cases} \pi_0(1-b_0) - \pi_1 a_{-1} = 0 \\ -\pi_0 b_1 + \pi_1(1-a_0) - \pi_2 a_{-1} = 0 \\ -\pi_i a_1 + \pi_{i+1}(1-a_0) - \pi_{i+2} a_{-1} = 0 \ \textit{per} \ i \geq 1 \end{cases}$$

il problema sopra è un sistema di equazioni alle differenze omogenee, calcoliamo il polinomio caratteristico $p(\lambda) = -a_{-1}\lambda^2 + (1-a_0)\lambda - a_1$ che ha come zeri $\lambda_{1,2} = 1, \frac{a_1}{a_{-1}}$

Ora se $\lambda_1 = \lambda_2$ *allora la soluzione del sistema è del tipo* $\pi_i = \alpha\lambda^{i-1} + (i-1)\beta\lambda^{i-1} = \alpha + (i-1)\beta$. *Un* π *di questa forma è certamente non sommabile*

Se invece abbiamo $\lambda_1 \neq \lambda_2$ *la soluzione ha la forma* $\pi_i = \alpha\lambda_1^{i-1} + \beta\lambda_2^{i-1} = \alpha + \beta\lambda_2^{i-1}$ *Possiamo interpretare* a_{-1} *come la "probabilità che la coda diminuisca" e* a_1 *come la "probabilità che aumenti". Aggiungendo le informazioni sui valori iniziali otteniamo* $\alpha = 0$ *e*

$$\begin{cases} \pi_i = \left(\dfrac{a_1}{a_{-1}}\right)^{i-1} \beta \text{ per } i \geq 1 \\ \\ \pi_0 = \dfrac{\pi_1 a_{-1}}{1 - b_0} = \dfrac{\beta a_{-1}}{1 - b_0} \end{cases}$$

da cui si capisce che la soluzione π *è sommabile se e solo se* $a_{-1} > a_1$*, in questo caso sceglieremo poi* β *in modo da normalizzare* $|\pi| = 1$

Nota 1.3.13. Possiamo interpretare a_1 come " probabilità che la coda aumenti" e a_{-1} come "probabilità che la coda diminuisca"

Capitolo 2

Teoria delle matrici non negative

Definizione 2.0.1. Ricordiamo che in questo contesto una matrice A si dice **non negativa** se $a_{ij} \geq 0 \; \forall i, j$

Definizione 2.0.2. A e B matrici di dimensioni compatibili, diciamo che $A \geq B$ se $A - B \geq 0$

Teorema 2.0.3 (Perron-Frobenius)**.** *$A \geq 0$ irriducibile (ovvero il grafo associato ad A è fortemente connesso) allora valgono le seguenti proprietà:*

- *$\rho(A)$ è autovalore*
- *$\rho(A) > 0$*
- *$\exists u \in \mathbb{R}^n \mid u > 0$ e $Au = \rho(A)u$*
- *se $B \geq A$ e $B \neq A$ allora $\rho(B) > \rho(A)$*
- *$\rho(A)$ ha molteplicità algebrica 1*

Esempio 2.0.4. *Supponiamo di avere una matrice stocastica P^T, allora $\rho(P^T) = 1 \Rightarrow \exists! \pi > 0$ tale che $P^T \pi = \pi \Rightarrow \pi^T P = \pi^T$ (ovvero la catena di Markov irriducibile associata è positiva ricorrente)*

Lemma 2.0.5. *$A \in M_{n \times n} \geq 0$ reale irriducibile allora $(I + A)^{n-1} > 0$*

Dimostrazione. Sia $x \in \mathbb{R}^n \geq 0$, vediamo che $(I + A)^{n-1} x > 0$.
Se $x > 0$ è ovvio, se $x \geq 0$ definiamo la successione

$$\begin{cases} x_0 = x \\ x_{k+1} = (I + A) x_k \end{cases}$$

Vediamo che il numero di entrate nulle di x_k è minore strettamente del numero di entrate nulle di x_{k-1}. Certamente vale il $\leq$, se per assurdo i due numeri fossero uguali avremmo che esiste una matrice Π di permutazione tale che $\Pi x_{k-1} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix}$

Con α un vettore > 0 di lunghezza h, allora

$$\Pi x_k = \Pi(I+A)\Pi^T \Pi x_{k-1} = \begin{pmatrix} V_1 & V_2 \\ V_3 & V_4 \end{pmatrix} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} V_1\alpha \\ V_3\alpha \end{pmatrix}$$

con $V_3\alpha = 0$ (dato che $V_1\alpha > 0$) ma allora $V_3 = 0$ contro l'ipotesi di irriducibilità

□

Nota 2.0.6. Sia $A \geq 0$ irriducibile, $x \geq 0$ e $x \neq 0$, poniamo

$$r_x := min_{x_i \neq 0}\{\frac{\sum_j a_{ij}x_j}{x_i}\} = min_i \frac{A_i x}{x_i}$$

$Ax \geq r_x x$, chiamiamo ora

$$r := sup_{x\neq 0, x\geq 0} r_x = sup_P r_x$$

dove $P := \{x \geq 0 \mid ||x||_1 = 1\}$ e $Q := \{y \mid y = (I+A)^{n-1}x \text{ con } x \in P\}$

P e Q compatti, ora siccome $Ax \geq r_x x \Rightarrow (I+A)^{n-1}Ax \geq r_x(I+A)^{n-1}x \Rightarrow A(A+I)^{n-1}x \geq r_x(I+A)^{n-1}x$.

quindi $r_y := min_{i,y\in Q}\frac{\sum_j a_{ij}y_j}{y_i} \geq min_i \frac{r_x y_i}{y_i} = r_x \Rightarrow r = sup_Q r_y = max_Q r_y \Rightarrow \exists w > 0$ tale che $r_w = r, Aw \geq rw$

Definizione 2.0.7. $z \geq 0$, $z \neq 0$ si dice estremale se $Az \geq rz$

Lemma 2.0.8. *$A \in M_{n\times n} \geq 0$ reale irriducibile, $r \geq 0$, z estremale $\Rightarrow Az = rz$ e $z > 0$*

Dimostrazione. $r > 0$ (considerare per esempio $x = (1)$), prendiamo $\eta := Az - rz$ (vettore residuo), $\eta \geq 0$. Se per assurdo $\eta \neq 0$ allora $(I+A)^{n-1}Az - r(I+A)^{n-1}z > 0$ per il lemma precedente $\Rightarrow Aw > rw \Rightarrow r_w > r$ assurdo
Vediamo adesso che $z > 0$. $A(I+A)^{n-1}z = r(I+A)^{n-1}z \Rightarrow (I+A)^{n-1}Az > 0$ $(z \neq 0)$ $\Rightarrow r(1+r)^{n-1}z > 0 \Rightarrow z > 0$ □

Lemma 2.0.9. *$A \in M_{n\times n} \geq 0$ reale irriducibile, $B \in \mathbb{C}_{n\times n}$ tale che $|B| \leq A$, allora se $\beta \in \sigma(B) \Rightarrow |\beta| \leq r$. Inoltre $\beta = r \Leftrightarrow |B| = A$ e $\beta = re^{i\theta}$, $B = e^{i\theta}DAD^{-1}$ dove $D = diag(\lambda_i), |\lambda_i| = 1$*

Corollario 2.0.10. *$r = \rho(A)$ (Lemma sopra con $B = A$)*

Lemma 2.0.11. *$A \in M_{n\times n} \geq 0$ irriducibile, B sottomatrice principale di A (cioè tolgo colonne e righe dello stesso indice) allora $\rho(B) < \rho(A)$*

Dimostrazione. (Punto 4 di Perron- Frobenius, $\rho(A)$ autovalore semplice)
Definiamo $\phi(t) := det(tI - A)$, $\phi(\rho(A)) = 0$, $\phi'(t) = \sum det(tI_{n-1} - A_i)$ dove le A_i sono sottomatrici di A ottenute eliminando la i-esima riga e la i-esima colonna, per uno dei lemmi sopra $\rho(A) > \rho(A_i)\forall i \Rightarrow \phi'(\rho(A)) \neq 0$

□

Teorema 2.0.12 (Perron-Frobenius debole). *Sia $A \geq 0$, allora valgono le seguenti affermazioni*

- $\rho(A)$ *è autovalore*
- $\exists u \in \mathbb{R}^n \mid u \geq 0$ *e* $Au = \rho(A)u$
- *se* $B \geq A$ *allora* $\rho(B) \geq \rho(A)$

Consideriamo il caso in cui $A \geq 0$ irriducibile, vediamo in quali ipotesi esistono più autovalori di modulo $\rho(A)$

Definizione 2.0.13. $A \geq 0$ si dice ciclica di indice k ($k \geq 2$) se ha k autovalori di modulo $\rho(A)$. $A \geq 0$ si dice primitiva altrimenti

Teorema 2.0.14. *$A \geq 0$ irriducibile ciclica di indice k allora gli autovalori di modulo $\rho(A)$ sono $\{e^{i\theta_j}\rho(A)\}_j$ con $\theta_j = \frac{2\pi j}{k}$*

Dimostrazione. Certamente gli autovalori hanno la forma $e^{i\theta_j}\rho(A)$ per qualche $0 = \theta_0 \leq ... \leq \theta_{k-1}$. Per Perron-Frobenius $0 < \theta_1$. Prendiamo adesso $A = e^{i\theta_j}D_jAD_j^{-1}$ come nel lemma (possiamo scrivere l'analogo per ciascuno dei θ_m) ,abbiamo quindi $A = e^{i\theta_j}D_jAD_j^{-1} = e^{i\theta_m}D_mAD_m^{-1} \Rightarrow A = e^{i(\theta_j-\theta_m)}D_m^{-1}D_jAD_j^{-1}D_m \Rightarrow e^{i(\theta_j-\theta_m)}\rho(A)$ è ancora autovalore. Con maneggiamenti analoghi si vede che anche $e^{-i\theta_j}\rho(A)$ e $e^{-i(\theta_j+\theta_m)}\rho(A)$ sono autovalori $\Rightarrow \{e^{i\theta_j}\rho(A)\}_j$ è un gruppo $\square$

Proposizione 2.0.15. *A ciclica di indice $k \Leftrightarrow \exists\Pi$ matrice di permutazione tale che*

$$\Pi^T A \Pi = \begin{pmatrix} 0 & A_{1,2} & 0 & 0 \\ 0 & 0 & A_{2,3} & 0 \\ 0 & 0 & 0 & A_{k-1,k} \\ A_{k,1} & 0 & 0 & 0 \end{pmatrix}$$

Con blocchi sopradiagonali $A_{i,j}$ quadrati

Esempio 2.0.16. $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

La matrice sopra ha polinomio caratteristico $p(\lambda) = \lambda^n - 1$

Nota 2.0.17. A ciclica, $\rho(A) = 1$ allora $A = VJV^{-1}$, $J = \begin{pmatrix} D & 0 \\ 0 & \tilde{J} \end{pmatrix}$ dove $D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ & & ... & \\ 0 & 0 & 0 & \lambda_k \end{pmatrix}$, $\tilde{J}$ forma di Jordan complessa con $|\lambda_i| = 1$ e $\rho(\tilde{J}) < 1$.

Osserviamo che

$$J^n = \begin{pmatrix} D^n & 0 \\ 0 & \tilde{J}^n \end{pmatrix}$$

con $D^n = \begin{pmatrix} \lambda_{n_1} & 0 & 0 & 0 \\ 0 & \lambda_{n_2} & 0 & 0 \\ & & ... & \\ 0 & 0 & 0 & \lambda_{n_k} \end{pmatrix}$ in particolare esiste il limite $\lim_{n\to\infty} A^n$ se e solo se $k = 1$

Nota 2.0.18. $\frac{1}{\rho(A)}A = \tilde{A}, \rho(\tilde{A}) = 1$

Nota 2.0.19. A catena di Markov finita irriducibile è periodica se e solo se è ciclica

Teorema 2.0.20. $A \geq 0$ *irriducibile allora*

- *se* $\sum_j a_{ij} = \alpha$ *per ogni* i *allora* $\rho(A) = \alpha$
- *se* $\alpha = min_i \sum_j a_{ij} < max_i \sum_j aij = \beta$ *allora* $\alpha < \rho(A) < \beta$

Dimostrazione. Il punto 1 segue banalmente da P-F. Per il punto 2 consideriamo una matrice $B \geq 0$ irriducibile, $B \leq A$, $\sum_j b_{ij} = \alpha$ per ogni i. Dal primo punto segue che $\rho(A) > \rho(B)$. Procedere analogamente per l'altra disuguaglianza □

Capitolo 3

M-matrici

Definizione 3.0.1. Siano $B \geq 0$, $s \in \mathbb{R} > 0$ tali che $\rho(B) \leq s$. Chiamiamo una matrice A della forma $A := sI - B$ una **M-matrice**
In particolare se $s = \rho(B)$ abbiamo una M-matrice singolare

Noi saremo particolarmente interessati al caso $A = I - P$, P matrice stocastica

3.0.1 Z-matrici

Definizione 3.0.2. Chiamiamo una matrice A della forma

$$A = \begin{pmatrix} * & & \leq 0 \\ & * & \\ \leq 0 & & * \end{pmatrix}$$

Una **Z-matrice**

Nota 3.0.3. Osserviamo che A M-matrice $\Rightarrow$ A Z-matrice

Teorema 3.0.4. *A una Z-matrice allora sono equivalenti le seguenti affermazioni*

- *A M-matrice non singolare*
- *$Re\lambda > 0$ $\forall\lambda$ autovalore*
- *A invertibile con inverso $A^{-1} \geq 0$*
- *$x \in \mathbb{R}^n$ tale che $Ax \geq 0 \Rightarrow x \geq 0$*
- *$\exists D = diag(d_i)$, $d_i > 0$ tale che $AD(1) > 0$*
- *$\exists D = diag(d_i)$, $d_i > 0$ tale che $a_{ii}d_i > -\sum_{j\neq i} a_{ij}d_j$*

Nota 3.0.5. $A = sI - B = s(I - \frac{B}{s}) \Rightarrow A^{-1} = s^{-1}(I - \frac{B}{s})^{-1} = s^{-1}\sum\left(\frac{B}{s}\right)^i \geq 0$ (nelle ipotesi $\rho\left(\frac{B}{s}\right) < 1$)

Localizzazione degli autovalori di A Osserviamo che λ è autovalore per $A = sI - B$ se e solo se $\lambda = s - \beta$, $\beta \in \sigma(B)$

Nota 3.0.6. Gli elementi diagonali sono strettamente positivi per le M-matrici non singolari

Nota 3.0.7. $\frac{a_{ii}d_i}{d_i} > -\sum_{j\neq i} \frac{a_{ij}d_j}{d_i} \Rightarrow a_{ii} > -\sum a_{ij}d_j d_i^{-1} \Rightarrow D^{-1}AD$ è dominante diagonale forte

Lemma 3.0.8. *A Z-matrice allora A è M-matrice se e solo se $A + \varepsilon I$ è M-matrice non singolare $\forall \varepsilon > 0$*

Teorema 3.0.9. *$A \in \mathbb{R}^{n\times n}$ allora A è M-matrice non singolare se e solo se $A + D$ è non singolare e $(A + D)^{-1} \geq 0$ $\forall D$ diagonale > 0*

3.1 Fattorizzazione LU di M-matrici

Sia A una M-matrice, ci chiediamo sotto quali ipotesi esiste la fattorizzazione LU

Lemma 3.1.1. *A M-matrice non singolare oppure singolare irriducibile allora ogni sottomatrice principale di A è M-matrice*

Dimostrazione. Caso 1 (A non singolare).

$$A = \begin{pmatrix} \hat{A} & \leq 0 \\ \leq 0 & + \end{pmatrix}$$

$\hat{A}$ sottomatrice di testa, $\hat{A} = s\hat{I} - \hat{B}$, siccome $\hat{B} \geq 0 \Rightarrow \rho(\hat{B}) \leq \rho(B)$
Caso 2 (A singolare irriducibile) $B \geq 0$ irriducibile, $\hat{A} = s\hat{I} - \hat{B} \Rightarrow \rho(\hat{B}) < \rho(B) = s$

□

Corollario 3.1.2. *Nelle ipotesi del lemma precedente esiste ed unica la fattorizzazione LU di A*

Teorema 3.1.3. *A M-matrice non singolare, $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ con $A_{11} \in M_{k\times k}$, $A_{22} \in M_{(n-k)\times(n-k)}$ e $1 \leq k \leq n-k$ allora A_{11} è m-matrice non singolare e $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$ (complemento di Schur di A_{11} in A) è M-matrice non singolare*

Dimostrazione. (idea: eliminazione Gaussiana a blocchi)

$$A = \begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & S \end{pmatrix} = LU$$

$A_{21} \leq 0$ fuori dalla diagonale, $A_{11}^{-1} \geq 0 \Rightarrow A_{21}A_{11}^{-1} \leq 0$ dato che A_{11} M-matrice non singolare, A_{22} M-matrice

$S = A_{22} - A_{21}A_{11}^{-1}A_{12} \Rightarrow S \leq 0$ fuori dalla diagonale. In particolare S è Z-matrice. Ora S è anche invertibile, infatti $det \begin{pmatrix} A_{11} & A_{12} \\ 0 & S \end{pmatrix} \neq 0 \Rightarrow detS \neq 0$, vediamo che S ha inversa non negativa.

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & * \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} I & 0 \\ * & I \end{pmatrix} = \begin{pmatrix} * & * \\ * & S^{-1} \end{pmatrix} \geq 0$$

□

Teorema 3.1.4. *A M-matrice non singolare allora esiste ed unica la fattorizzazione LU, inoltre L ed U sono M-matrici non singolari*

Dimostrazione. (per induzione su n) Caso base $n = 1$ ok, caso scalare

$n-1 \Rightarrow n$: $A = \begin{pmatrix} A_1 & c \\ b^T & a_{nn} \end{pmatrix} = \begin{pmatrix} I & 0 \\ b^T A_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_1 & c \\ 0 & s \end{pmatrix}$

dove A_1 è una M-matrice non singolare (per uno dei lemmi fatti in precedenza) $(n-1)\times(n-1)$ e $s = a_{nn} - b^T A_1^{-1} c$. Siccome A_1 è M-matrice non singolare possiamo applicare l'ipotesi induttiva $\Rightarrow A_1 = L_1 U_1$ con L_1, U_1 M-matrici non singolari

$$\begin{pmatrix} I & 0 \\ b^T A_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_1 & c \\ 0 & s \end{pmatrix} = \begin{pmatrix} I & 0 \\ b^T A_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} L_1 U_1 & c \\ 0 & s \end{pmatrix} = \begin{pmatrix} L_1 & 0 \\ b^T U_1^{-1} & 1 \end{pmatrix} \begin{pmatrix} U_1 & L_1^{-1} c \\ 0 & s \end{pmatrix} = LU$$

Notiamo che $b^T \leq 0$, $U_1^{-1} \geq 0$ in quanto U_1 M-matrice non singolare quindi

$\Rightarrow L = I - \begin{pmatrix} 0 & 0 \\ \geq 0 & 0 \end{pmatrix}$ dunque L è M-matrice non singolare

□

Nota 3.1.5. A M-matrice non singolare, consideriamo il problema $Ax = b$. Le strategie di pivoting distruggono la struttura di M-matrice, se sfruttiamo la decomposizione LU abbiamo

$\begin{cases} Ly = b \\ Ux = y \end{cases}$

$x = A^{-1} b \geq 0$

$$\begin{pmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ l_{ij} & & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ ... \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ ... \\ b_n \end{pmatrix}$$

Per cui $y_i = b_i - \sum L_{ij} y_j$. non c'è cancellazione

3.1.1 M-matrici singolari

Teorema 3.1.6. *A M-matrice singolare irriducibile, allora valgono le seguenti affermazioni*

- $rgA = n - 1$
- $\exists x > 0$ *tale che* $x \in KerA$
- $\exists A = LU$, *L e U M-matrici, U singolare*

Dimostrazione. I primi due punti seguono da Perron-Frobenius: $A = sI - B$, B irriducibile ≥ 0. Allora $\rho(B)$ è autovalore semplice di B, $s = \rho(B)$ ed esiste un autovettore > 0
Per quanto riguarda l'esistenza della fattorizzazione LU abbiamo una situazione analoga a quella trattata precedentemente per le M-matrici non singolari:

$$\hat{L} = \begin{pmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ * & & 1 \end{pmatrix}, \hat{U} = \begin{pmatrix} * & & \\ 0 & * & \\ 0 & 0 & * \end{pmatrix} \text{ quindi } A = \begin{pmatrix} \hat{L} & \\ & 1 \end{pmatrix} \begin{pmatrix} \hat{U} & \\ 0 & 0 \end{pmatrix} = LU(s=0)$$

le sottomatrici principali sono perciò non singolari.

L è non singolare per lo stesso argomento, U ha elementi > 0 sulla diagonale e $u_{nn} = 0$, inoltre si verifica che $U + \varepsilon I$ è M-matrice non singolare $\Rightarrow$ U M-matrice

□

Capitolo 4

Risoluzione del sistema lineare

Sia P una matrice stocastica irriducibile, siamo interessati a risolvere il sistema lineare

$$\pi^T(I-P)=0$$

ricordiamo inoltre che π quando esiste si chiama vettore invariante di probabilità ed è unico imponendo $\sum \pi_i = 1$, $\pi_i \geq 0$

4.1 Fattorizzazione LU

$I-P=LU$, abbiamo allora $\pi^T LU = 0$ da cui il sistema

$$\begin{cases} y^T U = 0 \\ \pi^T L = y^T \end{cases}$$

$y^T \begin{pmatrix} \tilde{U} & * \\ 0 & 0 \end{pmatrix} = (0)^T$ con $\tilde{U}$ triangolare superiore, pongo $y = \begin{pmatrix} 0 \\ ... \\ 1 \end{pmatrix}$: (un tale y funziona) devo solo risolvere il problema

$$\tilde{\pi}^T \tilde{L} = \begin{pmatrix} 0 & ... & 0 & 1 \end{pmatrix}$$

Per cui ottengo il sistema

$$\begin{cases} \tilde{\pi}_n = 1 \\ \tilde{\pi}_i = -\sum \tilde{\pi}_j l_{ij} \end{cases}$$

(sostituzione all'indietro), $\pi = \dfrac{1}{\sum \tilde{\pi}_i}\tilde{\pi}$

4.1.1 Calcolo fattorizzazione LU

Censoring Sia P stocastica irriducibile (possibilmente infinita), $E = E_1 \cup E_2$, $E_1 \cap E_2 = \emptyset$, $P_{ij} = (P_{hk})_{h \in E_i, k \in E_j}$ e sia Π una matrice di permutazione tale che

$$\Pi P \Pi^T = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}$$

Chiamo $X_n^{(c)} := X_n \mid_{E_1}$

$$P^{(E_1)} = P_{11} + P_{12}P_{22}P_{21} + P_{12}P_{22}^2P_{21} + .. = P_{11} + \sum_k P_{12}P_{22}^k P_{21} = P_{11} + P_{12}(I - P_{22})^{-1}P_{21}$$

$P^{(E_1)}$ è stocastica e inoltre $\pi^T P = \pi^T \Rightarrow \pi^{(E_1)}$ è tale che $\pi^{(E_1)T}P^{(E_1)} = \pi^{(E_1)T}$

Dimostrazione. Scriviamo (P a meno di applicare la permutazione sopra)

$$\begin{pmatrix} \pi^{(E_1)T} & \pi^{(E_2)T} \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{pmatrix} \pi^{(E_1)T}P_{11} + \pi^{(E_2)}P_{21} & \pi^{(E_1)T}P_{12} + \pi^{(E_2)T} \end{pmatrix} = \begin{pmatrix} \pi^{(E_1)T} & \pi^{(E_2)T} \end{pmatrix}$$

da cui il sistema

$$\begin{cases} \pi^{(E_1)T}P_{11} + \pi^{(E_2)T}P_{21} = \pi^{(E_1)T} \\ \pi^{(E_1)T}P_{12} + \pi^{(E_2)T}P_{22} = \pi^{(E_2)T} \end{cases}$$

dalla seconda equazione otteniamo

$$\pi^{(E_1)T}P_{12} = (I - P_{22})\pi^{(E_2)T} \Rightarrow \pi^{(E_1)T}P_{12}(I - P_{22})^{-1} = \pi^{(E_2)T}$$

sostituendo nella prima

$$\pi^{(E_1)T}P_{11} + \pi^{(E_1)T}P_{12}(I - P_{22})^{-1}P_{21} = \pi^{(E_1)T}$$

si ottiene poi un'espressione analoga per $\pi^{(E_2)T}$

$\square$

Vediamo i complementi di Schur

$I - P^{(E_1)} = 1 - P_{11} - P_{12}(I - P_{22})^{-1}P_{21}$, $I - P = \begin{pmatrix} (I - P_{11}) & -P_{12} \\ -P_{21} & (1 - P_{22}) \end{pmatrix}$, $(I - P)\underline{1} = 0 \Rightarrow$

$$\begin{cases} (I - P_{11})(1) - P_{12}(1) = 0 \\ -P_{21}(1) + (I - P_{22})(1) = 0 \end{cases}$$

per cui in particolare $(I - P_{22})(1) = 0 \Rightarrow P_{22}$ stocastica

$$\begin{pmatrix} \pi^{(E_1)} & \pi^{(E_2)} \end{pmatrix} \begin{pmatrix} I - P_{11} & -P_{12} \\ -P_{21} & (I - P_{22}) \end{pmatrix} = 0 \Rightarrow \pi^{(E_2)}(I - P^{(E_2)}) = 0$$

Eliminazione gaussiana (Caso $dim < \infty$, P irriducibile) $I - P = LU$

$$\begin{pmatrix} 1-p_{11} & ... & -p_{1n} \\ & A_1 & \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1-p_{11} & .. & -p_{1n} \\ 0 & B^{(1)} & \\ 0 & & \end{pmatrix}$$

(primo passo di eliminazione gaussiana) dove $B^{(1)} = A^{(1)} - \begin{pmatrix} p_{21} \\ p_{31} \\ ... \\ p_{n1} \end{pmatrix} \frac{1}{1-p_{11}} \begin{pmatrix} p_{12} & ... & p_{1n} \end{pmatrix} \Rightarrow$

$B^{(1)} = I_P^{(E_2)}$, (sappiamo già che $B^{(1)} M - matrice$) dove $E_2 = \{2, ..., n\}$.

Ora il calcolo numerico degli elementi di $B^{(1)} = (b_{ij})^{(1)}$ è stabile per $i \neq j$, tuttavia possiamo non calcolare gli elementi diagonali in quanto sappiamo già che la somma per righe deve fare 0: abbiamo quindi $b_{ii}^{(1)} = -\sum_{j \neq i} b_{ij}^{(1)}$ (diagonal adjustment)

4.2 Possibile approccio 2

Consideriamo di nuovo il problema

$$\pi^T (I - P) = 0$$

fisso $\pi_n = 1$

$$\begin{pmatrix} \pi_1 & ... & \pi_{n-1} & 1 \end{pmatrix} \begin{pmatrix} I_{n-1} - \hat{P} & v \\ u^T & * \end{pmatrix} = (0)$$

Da cui $\begin{pmatrix} \pi_1 & ... & \pi_{n-1} \end{pmatrix} (I_{n-1} - \hat{P}) = -u^T$

questa volta la matrice è non singolare $\tilde{\pi} LU = -u^T$, con L,U entrambe M-matrici non singolari, $-u > 0$ in quanto ultima riga di P: la soluzione è stabile numericamente ammesso di conoscere la fattorizzazione LU, si devono a questo punto risolvere due sistemi triangolari

Ancora una volta si evita di utilizzare strategie di pivoting per non modificare la struttura di M-matrice

4.3 Possibile approccio 3

Osserviamo che $rg(I - P) = n - 1$, con questo approccio includiamo la condizione di normalizzazione $\pi^T (1) = 1$ ponendo

$$\pi^T \begin{pmatrix} * & * & 1 \\ * & * & 1 \\ * & * & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & .. & 1 \end{pmatrix}$$

dove $\begin{pmatrix} * & * & 1 \\ * & * & 1 \\ * & * & 1 \end{pmatrix} = (I - P \mid (1))$

si ottiene così una matrice non singolare e non occorre rinormalizzare. La matrice così ottenuta non è più una M-matrice

4.4 GTH, accenno

Questo metodo si può generalizzare al problema $Ax = b$, A M-matrice non singolare e $b \geq 0$

4.5 Metodi iterativi

Supponiamo $Ax = b$, $detA \neq 0$ $A = M - N$ con M invertibile. Abbiamo allora il problema equivalente di punto fisso

$$(M - N)x = b \Rightarrow x = M^{-1}Nx + M^{-1}b$$

diciamo che il metodo è convergente se converge per ogni punto di partenza x_0, in particolare abbiamo che il metodo è convergente se e solo se $\rho(M^{-1}N) < 1$

Definizione 4.5.1. A M-matrice non singolare, diciamo che A=M-N è un regular splitting se $detM \neq 0$, $M^{-1} \geq 0$, $N \geq 0$

Teorema 4.5.2. *A M-matrice non singolare reale, $A^{-1} \geq 0$, $A = M - N$ regular splitting allora*

$$\rho(M^{-1}N) = \frac{\rho(A^{-1}N)}{1 + \rho(A^{-1}N)} < 1$$

e in particolare il metodo iterativo associato converge

Dimostrazione. Abbiamo

$$H := M^{-1}N = (A + N)^{-1}N = (A(I + A^{-1}N))^{-1}N = (I + A^{-1}N)^{-1}A^{-1}N$$

Sia ora $G := A^{-1}N$, supponiamo $Gx = \lambda x$, $x \neq 0$, ora $Hx = (I + G)^{-1}Gx = (I + G)^{-1}\lambda x = \frac{\lambda}{1+\lambda}x$, cioè $\mu = \frac{\lambda}{1+\lambda}$ è autovalore della matrice di iterazione con lo stesso autovettore.

Facciamo ora il procedimento inverso, sia $Hx = \mu x$, $(I + G)^{-1}Gx = \mu x \Rightarrow \mu(I + G)x = Gx \Rightarrow \mu x = (1 - \mu)Gx \Rightarrow Gx = \frac{\mu}{1-\mu}x$

Notiamo inoltre che $\mu = 1$ significherebbe $M^{-1}Nx = x \Rightarrow Nx = Mx \Rightarrow (M - N)x = 0$ contro l'ipotesi di non singolarità di A. Analogamente se avessimo $\lambda = -1$ allora esisterebbe x tale che $(I + A^{-1}N)x = 0 \Rightarrow A^{-1}(A + N) = M$ singolare contro l'ipotesi di regular splitting.

Le nostre ipotesi ed il teorema di Perron-Frobenius ci dicono che $H = M^{-1}N \geq 0 \Rightarrow 0 < \rho(M^{-1}N)$ è autovalore $\mu = \frac{\lambda}{1+\lambda}$ dove λ è autovalore di $G = A^{-1}N$. Mostriamo adesso che $\mu \geq 0$, $\lambda \geq 0$ ed esiste un autovettore $x \geq 0$. $Hx = \mu x$, per quanto osservato in precedenza abbiamo $Gx = \lambda x$, ora $A^{-1} \geq 0$ e $N \geq 0 \Rightarrow Gx \geq 0 \Rightarrow \lambda \geq 0$, notiamo adesso che la funzione $f(x) = \frac{x}{1+x}$ è strettamente crescente su $\mathbb{R}^+$ per cui assume il valore massimo per x massimo $\Rightarrow \rho(H) = max_{\lambda \in \sigma(G)} \dfrac{\lambda}{1+\lambda} = \dfrac{\rho(G)}{1+\rho(G)}$

□

Nota 4.5.3. La matrice di iterazione associata al metodo di Jacobi è data da $J = D^{-1}(L+U)$, dove $A = D-L-U$ mentre per Gauss-Seidel abbiamo $G = (D-L)^{-1}U$. Nel primo caso otteniamo la convergenza del metodo applicando il teorema sopra direttamente, nel secondo abbiamo che $(D-L)$ è M-matrice e $U \geq 0$ per cui di nuovo abbiamo convergenza grazie al teorema

Nota 4.5.4. Supponiamo di avere due regular splitting della matrice $A = M_1 - N_1 = M_2 - N_2$ e supponiamo $N_2 \geq N_1$. Allora abbiamo $A^{-1}N_2 \geq A^{-1}N_1 \Rightarrow \rho(A^{-1}N_2) \geq \rho(A^{-1}N_1)$ per P-F, quindi per il teorema sopra $\Rightarrow \rho(M_2^{-1}N_2) \geq \rho(M_1^{-1}N_1) \Rightarrow$ nel nostro caso meglio usare il metodo di Gauss-Seidel in quanto $N_J \geq N_{GS}$

Nota 4.5.5. Il raggio spettrale del metodo iterativo rappresenta la riduzione asintotica dell'errore

4.5.1 Caso A singolare

Supponiamo di avere $A = M - N$ un regular splitting di A singolare e sia $x \in KerA$, allora abbiamo $Ax = 0 = (M-N)x \Rightarrow x = M^{-1}Nx$, in particolare la matrice di iterazione H ha raggio spettrale $\rho(H) \geq 1$.

Definizione 4.5.6. $S \in \mathbb{C}^{n \times n}$ si dice semiconvergente se $\exists \lim_{k\to\infty} S^k$.

Nota 4.5.7. Notiamo che una condizione necessaria per avere semiconvergenza è certamente $\rho(S) \leq 1$

Nota 4.5.8. La condizione $\rho(S) \leq 1$ non è sufficiente, consideriamo ad esempio la matrice di raggio spettrale 1 $M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, osserviamo che $M^{2k} = Id$ mentre $M^{2k+1} = M$ per cui evidentemente non abbiamo semiconvergenza

Nota 4.5.9. Vediamo che anche la condizione di avere come unico autovalore di modulo 1 $\lambda = 1$ non è sufficiente, consideriamo ad esempio la matrice $M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, abbiamo che $M^k = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ con $* \to \infty$ per $k \to \infty$

Proposizione 4.5.10. *Una condizione necessaria su S per avere la semiconvergenza è che l'autovalore $\lambda = 1$ deve avere molteplicità algebrica uguale alla molteplicità geometrica $\mu_a = \mu_g$*

Nelle ipotesi della proposizione sopra notiamo che la matrice S ha forma di Jordan del tipo

$$S = VJV^{-1} = V \begin{pmatrix} I_r & 0 \\ 0 & \tilde{J} \end{pmatrix} V^{-1}$$

con $\rho(\tilde{J}) < 1$, per cui in questo caso

$$S^k \to V \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} V^{-1} = V \begin{pmatrix} I_r \\ 0 \\ ... \\ 0 \end{pmatrix} \begin{pmatrix} I & 0 & ... & 0 \end{pmatrix} V^{-1}$$

possiamo scrivere anche $HV = V \begin{pmatrix} I_r & 0 \\ 0 & \tilde{J} \end{pmatrix}$

Ci chiediamo per quale scelta di x_0 si ha convergenza a soluzioni non banali ($x^k \nrightarrow 0$). Un tale valore di partenza deve essere non ortogonale agli autovettori sinistri, ovvero $\begin{pmatrix} I_r & 0 & .. & 0 \end{pmatrix} V^{-1}x_0 \neq 0$

Teorema 4.5.11. *A M-matrice singolare irriducibile, $A = M - N$ regular splitting allora $\rho(M^{-1}N) = 1$ e $\lambda = 1$ è autovalore semplice di $M^{-1}N$, in generale non è l'unico autovalore di modulo 1*

Dimostrazione. Per quanto osservato prima abbiamo sicuramente $\rho(M^{-1}N) \geq 1$, vediamo che è esattamente 1. Per Perron-Frobenius $\exists x > 0$ tale che $Ax = 0$, x è anche autovettore di $M^{-1}N$ per quanto osservato prima per cui $M^{-1}Nx = x$ (ricordiamo inoltre che $M^{-1}N \geq 0$)

$$\begin{pmatrix} x_1 & & 0 \\ & ... & \\ 0 & & x_n \end{pmatrix}^{-1} M^{-1}N \begin{pmatrix} x_1 & & 0 \\ & ... & \\ 0 & & x_n \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ ... \\ 1 \end{pmatrix}$$

(ok perchè $x > 0$) per cui

$$D^{-1}M^{-1}ND(1) = (1)$$

allora $\rho(M^{-1}N) = \rho(D^{-1}M^{-1}ND) = 1$ ($D^{-1}M^{-1}ND$ stocastica)

λ è autovalore di $M^{-1}N$ e $M^{-1}Nx = x \Leftrightarrow Ax = 0$, $A = sI - B$ con B irriducibile ≥ 0

Voglio calcolare il vettore invariante di probabilità usando il regular splitting: $A = I - P^T$, P stocastica irriducibile,

$$\begin{cases} A\pi = 0 \\ \sum \pi_i = 1 \end{cases}$$

$I - P^T = M - N$,definiamo la successione

$$x^{k+1} = M^{-1}Nx^k$$

dove

$$M^{-1}N = V \begin{pmatrix} 1 & 0 \\ 0 & \tilde{J} \end{pmatrix} V^{-1}$$

$(1)(I - P^T) = 0$, $(1)^T(M - N) = 0 \Rightarrow (1)^T M(I - M^{-1}N) = 0$ con $(1)^T M \geq 0$

□

Nota 4.5.12. $\rho(H) = 1$, H semiconvergente, allora la riduzione asintotica dell'errore $||x^k - x^*|| \to 0$ in norma 2 è stimato dal raggio spettrale di $\rho(\tilde{J})$

Esempio 4.5.13. *Prendiamo $M = I$, $N = P^T$, $M^{-1}N = P^T$ (metodo delle potenze) e definiamo la successione $\pi^{k+1} = P^T\pi^k$. Ci chiediamo se esiste e non banale il limite*

$$\lim_{k\to\infty} \pi^k$$

Supponiamo ad esempio di trovarci nel caso $P^T = diag(\lambda_i)$, $v_1, ..., v_n$ autovettori rispettivamente per $|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_n|$ con l'autovalore dominante di modulo 1. Abbiamo il problema

$$\begin{cases} \pi_0 = \sum \alpha_i v_i \\ \pi^k = (P^T)^k \pi_0 = \sum \alpha_i \lambda_i^k v_i = \alpha_1 v_1 + \sum_{i>1} \alpha_i \lambda_i^k v_i \end{cases}$$

Osserviamo che in questo caso se $|\lambda_i| < 0$ per $i > 1$ il metodo converge

Supponiamo ora che $\lambda = 1$ sia l'unico autovalore di modulo massimo, in questo caso

$$P^T = V \begin{pmatrix} 1 & 0 \\ 0 & \tilde{J} \end{pmatrix} V^{-1} \to V \begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \end{pmatrix} \begin{pmatrix} 1 & ... & 0 \end{pmatrix} V^{-1} = \pi(1)^T$$

(autovalore destro)

$$\pi^k = (P^T)^k \pi_0 \to \pi(1)^T \pi_0$$

in particolare ci basta imporre $(1)^T\pi_0 \neq 0$. Per esempio possiamo porre $(1)^T\pi_0 = 1 \Rightarrow \pi^k \to \pi$

Nota 4.5.14. Il metodo delle potenze potrebbe essere molto lento a convergere se $\rho(\tilde{J}) \sim 1$

Esempio 4.5.15.

$$I - P^T = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & \frac{1}{2} & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & -1 & 1 \end{pmatrix}$$

Si verifica che in questo caso il metodo di Jacobi converge mentre Gauss-Seidel non converge

Esempio 4.5.16.

$$I - P^T = \begin{pmatrix} * & 0 & ... & * \\ * & * & 0 & 0 \\ 0 & * & * & 0 \\ 0 & 0 & * & * \end{pmatrix}$$

In questo caso gli autovalori della matrice di iterazione di Jacobi sono le radici n-esime dell'unità per cui certamente il metodo di Jacobi non converge. Si verifica invece che il metodo di Gauss-Seidel porta a convergenza in 1 passo

Teorema 4.5.17. *A M-matrice singolare irriducibile, $\varepsilon > 0$*

$$A = (D + \varepsilon I - L) - (U + \varepsilon I) := M_\varepsilon - N_\varepsilon$$

Con L, U triangolari inferiore e superiore, allora

$$H_\varepsilon = M_\varepsilon^{-1} N$$

è semiconvergente

Nota 4.5.18. $A = (D + \varepsilon I - L) - (U + \varepsilon I) := M_\varepsilon - N_\varepsilon$ è un regular splitting: M_ε è M-matrice non singolare e $N_\varepsilon \geq 0$, in particolare $\rho(M_\varepsilon^{-1} N_\varepsilon) = 1$ e $\lambda = 1$ è autovalore semplice

Dimostrazione. Per l'osservazione sopra dobbiamo solo mostrare che $\lambda = 1$ è l'unico autovalore di modulo massimo

Nota 4.5.19. A ciclica di indice k allora $\exists \Pi$ matrice di permutazione tale che

$$\Pi A \Pi^T = \begin{pmatrix} 0 & & & * \\ * & 0 & & \\ & * & 0 & \\ & & * & 0 \end{pmatrix}$$

Vediamo che $M_\varepsilon^{-1} N_\varepsilon$ non può essere messa in questa forma ($\Rightarrow$ l'unico autovalore di modulo massimo è 1) :

$$M_\varepsilon^{-1} = \begin{pmatrix} + & 0 & 0 & 0 \\ & + & 0 & 0 \\ & & + & 0 \\ \geq 0 & & & + \end{pmatrix}, \; N_\varepsilon = \begin{pmatrix} \varepsilon & & & \geq 0 \\ & \varepsilon & & \\ & & \varepsilon & \\ 0 & & & \varepsilon \end{pmatrix} \Rightarrow$$

$H_\varepsilon = \begin{pmatrix} + & & & \\ & + & & \\ & & + & \\ & & & + \end{pmatrix}$ che non può essere ciclica in quanto la traccia è invariante di similitudine.

□

Nota 4.5.20. Se prendo $\varepsilon >> 1$ sopprimo l'informazione della matrice

Teorema 4.5.21. *A M-matrice singolare irriducibile, $A = M - N$ regular splitting, $H = M^{-1}N$, $0 < \alpha < 1$ allora*

$$H_\alpha := (1 - \alpha)I + \alpha H$$

è semiconvergente

Nota 4.5.22. Sia x tale che $H_\alpha x = x$, allora $\Rightarrow x = (1 - \alpha)x + \alpha H x \Rightarrow x = Hx$ dunque se abbiamo convergenza del metodo la abbiamo alla quantità giusta, inoltre notiamo che per lo stesso calcolo λ è autovalore di H $\Leftrightarrow (1 - \alpha) + \alpha\lambda$ è autovalore di H_α.

Nota 4.5.23. $(1 - \alpha) + \alpha\lambda = 1$ se $\lambda = 1$, $(1 - \alpha) + \alpha\lambda < 1$ se $\lambda \neq 1 \Rightarrow$ semiconvergenza

4.5.2 Altro possibile approccio

Ci riconduciamo al caso A non singolare nel seguente modo:

$$\left(I - P^T\right) \begin{pmatrix} \pi_1 \\ ... \\ \pi_n \end{pmatrix} = (0)$$

fisso $\pi_n = 1$, a questo punto

$$(I_{n-1} - \hat{P}^T) \begin{pmatrix} \pi_1 \\ ... \\ \pi_{n-1} \end{pmatrix} = \pi_n \begin{pmatrix} p_{1,n} \\ ... \\ p_{n-1,n} \end{pmatrix}$$

dove $(I_{n-1} - \hat{P}^T)$ è una M-matrice non singolare

4.6 Metodi IAD (Iterative Aggregation/Disaggregation)

Sia P una matrice stocastica irriducibile,

$$P = (P_{ij})$$

con P_{ii} quadrati di dimensione $h_i \times h_i$, $i = 1...N$ sia π il vettore invariante di probabilità $\pi^T P = \pi^T$ con

$\pi^T = \begin{pmatrix} \pi_1^T & ... & \pi_N^T \end{pmatrix}$, $dim(\pi_i) = h_i$. Definiamo la quantità $\phi_i := \dfrac{\pi_i}{||\pi_i||_1}$ e $A = (a_{ij})$ come $a_{ij} = \phi_i^T P_{ij}(1)$ per $i, j = 1...N$

Lemma 4.6.1. *A definita come sopra è stocastica irriducibile*

Dimostrazione. $A \geq 0$ evidentemente per come è definita, $\sum_j a_{ij} = \phi_i^T \left(\sum_j P_{ij}\right)(1) = 1$ perchè P stocastica per ipotesi e $||\phi_i||_1 = 1$. Vediamo l'irriducibilità: $a_{ij} = 0 \Leftrightarrow \phi_i^T P_{ij}(1) = 0 \Leftrightarrow P_{ij} = 0 \Rightarrow A$ irriducibile perchè lo è P $\square$

Per Perron-Frobenius allora $\exists \xi > 0$ tale che $\xi^T A = \xi^T$ e $\xi^T(1) = 1$

Teorema 4.6.2. *π vettore invariante per P è definito da*

$$\pi^T = \begin{pmatrix} \xi_1 \phi_1^T & ... & \xi_N \phi_N^T \end{pmatrix}$$

Definiamo

$$S_i := P_{ii} + \begin{pmatrix} * & * & * \end{pmatrix} \left(I - \hat{P}_i\right) \begin{pmatrix} * \\ * \\ * \end{pmatrix}$$

dove $\begin{pmatrix} * & * & * \end{pmatrix}$ *è la i-esima riga a blocchi togliendo l'elemento diagonale di* P *e* $\begin{pmatrix} * \\ * \\ * \end{pmatrix}$ *è la i-esima colonna a blocchi togliendo l'elemento diagonale di* P*, chiamiamo inoltre* $\hat{P}_i$ *la sottomatrice principale di* P *ottenuta togliendo la i-esima riga e la i-esima colonna, abbiamo ora*

$$\phi_i^T S_i = \phi_i^T$$

Dimostrazione. $\xi^T A = \xi^T$, $\xi^T(1) = 1$ vediamo che $\xi^T = \begin{pmatrix} ||\pi_1|| & ... & ||\pi_N|| \end{pmatrix}$

$$\xi^T A = \xi^T \begin{pmatrix} \phi_1^T & & 0 \\ & & \\ 0 & & \phi_N^T \end{pmatrix} P \begin{pmatrix} 1 & & \\ & & \\ & & 1 \end{pmatrix} = \begin{pmatrix} \pi_1^T & ... & \pi_N^T \end{pmatrix} P \begin{pmatrix} & \\ & \\ & \end{pmatrix} = \pi^T P() = \begin{pmatrix} ||\pi_1|| & ... & ||\pi_N|| \end{pmatrix}$$

□

In sintesi i metodi IAD consistono dei seguenti passi

- Calcolo delle S_i
- Calcolo delle ϕ come $\phi_i^T S_i = \phi_i^T$
- costruzione di A
- calcolo di ξ come $\xi^T A \xi^T$, $\xi^T(1) = 1$
- $\pi^T = \begin{pmatrix} \xi_1 \phi_1^T & ... & \xi_N \phi_N^T \end{pmatrix}$

Possibili difficoltà computazionali: n inversioni al punto 1

Nota 4.6.3. Per ovviare a queste difficoltà si implementa un metodo iterativo

Nota 4.6.4. Normalmente si utilizza questo metodo per matrici NCD (nearly completely decomposible)

Capitolo 5

Matrici M/G/1

5.1 Definizione e generalità

Il termine **M/G/1** si riferisce alle matrici stocastiche di tipo $P = \begin{pmatrix} B_0 & B_1 & B_2 & ... & & \\ A_{-1} & A_0 & A_1 & ... & & \\ 0 & A_{-1} & A_0 & A_1 & ... & \\ 0 & 0 & A_{-1} & A_0 & A_1 & ... \\ 0 & 0 & 0 & ... & & \end{pmatrix}$

che modellizzano code markoviane (Markovian (queue) , General (server), 1= numero di code). La matrice sopra ha forma di Hessemberg superiore a blocchi e Toeplitz a blocchi. Considereremo matrici in cui i blocchi sono della forma $A_i, B_i \in M_{m,m}$, in questo caso lo spazio degli eventi è $E = \{(i, j), i \in \mathbb{N}\ j \in \{1, .., m\}\}$
Assumiamo come ipotesi di comodo P irriducibile e $A := \sum_{i \geq -1} A_i$ irriducibile (convergente in quanto P stocastica)

Definizione 5.1.1. Chiamiamo la funzione polinomiale matriciale $S(z) := \sum_{i \geq -1} z^i A_i$ serie di Laurent e $A(z) := \sum_{i \geq -1} z^{i+1} A_i = zS(z)$ serie di potenze associata

Nota 5.1.2. $A(z)$ è convergente per $|z| \leq 1$ e analitica per $|z| < 1$

Definizione 5.1.3. Definiamo formalmente il drift di una M.C: chiamiamo $\underline{a} := \sum_{i \geq -1} iA_i(1)$ e α tale che $\alpha^T A = \alpha^T$, $\alpha^T(1) = (1)$, chiamiamo $\mu := \alpha^T \underline{a}$ il drift della catena di Markov

Nota 5.1.4. Abbiamo osservato nel primo capitolo che nel caso

$$P = \begin{pmatrix} b & & & \\ a_{-1} & a_0 & a_1 & 0 \\ 0 & a_{-1} & a_0 & a_1 \\ 0 & 0 & a_{-1} & ... \end{pmatrix}$$

Siamo nel caso di catena positivo ricorrente se e solo se $a_{-1} > a_1$

Teorema 5.1.5. *Supponiamo che* $\sum_{i\geq -1}(i+1)A_i < \infty$ *(ovvero la derivata prima esiste ed analitica nel disco unitario aperto) allora valgono le seguenti proprietà*

- *se* $\mu \leq 0$ *allora MC è ricorrente*
- *se* $\mu > 0$ *allora MC è transiente*
- *se* $\mu < 0$ *e* $\sum_j jB_j$ *convergente allora MC positivo ricorrente*

Nota 5.1.6. Nel caso scalare che abbiamo visto in precedenza abbiamo $\alpha = 1$, $a = -a_{-1} + a_1$, $\mu = -a_{-1} + a_1 = a$

5.2 Soluzione dell'equazione matriciale

Consideriamo ora l'equazione matriciale $X = \sum_{i\geq -1}^{\infty} A_i X^{i+1}$ con $A = \sum A_i$ stocastica e $A_i \geq 0$

Nota 5.2.1. In generale potrebbe non esistere soluzione o non essere unica nel caso di una generica equazione nella forma sopra anche nel caso di un numero finito di indici:

Esempio 5.2.2. $X^2 = Id$ *ha come soluzione una qualsiasi matrice simmetrica ortogonale*
$X^2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ *non ha soluzione*

Teorema 5.2.3. *Le ipotesi* $A = \sum A_i$ *stocastica e* $A_i \geq 0$ *ci garantiscono l'esistenza di una soluzione sottostocastica ed in particolare esiste una matrice* G *che risolve l'equazione* $X = \sum_{i=-1}^{\infty} A_i X^{i+1}$ *minimale tra le soluzioni non negative, ovvero se* $Y \geq 0$ *e* $Y = \sum_{i=-1}^{\infty} A_i Y^{i+1}$ *allora* $G \leq Y$. *Inoltre* G *è sottostocastica ed è stocastica sse* $\mu \leq 0$ *(caso ricorrente)*

Nota 5.2.4. $Xv = (\sum A_i X^{i+1})v$ e $Xv = \lambda v$ allora $\lambda v = \sum \lambda^{i+1} A_i v \Rightarrow (\lambda Id - \sum \lambda^{i+1} A_i)v = 0$ $\Rightarrow det(zId - S(z)) = 0$

Dimostrazione. (Natural iteration) Definiamo $\{X_n\}$ come segue

$$\begin{cases} X_0 = 0 \\ X_{n+1} = \sum_{i\geq -1} A_i X_n^{i+1} \end{cases}$$

vediamo che $X_n \geq X_{n-1}$ e $X_n(1) \leq (1)$. Se dimostro questo infatti la successione X_n è limitata superiormente ed esiste perciò $\lim_{n\to\infty} X_n = G$. Procediamo per induzione su n
$X_1 = A_{-1} \geq 0 = X_0$, supponiamo allora $X_n \geq X_{n-1}$, allora $X_n^j \geq X_{n-1}^j$ per cui posso scrivere $X_{n+1} = \sum A_i X_n^{i+1} \geq \sum A_i X_{n-1}^{i+1} = X_n$
Vediamo ora la sottostocasticità
$X_0(1) = (0) \leq (1)$ evidentemente, supponiamo ora $X_n(1) \leq (1)$

$$X_{n+1}(1) = \sum A_i X_n^{i+1}(1) \leq \left(\sum A_i\right)(1) = (1)$$

Vediamo adesso l'unicità: Sia $Y \geq 0$ soluzione, $G = \lim_{n\to\infty} X_n$. Ora siccome la successione X_n è crescente vediamo che $Y \geq X_n$ per ogni n. Vero banalmente per $n = 0$, supponiamo $Y \geq X_n$. Allora $X_{n+1} = \sum A_i X_n^{i+1} \leq \sum A_i Y^{i+1} = Y \Rightarrow Y \geq X_{n+1}$.

G è evidentemente sottostocastica, vediamo adesso che se $\mu \leq 0 \Rightarrow G$ stocastica
Provo a costruire una successione di matrici stocastiche. Prendiamo $X_0 = S$ stocastica, come prima $X_{n+1} = \sum A_i X_n^{i+1}$. Tutti i membri di questa successione sono stocastici, infatti se X_n stocastico allora $X_{n+1}(1) = \sum A_i X_n^{i+1}(1) = \sum A_i(1) = (1)$, inoltre $X_n \geq Y_n$ $(Y_0 = 0)$. Siccome $||X_n||_\infty = 1$ $\forall n \Rightarrow$ esiste una sottosuccesione convergente. Sia allora X^* punto di accumulazione per X_n, ora però $X^* \geq \lim_{n\to\infty} Y_n = G$

□

Nota 5.2.5. Altri possibili metodi di iterazione per il problema sopra sono i seguenti

Traditional Iteration

$$\begin{cases} X_0 = 0 \\ X_{n+1} = (I - A_0)^{-1} \left(A_{-1} + \sum_{i\geq 1} A_i X_n^{i+1}\right) \end{cases}$$

U-based Iteration

$$\begin{cases} X_0 = 0 \\ X_{n+1} = \left(I - \sum_{i\geq 0} A_i X_n^{i+1}\right)^{-1} A_{-1} \end{cases}$$

Continuo ad avere non negatività, posso perciò dimostrare la convergenza monotona come nel caso di natural iteration, osserviamo inoltre che si è dimostrata la convergenza globale
Nel caso di U-based iteration si può riscrivere la successione come

$$X_{n+1} = \sum A_i X_n^i X_{n+1} ...$$

Proposizione 5.2.6. *Riguardo i 3 metodi di iterazioni descritti nella nota sopra valgono le seguenti disuguaglianze*

$$0 \leq X_n^{(N)} \leq X_n^{(T)} \leq X_n^{(U)} \leq G$$

e inoltre la convergenza è circa lineare in tutti e 3 i casi

Dimostrazione. Dimostriamo che $X_n^{(N)} \leq X_n^{(T)}$ per induzione su n. Per $n = 0$ ok perchè abbiamo che vale l'uguaglianza, supponiamo quindi $X_n^{(N)} \leq X_n^{(T)}$. Per definizione abbiamo

$$X_{n+1}^{(T)} = (I - A_0)^{-1} \left(A_{-1} + \sum_{i \geq 1} A_i (X_n^{(T)})^{i+1} \right) = (I + A_0(I - A_0)^{-1}) \left(A_{-1} + \sum_{i \geq 1} A_i (X_n^{(T)})^{i+1} \right) =$$
$$A_{-1} + \sum A_i (X_n^{(T)})^{i+1} + A_0 X_{n+1}^{(T)} \geq A_{-1} + \sum_{i \geq 1} A_i (X_n^{(N)})^{i+1} + A_0 X_n^{(N)} = X_{n+1}^{(N)}$$

$\square$

5.2.1 Stima dell'errore

Consideriamo adesso il metodo di iterazione naturale e poniamo quindi $X_n^{(N)} := X_n$, definiamo la matrice errore

$$E_n := G - X_n$$

Osserviamo che per la convergenza monotona delle X_n si ha sempre $E_n \geq 0$, consideriamo ora il seguente calcolo

$$\begin{aligned} E_{n+1} = G - X_{n+1} &= \sum_{i \geq -1} A_i G^{i+1} - \sum_{i \geq -1} A_i X_n^{i+1} \\ &= \sum_{i \geq 0} A_i (G^{i+1} - X_n^{i+1}) = \sum_{i \geq 0} A_i \sum_{j=0}^{i} G^j (G - X_n) X_n^{i-j} = \sum_{i \geq 0} A_i \sum_{j=0}^{i} G^j E_n X_n^{i-j} \end{aligned}$$

notiamo in particolare che c'è una relazione di tipo lineare tra E_n e E_{n+1}

Definiamo adesso la quantità $\varepsilon_n = E_n(1)$, abbiamo

$$\begin{aligned} \varepsilon_{n+1} = E_{n+1}(1) &= \left(\sum_{i \geq 0} A_i \sum_{j=0}^{i} G^j (G - X_n) X_n^{i-j} \right) (1) \\ &\leq \sum_i A_i \sum_j G^j E_n \, (1) = \left(\sum A_i \sum_j G^j \right) \varepsilon_n := R \varepsilon_n \end{aligned}$$

dove la quantità R non dipende da n, le disuguaglianze sopra ci dicono inoltre che possiamo stimare l'errore al passo n come

$$0 \leq \varepsilon_n \leq R^n \varepsilon_0$$

in particolare (ricordiamo che R è una matrice)

$$\lim_{n \to \infty} \sqrt[n]{|\varepsilon_n|} \leq \rho(R)$$

Abbiamo inoltre

$$R = A_0 + A_1(I + G) + A_2(I + G + G^2) ... = \sum_{i \geq 0} A_i^*$$

dove $A_i^* = \sum_{j=i}^{\infty} A_j G^{i-j}$

Nota 5.2.7. Si ottengono risultati simili per i metodi (T) e (U)

$$\varepsilon_n^{(T)} \leq R^{(T)n} \varepsilon_0^{(T)}$$
$$\varepsilon_n^{(U)} \leq R^{(U)n} \varepsilon_0^{(U)}$$

Inoltre

$$R^{(T)} = (M^{(T)})^{-1} N^{(T)}$$
$$R^{(U)} = (M^{(U)})^{-1} N^{(U)}$$

dove

$$M^{(U)} - N^{(U)} = M^{(T)} - N^{(T)} = I - \sum_{i \geq 0} A_i^* = M^{(N)} - N^{(N)} = I - R^{(N)}$$

e sono tutti regular splitting

Nota 5.2.8. Risulta

$$\rho(R^{(U)}) \leq \rho(R^{(T)}) \leq \rho(R^{(N)})$$

e in particolare nel caso $\mu = 0$ si ha

$$\rho(R^{(U)}) = \rho(R^{(T)}) = \rho(R^{(N)}) = 1$$

e quindi convergenza sublineare

Nota 5.2.9. Se il drift $\mu < 0$ allora $\rho(R) < 1$

Nota 5.2.10. (Interpretazione probabilistica di G)

$$G = A_{-1} + A_0 G + A_1 G^2 + ...$$

$G_{ij} = \{(1, i) \rightsquigarrow (0, j)$ e $(0, j)$ è il primo che visito nel livello 0 $\}$

Notiamo inoltre che per la struttura Toeplitz di

$$\begin{pmatrix} B_0 & B_1 & B_2 & \\ A_{-1} & A_0 & A_1 & 0 \\ 0 & A_{-1} & A_0 & A_1 \\ 0 & 0 & A_{-1} & ... \end{pmatrix}$$

è equivalente a scendere di un qualsiasi livello $k \rightsquigarrow k - 1$

5.2.2 Successione stocastica

Prendiamo adesso in esame $\mu \leq 0 \Rightarrow G(1) = (1)$ (ovvero G stocastica), proviamo adesso a costruire una successione di matrici stocastiche partendo da $Y_0 = S$ stocastica. Abbiamo la seguente situazione (natural iteration):

$$\begin{cases} Y_0 \geq 0 \\ Y_n(1) = (1) \end{cases}$$

$$Y_{n+1}(1) = \sum A_i Y_n^{i+1}(1) = 1$$

inoltre

$$Y_n \geq X_n$$

(X_n successione generata allo stesso modo partendo da $X_0 = (0)$). Notiamo adesso che $Y_n(1) = (1)$ e $||Y_n||_\infty = 1$ $\forall n \Rightarrow Y_n$ ammette una sottosuccessione convergente, sia allora Y^* un punto di accumulazione,

$$Y^* \geq \lim_{n\to\infty} X_n = G$$

dal momento che G e Y^* sono entrambe stocastiche $\Rightarrow Y^* = G$.

Vediamo adesso cosa possiamo dire sulla matrice errore generata in questo caso

$$E_n = G - Y_n$$

non è necessariamente ≥ 0 a differenza del caso che abbiamo visto in precedenza, inoltre

$$E_n(1) = (1) - (1) = (0)$$

ovvero in questo spazio la soluzione è già approssimata, abbiamo inoltre una riduzione asintotica dell'errore del tipo

$$\lim_{n\to\infty} \sqrt[n]{||E_n||} \leq \rho_2(R)$$

dove $\rho_2(R)$ rappresenta il modulo del secondo autovalore di modulo massimo

5.3 Localizzazione degli autovalori di G

Osserviamo che se G risolve l'equazione $Gv = \lambda v$ e $G = \sum_{i\geq -1} A_i G^{i+1}$ allora abbiamo ($m = dimG$)

$$\lambda v = \sum A_i \lambda^{i+1} v \Rightarrow \left(\lambda Id - \sum A_i \lambda^{i+1}\right) v = 0$$

in particolare ponendo $a(z) = det\left(zId - \sum A_i z^{i+1}\right)$ abbiamo $a(\lambda) = 0$. Notiamo inoltre che $a(z)$ ha almeno m zeri nel disco unitario chiuso $\{|z| \leq 1\}$ contati con molteplicità (autovalori di G)

Nota 5.3.1. Osserviamo che $a(z)$ ben definito (serie convergente) su un disco $\{|z| < r\}$, $r > 1$.

Teorema 5.3.2. *$\mu \leq 0 \Rightarrow a(z)$ ha $m - k$ zeri di modulo < 1 e k di modulo 1 che coincidono con le radici k-esime dell'unità. Tale k dipende dalle proprietà di ciclicità di G. In particolare se $\mu = 0$ $a(z)$ i k zeri di modulo 1 hanno ciascuno molteplicità algebrica 2.*

Teorema 5.3.3. *$\mu > 0 \Rightarrow a(z)$ ha m zeri di modulo < 1 e k zeri di modulo 1 che sono le radici k-esime dell'unità*

Esempio 5.3.4. *(Possibile metodo per calcolare G) Supponiamo di avere gli autovalori $\lambda_1, ..., \lambda_n$ con $\lambda_i \neq \lambda_j$, proviamo a risolvere il problema omogeneo*

$$\left(\lambda_j I - \sum A_i \lambda_j^{i+1}\right) v_j = 0$$

e poniamo

$$V = \begin{pmatrix} v_1 & & v_n \end{pmatrix}$$

$$GV = V \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & ... & \\ 0 & & & \lambda_n \end{pmatrix}$$

per cui

$$G = V \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & ... & \\ 0 & & & \lambda_n \end{pmatrix} V^{-1}$$

abbiamo però un problema in quanto la matrice V potrebbe essere mal condizionata

I teoremi sopra ci dicono in particolare che se G è soluzione dell'equazione matriciale

$$X = \sum A_i X^{i+1}$$

$\lambda_1, ..., \lambda_m$ sono gli m zeri più piccoli in modulo di $a(z)$, dove ricordiamo che

$$a(z) = det\left(zI - \sum_{i \geq -1} A_i z^{i+1}\right)$$

Teorema 5.3.5 (Gail-Hantler-Taylor)**.** *Nelle ipotesi sopra (caso $\mu < 0$) esiste $\xi \in \mathbb{R}$ tale che $1 < \xi < r$, $a(\xi) = 0$, $\xi = min\{1 < |z| < r : a(z) = 0\}$*
In particolare $a(z)$ non si annulla nell'anello $\{1 < |z| < \xi\}$, inoltre per $\mu \to 0^-$ abbiamo $\xi \to 1^+$

Nota 5.3.6. Risulta che ξ è l'unico valore $(1 < \xi < r)$ tale che

$$\rho(A(\xi)) = \xi$$

Consideriamo adesso di nuovo

$$A(z) = \sum_{i=-1} A_i z^{i+1}$$

(mettiamoci ancora nel caso $\mu < 0$) e facciamo delle considerazioni su $\rho(A(z))$ $(0 < z < r)$. Nelle notazioni precedenti abbiamo che se λ è autovalore di G allora $A(\lambda) - \lambda Id = 0$

Nota 5.3.7. Abbiamo che $z_1 > z_2 \Rightarrow A(z_1) > A(z_2)$ per come sono definiti ed il teorema di Perron-Frobenius ci dice che allora $\rho(A(z_1)) > \rho(A(z_2))$

5.4 Fattorizzazione di Wiener-Hopf

consideriamo

$$\varphi(z) = I - \sum_{i \geq -1} z^i A_i = \left(Id - \sum z^i A_i^*\right)\left(Id - z^{-1}G\right) = U(z)L(z^{-1})$$

(verifica diretta mediante calcoli) dove $A_i^* = \sum_{j \geq i} A_j G^{j-i}$

Osserviamo che $detU(z) \neq 0$ se $|z| \leq 1$, $detL(z^{-1}) \neq 0$ se $|z| \geq 1$ e $detL(z^{-1}) = 0$ se z autovalore di G

Le matrici associate a questa fattorizzazione sono

$$\begin{pmatrix} I - A_0 & -A_1 & A_2 & \\ -A_{-1} & I - A_0 & A_1 & \\ ... & & & \\ & & & I - A_0 \end{pmatrix} = U(z)L(z^{-1}) = \begin{pmatrix} I - A_0^* & -A_1^* & \\ & I - A_0^* & \\ & & \\ 0 & & \end{pmatrix} \begin{pmatrix} I & & 0 \\ -G & I & \\ & & \end{pmatrix}$$

Calcolo di π

$$\begin{pmatrix} \pi_0^T & \pi_2^T & ... & \end{pmatrix} \begin{pmatrix} I - B_0 & -B_1 & -B_2 & \\ -A_{-1} & I - A_0 & & \\ 0 & & & \\ 0 & & & \end{pmatrix} = 0$$

chiamo adesso $H = \begin{pmatrix} I - A_0 & & \\ & I - A_0 & \\ & & I - A_0 \end{pmatrix}$

supponendo di conoscere π_0 possiamo allora scrivere

$$\begin{pmatrix} \pi_1^T & & ... \end{pmatrix} H = \pi_0^T \begin{pmatrix} B_1 & B_2 & ... \end{pmatrix}$$

moltiplichiamo l'espressione sopra a destra per $L^{-1} = \begin{pmatrix} I & & & 0 \\ G & I & & \\ G^2 & G & I & \\ ... & & & I \end{pmatrix}$

e otteniamo l'espressione

$$\begin{pmatrix} \pi_1^T & .. & \end{pmatrix} U = \pi_0^T \begin{pmatrix} B_1 & B_2 & ... \end{pmatrix} L^{-1}$$

Possiamo applicare adesso la sostituzione in avanti[1]

$$\pi_1^T = \pi_0^T B_1^*(I - A_0^*)^{-1}$$

$$\pi_i^T = \left(\sum_{j=1}^{i-1} \pi_j^T A_{i-j}^* + \pi_0^T B_i^* \right) (I - A_0^*)^{-1}$$

dove $B_i^* = \sum_{j=i}^{\infty} B_j G^{j-i}$

Nota 5.4.1. G stocastica $\Rightarrow G^i$ stocastica

Ci resta quindi solo da ricavare π_0. L'idea è utilizzare l'eliminazione gaussiana

5.4.1 Calcolo di π_0

Abbiamo che π_0 risolve il problema

$$\pi_0^T \left(I - B_0 - \begin{pmatrix} B_1 & B_2 & ... \end{pmatrix} H^{-1} \begin{pmatrix} A_{-1} \\ 0 \\ ... \end{pmatrix} \right) = 0$$

che possiamo riscrivere come

$$\pi_0^T (I - \tilde{P}) = 0$$

dove $\tilde{P}$ è stocastica e si scrive come

$$\tilde{P} = B_0 + \begin{pmatrix} B_1 & B_2 & ... \end{pmatrix} L^{-1} U^{-1} \begin{pmatrix} A_{-1} \\ 0 \\ ... \end{pmatrix}$$

e H è fattorizzata come $H = UL$, dove U triangolare superiore e L triangolare inferiore, ricordiamo che

[1] Ramaswamj

$$H = UL = \begin{pmatrix} I-A_0^* & -A_1^* & \dots & \\ & I-A_0^* & \dots & \\ & & I-A_0^* & \\ 0 & & & I-A_0^* \end{pmatrix} \begin{pmatrix} I & & & 0 \\ -G & I & & \\ & -G & I & \\ & & -G & I \end{pmatrix}$$

risultà perciò $L^{-1} = \begin{pmatrix} I & & & 0 \\ G & I & & \\ G^2 & & I & \\ \dots & & & I \end{pmatrix}$

Osserviamo adesso che di U^{-1} ci interessa in effetti solo la prima colonna, ovvero la parte che interessa A_{-1}, quindi

$$U^{-1} \begin{pmatrix} A_{-1} \\ 0 \\ \dots \end{pmatrix} \rightsquigarrow \begin{pmatrix} (I-A_0^*)^{-1} \\ 0 \\ \dots \end{pmatrix} \begin{pmatrix} A_{-1} \\ 0 \\ \dots \end{pmatrix}$$

da cui

$$\tilde{P} = B_0 + \sum_{i \geq 1} B^i G^{i-1} (I - A_0^*)^{-1} A_{-1}$$

adesso dobbiamo calcolare G. Sappiamo che deve valere la relazione $\pi_0^T(1) + \sum_{i\geq 1} \pi_i^T(1) = 1$, scriviamo allora[2]

$$\begin{pmatrix} \pi_1^T & \dots & \end{pmatrix} \begin{pmatrix} 1 \\ \dots \\ 1 \end{pmatrix} = \pi_0^T \begin{pmatrix} B_1 & .. & \end{pmatrix} L^{-1} U^{-1} (1 \dots 1) = \pi_0^T \sum_{i \geq 1} B_i^* (I - A_0^*)^{-1} (1) = 1 - \pi_0^T (1)$$

da cui, chiamando $\sum_{i\geq 1} B_i^* (I - A_0^*)^{-1}(1) := r$

$$\pi_0^T (r + 1) = 1$$

Nota 5.4.2. Abbiamo che $\sum A_i$, stocastica, per fare il calcolo nella pratica si tronca A_i^* in modo che sia "abbastanza stocastica", per esempio a meno della precisione di macchina ponendo $A_i^* = \sum_i^k A_j G^{j-i}$, in particolare $A_i^* = 0$ se $i > k$, $A_k^* = A_k$, $A_i^* = A_{i+1}^* G + A_i$ per $i = k-1, ..., 0$[3]

[2] Ricordiamo che $A_i^* = \sum_{j=i}^{\infty} A_j G^{j-i}$

[3] Ruffini-Horner

5.4.2 QBD (quasi birth-death processes)

Esamiano adesso il caso in cui la matrice iniziale è tridiagonale a blocchi (QBD), prendiamo la matrice di transizione

$$P = \begin{pmatrix} B_0 & B_1 & & 0 \\ A_{-1} & A_0 & A_1 & \\ & A_{-1} & A_0 & A_1 \\ 0 & & A_{-1} & A_0 \end{pmatrix}$$

in questo caso il problema da risolvere diventa

$$G = A_{-1} + A_0 G + A_1 G^2$$

ribaltando i coefficienti troviamo

$$R = R^2 A_{-1} + R A_0 + A_1$$

che ha una soluzione minimale non negativa e risulta inoltre

$$R = \lim_{k\to\infty} X_k$$

dove X_k è la successione definita da

$$\begin{cases} X_0 = 0 \\ X_{k+1} = X_k^2 A_{-1} + X_k A_0 + A_{-1} \end{cases}$$

Tale R può essere interpretato come la matrice dei "tempi medi di attesa" nell'ambito della teoria delle code

Nota 5.4.3. Valgono i seguenti fatti

- $\rho(R) < 1 \Leftrightarrow \mu < 0$
- $\rho(R) = 1 \Leftrightarrow \mu \geq 0$

Nelle notazioni precedenti prendiamo di nuovo la fattorizzazione di Wiener-Hopf del problema come

$$I - \sum_{i=-1}^{1} z^i A_i = (I - zR)(I - U)(I - z^{-1}G) = (I - A_0^* - A_1^* z)(I - z^{-1}G) = (I - A_1^*(I - A_0^*)^{-1} z)(I - A_0^*)(I - z^{-1}G)$$

per cui
$$R = A_1^*(I - A_0^*)^{-1} = A_1(I - A_0 - A_1G)^{-1}$$

Nota 5.4.4. $det(I - z^{-1}G) = 0 \Leftrightarrow det(zI - G) = 0$, ora $\rho(G) \leq 1 \Rightarrow det(I - z^{-1}G) \neq 0$ se $z > 1$

$det(I - zR) = 0 \Leftrightarrow z^{-1} \in \sigma(R)$ ma $\rho(R) < 1 \Rightarrow det(I - zR) \neq 0$ se $|z| < 1$

Come prima scriviamo

$$\begin{pmatrix}\pi_1^T & ...\end{pmatrix}\begin{pmatrix}I - A_0 & -A_1 & & \\ -A_{-1} & I - A_0 & -A_1 & \\ & & I - A_0 & \\ & & & I - A_0\end{pmatrix} = \pi_0^T\begin{pmatrix}B_1 & ...\end{pmatrix} \Rightarrow$$

$$\begin{pmatrix}\pi_1^T & ...\end{pmatrix}\begin{pmatrix}I & -R & & \\ & I & -R & \\ & & I & -R \\ 0 & & & I\end{pmatrix}\begin{pmatrix}I - U & & & 0 \\ & I - U & & \\ & & I - U & \\ 0 & & & I - U\end{pmatrix}\begin{pmatrix}I & & 0 \\ -G & I & \\ ... & & I \\ & & & I\end{pmatrix} = \pi_0^T\begin{pmatrix}B_1 & ...\end{pmatrix} \Rightarrow$$

$$\begin{pmatrix}\pi_1^T & ...\end{pmatrix}\begin{pmatrix}I & -R & & \\ & I & -R & \\ & & I & -R \\ 0 & & & I\end{pmatrix} = \pi_0^T\begin{pmatrix}B_1 & ...\end{pmatrix}\begin{pmatrix}I & & \\ G & ... & \\ G^2 & & \\ ... & & \end{pmatrix}\begin{pmatrix}(I - U)^{-1} & & 0 \\ & (I - U)^{-1} & \\ & & ... \\ 0 & & ...\end{pmatrix} =$$

$$\pi_0^T\begin{pmatrix}B_1(I - U)^{-1} & 0 & ...\end{pmatrix}$$

Possiamo quindi ricavare i π_i come[4]

$$\begin{cases}\pi_1^T = \pi_0(B_1(I - U)^{-1}) \\ \\ \pi_i^T = \pi_{i-1}^T R \text{ per } i \geq 2 \Rightarrow \pi_i = \pi_1 R^{i-1}\ ,\ i \geq 2\end{cases}$$

per cui le componenti tendono a 0 come $\rho(R)$

Nota 5.4.5.
$$G = B_{-1}^{(0)} + \sum_{i=1}^{n}\left(\prod_{j=0}^{i-1} B_1^{(j)}\right) B_{-1}^{(i)} + \prod_{j=0}^{n} B_1^{(j)} G^{2^{n+1}}$$

dove notiamo che il terzo termine converge a 0 con velocità doppiamente esponenziale

[4] ci siamo messi nel caso $\rho(R) < 1$

Consideriamo il seguente sistema

$$\begin{pmatrix} I & -B_1^{(0)} & & 0 \\ -B_{-1}^{(0)} & I & -B_1^{(0)} & \\ & -B_{-1}^{(0)} & I & -B_1^{(0)} \\ & & -B_{-1}^{(0)} & I \end{pmatrix} \begin{pmatrix} G \\ G^2 \\ ... \end{pmatrix} = \begin{pmatrix} B_{-1}^{(0)} \\ 0 \\ ... \end{pmatrix}$$

che si può riscrivere come

$$(-B_{-1}^{(0)} + G - B_1^{(0)} G^2) G^j = 0$$

applicando una permutazione in modo da separare le potenze pari e le potenze dispari di G troviamo

$$\begin{pmatrix} Id & V \\ U & Id \end{pmatrix} \begin{pmatrix} G^3 \\ G^5 \\ ... \\ G^2 \\ G^4 \\ ... \end{pmatrix} = \begin{pmatrix} B_{-1}^{(0)} \\ 0 \\ ... \\ 0 \\ ... \end{pmatrix}$$

Dove $Id = \begin{pmatrix} I & & & 0 \\ & I & & \\ & & I & \\ 0 & & & I \end{pmatrix}$, $V = \begin{pmatrix} -B_1^{(0)} & & & 0 \\ -B_{-1}^{(0)} & -B_1^{(0)} & & \\ & -B_{-1}^{(0)} & -B_1^{(0)} & \\ 0 & & -B_{-1}^{(0)} & B_1^{(0)} \end{pmatrix}$, $U = \begin{pmatrix} -B_{-1}^{(0)} & -B_1^{(0)} & & 0 \\ & -B_{-1}^{(0)} & -B_1^{(0)} & \\ & & -B_{-1}^{(0)} & -B_1^{(0)} \\ 0 & & & -B_{-1}^{(0)} \end{pmatrix}$

da cui

$$\begin{cases} \begin{pmatrix} G^3 \\ G^5 \\ ... \end{pmatrix} + V \begin{pmatrix} G^2 \\ G^4 \\ ... \end{pmatrix} = \begin{pmatrix} B_{-1}^{(0)} \\ 0 \\ ... \end{pmatrix} \\ \\ U \begin{pmatrix} G^3 \\ G^5 \\ ... \end{pmatrix} + \begin{pmatrix} G^2 \\ G^4 \\ ... \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ ... \end{pmatrix} \end{cases} \Rightarrow \left(I - UV\right) \begin{pmatrix} G^2 \\ G^4 \\ ... \end{pmatrix} = U \begin{pmatrix} B_{-1}^{(0)} \\ 0 \\ ... \end{pmatrix} \Rightarrow$$

$$I - \begin{pmatrix} * & * & & 0 \\ & * & * & \\ & & * & * \\ 0 & & & * \end{pmatrix} \begin{pmatrix} * & & & 0 \\ * & * & & \\ & * & * & \\ 0 & & * & * \end{pmatrix} = \begin{pmatrix} (I - B_{-1}B_1 - B_1 B_{-1}) & -B_1^2 & & 0 \\ -B_{-1}^2 & ... & -B_1^2 & \\ & ... & ... & \\ 0 & & & (I - B_{-1}B_1 - B_1 B_{-1}) \end{pmatrix}$$

otteniamo quindi

$$(I - B_{-1}B_1 - B_1B_{-1})G^2 = B_1^2G^4 + B_{-1}^2$$

$$\begin{cases} B_1^{(1)} = (C^{(1)})^{-1}B_1^2 \\ B_{-1}^{(1)} = (C^{(1)})^{-1}B_{-1}^2 \end{cases}$$

dove $C^{(1)} =: I - B_{-1}B_1 - B_1B_{-1}$, l'espressione per i coefficienti B,C,G diventa (verifica mediante calcoli diretti)

$$\begin{cases} G^{2k} = B_{-1}^{(k)} + B_1^{(k)}G^{2^{k+1}} \\ \\ C^{(k)} = I - B_{-1}^{(k)}B_1^{(k)} - B_1^{(k)}B_{-1}^{(k)} \\ \\ B_{-1}^{(k+1)} = (C^{(k)})^{-1}B_{-1}^{(k)2} \\ \\ B_1^{(k+1)} = (C^{(k)})^{-1}B_1^{(k)2} \end{cases}$$

e ricordando che $G = B_{-1}^{(0)} + \sum_{i=1}^{n}\left(\prod_{j=0}^{i-1} B_1^{(j)}\right) B_{-1}^{(i)} + \prod_{j=0}^{n} B_1^{(j)} G^{2^{n+1}}$ troviamo

$$\begin{cases} U^{(0)} = I \\ \\ U^{(k+1)} = U(k)B_1(k+1) \\ \\ G^{(0)} = B_{-1}^{(0)} \\ \\ G^{k+1} = G^k + U^{(k)} + U^{(k)}B_{-1}^{(k+1)} \end{cases}$$

inoltre

$$\begin{cases} B_{-1}^{(0)} = (I - A_0)^{-1}A_{-1} \geq 0 \\ \\ B_1^{(0)} = (I - A_0)^{-1}A_1 \geq 0 \end{cases}$$

vale $(B_{-1}^{(0)} + B_1^{(0)})(1) = (1)$[5], $C^{(k)}$ risulta essere una M-matrice non singolare e troviamo inoltre

$$G^{k+1} \geq G^k$$

$$G^k(1) \leq (1)$$

[5] questa relazione è in effetti valida al generico passo (k)

5.5 Complementi

Teorema 5.5.1. *Supponiamo*[6] $\mu < 0$, $||.||$ *norma matriciale indotta,* $\varepsilon > 0$ *tale che* $\varepsilon + \eta < 1$ *dove* $a(z) = det(zI - (A_{-1} + zA_0 + z^2A_1))$ *e* $\eta = max\{|z|$ *tale che* $|z| < 1$ *e* $a(z) = 0\}$, *allora esiste* $\gamma > 0$ *tale che*

$$||B_1^{(k)}|| \leq \gamma\xi^{-2k}$$

dove $\xi = min\{|z|$ *tale che* $|z| < 1$ *e* $a(z) = 0\}$, *inoltre posto* g *tale che* $\begin{cases} g^TG = g^T \\ g^T(1) = (1) \end{cases}$ *abbiamo che*[7]

$$||B_{-1}^{(k)} - (1)g^T|| \leq \gamma(\xi^{-2^k} + (\eta + \varepsilon)^{2^k})$$

$$||I - C^{(k)}|| \leq \gamma\xi^{-2^k}$$

$$||G - G^k|| \leq \gamma\xi^{-2^k}$$

Nota 5.5.2. $\varphi(z) = A_{-1} + zA_0 + z^2A_1 \rightsquigarrow R = R^2A_{-1} + RA_0 + A_1$ allora $\xi = \rho(R)^{-1}$

Dimostrazione. (idea) Si verifica direttamente che

$$R^{2^k} = R^{2^{k+1}}B_{-1}^{(k)} + B_1^{(k)}$$

per cui

$$B_1^{(k)} = R^{2^k} - R^{2^{k+1}}B_{-1}^{(k)} \leq R^{2^k}$$

ora $B_{-1}^{(k)} \geq 0$, $R \geq 0$ per cui

$$||B_1^{(k)}||_\infty \leq ||R^{2^k}||_\infty \longrightarrow 0$$

in quanto siamo nel caso $\rho(R) < 1$ ed esiste una norma matriciale indotta tale che[8] $||R||_* = \rho(R) = \xi^{-1}$ per cui

$$||R^{2^k}||_\infty \leq cost(\xi^{-1})^{2^k} = cost(\xi^{-2^k})$$

analogamente abbiamo in questa norma

$$cost_2||B_1^{(k)}||_* \leq ||B_1^{(k)}||_\infty \leq cost||B_1^{(k)}||_*$$

$$B_{-1}^{(k)} = G^{2^k} - B_1G^{2^{k+1}}$$

$$B_{-1}^{(k)} - (1)g^T = G^{2^k} - (1)g^T - B_1G^{2^{k+1}}$$

[6] ci mettiamo nelle notazioni usate in precedenza
[7] ricordiamo che in questo caso $G = A_{-1} + A_0G + A_1G^2$
[8] ricordiamo che $\rho(R)$ è autovalore semplice

$$(1)g^T = \lim_{n\to\infty} G^n$$

da cui

$$||B_{-1}^{(k)} - (1)g^T|| \leq ||G^{2^k} - (1)g^T|| + ||B_1 G^{2^{k+1}}||$$

$||G^{2^k} - (1)g^T|| \longrightarrow 0$ come $(\eta+\varepsilon)^{2^k}$ e $||B_1G^{2^{k+1}}|| \longrightarrow 0$ come ξ^{-2^k} per cui abbiamo la disuguaglianza richiesta, inoltre

$$||G - G^k|| = ||\prod_{j=0}^{k} B_1^{(j)} G^{2^{k+1}}|| \leq ||\prod_{j=0}^{k} R^{2j} G^{2^{k+1}}|| = ||R^{2^{k+1}-1} G^{2^{k+1}}||$$

dove $G^{2^{k+1}}$ è limitato e $||R^{2^{k+1}-1}|| \longrightarrow 0$ come ξ^{-2^k}

per cui abbiamo provato la convergenza doppiamente esponenziale

□

Nota 5.5.3. Il costo computazionale delle operazioni sopra descritte è quantificato da 8 moltiplicazioni di matrici ed una inversione

Nota 5.5.4. Si può dimostrare che nel caso $\mu = 0$ la convergenza diventa lineare

Nota 5.5.5. Le iterazioni del tipo
$$\begin{cases} X_0 \\ X_n = F(X_{n-1}) \end{cases}$$
sono self-contracting: gli errori numerici sono cancellati dalla nuova applicazione di F, nell'algoritmo utilizzato in precedenza (L-R) invece l'errore di arrotondamento si accumula

5.5.1 Ciclic-reduction (accenno)

Consideriamo di nuovo il caso di matrice QBD tridiagonale a blocchi, come abbiamo visto dobbiamo risolvere il sistema

$$G = A_{-1} + A_0 G + A_1 G^2$$

che riscriviamo come $-A_{-1} + (I - A_0)G - A_1G^2 = 0$, consideriamo di nuovo la permutazione

$$\begin{pmatrix} (I-A_0) & -A_1 & & 0 \\ -A_{-1} & (I-A_0) & \dots & \\ & & (I-A_0) & \\ 0 & & & (I-A_0) \end{pmatrix} \begin{pmatrix} G \\ G^2 \\ \dots \end{pmatrix} \begin{pmatrix} G \\ G^2 \\ G^3 \\ \dots \end{pmatrix} \rightsquigarrow \begin{pmatrix} U & V \\ T & S \end{pmatrix} \begin{pmatrix} G^2 \\ G^4 \\ \dots \\ G \\ G^3 \\ \dots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ A_{-1} \\ 0 \\ \dots \end{pmatrix}$$

dove $S = U = \begin{pmatrix} (I - A_0) & 0 \\ 0 & (I - A_0) \end{pmatrix}$, $V = \begin{pmatrix} -A_{-1} & -A_1 \\ 0 & -A_{-1} \end{pmatrix}$, $T = \begin{pmatrix} -A_1 & 0 \\ -A_{-1} & -A_1 \end{pmatrix}$

e otteniamo[9]

$$S - TU^{-1}V \begin{pmatrix} G \\ G^3 \\ \dots \end{pmatrix} = \begin{pmatrix} A_{-1} \\ 0 \\ \dots \end{pmatrix}$$

[9] complemento di Schur