



UNIVERSITÀ DEGLI STUDI DI FERRARA

**FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E
NATURALI**

Corso di Laurea Triennale in Matematica
Indirizzo Matematica Applicata

**CATENE DI MARKOV FINITE E
TEORIA DI PERRON-FROBENIUS**

Relatore:
**Chiar.mo Prof.
Josef Eschgfäller**

Laureanda:
Elena Testoni

Anno Accademico 2008-2009

Indice

Introduzione	1
I. CATENE DI MARKOV FINITE	
1. Variabili aleatorie elementari	5
2. Il teorema di Radon-Nikodým discreto	14
3. Regole per probabilità composte	19
4. Sequenze di Markov	20
5. Catene di Markov finite	25
6. La matrice di transizione	29
7. Esempi di catene di Markov	31
8. Un programma di simulazione	38
9. Una formula per le potenze	40
10. Il teorema ergodico matriciale	45
II. MATRICI NON NEGATIVE	
11. Sistemi dinamici stocastici finiti	59
12. Una stima per il raggio spettrale	63
13. Il teorema di Perron	67
14. Matrici irriducibili	73
15. Il teorema di Frobenius	77
16. Il teorema di Wielandt	79
17. Matrici primitive	82
18. Un modello demografico	91
Bibliografia	92

Introduzione

La tesi è suddivisa in due parti: la prima dedicata alla teoria ergodica delle catene di Markov finite, la seconda allo studio delle matrici non negative.

Nei primi tre capitoli vengono forniti alcuni richiami alla teoria della probabilità, con la definizione di σ -algebra, variabile aleatoria elementare e alcune proprietà che le caratterizzano. Viene anche enunciato e dimostrato il teorema di Radon-Nikodým discreto per la media condizionata di una variabile aleatoria elementare.

Il quarto capitolo introduce le sequenze di Markov, un concetto ausiliario per rendere più trasparenti i ragionamenti probabilistici usati nei capitoli seguenti.

Nel quinto capitolo vengono introdotte le catene di Markov finite, ossia processi stocastici finiti per i quali l'applicazione $\bigcirc_{n,i} (X_n = i)$

è una sequenza di Markov.

Nel capitolo 6 vengono approfondite alcune proprietà della matrice di transizione T con cui può essere descritta una catena di Markov finita. Per il teorema di Kolmogorov (di cui una dimostrazione si trova nella monografia di Woess) per ogni matrice stocastica T ed ogni vettore stocastico p_0 esiste una catena di Markov $\bigcirc_n X_n$ finita ed omogenea tale che la matrice di transizione associata sia T e tale che l'elemento i -esimo di p_0 equivale alla probabilità che X_0 assuma tale valore i .

Alcuni esempi di catene di Markov e un programma di simulazione sono contenuti nei capitoli 7 e 8.

Il capitolo 9 contiene una formula che ci permette di calcolare la potenza n -esima di una matrice qualunque utilizzando il polinomio di interpolazione di Hermite calcolato mediante lo schema alle differenze. Questa formula, numericamente instabile e complicata, è più di interesse teorico, e in pratica per le matrici stocastiche e non negative si useranno invece le formule al limite che deriveremo nel seguito.

Nel capitolo 10, attraverso diversi risultati dell'algebra lineare numerica, si studiano le proprietà di una matrice stocastica T .

1 è autovalore di T e coincide con il raggio spettrale. Il teorema ergodico matriciale (teorema 10.9) afferma che $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n = P$, dove P è la proiezione su $\text{Fix } T$, nella decomposizione $\mathbb{R}^q = \text{Fix } T \oplus \text{Im}(T - \delta)$.

In generale la successione $\bigcirc_n T^n$ stessa non converge, ma soltanto nel caso in cui 1 sia l'unico autovalore di modulo massimo (teorema 10.34). In questo caso si ha allora che $\lim_{n \rightarrow \infty} T^n = P$ (corollario 10.44). Alla fine del capitolo la teoria viene illustrata con alcuni esempi calcolati con Sage.

La seconda parte inizia con il capitolo 11 nel quale vengono definiti sistemi dinamici stocastici finiti ad ognuno dei quali viene associata in modo naturale una matrice stocastica, quindi una catena di Markov finita. Tali sistemi possono essere utilizzati per simulare il comportamento di un insieme di cellule in un tessuto che potrebbero passare da uno stato ad un altro. Facciamo vedere come da questo modello si arrivi in modo intuitivo al concetto di matrice non negativa.

Nel capitolo 12 lo studio delle proprietà delle matrici non negative inizia con alcune importanti stime del raggio spettrale. Esso è ad esempio compreso tra il minimo e il massimo delle somme di riga e tra il minimo e il massimo delle somme di colonna.

Nei capitoli successivi viene esposta la teoria di Perron-Frobenius. Questa teoria inizia con il teorema di Perron. Si dimostra che se A è una matrice positiva, allora $\rho(A)$ è un autovalore di A di molteplicità algebrica uguale a 1 e ad esso corrisponde un autovettore positivo (vettore di Perron) e che tale autovettore è unico se chiediamo che sia stocastico. Un altro risultato importante è il teorema 13.21 che permette di calcolare il $\lim_{n \rightarrow \infty} \left(\frac{1}{\rho(A)} A^n\right)$.

Nel capitolo 14, invece, vengono introdotte le matrici (non negative) irriducibili e alcune proprietà che le caratterizzano: una matrice non negativa è irriducibile se e solo se esiste $m \in \mathbb{N}$ tale che $(A + \delta)^m$ sia positiva.

Il capitolo 15 contiene il teorema di Frobenius per matrici non negative ed irriducibili. Si dimostra che se A è una tale matrice, il suo raggio spettrale $\rho(A)$ è autovalore positivo di A , con molteplicità algebrica uguale a 1 ed esiste inoltre un autovettore positivo rispetto all'autovalore $\rho(A)$ (teorema 15.4).

Il teorema di Wielandt nel capitolo 16 stabilisce che se A è una matrice non negativa ed irriducibile, l'insieme degli autovalori di modulo massimo è dato da $\mathcal{M}(A) = \{e^{\frac{2\pi i}{h} j} \rho(A) \mid j = 0, \dots, h-1\}$, dove h è la cardinalità di $\mathcal{M}(A)$. Esso stabilisce inoltre che ogni elemento di $\mathcal{M}(A)$ ha molteplicità algebrica 1 come autovalore di A .

Il capitolo 17 introduce le matrici primitive, che costituiscono un'ulteriore sottoclasse delle matrici non negative ed irriducibili, e vengono esposte le proprietà che le caratterizzano. Esse sono molto importanti poiché alle matrici primitive si generalizza in modo soddisfacente il teorema di Perron per le matrici positive, come mostrato nel teorema 17.5. Vengono introdotte anche le matrici di incidenza e, attraverso alcuni semplici programmi in Sage, elencate tutte le matrici di incidenza 2×2 e 3×3 primitive e quelle primitive minimali.

Il capitolo 18 mostra in un semplice e popolare modello (noto come modello di Leslie) come si può applicare la teoria in demografia. In esso si considera una popolazione suddivisa in fasce d'età, che però possono anche essere interpretate come compartimenti funzionali delle cellule di un tessuto. A questa situazione possono essere applicati i teoremi di convergenza ottenuti nei capitoli precedenti.

La teoria di Perron-Frobenius è stata originata da Perron nell'ambito della teoria ergodica delle frazioni continue (cfr. Iosifescu/Kraaikamp).

Si è però rivelata, come mostra il contenuto di questa tesi, utile in molti altri campi. Negli anni '60 è stata generalizzata a operatori e semigrupp di operatori positivi su reticoli di Banach; si vedano ad esempio i lavori citati di Lotz, Greiner e Keicher/Nagel, il libro di Nagel e coll. (Springer LN 1184) e l'ultimo capitolo di Engel/Nagel. Queste generalizzazioni richiedono tecniche e concetti molto avanzati dell'analisi funzionale.

A livello più elementare, ma piuttosto interessanti, si presentano le applicazioni della teoria delle matrici non negative in dinamica simbolica e combinatoria delle parole. Siano $S := \{1, \dots, q\}$ ed S^* il monoide libero delle parole formate con lettere da S . Per $w \in S^*$ e $j \in S$ sia $f(j, w)$ il numero delle volte che j appare in w .

Per ogni endomorfismo $\varphi : S^* \rightarrow S^*$ possiamo allora formare la matrice

$$T := \begin{pmatrix} f(1, \varphi(1)) & \dots & f(1, \varphi(q)) \\ \vdots & & \vdots \\ f(q, \varphi(1)) & \dots & f(q, \varphi(q)) \end{pmatrix}$$

Se con $Uw := \begin{pmatrix} f(1, w) \\ \vdots \\ f(q, w) \end{pmatrix}$ denotiamo il vettore di Parikh di w , otteniamo un diagramma commutativo

$$\begin{array}{ccc} S^* & \xrightarrow{\varphi} & S^* \\ U \downarrow & & \downarrow U \\ \mathbb{N}^q & \xrightarrow{T} & \mathbb{N}^q \end{array}$$

per cui $U\varphi^n(j) = T^n Uj = (T^n)_{(j)}$ per ogni $n \in \mathbb{N}$ ed ogni $j \in S$, essendo evidentemente $Uj = \delta_j$. Cfr. Brin/Stuck, pagg. 64-66, Queffélec, pagg. 87-95, Lind/Marcus, pagg. 106-135, per le interpretazioni nell'ambito della dinamica simbolica e Lothaire per le applicazioni in combinatoria delle parole.

La tesi contiene, come già osservato, nel capitolo 11 alcune idee per la modellazione di sistemi biologici (rappresentati ad esempio tramite reti di Petri) mediante matrici non negative. Tali modelli sono molto versatili e dovrebbero avere molte applicazioni.

I. CATENE DI MARKOV FINITE

1. Variabili aleatorie elementari

Definizione 1.1. Sia Ω un insieme. Un sistema di insiemi $\mathcal{A} \subset \mathcal{P}(\Omega)$ si chiama una σ -algebra su Ω , se sono verificate le seguenti condizioni:

- (1) $\Omega \in \mathcal{A}$.
- (2) Se $A \in \mathcal{A}$, allora $\Omega \setminus A \in \mathcal{A}$.
- (3) Se $A_1, A_2, \dots \in \mathcal{A}$, allora $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Osservazione 1.2. Sia Ω un insieme e \mathcal{A} una σ -algebra su Ω . Allora:

- (1) $\emptyset \in \mathcal{A}$.
- (2) Se $A_1, A_2, \dots \in \mathcal{A}$, allora $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.
- (3) Se $A_1, \dots, A_k \in \mathcal{A}$, allora $\bigcup_{n=1}^k A_n \in \mathcal{A}$ e $\bigcap_{n=1}^k A_n \in \mathcal{A}$.
- (4) Se $A, B \in \mathcal{A}$, allora $A \setminus B \in \mathcal{A}$.

Dimostrazione.

- (1) $\emptyset = \Omega \setminus \Omega \in \mathcal{A}$.
- (2) $\bigcap_{n=1}^{\infty} A_n = \Omega \setminus (\Omega \setminus \bigcap_{n=1}^{\infty} A_n) \in \mathcal{A}$.
- (3) Ponendo $A_n = \emptyset$ per $k > n$ abbiamo: $\bigcup_{n=1}^k A_n = \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.
- Ponendo $A_n = \Omega$ per $k > n$ abbiamo: $\bigcap_{n=1}^k A_n = \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.
- (4) $A \setminus B = A \cap (\Omega \setminus B) \in \mathcal{A}$.

Definizione 1.3. Uno spazio misurabile è una coppia (Ω, \mathcal{A}) , dove Ω è un insieme e \mathcal{A} una σ -algebra su Ω .

Definizione 1.4. Sia (Ω, \mathcal{A}) uno spazio misurabile. Una misura su \mathcal{A} è un'applicazione $\mu : \mathcal{A} \rightarrow [0, \infty]$ con le seguenti proprietà:

- (1) $\mu(\emptyset) = 0$.
- (2) Se $A_1, A_2, \dots \in \mathcal{A}$ sono disgiunti, allora $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

L'addizione in $[0, \infty]$ è definita in modo che $a + \infty = \infty$ per ogni a .
È chiaro che in questo modo sono anche definite le somme infinite.

Osservazione 1.5. Siano (Ω, \mathcal{A}) uno spazio misurabile e μ una misura su \mathcal{A} . Se $A_1, \dots, A_k \in \mathcal{A}$ sono disgiunti, allora $\mu(\bigcup_{n=1}^k A_n) = \sum_{n=1}^k \mu(A_n)$.

Dimostrazione. Siccome $\mu(\emptyset) = 0$ è sufficiente porre $A_n = \emptyset$ per $n > k$.

Definizione 1.6. Sia (Ω, \mathcal{A}) uno spazio misurabile. Una *misura di probabilità* su \mathcal{A} è una misura p su \mathcal{A} tale che $p(\Omega) = 1$.

È chiaro che allora $p(\Omega \setminus A) = 1 - p(A)$ per ogni $A \in \mathcal{A}$.

Osservazione 1.7. (Ω, \mathcal{A}) sia uno spazio misurabile e μ una misura su \mathcal{A} . Allora:

- (1) Se $A, B \in \mathcal{A}$ e $A \subset B$, allora $\mu(A) \leq \mu(B)$.
- (2) Se μ è una misura di probabilità, allora $0 \leq p(A) \leq 1$ per ogni $A \in \mathcal{A}$.

Dimostrazione.

(1) Abbiamo $B = A \dot{\cup} (B \setminus A)$ con $B \setminus A \in \mathcal{A}$.

Perciò $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.

(2) Immediato dal punto (1) perché $A \subset \Omega$ e $p(\Omega) = 1$ per ipotesi.

Definizione 1.8. Uno *spazio di misura* è una tripla $(\Omega, \mathcal{A}, \mu)$, dove (Ω, \mathcal{A}) è uno spazio misurabile con $\Omega \neq \emptyset$ e μ è una misura su \mathcal{A} .

Uno *spazio di probabilità* è una tripla (Ω, \mathcal{A}, p) , dove (Ω, \mathcal{A}) è uno spazio misurabile e p una misura di probabilità. Gli elementi di \mathcal{A} sono spesso detti *eventi*.

Osservazione 1.9. Ω sia un insieme e ρ un insieme di σ -algebre su Ω . Allora $\bigcap_{\mathcal{A} \in \rho} \mathcal{A}$ è una σ -algebra su Ω .

Dimostrazione. Immediata.

Definizione 1.10. Siano Ω un insieme, $\mathcal{S} \subset \mathcal{P}(\Omega)$ e ρ l'insieme di tutte le σ -algebre \mathcal{A} su Ω con $\mathcal{S} \subset \mathcal{A}$. Allora

$$\sigma\text{-algebra}(\mathcal{S}) := \bigcap_{\mathcal{A} \in \rho} \mathcal{A}$$

è per l'oss. 1.9 una σ -algebra su Ω che evidentemente è la più piccola σ -algebra su Ω che contiene \mathcal{S} .

Definizione 1.11. Sia X uno spazio topologico. Gli elementi di

$$\text{borel}(X) = \sigma\text{-algebra}(\text{aperti}(X)) = \sigma\text{-algebra}(\text{chiusi}(X))$$

sono detti *insiemi di Borel* di X , $\text{borel}(X)$ si chiama l'*algebra di Borel* di X .

Definizione 1.12. Sia $(\Omega, \mathcal{A}, \mu)$ uno spazio di misura. Poniamo

$$(\mu = 0) := \{A \in \mathcal{A} \mid \mu(A) = 0\}$$

Gli elementi di $(\mu = 0)$ sono detti *insiemi di misura nulla*; essi sono, per definizione, automaticamente misurabili.

Definiamo inoltre

$$(\mu = 0)^* = \{N \subset \Omega \mid \text{esiste } A \in (\mu = 0) \text{ con } N \subset A\}$$

Definizione 1.13. Uno spazio di misura $(\Omega, \mathcal{A}, \mu)$ si dice *completo*, se è soddisfatta una delle seguenti condizioni equivalenti:

- (1) $A \in (\mu = 0)$ ed $N \subset A \implies N \in \mathcal{A}$.
- (2) $(\mu = 0)^* \subset \mathcal{A}$.
- (3) $(\mu = 0)^* = (\mu = 0)$.

In uno spazio di misura completo perciò ogni sottoinsieme di un insieme di misura nulla (automaticamente misurabile) è ancora misurabile e di misura nulla.

Dimostrazione. Dobbiamo dimostrare che le tre condizioni sono equivalenti.

(1) \iff (2): Chiaro.

(2) \implies (3): È chiaro che $(\mu = 0) \subset (\mu = 0)^*$.

Sia $N \in (\mu = 0)^*$. Ciò significa che esiste $A \in (\mu = 0)$ tale che $N \subset A$. Abbiamo supposto che $(\mu = 0)^* \subset \mathcal{A}$, perciò $N \in \mathcal{A}$. Dall'oss. 1.7 segue $\mu(N) \leq \mu(A) = 0$ e quindi $\mu(N) = 0$, cosicché $N \in (\mu = 0)$.

(3) \implies (2): Chiaro, perché $(\mu = 0) \subset \mathcal{A}$.

Nota 1.14. In $\text{borel}(\mathbb{R}^m)$ si può definire in modo naturale una misura che per ogni intervallo m -dimensionale coincide con il volume di quell'intervallo. Lo spazio di misura che così si ottiene però non è completo. Si può dimostrare che esiste un completamento $(\mathbb{R}^m, \text{leb}(\mathbb{R}^m, \mu_{\text{leb}}))$.

Gli elementi di $(\text{leb}(\mathbb{R}^m))$ si chiamano *insiemi misurabili nel senso di Lebesgue* e μ_{leb} è detta *misura di Lebesgue*.

Per costruzione $\text{borel}(\mathbb{R}^m) \subset \text{leb}(\mathbb{R}^m)$; in particolare ogni aperto e ogni chiuso di \mathbb{R}^m è misurabile nel senso di Lebesgue.

Definizione 1.15. X sia uno spazio topologico ed $E \in X$.

E si dice un G_δ , se esistono aperti A_1, A_2, \dots di X tali che $E = \bigcap_{n=1}^{\infty} A_n$.

E si dice un F_σ , se esistono chiusi F_1, F_2, \dots di X tali che $E = \bigcup_{n=1}^{\infty} F_n$.

È chiaro che ogni G_δ e ogni F_σ appartiene a $\text{borel}(X)$.

Osservazione 1.16. Ogni G_δ o F_σ di \mathbb{R}^m appartiene a $\text{leb}(\mathbb{R}^m)$.

Teorema 1.17. Per $A \subset \mathbb{R}^m$ sono equivalenti:

- (1) $A \in \text{leb}(\mathbb{R}^m)$.
- (2) Per ogni $\varepsilon > 0$ esistono un chiuso C ed un aperto U con $C \subset A \subset U$ tali che $\mu_{\text{leb}}(U \setminus C) < \varepsilon$.
- (3) Esistono un F_σ F e un G_δ G con $F \subset A \subset G$ tali che $\mu_{\text{leb}}(G \setminus F) = 0$.

Dimostrazione. Elstrodt, pag. 67.

Definizione 1.18. Siano (Ω, \mathcal{A}) e (Ω', \mathcal{A}') due spazi misurabili.

Un'applicazione $f : \Omega \rightarrow \Omega'$ si dice *misurabile* (più precisamente \mathcal{A} - \mathcal{A}' -misurabile) se $f^{-1}(A') \in \mathcal{A}$ per ogni $A' \in \mathcal{A}'$.

Definizione 1.19. (Ω, \mathcal{A}) sia uno spazio misurabile. Un'applicazione $X : \Omega \rightarrow \mathbb{R}$ si dice *misurabile*, se è \mathcal{A} - $\text{borel}(\mathbb{R})$ -misurabile.

Definizione 1.20. Siano Ω un insieme ed f, g due applicazioni. Poniamo allora:

$$(f = g) := \{\omega \in \Omega \mid f(\omega) = g(\omega)\}$$

$$(f \leq g) := \{\omega \in \Omega \mid f(\omega) \leq g(\omega)\}$$

Per $U \subset \mathbb{R}$ ed $a \in \mathbb{R}$ siano inoltre:

$$(f \in U) := \{\omega \in \Omega \mid f(\omega) \in U\} = f^{-1}(U)$$

$$(f = a) := \{\omega \in \Omega \mid f(\omega) = a\} = f^{-1}(\{a\})$$

$$(f < a) := \{\omega \in \Omega \mid f(\omega) < a\} = f^{-1}((-\infty, a))$$

$$(f \leq a) := \{\omega \in \Omega \mid f(\omega) \leq a\} = f^{-1}((-\infty, a])$$

ecc.

Definizione 1.21. Sia Ω un insieme. Date applicazioni $f_1, \dots, f_m : \Omega \rightarrow \mathbb{R}$ e insiemi $U_1, \dots, U_m \subset \mathbb{R}$ definiamo

$$(f_1 \in U_1, \dots, f_m \in U_m) := (f_1 \in U_1) \cap \dots \cap (f_m \in U_m)$$

Proposizione 1.22. (Ω, \mathcal{A}) sia uno spazio misurabile. Per una funzione $X : \Omega \rightarrow \mathbb{R}$ sono equivalenti:

- (1) X è misurabile.
- (2) $(X \in U) \in \mathcal{A}$ per ogni aperto U di \mathbb{R} .
- (3) $(X \in V) \in \mathcal{A}$ per ogni chiuso V di \mathbb{R} .
- (4) $(X < a) \in \mathcal{A}$ per ogni $a \in \mathbb{R}$.
- (5) $(X \leq a) \in \mathcal{A}$ per ogni $a \in \mathbb{R}$.
- (6) $(X > a) \in \mathcal{A}$ per ogni $a \in \mathbb{R}$.
- (7) $(X \geq a) \in \mathcal{A}$ per ogni $a \in \mathbb{R}$.

Dimostrazione. Verifica tecnica.

Definizione 1.23. Una *variabile aleatoria* (o casuale) reale è un'applicazione misurabile $X : \Omega \rightarrow \mathbb{R}$, dove (Ω, \mathcal{A}) è uno spazio misurabile che spesso rimane sottinteso e non viene specificato in dettaglio.

Quando necessario diciamo anche che X è definita su (Ω, \mathcal{A}) .

Definizione 1.24. Sia (Ω, \mathcal{A}) uno spazio misurabile. Una variabile aleatoria reale $X : \Omega \rightarrow \mathbb{R}$ si dice *elementare*, se assume solo un numero finito di valori, cioè se l'immagine $V := X(\Omega)$ è finita.

In tal caso $(X = a) = \emptyset$ per ogni $a \in \mathbb{R} \setminus V$.

Quando necessario diciamo anche che X è \mathcal{A} -elementare.

Lemma 1.25. Siano (Ω, \mathcal{A}) uno spazio misurabile ed $X : \Omega \rightarrow \mathbb{R}$ un'applicazione che assume solo un numero finito di valori. Allora sono equivalenti:

- (1) X è misurabile.
- (2) $(X = a) \in \mathcal{A}$ per ogni $a \in \mathbb{R}$.
- (3) X è $\mathcal{A}\text{-}\mathcal{P}(\mathbb{R})$ -misurabile.

Dimostrazione.

(1) \implies (2): Chiaro, perché $(X = a) = X^{-1}(\{a\})$ ed $\{a\} \in \text{borel}(\mathbb{R})$.

(2) \implies (3): Sia $B \subset \mathbb{R}$. Allora $(X \in B) = \bigcup_{b \in B} (X = b)$. Siccome X assume solo un numero finito di valori, $(X = b) \neq \emptyset$ solo per un numero finito di $b \in B$. Siccome per ipotesi ogni $(X = b) \in \mathcal{A}$, anche $(X \in B) \in \mathcal{A}$.

(3) \implies (1): Chiaro.

Definizione 1.26. Ω sia un insieme ed $A \subset \Omega$. La funzione caratteristica $\chi_A : \Omega \rightarrow \mathbb{R}$ associa il valore 1 a tutti gli elementi di A e il valore 0 agli elementi di $\Omega \setminus A$.

Essa assume quindi al massimo due valori.

Osservazione 1.27. (Ω, \mathcal{A}) sia uno spazio misurabile ed $A \subset \Omega$. Allora sono equivalenti:

- (1) χ_A è misurabile.
- (2) $A \in \mathcal{A}$.
- (3) χ_A è elementare.

Dimostrazione.

(1) \implies (2): χ_A sia misurabile. Allora: $A = (\chi_A = 1) \in \mathcal{A}$.

(2) \implies (1): Possiamo usare il lemma 1.25. Però

$(\chi_A = 1) = A \in \mathcal{A}$ e $(\chi_A = 0) = \Omega \setminus A \in \mathcal{A}$.

(2) \iff (3): Chiaro. Infatti χ_A assume solo i valori 0 e 1.

Lemma 1.28. Ω sia un insieme ed $X, Y : \Omega \rightarrow \mathbb{R}$ due applicazioni. Siano $a, \alpha \in \mathbb{R}$. Allora:

$$(1) (X + Y = a) = \dot{\bigcup}_{b \in \mathbb{R}} (X = b, Y = a - b).$$

$$(2) (XY = a) = \dot{\bigcup}_{b \in \mathbb{R} \setminus 0} (X = b, Y = \frac{a}{b}) \text{ se } a \neq 0.$$

$$(XY = 0) = (X = 0) \cup (Y = 0).$$

$$(3) (\alpha X = a) = (X = \frac{a}{\alpha}) \text{ per } \alpha \neq 0.$$

$$(0X = a) = \begin{cases} \Omega & \text{per } a = 0 \\ \emptyset & \text{per } a \neq 0 \end{cases}$$

$$(4) (|X| = a) = (X = a) \dot{\cup} (X = -a).$$

Dimostrazione. Chiaro.

Proposizione 1.29. (Ω, \mathcal{A}) sia uno spazio misurabile ed $X, Y : \Omega \rightarrow \mathbb{R}$ due variabili aleatorie elementari. Sia $\alpha \in \mathbb{R}$. Allora $X + Y, XY, \alpha X$ e $|X|$ sono ancora variabili aleatorie elementari.

Dimostrazione. (1) Dimostriamo prima che queste funzioni assumono solo un numero finito di valori. Ma gli insiemi

$$(X + Y)(\Omega) \subset \{a + b \mid a \in X(\Omega), b \in Y(\Omega)\}$$

$$(XY)(\Omega) \subset \{ab \mid a \in X(\Omega), b \in Y(\Omega)\}$$

$$\alpha X(\Omega) = \{\alpha a \mid a \in X(\Omega)\}$$

$$|X|(\Omega) = \{|a| \mid a \in X(\Omega)\}$$

sono finiti.

(2) Possiamo perciò applicare il lemma 1.25. L'enunciato segue adesso dal lemma 1.28 in cui le unioni contengono solo un numero finito di termini $\neq \emptyset$. Infatti

$$\bigcup_{b \in \mathbb{R}} (X = b, Y = a - b) = \bigcup_{b \in \mathbb{R}} (X = b) \cap (Y = a - b)$$

$$\bigcup_{b \in \mathbb{R} \setminus 0} (X = b, Y = \frac{a}{b}) = \bigcup_{b \in \mathbb{R} \setminus 0} (X = b) \cap (Y = \frac{a}{b})$$

Definizione 1.30. Siano (Ω, \mathcal{A}, p) uno spazio di probabilità ed $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria elementare. Allora

$$MX := \sum_{a \in \mathbb{R}} ap(X = a)$$

si chiama la *media* di X .

Osserviamo che questa somma contiene solo un numero finito di sommandi diverso da zero, perché l'insieme $\{a \in \mathbb{R} \mid (X = a) \neq \emptyset\}$ è finito per ipotesi quindi anche l'insieme $\{a \in \mathbb{R} \mid p(X = a) \neq 0\}$ è finito.

Lemma 1.31. Siano (Ω, \mathcal{A}, p) uno spazio di probabilità e $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria elementare. Sia $A \in \mathcal{A}$. Allora

$$\sum_{a \in \mathbb{R}} p(A, X = a) = p(A).$$

Dimostrazione. Infatti $A = \dot{\bigcup}_{a \in \mathbb{R}} (A \cap (X = a))$.

Teorema 1.32. (Ω, \mathcal{A}, p) sia uno spazio di probabilità ed $X, Y : \Omega \rightarrow \mathbb{R}$ due variabili aleatorie elementari. Sia $\alpha \in \mathbb{R}$. Allora:

$$\begin{aligned} M(\alpha X) &= \alpha MX \\ M(X + Y) &= MX + MY \end{aligned}$$

Dimostrazione. Scriviamo dappertutto \sum_a invece di $\sum_{a \in \mathbb{R}}$. Si noti che in tutte le somme appare solo un numero finito di sommandi non nullo.

(1) L'enunciato è ovvio per $\alpha = 0$

Sia $\alpha \neq 0$. Allora

$$\begin{aligned} M(\alpha X) &= \sum_a ap(\alpha X = a) = \sum_a ap(X = a/\alpha) \\ &= \alpha \sum_a \frac{a}{\alpha} p(X = a/\alpha) = \alpha \sum_a ap(X = a) = \alpha M(X) \end{aligned}$$

(2) Utilizzando il lemma 1.28 abbiamo

$$\begin{aligned} M(X + Y) &= \sum_a ap(X + Y = a) \\ &= \sum_a \sum_b ap(X = b, Y = a - b) = \sum_b \sum_a ap(X = b, Y = a - b) \\ &= \sum_b \sum_a (a + b)p(X = b, Y = a) \\ &= \sum_b \sum_a ap(X = b, Y = a) + \sum_b \sum_a bp(X = b, Y = a) \\ &= \sum_a a \sum_b p(X = b, Y = a) + \sum_b b \sum_a p(X = b, Y = a) \\ &\stackrel{1.31}{=} \sum_a ap(X = a) + \sum_b bp(Y = b) = MX + MY \end{aligned}$$

Osservazione 1.33. (Ω, \mathcal{A}, p) sia uno spazio di probabilità ed $X, Y : \Omega \rightarrow \mathbb{R}$ due variabili aleatorie elementari. Allora:

(1) $X \geq 0 \implies MX \geq 0$.

(2) $X \leq Y \implies MX \leq MY$.

Dimostrazione. (1) Se $X \geq 0$, allora

$$MX = \sum_{a \in \mathbb{R}} ap(X = a) = \sum_{a \in [0, \infty)} ap(X = a) \geq 0.$$

(2) L'ipotesi implica $Y - X \geq 0$, per cui $MY - MX = M(Y - X) \geq 0$.

Osservazione 1.34. Siano (Ω, \mathcal{A}) , (Ω', \mathcal{A}') , $(\Omega'', \mathcal{A}'')$ tre spazi misurabili. Siano inoltre $X : \Omega \rightarrow \Omega'$ \mathcal{A} - \mathcal{A}' -misurabile e $f : \Omega' \rightarrow \Omega''$ \mathcal{A}' - \mathcal{A}'' -misurabile. Allora $f(X) := f \circ X : \Omega \rightarrow \Omega''$ è \mathcal{A} - \mathcal{A}'' -misurabile.

Dimostrazione. Sia $A'' \in \mathcal{A}''$. Allora $f \circ X)^{-1}(A'') = X^{-1}(f^{-1}(A''))$. Per ipotesi $f^{-1}(A'') \in \mathcal{A}'$ e quindi $X^{-1}(f^{-1}(A'')) \in \mathcal{A}$.

Corollario 1.35. Siano (Ω, \mathcal{A}) uno spazio misurabile ed $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria reale. La funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ sia borel(\mathbb{R})-borel(\mathbb{R})-misurabile, ad esempio una funzione continua. Allora

$f(X) : \Omega \rightarrow \mathbb{R}$ è una variabile aleatoria reale. Se X è elementare, anche $f(X)$ è elementare.

Dimostrazione. La prima parte segue dall'oss. 1.34.

Se X è elementare, allora $X(\Omega)$ è finito per cui è finito anche l'insieme $f(X)(\Omega) = f(X(\Omega))$.

Teorema 1.36. (Ω, \mathcal{A}, p) sia uno spazio di probabilità ed $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria elementare. Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ borel(\mathbb{R})-borel(\mathbb{R})-misurabile. Allora

$$Mf(X) = \sum_{a \in \mathbb{R}} f(a)p(X = a)$$

Dimostrazione. Γ sia il grafico di f , cioè

$$\Gamma = \{(a, t) \in \mathbb{R} \times \mathbb{R} \mid f(a) = t\} = \{(a, f(a)) \mid a \in \mathbb{R}\}$$

Per ogni $t \in \mathbb{R}$ abbiamo $(f(X) = t) = \dot{\bigcup}_{a \in f^{-1}(t)} (X = a)$, per cui

$$\begin{aligned} Mf(X) &= \sum_t tp(f(X) = t) = \sum_t t \sum_{a \in f^{-1}(t)} p(X = a) \\ &= \sum_{(a,t) \in \Gamma} tp(X = a) = \sum_a f(a)p(X = a) \end{aligned}$$

Esempio 1.37. Nella situazione del teorema 1.36 siano $X(\Omega) = \{0, 1, 2, 3, 4\}$ ed $\mathcal{A} = \mathcal{P}(\Omega)$.

La funzione f su $X(\Omega)$ sia data dalla tabella

0	17
1	30
2	30
3	17
4	30

La funzione $f(X)$ assume solo i valori 17 e 30

per cui

$$Mf(X) = 17p(f(X) = 17) + 30p(f(X) = 30).$$

Inoltre

$$(f(X) = 17) = (X = 0) \dot{\cup} (X = 3)$$

$$(f(X) = 30) = (X = 1) \dot{\cup} (X = 2) \dot{\cup} (X = 4)$$

per cui

$$\begin{aligned} Mf(X) &= 17[p(X = 0) + p(X = 3)] + 30[p(X = 1) + p(X = 2) + p(X = 4)] \\ &= f(0)p(X = 0) + f(3)p(X = 3) + f(1)p(X = 1) + f(2)p(X = 2) + f(4)p(X = 4) \end{aligned}$$

in accordo con l'enunciato del teorema 1.36.

Osservazione 1.38. (Ω, \mathcal{A}, p) sia uno spazio di probabilità ed $A \in \mathcal{A}$. Allora χ_A è una variabile aleatoria elementare (per l'oss. 1.27) e si ha $M\chi_A = p(A)$.

Dimostrazione. χ_A assume solo i valori 0 e 1, perciò

$$M\chi_A = \sum_{a \in \mathbb{R}} ap(\chi_A = a) = p(\chi_A = 1) = p(A)$$

Definizione 1.39. (Ω, \mathcal{A}, p) sia uno spazio di probabilità, $X : \Omega \rightarrow \mathbb{R}$ una variabile aleatoria elementare ed $E \in \mathcal{A}$. Poniamo allora

$$\int_E Xp := M(X\chi_E)$$

In particolare quindi $MX = \int_{\Omega} Xp$.

Osservazione 1.40. Nella situazione della def. 1.39 si ha

$$\int_E Xp = \sum_{a \in \mathbb{R}} ap((X = a) \cap E).$$

Dimostrazione. È sufficiente dimostrare che

$$\int_E Xp = \sum_{a \in \mathbb{R} \setminus 0} ap((X = a) \cap E)$$

Per definizione però

$$\int_E Xp = \sum_{a \in \mathbb{R}} ap(X\chi_E = a) = \sum_{a \in \mathbb{R} \setminus 0} ap(X\chi_E = a)$$

Per $a \neq 0$ ed $\omega \in \Omega$ si ha però

$$\omega \in (X\chi_E = a) \iff \omega \in E \text{ e } X(\omega) = a \iff \omega \in (X = a) \cap E.$$

2. Il teorema di Radon-Nikodým discreto

Situazione 2.1. (Ω, \mathcal{A}, p) sia uno spazio di probabilità.

Definizione 2.2. Per $A, E \in \mathcal{A}$ definiamo la *probabilità condizionata* $p(A|E)$ ponendo

$$p(A|E) := \begin{cases} 0 & \text{se } p(E) = 0 \\ \frac{p(A \cap E)}{p(E)} & \text{se } p(E) > 0 \end{cases}$$

Si noti che $p(A|E) = 0$ anche quando solo $p(A \cap E) = 0$.

Osservazione 2.3. Siano $A, E \in \mathcal{A}$. Allora $p(A|E)p(E) = p(A \cap E)$.

Dimostrazione. Ciò è chiaro per $p(E) > 0$.

Se invece $p(E) = 0$, allora anche $p(A \cap E) = 0$ e l'uguaglianza rimane valida.

Lemma 2.4 (formula della probabilità totale). Sia E_1, E_2, \dots una successione finita o infinita di elementi disgiunti di \mathcal{A} tali che $\bigcup_n E_n = \Omega$. Allora per ogni $A \in \mathcal{A}$ vale

$$p(A) = \sum_n p(A|E_n)p(E_n)$$

Dimostrazione. L'ipotesi implica $A = \bigcup_n (A \cap E_n)$, per cui

$$p(A) = \sum_n p(A \cap E_n) = \sum_n p(A|E_n)p(E_n)$$

Proposizione 2.5 (formula di Bayes). Sia E_1, E_2, \dots una successione finita o infinita di elementi disgiunti di \mathcal{A} tali che $\bigcup_n E_n = \Omega$. Allora per ogni $A \in \mathcal{A}$ con $p(A) > 0$ ed ogni k vale

$$p(E_k|A) = \frac{p(A|E_k)p(E_k)}{\sum_n p(A|E_n)p(E_n)}$$

Dimostrazione. Per il lemma 2.4 $\sum_n p(A|E_n)p(E_n) = p(A) > 0$, per cui

$$p(E_k|A) = \frac{p(E_k \cap A)}{p(A)} \stackrel{2.4}{=} \frac{p(A|E_k)p(E_k)}{\sum_n p(A|E_n)p(E_n)}$$

Definizione 2.6. $X : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare. Per $E \in \mathcal{A}$ definiamo

$$M(X|E) := \sum_{a \in \mathbb{R}} ap(X = a|E)$$

$M(X|E)$ si chiama la *media di X su E* (cfr. oss. 2.7) oppure la *media condizionata di X rispetto ad E*.

Osservazione 2.7. $X : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare ed $E \in \mathcal{A}$. Allora

$$\int_E Xp = M(X|E)p(E)$$

Per $p(E) > 0$ si ha quindi $M(X|E) = \frac{1}{p(E)} \int_E Xp$.

Dimostrazione. Abbiamo

$$\int_E Xp \stackrel{1.40}{=} \sum_{a \in \mathbb{R}} ap((X = a) \cap E) = \sum_{a \in \mathbb{R}} ap(X = a|E)p(E) = p(E)M(X|E)$$

Osservazione 2.8. $X : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare ed $E \in \mathcal{A}$ con $p(E) = 0$. Allora $M(X|E) = 0$.

Dimostrazione. Siccome $p(E) = 0$, ogni sommando in $M(X|E) = \sum_{a \in \mathbb{R}} ap(X = a|E)$ si annulla.

Corollario 2.9. Siano $A, E \in \mathcal{A}$. Allora $p(A|E) = M(\chi_A|E)$.

Dimostrazione. Siccome χ_A assume solo i valori 0 e 1, abbiamo

$$M(\chi_A|E) = \sum_{a \in \mathbb{R}} ap(\chi_A = a|E) = 1 \cdot p(\chi_A = 1|E) = p(A|E)$$

Osservazione 2.10. Siano $E_1, \dots, E_n \subset \Omega$ tali che $\Omega = E_1 \dot{\cup} \dots \dot{\cup} E_n$. Allora

$$\sigma\text{-algebra}(\{E_1, \dots, E_n\}) = \left\{ \bigcup_{k \in K} E_k \mid K \subset \{1, \dots, n\} \right\} =: \mathcal{F}.$$

Se gli E_i appartengono tutti ad \mathcal{A} , allora \mathcal{F} è una sotto- σ -algebra finita di \mathcal{A} .

Dimostrazione. È chiaro che $\{E_1, \dots, E_n\} \subset \mathcal{F} \subset \sigma\text{-algebra}(\{E_1, \dots, E_n\})$, per cui è sufficiente dimostrare che \mathcal{F} è una σ -algebra.

(1) $\Omega \in \mathcal{F}$ perché $\Omega = E_1 \cup \dots \cup E_n$.

(2) È chiaro che l'unione finita di un numero finito di elementi di \mathcal{F} appartiene ancora ad \mathcal{F} . Siccome l'insieme \mathcal{F} è finito, ogni unione di un numero numerabile di elementi di \mathcal{F} è in verità un'unione finita.

(3) È a questo punto sufficiente dimostrare che il complemento di un elemento di \mathcal{F} appartiene ancora ad \mathcal{F} . Sia $K \subset \{1, \dots, n\}$ e $P := \bigcup_{k \in K} E_k$. Siccome gli E_j sono disgiunti, abbiamo semplicemente

$$\Omega \setminus P = \bigcup_{j \in \{1, \dots, n\} \setminus K} E_j \in \mathcal{F}.$$

Definizione 2.11. \mathcal{F} sia una σ -algebra su Ω . Un elemento $E \in \mathcal{F}$ si dice *minimale* (in \mathcal{F}), se E è minimale tra gli elementi non vuoti di \mathcal{F} . Denotiamo con $\text{Min } \mathcal{F}$ l'insieme degli elementi minimali di \mathcal{F} .

Osservazione 2.12. \mathcal{F} sia una σ -algebra finita su Ω . Allora ogni elemento non vuoto di \mathcal{F} contiene un elemento minimale di \mathcal{F} .

Lemma 2.13. \mathcal{F} sia una σ -algebra finita su Ω .

(1) Siano $A \in \mathcal{F}$ ed $E \in \text{Min } \mathcal{F}$. Allora $E \subset A$ oppure $E \cap A = \emptyset$.

(2) Siano $E, F \in \text{Min } \mathcal{F}$ ed $E \neq F$. Allora $E \cap F = \emptyset$.

Dimostrazione.

(1) Infatti $E \cap A \in \mathcal{F}$ è contenuto in E . Se $E \cap A \neq \emptyset$, per la minimalità di E si deve avere $E \cap A = E$, cioè $E \subset A$.

(2) Segue immediatamente da (1).

Osservazione 2.14. \mathcal{F} sia una σ -algebra finita su Ω . In questo caso anche $\text{Min } \mathcal{F}$ è un insieme finito, ad esempio $\text{Min } \mathcal{F} = \{E_1, \dots, E_n\}$ con gli E_j tutti distinti. Allora $\Omega = E_1 \dot{\cup} \dots \dot{\cup} E_n$.

Ogni σ -algebra finita su Ω è quindi della forma indicata nell'oss. 2.10.

Dimostrazione. Dal lemma 2.13 sappiamo che $E_i \cap E_j = \emptyset$ per ogni $i \neq j$. Dobbiamo perciò solo dimostrare che $E_1 \cup \dots \cup E_n = \Omega$.

Sia $B := \Omega \setminus (E_1 \cup \dots \cup E_n) \neq \emptyset$. Allora $B \in \mathcal{F}$ e per l'oss. 2.12 esiste un j tale che $E_j \subset B$. Ma $B \subset \Omega \setminus E_j$, per cui si arriva ad $E_j \subset \Omega \setminus E_j$, cioè $E_j = E_j \cap E_j = \emptyset$, una contraddizione.

Definizione 2.15. $X : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare ed $E_1, \dots, E_n \in \mathcal{A}$ tali che $\Omega = E_1 \dot{\cup} \dots \dot{\cup} E_n$. Sia $\mathcal{F} := \sigma\text{-algebra}(\{E_1, \dots, E_n\})$.

Per $\omega \in \Omega$ definiamo allora:

$$M(X|\mathcal{F})(\omega) := M(X|E_i)$$

se $\omega \in E_i$. Otteniamo così un'applicazione $M(X|\mathcal{F}) : \Omega \rightarrow \mathbb{R}$ che possiamo scrivere nella forma $M(X|\mathcal{F}) = \sum_{k=1}^n M(X|E_k)\chi_{E_k}$. Questa rappresentazione mostra che $M(X|\mathcal{F})$ è \mathcal{F} -borel(\mathbb{R})-misurabile e quindi anche \mathcal{A} -borel(\mathbb{R})-misurabile.

$M(X|\mathcal{F})$ assume solo i valori $M(X|E_1), \dots, M(X|E_n)$ ed è quindi \mathcal{F} -elementare e \mathcal{A} -elementare.

Osservazione 2.16. $Y : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare ed $E \in \mathcal{A}$. Y sia costante su E , ad esempio $Y(\omega) = c$ per ogni $\omega \in E$ per qualche $c \in \mathbb{R}$. Allora $\int_E Y p = c \cdot p(E)$.

Dimostrazione. Per l'oss. 1.40 abbiamo $\int_E Y p = \sum_{a \in \mathbb{R}} a \cdot p((Y = a) \cap E)$.

Per ipotesi però $(Y = a) \cap E = \emptyset$ per $a \neq c$, mentre $(Y = c) \cap E = E$, cosicché $\int_E Y p = c \cdot p((Y = c) \cap E) = c \cdot p(E)$.

Teorema 2.17. Nella situazione della def. 2.15 per ogni $k = 1, \dots, n$ si ha

$$\int_{E_k} M(X|\mathcal{F}) p = \int_{E_k} X p$$

oppure, equivalentemente,

$$M(M(X|\mathcal{F})|E_k) = M(X|E_k)$$

Questa è la versione discreta del teorema di Radon-Nikodým.

Dimostrazione. Siccome $M(X|\mathcal{F})$ è costante su E_k ed uguale a $M(X|E_k)$ dall'oss. 2.16 segue

$$\int_{E_k} M(X|\mathcal{F})p = M(X|E_k)p(E_k) \stackrel{2.7}{=} \int_{E_k} Xp$$

Esempio 2.18. Siano $\Omega = \{1, \dots, 12\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ e $p(\omega) = \frac{1}{12}$ per ogni $\omega \in \Omega$. La variabile aleatoria elementare X sia definita tramite la tabella

	ω	$X(\omega)$
E_1	1	2
	2	1
	3	2
	4	3
E_2	5	7
	6	1
	7	1
E_3	8	5
	9	3
	10	5
	11	5
	12	2

Abbiamo allora

$$\begin{aligned} M(X|E_1) &= 1 \cdot p(X = 1|E_1) + 2 \cdot p(X = 2|E_1) + 3 \cdot p(X = 3|E_1) \\ &= \frac{1/12 + 2 \cdot 2/12 + 3/12}{4/12} = \frac{8}{4} = 2 \end{aligned}$$

$$\begin{aligned} M(X|E_2) &= 1 \cdot p(X = 1|E_2) + 7 \cdot p(X = 7|E_2) \\ &= \frac{2/12 + 7 \cdot 1/12}{3/12} = \frac{9}{3} = 3 \end{aligned}$$

$$\begin{aligned} M(X|E_3) &= 2 \cdot p(X = 2|E_3) + 3 \cdot p(X = 3|E_3) + 5 \cdot p(X = 5|E_3) \\ &= \frac{2 \cdot 1/12 + 3 \cdot 1/12 + 5 \cdot 3/12}{5/12} = \frac{20}{5} = 4 \end{aligned}$$

Con $\mathcal{F} := \sigma\text{-algebra}(\{E_1, E_2, E_3\})$ si ha

$$M(X|\mathcal{F})(\omega) = \begin{cases} 2 & \text{per } \omega \in E_1 \\ 3 & \text{per } \omega \in E_2 \\ 4 & \text{per } \omega \in E_3 \end{cases}$$

Definizione 2.19. $Y : \Omega \rightarrow \mathbb{R}$ sia una variabile aleatoria elementare ed $Y(\Omega) = \{t_1, \dots, t_n\}$ con t_1, \dots, t_n tutti distinti. Poniamo

$$\sigma\text{-algebra}(Y) := \sigma\text{-algebra}(\{(Y = t_1), \dots, (Y = t_n)\})$$

Per una variabile aleatoria elementare $X : \Omega \rightarrow \mathbb{R}$ sia

$$M(X|Y) := M(X|\sigma\text{-algebra}(Y)) = \sum_{k=1}^n M(X|Y = t_k)\chi_{(Y=t_k)}$$

3. Regole per probabilità composte

Situazione 3.1. (Ω, \mathcal{A}, p) sia uno spazio di probabilità.

Lemma 3.2. Siano $A, B, C \in \mathcal{A}$. Allora

$$p(A \cap B|C) = p(A|B \cap C)p(B|C)$$

Dimostrazione. Se $p(B \cap C) = 0$, allora anche $p(A \cap B \cap C) = 0$, per cui le espressioni sia a sinistra che a destra sono uguali a zero.

Sia quindi $p(B \cap C) \neq 0$. Allora anche $p(C) \neq 0$, cosicché

$$\begin{aligned} p(A|B \cap C)p(B|C) &= \frac{p(A \cap B \cap C)}{p(B \cap C)} \frac{p(B \cap C)}{p(C)} = \frac{p(A \cap B \cap C)}{p(C)} = \\ &= p(A \cap B|C) \end{aligned}$$

Lemma 3.3. Siano $n \in \mathbb{N} + 1$ ed $A_0, A_1, \dots, A_n \in \mathcal{A}$. Allora:

$$\begin{aligned} p(A_n \cap \dots \cap A_1|A_0) &= p(A_n|A_{n-1} \cap \dots \cap A_0)p(A_{n-1}|A_{n-2} \cap \dots \cap A_0) \dots p(A_1|A_0) \\ p(A_n \cap \dots \cap A_0) &= p(A_n|A_{n-1} \cap \dots \cap A_0)p(A_{n-1}|A_{n-2} \cap \dots \cap A_0) \dots p(A_1|A_0)p(A_0) \end{aligned}$$

Dimostrazione. (1) Dimostriamo la prima equazione per induzione su n .

$n = 1$: Lemma 3.2.

$n \rightarrow n + 1$:

$$\begin{aligned} p(A_{n+1} \cap \dots \cap A_1|A_0) &= p(A_{n+1} \cap (A_n \cap \dots \cap A_1)|A_0) \\ &\stackrel{3.2}{=} p(A_{n+1}|A_n \cap \dots \cap A_0)p(A_n \cap \dots \cap A_1|A_0) \\ &\stackrel{ind.}{=} p(A_{n+1}|A_n \cap \dots \cap A_0)p(A_n|A_{n-1} \cap \dots \cap A_0) \dots p(A_1|A_0) \end{aligned}$$

(2) Moltiplicando la prima equazione con $p(A_0)$ e utilizzando l'oss. 2.2 si ottiene la seconda.

Corollario 3.4. Siano $n, k \in \mathbb{N}$ con $k < n$ ed $A_0, \dots, A_n \in \mathcal{A}$. Allora:

$$\begin{aligned} p(A_n \cap \dots \cap A_1|A_0) &= p(A_n \cap \dots \cap A_{k+1}|A_k \cap \dots \cap A_0)p(A_k|A_{k-1} \cap \dots \cap A_0) \dots p(A_1|A_0) \\ p(A_n \cap \dots \cap A_0) &= p(A_n \cap \dots \cap A_{k+1}|A_k \cap \dots \cap A_0)p(A_k|A_{k-1} \cap \dots \cap A_0) \dots p(A_1|A_0)p(A_0) \end{aligned}$$

Dimostrazione. Segue direttamente dal lemma 3.3.

4. Sequenze di Markov

Situazione 4.1. Siano (Ω, \mathcal{A}, p) uno spazio di probabilità, $q \in \mathbb{N} + 1$ fissato ed $S := \{1, \dots, q\}$. Scriviamo spesso \bigcirc_n invece di $\bigcirc_{n \in \mathbb{N}}$.

Il concetto di sequenza di Markov è puramente notazionale (in virtù del teorema di Kolmogorov) allo scopo di rendere trasparenti le dimostrazioni.

Definizione 4.2. $\bigcirc_{n,i} E_n^i : \mathbb{N} \times S \longrightarrow \mathcal{A}$ sia un'applicazione tale che per ogni $n \in \mathbb{N}$ si abbia $\Omega = E_n^1 \dot{\cup} \dots \dot{\cup} E_n^q$. Diciamo allora che $\bigcirc_{n,i} E_i$ è una sequenza di S -partizioni \mathcal{A} -misurabili.

Una successione $\bigcirc_n A_n : \mathbb{N} \longrightarrow \mathcal{A}$ si chiama una *realizzazione* della sequenza, se per ogni $n \in \mathbb{N}$ esiste $i \in S$ tale che $A_n = E_n^i$.

Definizione 4.3. Una *sequenza di Markov* è una sequenza $\bigcirc_{n,i} E_n^i$ di S -partizioni \mathcal{A} -misurabili tali che per ogni realizzazione $\bigcirc_n A_n$ della sequenza ed ogni $n \in \mathbb{N}$ per cui $p(A_n \cap \dots \cap A_0) > 0$ si abbia

$$p(A_{n+1} | A_n \cap \dots \cap A_0) = p(A_{n+1} | A_n)$$

Evidentemente è sufficiente chiedere questa condizione solo per $n \in \mathbb{N} + 1$.

Osservazione 4.4. $\bigcirc_{n,i} E_n^i$ sia una sequenza di S -partizioni \mathcal{A} -misurabili.

Allora sono equivalenti:

- (1) $\bigcirc_{n,i} E_n^i$ è una sequenza di Markov.
- (2) Per ogni sua realizzazione $\bigcirc_n A_n$ e per ogni $n \in \mathbb{N} + 1$ vale

$$p(A_{n+1} \cap \dots \cap A_0) = p(A_{n+1} \cap A_n) \cdot p(A_{n-1} \cap \dots \cap A_0 | A_n).$$

Dimostrazione. $\bigcirc_n A_n$ sia una realizzazione della sequenza data.

Fissato $n \in \mathbb{N} + 1$ poniamo $E := A_{n-1} \cap \dots \cap A_0$.

(1) \implies (2): Dobbiamo dimostrare che

$$p(A_{n+1} \cap A_n \cap E) = p(A_{n+1} \cap A_n) \cdot p(E | A_n)$$

Se $p(E \cap A_n) = 0$, entrambi i lati dell'equazione si annullano.

Sia quindi $p(E \cap A_n) \neq 0$. L'ipotesi (1) implica allora che $p(A_{n+1} | A_n \cap E) = p(A_{n+1} | A_n)$, per cui

$$\begin{aligned} p(A_{n+1} \cap A_n \cap E) &= p(A_{n+1} | A_n \cap E) \cdot p(A_n \cap E) = p(A_{n+1} | A_n) \cdot p(A_n \cap E) \\ &= \frac{p(A_{n+1} \cap A_n)}{p(A_n)} \cdot p(A_n \cap E) = p(A_{n+1} \cap A_n) \cdot p(E | A_n) \end{aligned}$$

(2) \implies (1): Sia $p(A_n \cap E) > 0$. Dobbiamo dimostrare che

$p(A_{n+1}|A_n \cap E) = p(A_{n+1}|A_n)$. Per ipotesi abbiamo però

$p(A_{n+1} \cap A_n \cap E) = p(A_{n+1} \cap A_n)(E|A_n)$, quindi

$$\begin{aligned} p(A_{n+1}|A_n \cap E) &= \frac{p(A_{n+1} \cap A_n \cap E)}{p(A_n \cap E)} = \frac{p(A_{n+1} \cap A_n) \cdot p(E|A_n)}{p(A_n \cap E)} \\ &= \frac{p(A_{n+1} \cap A_n) \cdot p(E \cap A_n)}{p(E \cap A_n) \cdot p(A_n)} = p(A_{n+1}|A_n) \end{aligned}$$

Lemma 4.5. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov e $\bigcirc_n A_n$ una sua realizzazione. Siano $n, m \in \mathbb{N}$ con $0 \leq m \leq n$ tali che $p(A_n \cap \dots \cap A_m) \neq 0$. Allora

$$p(A_{n+1}|A_n \cap \dots \cap A_m) = p(A_{n+1}|A_n)$$

Dimostrazione. (1) Per $m = 0$ ciò è quanto richiesto dalla def. 4.3.

(2) Sia quindi $m > 0$. Per $\alpha = (i_{m-1}, \dots, i_0) \in S^m$ poniamo

$E^\alpha := E_{m-1}^{i_{m-1}} \cap \dots \cap E_0^{i_0}$. Inoltre sia $S^{(m)} := \{\alpha \in S^m \mid p(A_n \cap \dots \cap A_m \cap E^\alpha) \neq 0\}$

Allora

$$\begin{aligned} p(A_{n+1} \cap \dots \cap A_m) &= \sum_{\alpha \in S^m} p(A_{n+1} \cap \dots \cap A_m \cap E^\alpha) \\ &= \sum_{\alpha \in S^{(m)}} p(A_{n+1} \cap \dots \cap A_m \cap E^\alpha) \\ &= \sum_{\alpha \in S^{(m)}} p(A_{n+1}|A_n \cap \dots \cap A_m \cap E^\alpha) \cdot p(A_n \cap \dots \cap A_m \cap E^\alpha) \\ &\stackrel{Markov}{=} \sum_{\alpha \in S^{(m)}} p(A_{n+1}|A_n) \cdot p(A_n \cap \dots \cap A_m \cap E^\alpha) \\ &= p(A_{n+1}|A_n) \sum_{\alpha \in S^{(m)}} p(A_n \cap \dots \cap A_m \cap E^\alpha) \\ &= p(A_{n+1}|A_n) p(A_n \cap \dots \cap A_m) \end{aligned}$$

per cui

$$p(A_{n+1}|A_n) = \frac{p(A_{n+1} \cap \dots \cap A_m)}{p(A_n \cap \dots \cap A_m)} = p(A_{n+1}|A_n \cap \dots \cap A_m)$$

Proposizione 4.6. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov e $\bigcirc_n A_n$ una sua realizzazione. Siano $n \in \mathbb{N}$, $r \in \mathbb{N} + 1$, $m \in \mathbb{N}$ con $0 \leq m \leq n$ tali che $p(A_n \cap \dots \cap A_m) \neq 0$. Allora

$$p(A_{n+r} \cap \dots \cap A_{n+1}|A_n \cap \dots \cap A_m) = p(A_{n+r} \cap \dots \cap A_{n+1}|A_n)$$

Dimostrazione. Possiamo assumere che $p(A_{n+r-1} \cap \dots \cap A_n) \neq 0$, perché altrimenti entrambi i lati dell'equazione si annullano. Allora

$$\begin{aligned}
& p(A_{n+r} \cap \dots \cap A_{n+1} | A_n \cap \dots \cap A_m) \\
& \stackrel{3.4}{=} \frac{p(A_{n+r} \cap \dots \cap A_{m+1} | A_m)}{p(A_n | A_{n-1} \cap \dots \cap A_m) \cdots p(A_{m+1} | A_m)} \\
& \stackrel{4.5}{=} \frac{p(A_{n+r} \cap \dots \cap A_{m+1} | A_m)}{p(A_n | A_{n-1}) \cdot p(A_{n-1} | A_{n-2}) \cdots p(A_{m+1} | A_m)} \\
& \stackrel{3.3}{=} \frac{p(A_{n+r} | A_{n+r-1} \cap \dots \cap A_m) \cdot p(A_{n+r-1} | A_{n+r-2} \cap \dots \cap A_m) \cdots p(A_{m+1} | A_m)}{p(A_n | A_{n-1}) \cdot p(A_{n-1} | A_{n-2}) \cdots p(A_{m+1} | A_m)} \\
& \stackrel{4.5}{=} \frac{p(A_{n+r} | A_{n+r-1}) \cdot p(A_{n+r-1} | A_{n+r-2}) \cdots p(A_{m+1} | A_m)}{p(A_n | A_{n-1}) \cdot p(A_{n-1} | A_{n-2}) \cdots p(A_{m+1} | A_m)} \\
& = p(A_{n+r} | A_{n+r-1}) \cdot p(A_{n+r-1} | A_{n+r-2}) \cdots p(A_{n+1} | A_n) \\
& \stackrel{4.5}{=} p(A_{n+r} | A_{n+r-1} \cap \dots \cap A_n) \cdot p(A_{n+r-1} | A_{n+r-2} \cap \dots \cap A_n) \cdots p(A_{n+1} | A_n) \\
& \stackrel{3.3}{=} p(A_{n+r} \cap \dots \cap A_{n+1} | A_n)
\end{aligned}$$

Definizione 4.7. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov. Per $t \in \mathbb{N} + 1$, $n_1, \dots, n_t \in \mathbb{N}$ e $K \subset S^t$, poniamo $E_{n_1, \dots, n_t}^K := \bigcup_{(k_1, \dots, k_t) \in K} E_{n_1}^{k_1} \cap \dots \cap E_{n_t}^{k_t}$.

Teorema 4.8. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov. Siano $n \in \mathbb{N} + 1$, $i \in S$, $J \in S^r$, $K \subset S^n$ tali che $p(E_n^i \cap E_{n-1, \dots, 0}^K) \neq 0$. Allora

$$p(E_{n+r, \dots, n+1}^J | E_n^i \cap E_{n-1, \dots, 0}^K) = p(E_{n+r, \dots, n+1}^J | E_n^i)$$

Dimostrazione. $p(E_{n+r, \dots, n+1}^J \cap E_n^i \cap E_{n-1, \dots, 0}^K)$

$$\begin{aligned}
& = \sum_{(j_r, \dots, j_1) \in J} \sum_{(k_{n-1}, \dots, k_0) \in K} p(E_{n+r}^{j_r} \cap \dots \cap E_{n+1}^{j_1} \cap E_n^i \cap E_{n-1}^{k_{n-1}} \cap \dots \cap E_0^{k_0}) \\
& = \sum_{(j_r, \dots, j_1) \in J} \sum_{(k_{n-1}, \dots, k_0) \in K} p(E_{n+r}^{j_r} \cap \dots \cap E_{n+1}^{j_1} | E_n^i \cap E_{n-1}^{k_{n-1}} \cap \dots \cap E_0^{k_0}) \\
& \quad \cdot p(E_n^i \cap E_{n-1}^{k_{n-1}} \cap \dots \cap E_0^{k_0}) \\
& \stackrel{Markov}{=} \sum_{(j_r, \dots, j_1) \in J} \sum_{(k_{n-1}, \dots, k_0) \in K} p(E_{n+r}^{j_r} \cap \dots \cap E_{n+1}^{j_1} | E_n^i) \\
& \quad \cdot p(E_n^i \cap E_{n-1}^{k_{n-1}} \cap \dots \cap E_0^{k_0}) \\
& = \sum_{(j_r, \dots, j_1) \in J} p(E_{n+r}^{j_r} \cap \dots \cap E_{n+1}^{j_1} | E_n^i) \\
& \quad \cdot \sum_{(k_{n-1}, \dots, k_0) \in K} p(E_n^i \cap E_{n-1}^{k_{n-1}} \cap \dots \cap E_0^{k_0}) \\
& = \sum_{(j_r, \dots, j_1) \in J} p(E_{n+r}^{j_r} \cap \dots \cap E_{n+1}^{j_1} | E_n^i) \cdot p(E_n^i \cap E_{n-1, \dots, 0}^K) \\
& = p(E_{n+r, \dots, n+1}^J | E_n^i) \cdot p(E_n^i \cap E_{n-1, \dots, 0}^K)
\end{aligned}$$

e ciò implica l'enunciato.

Corollario 4.9. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov e $\bigcirc_n A_n$ una sua realizzazione. Siano $s, t \in \mathbb{N}$ con $s < t$ ed $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ tali che $p(A_{n_s} \cap \dots \cap A_{n_{s+1}}) \neq 0$. Allora

$$p(A_{n_t} \cap \dots \cap A_{n_{s+1}} | A_{n_s} \cap \dots \cap A_{n_0}) = p(A_{n_t} \cap \dots \cap A_{n_{s+1}} | A_{n_s})$$

Corollario 4.10. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov e $\bigcirc_n A_n$ una sua realizzazione. Siano $t \in \mathbb{N} + 1$ ed $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ tali che $p(A_{n_{t-1}} \cap \dots \cap A_{n_0}) \neq 0$. Allora

$$\begin{aligned} p(A_{n_t} \cap \dots \cap A_{n_1} | A_{n_0}) &= p(A_{n_t} | A_{n_{t-1}}) \cdots p(A_{n_1} | A_{n_0}) \\ p(A_{n_t} \cap \dots \cap A_{n_0}) &= p(A_{n_t} | A_{n_{t-1}}) \cdots p(A_{n_1} | A_{n_0}) \cdot p(A_{n_0}) \end{aligned}$$

Dimostrazione. In virtù del lemma 3.3 è sufficiente dimostrare la prima uguaglianza. Abbiamo

$$\begin{aligned} p(A_{n_t} \cap \dots \cap A_{n_1} | A_{n_0}) &\stackrel{3.3}{=} p(A_{n_t} | A_{n_{t-1}} \cap \dots \cap A_{n_0}) \cdots p(A_{n_1} | A_{n_0}) \\ &\stackrel{4.9}{=} p(A_{n_t} | A_{n_{t-1}}) \cdots p(A_{n_1} | A_{n_0}) \end{aligned}$$

Proposizione 4.11. $\bigcirc_{n,i} E_n^i$ sia una sequenza di Markov e $\bigcirc_n A_n$ una sua realizzazione. Siano $s, t \in \mathbb{N}$ con $s < t$ ed $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ tali che $p(A_{n_s} \cap \dots \cap A_{n_0}) \neq 0$. Allora

$$\begin{aligned} p(A_{n_t} \cap \dots \cap A_{n_1} | A_{n_0}) \\ = p(A_{n_t} \cap \dots \cap A_{n_{s+1}} | A_{n_s}) \cdot p(A_{n_s} | A_{n_{s-1}}) \cdots p(A_{n_1} | A_{n_0}) \end{aligned}$$

Dimostrazione. Infatti

$$\begin{aligned} p(A_{n_t} \cap \dots \cap A_{n_1} | A_{n_0}) \\ \stackrel{3.4}{=} p(A_{n_t} \cap \dots \cap A_{n_{s+1}} | A_{n_s} \cap \dots \cap A_{n_0}) \cdot p(A_{n_s} | A_{n_{s-1}} \cap \dots \cap A_{n_0}) \cdots p(A_{n_1} | A_{n_0}) \\ \stackrel{4.9}{=} p(A_{n_t} \cap \dots \cap A_{n_{s+1}} | A_{n_s}) \cdot p(A_{n_s} | A_{n_{s-1}}) \cdots p(A_{n_1} | A_{n_0}) \end{aligned}$$

Proposizione 4.12. $\bigcirc_{n,i} E_n^i$ sia una sequenza di S -partizioni \mathcal{A} -misurabili.

Allora sono equivalenti:

- (1) $\bigcirc_{n,i} E_n^i$ è una sequenza di Markov.
- (2) Per ogni realizzazione $\bigcirc_n A_n$ di $\bigcirc_{n,i} E_n^i$ ed ogni $n \in \mathbb{N}$ tali che $p(A_n \cap \dots \cap A_0) \neq 0$ vale

$$p(A_{n+1} \cap \dots \cap A_1 | A_0) = p(A_{n+1} | A_n) \cdots p(A_1 | A_0).$$

Dimostrazione.

- (1) \implies (2): Corollario 4.10.

(2) \implies (1):

$$\begin{aligned} p(A_{n+1}|A_n \cap \dots \cap A_0) &= \frac{p(A_{n+1} \cap \dots \cap A_0)}{p(A_n \cap \dots \cap A_0)} \\ &= \frac{p(A_{n+1}|A_n) \cdot p(A_n|A_{n-1}) \cdots p(A_1|A_0)}{p(A_n|A_{n-1}) \cdots p(A_1|A_0)} \\ &= p(A_{n+1}|A_n) \end{aligned}$$

Definizione 4.13. Una successione $\bigcirc_n A_n$ di elementi di \mathcal{A} si chiama *indipendente* se per ogni $n_1, \dots, n_t \in \mathbb{N}$ con $n_1 < \dots < n_t$ si ha

$$p(A_{n_1} \cap \dots \cap A_{n_t}) = p(A_{n_1}) \cdots p(A_{n_t})$$

5. Catene di Markov finite

Situazione 5.1. Siano (Ω, \mathcal{A}, p) uno spazio di probabilità, $q \in \mathbb{N} + 1$ fissato ed $S := \{1, \dots, q\}$.

Definizione 5.2. Un *processo stocastico finito* (con spazio degli stati S ed a tempo discreto) è una successione $\bigcirc_n X_n$ di variabili aleatorie elementari tali che $X_n(\Omega) \subset S$ per ogni $n \in \mathbb{N}$. Gli elementi di S si chiamano *stati del processo*. Si noti che le nostre ipotesi implicano $\sum_{i \in S} p(X_n = i) = 1$ per ogni n .

Riformuliamo adesso i risultati del capitolo precedente nel linguaggio dei processi stocastici.

Definizione 5.3. Nel seguito, in analogia con la def. 1.21, usiamo la notazione

$$p(E_1, \dots, E_m | F_1, \dots, F_l) := p(E_1 \cap \dots \cap E_m | F_1 \cap \dots \cap F_l)$$

per $E_1, \dots, E_m, F_1, \dots, F_l \in \mathcal{A}$, soprattutto per eventi della forma $(X_n = k)$ per un processo stocastico $\bigcirc_n X_n$.

Osservazione 5.4. $\bigcirc_n X_n$ sia una processo stocastico finito.

Per $n \in \mathbb{N}$ ed $i \in S$ poniamo $E_n^i := (X_n = i)$. Allora $\bigcirc_{n,i} E_n^i$ è una sequenza di S -partizioni \mathcal{A} -misurabili. Per ogni successione $\bigcirc_n i_n : \mathbb{N} \rightarrow S$ la successione $\bigcirc_n (X_n = i_n)$ è una realizzazione della sequenza.

Definizione 5.5. Una *catena di Markov finita* (con spazio degli stati S ed a tempo discreto) è un processo stocastico finito $\bigcirc_n X_n$ per il quale l'applicazione $\bigcirc_{n,i} (X_n = i) : \mathbb{N} \times S \rightarrow \mathcal{A}$ è una sequenza di Markov.

Chiediamo in altre parole che per $n \in \mathbb{N}$ ed $i_{n+1}, \dots, i_0 \in S$ con $p(X_n = i_n, \dots, X_0 = i_0) \neq 0$ si abbia sempre

$$p(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) = p(X_{n+1} = i_{n+1} | X_n = i_n)$$

Osservazione 5.6. $\bigcirc_n X_n$ sia un processo stocastico finito. Allora sono equivalenti:

- (1) $\bigcirc_n X_n$ è una catena di Markov.
- (2) Per $n \in \mathbb{N}$ ed $i_{n+1}, \dots, i_0 \in S$ vale

$$\begin{aligned} p(X_{n+1} = i_{n+1}, \dots, X_0 = i_0) \\ = p(X_{n+1} = i_{n+1}, X_n = i_n) \cdot p(X_{n-1} = i_{n-1}, \dots, X_0 = i_0 | X_n = i_n). \end{aligned}$$

Dimostrazione. Oss. 4.4.

Lemma 5.7. $\bigcirc_n X_n$ sia una catena di Markov finita.

Siano $n, m \in \mathbb{N}$ con $0 \leq m \leq n$ ed $i_{n+1}, \dots, i_m \in S$ tali che $p(X_n = i_n, \dots, X_m = i_m) \neq 0$. Allora

$$p(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_m = i_m) = p(X_{n+1} = i_{n+1} | X_n = i_n)$$

Dimostrazione. Lemma 4.5.

Proposizione 5.8. $\bigcirc_n X_n$ sia una catena di Markov finita. Siano $n \in \mathbb{N}$, $r \in \mathbb{N} + 1$, $m \in \mathbb{N}$ con $0 \leq m \leq n$ ed $i_{n+r}, \dots, i_m \in S$ tali che $p(X_n = i_n, \dots, X_m = i_m) \neq 0$. Allora

$$\begin{aligned} p(X_{n+r} = i_{n+r}, \dots, X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_m = i_m) \\ = p(X_{n+r} = i_{n+r}, \dots, X_{n+1} = i_{n+1} | X_n = i_n) \end{aligned}$$

Dimostrazione. Prop. 4.6.

Teorema 5.9. $\bigcirc_n X_n$ sia una catena di Markov finita.

Siano $n \in \mathbb{N}$, $r \in \mathbb{N} + 1$ ed $i \in S$, $J \subset S^r$, $K \subset S^n$ tali che $p(X_n = i, (X_{n-1}, \dots, X_0) \in K) \neq 0$. Allora

$$\begin{aligned} p((X_{n+r}, \dots, X_{n+1}) \in J | X_n = i, (X_{n-1}, \dots, X_0) \in K) = \\ p((X_{n+r}, \dots, X_{n+1}) \in J | X_n = i) \end{aligned}$$

Dimostrazione. Teorema 4.8.

Corollario 5.10. $\bigcirc_n X_n$ sia una catena di Markov finita.

Siano $s, t \in \mathbb{N}$ con $s < t$, $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ ed $i_0, \dots, i_t \in S$ tali che $p(X_{n_s} = i_s, \dots, X_{n_0} = i_0) \neq 0$. Allora

$$\begin{aligned} p(X_{n_t} = i_t, \dots, X_{n_{s+1}} = i_{s+1} | X_{n_s} = i_s, \dots, X_{n_0} = i_0) \\ = p(X_{n_t} = i_t, \dots, X_{s+1} = i_{s+1} | X_{n_s} = i_s). \end{aligned}$$

Dimostrazione. Cor. 4.9.

Corollario 5.11. $\bigcirc_{n,i} X_n$ sia una catena di Markov finita.

Siano $t \in \mathbb{N} + 1$, $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ ed $i_0, \dots, i_t \in S$ tali che $p(X_{n_{t-1}} = i_{t-1}, \dots, X_{n_0} = i_0) \neq 0$. Allora

$$\begin{aligned} p(X_{n_t} = i_t, \dots, X_{n_1} = i_1 | X_{n_0} = i_0) \\ = p(X_{n_t} = i_t | X_{n_{t-1}} = i_{t-1}) \cdots p(X_{n_1} = i_1 | X_{n_0} = i_0) \end{aligned}$$

$$\begin{aligned} p(X_{n_t} = i_t, \dots, X_{n_0} = i_0) \\ = p(X_{n_t} = i_t | X_{n_{t-1}} = i_{t-1}) \cdots p(X_{n_1} = i_1 | X_{n_0} = i_0) \cdot p(X_{n_0} = i_0) \end{aligned}$$

Dimostrazione. Cor. 4.10.

Proposizione 5.12. $\bigcirc_{n,i} X_n$ sia una catena di Markov finita.

Siano $s, t \in \mathbb{N}$ con $s < t$, $n_0, \dots, n_t \in \mathbb{N}$ con $n_0 < \dots < n_t$ ed $i_0, \dots, i_t \in S$ tali che $p(X_{n_s} = i_s, \dots, X_{n_0} = i_0) \neq 0$. Allora

$$\begin{aligned}
& p(X_{n_t} = i_t, \dots, X_{n_1} = i_1 | X_{n_0} = i_0) \\
&= p(X_{n_t} = i_t, \dots, X_{n_{s+1}} = i_{s+1} | X_{n_s} = i_s) \\
&\quad \cdot p(X_{n_s} = i_s | X_{n_{s-1}} = i_{s-1}) \cdots p(X_{n_1} = i_1 | X_{n_0} = i_0)
\end{aligned}$$

Dimostrazione. Prop. 4.11.

Proposizione 5.13. $\bigcirc_{n,i} X_n$ sia un processo stocastico finito. Allora sono equivalenti:

- (1) $\bigcirc_{n,i} X_n$ è una catena di Markov.
- (2) Per ogni $n \in \mathbb{N}$ ed ogni $i_{n+1}, \dots, i_0 \in S$ tali che $p(X_n = i_n, \dots, X_0 = i_0) \neq 0$ vale
$$\begin{aligned}
& p(X_{n+1} = i_{n+1}, \dots, X_1 = i_1 | X_0 = i_0) \\
&= p(X_{n+1} = i_{n+1} | X_n = i_n) \cdots p(X_1 = i_1 | X_0 = i_0).
\end{aligned}$$

Dimostrazione. Prop. 4.12.

Definizione 5.14. $\bigcirc_{n,i} X_n$ sia un processo stocastico finito. Le variabili aleatorie $\bigcirc_n X_n$ si dicono *indipendenti*, se per ogni successione $\bigcirc_n i_n : \mathbb{N} \rightarrow S$ la successione $\bigcirc_n (X_n = i_n)$ è indipendente. Chiediamo in altre parole che per ogni $n_1, \dots, n_t \in \mathbb{N}$ con $n_1 < \dots < n_t$ si abbia

$$p(X_{n_1} = i_{n_1}, \dots, X_{n_t} = i_{n_t}) = p(X_{n_1} = i_{n_1}) \cdots p(X_{n_t} = i_{n_t})$$

Osservazione 5.15. $\bigcirc_n X_n$ sia un processo stocastico finito. Le variabili aleatorie $\bigcirc_n X_n$ siano indipendenti. Allora $\bigcirc_n X_n$ è una catena di Markov.

Dimostrazione. Siano $n \in \mathbb{N}$ ed $i_{n+1}, \dots, i_0 \in S$ tali che $p(X_n = i_n, \dots, X_0 = i_0) \neq 0$. Allora

$$\begin{aligned}
p(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) &= \frac{p(X_{n+1} = i_{n+1}, \dots, X_0 = i_0)}{p(X_n = i_n, \dots, X_0 = i_0)} \\
&= \frac{p(X_{n+1} = i_{n+1}) \cdots p(X_0 = i_0)}{p(X_n = i_n) \cdots p(X_0 = i_0)} \\
&= \frac{p(X_{n+1} = i_{n+1}) \cdot p(X_n = i_n)}{p(X_n = i_n)} \\
&= p(X_{n+1} = i_{n+1} | X_n = i_n)
\end{aligned}$$

Nota 5.16. L'insieme S sia dotato di una struttura di gruppo. $\bigcirc_n Y_n$ sia un processo stocastico finito e le variabili aleatorie Y_n indipendenti. Per $n \in \mathbb{N}$ definiamo $X_n := Y_n \cdot Y_{n-1} \cdots Y_0$. Più esplicitamente per ogni $\omega \in \Omega$ si ha $X_n(\omega) = Y_n(\omega) \cdot Y_{n-1}(\omega) \cdots Y_0(\omega)$ usando il prodotto definito nel gruppo S . Allora $\bigcirc_n X_n$ è una catena di Markov.

Dimostrazione. Siano $n \in \mathbb{N}$ ed $i_0, \dots, i_{n+1} \in S$. Allora

$$\begin{aligned}
 p(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_0 = i_0) &= \frac{p(X_{n+1} = i_{n+1}, \dots, X_0 = i_0)}{p(X_n = i_n, \dots, X_0 = i_0)} \\
 &= \frac{p(Y_0 = i_0, Y_1 = i_1 i_0^{-1}, \dots, Y_{n+1} = i_{n+1} i_n^{-1})}{p(Y_0 = i_0, Y_1 = i_1 i_0^{-1}, \dots, Y_n = i_n i_{n-1}^{-1})} \\
 &= \frac{p(Y_0 = i_0) \cdots p(Y_{n+1} = i_{n+1} i_n^{-1})}{p(Y_0 = i_0) \cdots p(Y_n = i_n i_{n-1}^{-1})} \\
 &= p(Y_{n+1} = i_{n+1} i_n^{-1})
 \end{aligned}$$

D'altra parte però

$$\begin{aligned}
 p(X_{n+1} = i_{n+1}, X_n = i_n) &= \sum_{(i_{n-1}, \dots, i_0) \in S^n} p(X_{n+1} = i_{n+1}, X_n = i_n, \dots, X_0 = i_0) \\
 &= \sum_{(i_{n-1}, \dots, i_0) \in S^n} p(Y_0 = i_0, \dots, Y_{n+1} = i_{n+1} i_n^{-1}) \\
 &= \sum_{(i_{n-1}, \dots, i_0) \in S^n} p(Y_0 = i_0) \cdots p(Y_{n+1} = i_{n+1} i_n^{-1}) \\
 &= p(Y_{n+1} = i_{n+1} i_n^{-1}) \cdot \sum_{(i_{n-1}, \dots, i_0) \in S^n} p(Y_0 = i_0) \cdots p(Y_n = i_n i_{n-1}^{-1}) \\
 &= p(Y_{n+1} = i_{n+1} i_n^{-1}) \cdot p(X_n = i_n)
 \end{aligned}$$

e ciò implica l'enunciato.

Definizione 5.17. Una catena di Markov $\bigcirc_n X_n$ si dice *omogenea*, se sono soddisfatte le seguenti condizioni:

- (1) Per ogni $j \in S$ esiste $n \in \mathbb{N}$ tale che $p(X_n = j) > 0$.
 Poniamo $\text{est}(j) := \{n \in \mathbb{N} \mid p(X_n = j) > 0\}$.
 Gli elementi di $\text{est}(j)$ si dicono *tempi di esistenza di j*.
- (2) Per ogni $i, j \in S$ le probabilità condizionate $p(X_{n+1} = i | X_n = j)$ per $n \in \text{est}(j)$ sono tutte uguali.
 Poniamo $T_{(j)}^{(i)} := p(X_{n+1} = i | X_n = j)$ per $n \in \text{est}(j)$.

Otteniamo così una matrice $T \in [0, 1]_q^q$ che chiameremo *matrice di transizione* della catena di Markov data. Indichiamo gli indici tra parentesi per distinguerli dagli esponenti delle potenze di T.

6. La matrice di transizione

Situazione 6.1. Sia (Ω, \mathcal{A}, p) uno spazio di probabilità, $q \in \mathbb{N}+1$ fissato ed $S := \{1, \dots, q\}$. $\bigcirc_n X_n$ sia una catena di Markov finita (con spazio degli stati S) ed omogenea con matrice di transizione T .

Definizione 6.2. Per $i \in S$ ed $n \in \mathbb{N}$ poniamo $p_n^{(i)} := p(X_n = i)$. Otteniamo così per ogni $n \in \mathbb{N}$ un vettore $p_n \in [0, 1]_q^q$.

Osservazione 6.3. Sia $n \in \mathbb{N}$. Allora:

- (1) $p_{n+1} = T p_n$.
- (2) $p_n = T^n p_0$.

Dimostrazione. Dimostriamo la prima uguaglianza di cui la seconda è una conseguenza immediata. Sia $R := \{k \in S \mid p(X_n = k) \neq 0\}$. Per ogni $i \in S$ abbiamo

$$\begin{aligned} p_{n+1}^{(i)} &= p(X_{n+1} = i) = \sum_{k \in S} p(X_{n+1} = i, X_n = k) \\ &= \sum_{k \in S} p(X_{n+1} = i | X_n = k) \cdot p(X_n = k) \\ &= \sum_{k \in R} p(X_{n+1} = i | X_n = k) \cdot p(X_n = k) \\ &= \sum_{k \in R} T_{(k)}^{(i)} p_n^{(k)} = \sum_{k \in S} T_{(k)}^{(i)} p_n^{(k)} = (T p_n)^{(i)} \end{aligned}$$

Corollario 6.4. Siano $n, n_1, n_0 \in \mathbb{N}$ con $n_0 \leq n_1 \leq n$. Allora

$$p_n = T^{n-n_1} \cdot T^{n_1-n_0} \cdot p_0 = T^{n-n_1} \cdot p_{n_1-n_0}$$

Teorema 6.5 (equazioni di Chapman/Kolmogorov). Siano $i, j \in S$ ed $n, m \in \mathbb{N}$ tali che $p(X_n = j) \neq 0$. Allora

$$p(X_{n+m} = i | X_n = j) = (T^m)_{(j)}^{(i)}$$

Dimostrazione. L'enunciato è banale per $m = 0$.

Procediamo per induzione su $m \in \mathbb{N} + 1$.

$m = 1$: Definizione di T .

$m \implies m + 1$: Sia $R := \{k \in S \mid p(X_{n+m} = k, X_n = j) \neq 0\}$.

Si osservi che per $k \notin R$

$$(T^m)_{(j)}^{(k)} \stackrel{ind}{=} p(X_{n+m} = k | X_n = j) = 0$$

Abbiamo

$$\begin{aligned}
p(X_{n+m+1} = i | X_n = j) &= \sum_{k \in S} p(X_{n+m+1} = i, X_{n+m} = k | X_n = j) \\
&\stackrel{3.2}{=} \sum_{k \in S} p(X_{n+m+1} = i | X_{n+m} = k, X_n = j) p(X_{n+m} = k | X_n = j) \\
&= \sum_{k \in R} p(X_{n+m+1} = i | X_{n+m} = k, X_n = j) p(X_{n+m} = k | X_n = j) \\
&\stackrel{5.10}{=} \sum_{k \in R} p(X_{n+m+1} = i | X_{n+m} = k) p(X_{n+m} = k | X_n = j) \\
&\stackrel{ind}{=} \sum_{k \in R} T_{(k)}^{(i)} (T^m)_{(j)}^{(k)} = \sum_{k \in S} T_{(k)}^{(i)} (T^m)_{(j)}^{(k)} = (T^{m+1})_{(j)}^{(i)}
\end{aligned}$$

Definizione 6.6. Un vettore $v \in \mathbb{R}^q$ si dice *stocastico*, se $v \in [0, 1]^q$ e $\sum_{k=1}^q v^{(k)} = 1$.

Una matrice $A \in \mathbb{R}_q^q$ si dice *stocastica* (per colonne), se ogni colonna di A è un vettore stocastico.

Osservazione 6.7. (1) Per ogni $n \in \mathbb{N}$ il vettore p_n è stocastico.

(2) La matrice T è stocastica.

Dimostrazione. (1) $\sum_{i \in S} p_n^{(i)} = \sum_{i \in S} p(X_n = i) = 1$.

(2) Per $j \in S$ scegliamo $n \in \text{est}(j)$. Allora

$$\sum_{i \in S} T_{(j)}^{(i)} = \sum_{i \in S} p(X_{n+1} = i | X_n = j) = 1 \text{ perché } p(X_n = j) > 0.$$

Osservazione 6.8. Siano A, B due matrici stocastiche e v un vettore stocastico. Allora AB è una matrice stocastica e Av è un vettore stocastico.

Dimostrazione. La verifica è immediata (ad es. Castaldi, pag. 14, o Huppert, pag. 384).

Corollario 6.9. Le potenze T^n di T sono matrici stocastiche.

Osservazione 6.10. Dal teorema di esistenza di Kolmogorov si può dedurre che per ogni matrice stocastica \tilde{T} ed ogni vettore stocastico \tilde{p}_0 esiste una catena di Markov $\bigcirc_n \tilde{X}_n$ finita e omogenea con matrice di transizione \tilde{T} e tale che $\tilde{p}_0^{(i)} = p(\tilde{X}_0 = i)$ per ogni i .

Per questa ragione nelle applicazioni si parte spesso da una matrice stocastica T i cui coefficienti $T_{(j)}^{(i)}$ vengono interpretati probabilisticamente secondo una dizione della forma

$$T_{(j)}^{(i)} = p(X_{n+1} = i | X_n = j) \text{ se } p(X_n = j) \neq 0$$

Dimostrazione. Iosifescu/Tautu, pagg. 14-15.

7. Esempi di catene di Markov.

Lemma 7.1. Siano $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}_2^2$ e λ, μ gli autovalori di A , $n \in \mathbb{N}$.

(1) Sia $\lambda \neq \mu$. Allora

$$A^n = \frac{1}{\lambda - \mu} \begin{pmatrix} \lambda - d & b \\ c & d - \mu \end{pmatrix} \lambda^n + \frac{1}{\lambda - \mu} \begin{pmatrix} d - \mu & -b \\ -c & \lambda - d \end{pmatrix} \mu^n$$

(2) Sia $\lambda = \mu$. Allora

$$A^n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \lambda^n + \begin{pmatrix} \lambda - d & b \\ c & d - \lambda \end{pmatrix} n \lambda^{n-1}$$

Si osservi che in entrambi i casi $\lambda + \mu = a + d$, per cui nell'ipotesi $\lambda = \mu$ vale $\lambda = \frac{a+d}{2}$.

Dimostrazione. Lugli, teorema 2.7 e corollario 2.15.

Nota 7.2. T sia una matrice stocastica 2×2 . Allora T può essere scritta nella forma $T = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}$ con $a, b \in [0, 1]$. Inoltre:

(1) $|a - b| \leq 1$.

(2) Gli autovalori di T sono 1 ed $a - b$.

(3) Gli autovalori di T coincidono quindi se e solo se $a - b = 1$.

In tal caso $T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} =: \delta$ e si ha $T^n = \delta$ per ogni $n \in \mathbb{N}$.

(4) Se $a - b = -1$, allora $T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, cosicché $T^n = \delta$

per $n \in 2\mathbb{N}$ e $T^n = T$ per $n \in 2\mathbb{N} + 1$, per cui in questo caso $\lim_{n \rightarrow \infty} T^n$ non esiste.

(5) Sia invece $|a - b| < 1$. Allora $1 - a + b \neq 0$ e

$$\lim_{n \rightarrow \infty} T^n = \frac{1}{1 - a + b} \begin{pmatrix} b & b \\ 1 - a & 1 - a \end{pmatrix}$$

Se poniamo $c := 1 - a$ e $d := 1 - b$, abbiamo quindi

$$T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$T^n = \frac{1}{b+c} \begin{pmatrix} b & b \\ c & c \end{pmatrix} + \frac{1}{b+c} \begin{pmatrix} c & -b \\ -c & b \end{pmatrix} (1 - b - c)^n$$

$$\lim_{n \rightarrow \infty} T^n = \frac{1}{b+c} \begin{pmatrix} b & b \\ c & c \end{pmatrix}$$

Si osservi che dai punti precedenti segue che $|a - b| < 1$ se e solo se $0 < b < 1$ e $0 < c < 1$. Infine per un vettore stocastico

$$p_0 \in [0, 1]^2 \text{ nel nostro caso il limite } \lim_{n \rightarrow \infty} T^n p_0 = \frac{1}{b+c} \begin{pmatrix} b \\ c \end{pmatrix}$$

non dipende da p_0 .

Dimostrazione. (1) Ovvio, poiché $a, b \in [0, 1]$.

(2) Il determinante di T è uguale ad $a(1 - b) - b(1 - a) = a - b$, per cui il polinomio caratteristico di T è dato da

$$x^2 - x \operatorname{tr} A + \det A = x^2 - (1 + a - b)x + a - b = (x - 1)(x - a + b)$$

(3) $a, b \in [0, 1]$ implica che $a - b = 1$ è possibile solo se $b = 0, a = 1$.

(4) Similmente $a - b = -1$ è possibile solo se $a = 0, b = 1$.

(5) Sia $|a - b| < 1$. Per il punto (2) gli autovalori di T sono $\lambda = 1$ e $\mu = a - b$. L'ipotesi implica $\lambda \neq \mu$, inoltre abbiamo $\lambda - \mu = 1 - a + b$, $\lambda - d = b, d - \mu = 1 - a = c$. Dall'oss. 7.1 otteniamo

$$T^n = \frac{1}{b+c} \begin{pmatrix} b & b \\ c & c \end{pmatrix} + \frac{1}{b+c} \begin{pmatrix} c & -b \\ -c & b \end{pmatrix} \mu^n$$

Siccome però $|\mu| < 1$ il secondo sommando tende a zero, per cui

$$\lim_{n \rightarrow \infty} T^n = \frac{1}{1 - a + b} \begin{pmatrix} b & b \\ c & c \end{pmatrix}$$

Sia infine $p_0 = \begin{pmatrix} p_0^{(1)} \\ p_0^{(2)} \end{pmatrix}$ un vettore stocastico. Allora

$$\frac{1}{b+c} \begin{pmatrix} b & b \\ c & c \end{pmatrix} p_0 = \frac{1}{b+c} \begin{pmatrix} bp_0^{(1)} + bp_0^{(2)} \\ cp_0^{(1)} + cp_0^{(2)} \end{pmatrix} = \frac{1}{b+c} \begin{pmatrix} b \\ c \end{pmatrix}$$

Corollario 7.3. $\bigcirc_n X_n$ sia una catena di Markov omogenea con spazio

degli stati $S = \{1, 2\}$ e matrice di transizione $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Abbiamo quindi $p(X_{n+1} = 1 | X_n = 2) = b$ per $p(X_n = 2) \neq 0$ e

$p(X_{n+1} = 2 | X_n = 1) = c$ per $p(X_n = 1) \neq 0$.

Assumiamo $0 < b < 1$ e $0 < c < 1$. Allora:

(1) $\lim_{n \rightarrow \infty} T^n = \frac{1}{b+c} \begin{pmatrix} b & b \\ c & c \end{pmatrix}$.

(2) Le probabilità limite

$$\lim_{n \rightarrow \infty} p(X_n = 1) = \frac{b}{b+c}$$

$$\lim_{n \rightarrow \infty} p(X_n = 2) = \frac{c}{b+c}$$

non dipendono dalle probabilità iniziali $p(X_0 = 1)$ e $p(X_0 = 2)$.

Esempio 7.4. In un'elezione si presentano i candidati A e B. La campagna elettorale dura venti giorni. Siano

$$p(X_{n+1} = A | X_n = A) = 0.8$$

$$p(X_{n+1} = B | X_n = A) = 0.2$$

$$p(X_{n+1} = A | X_n = B) = 0.3$$

$$p(X_{n+1} = B | X_n = B) = 0.7$$

quando $p(X_n = A) \neq 0$ e $p(X_n = B) \neq 0$. La matrice di transizione è perciò

$$T = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}$$

Gli autovalori di T sono $\lambda = 1, \mu = 0.5$. Per la nota 7.2 abbiamo

$$\begin{aligned} T^{20} &= \frac{1}{0.5} \begin{pmatrix} 0.3 & 0.3 \\ 0.2 & 0.2 \end{pmatrix} + \frac{1}{0.5} \begin{pmatrix} 0.2 & -0.3 \\ -0.2 & 0.3 \end{pmatrix} 0.5^{20} \\ &= \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix} + \begin{pmatrix} 0.4 & -0.6 \\ -0.6 & 0.6 \end{pmatrix} \frac{1}{2^{20}} \\ &\sim \begin{pmatrix} 0.6 & 0.6 \\ 0.4 & 0.4 \end{pmatrix} \end{aligned}$$

Dopo venti giorni il 60% della popolazione del paese X voterà per il candidato A anche se all'inizio tutti erano a favore del candidato B.

Allo stesso modo si possono pronosticare anche cambiamenti dal punto di vista della condizione sociale, come ad esempio:

ricco	povero
sano	malato
contadino	altra professione
cristiano	non cristiano
diligente	pigro

Esempio 7.5. Sia $S = \{1, \dots, q\}$. Assumiamo che delle particelle si muovano sui punti di S con probabilità a verso sinistra (cioè da j a $j - 1$), con probabilità b verso destra (da j a $j + 1$) e rimangono nella posizione che occupano con probabilità r , tranne nelle due estremità dove rimangono intrappolate se le raggiungono. Ciò corrisponde ad una matrice di transizione della forma

$$T = \begin{pmatrix} 1 & a & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & r & a & 0 & \dots & 0 & 0 & 0 \\ 0 & b & r & a & \dots & 0 & 0 & 0 \\ 0 & 0 & b & r & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & r & a & 0 \\ 0 & 0 & 0 & 0 & \dots & b & r & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & b & 1 \end{pmatrix}$$

(1) Siano $a = 0.6, r = 0, b = 0.4$ per $q = 4$. Allora, arrotondando a due cifre decimali,

$$T = \begin{pmatrix} 1 & 0.6 & 0 & 0 \\ 0 & 0 & 0.6 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 1 \end{pmatrix}$$

$$T^8 = \begin{pmatrix} 1 & 0.79 & 0.47 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.21 & 0.52 & 1 \end{pmatrix} \simeq T^{20}$$

Vediamo che già dopo otto passaggi praticamente tutte le particelle si trovano alle due estremità: delle particelle che inizialmente si trovano in posizione 2, il 79% è rimasto intrappolato nell'estremità sinistra e il 21% nell'estremità destra; di quelle che inizialmente si trovano in posizione 3, il 47% è finito in 1 e il 53% in 4.

(2) Siano $a = 0.5$, $r = 0.2$, $b = 0.3$ per $q = 4$. Allora

$$T = \begin{pmatrix} 1 & 0.5 & 0 & 0 \\ 0 & 0.2 & 0.5 & 0 \\ 0 & 0.3 & 0.2 & 0 \\ 0 & 0 & 0.3 & 1 \end{pmatrix}$$

$$T^8 = \begin{pmatrix} 1 & 0.81 & 0.50 & 0 \\ 0 & 0.01 & 0.01 & 0 \\ 0 & 0.01 & 0.01 & 0 \\ 0 & 0.18 & 0.48 & 1 \end{pmatrix}$$

$$T^{20} = \begin{pmatrix} 1 & 0.82 & 0.51 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.18 & 0.49 & 1 \end{pmatrix}$$

In questo caso la convergenza è più lenta, soprattutto per la presenza del termine $r = 0.2$.

Esempio 7.6. Ci poniamo nella stessa situazione dell'es. 7.5, però con la possibilità di tornare indietro una volta arrivati alle due estremità, quindi con

$$T = \begin{pmatrix} 1-b & a & 0 & 0 & \dots & 0 & 0 & 0 \\ b & r & a & 0 & \dots & 0 & 0 & 0 \\ 0 & b & r & a & \dots & 0 & 0 & 0 \\ 0 & 0 & b & r & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & r & a & 0 \\ 0 & 0 & 0 & 0 & \dots & b & r & a \\ 0 & 0 & 0 & 0 & \dots & 0 & b & 1-a \end{pmatrix}$$

Siano $a = 0.5$, $r = 0.2$, $b = 0.3$ per $q = 4$. Allora

$$T = \begin{pmatrix} 0.7 & 0.5 & 0 & 0 \\ 0.3 & 0.2 & 0.5 & 0 \\ 0 & 0.3 & 0.2 & 0.5 \\ 0 & 0 & 0.3 & 0.5 \end{pmatrix}$$

$$T^8 = \begin{pmatrix} 0.49 & 0.46 & 0.42 & 0.38 \\ 0.28 & 0.28 & 0.27 & 0.27 \\ 0.15 & 0.16 & 0.19 & 0.21 \\ 0.08 & 0.10 & 0.12 & 0.15 \end{pmatrix}$$

$$T^{20} = \begin{pmatrix} 0.46 & 0.46 & 0.46 & 0.46 \\ 0.28 & 0.28 & 0.28 & 0.28 \\ 0.17 & 0.17 & 0.17 & 0.17 \\ 0.10 & 0.10 & 0.10 & 0.10 \end{pmatrix}$$

$$T^{100} = \begin{pmatrix} 0.46 & 0.46 & 0.46 & 0.46 \\ 0.28 & 0.28 & 0.28 & 0.28 \\ 0.17 & 0.17 & 0.17 & 0.17 \\ 0.1 & 0.1 & 0.1 & 0.1 \end{pmatrix}$$

In questo caso tutti i coefficienti di T^8 sono positivi e per un teorema generale (che verrà formulato e dimostrato più avanti) la matrice T^n converge ad una matrice le cui colonne sono uguali tra loro.

Esempio 7.7. Sia $S := \{1, \dots, q\}$ con $q \geq 2$. Ci poniamo nella stessa situazione dell'es. 7.6, però supponiamo che il movimento delle particelle sia ciclico, cioè con la possibilità di passaggi da 1 a q e viceversa. La matrice di transizione ha allora la forma

$$T = \begin{pmatrix} r & a & 0 & 0 & \dots & 0 & 0 & b \\ b & r & a & 0 & \dots & 0 & 0 & 0 \\ 0 & b & r & a & \dots & 0 & 0 & 0 \\ 0 & 0 & b & r & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & r & a & 0 \\ 0 & 0 & 0 & 0 & \dots & b & r & a \\ a & 0 & 0 & 0 & \dots & 0 & b & r \end{pmatrix}$$

Siano $a = 0.5$, $r = 0.2$, $b = 0.3$ on $q = 4$. Allora

$$T = \begin{pmatrix} 0.2 & 0.5 & 0 & 0.3 \\ 0.3 & 0.2 & 0.5 & 0 \\ 0 & 0.3 & 0.2 & 0.5 \\ 0.5 & 0 & 0.3 & 0.2 \end{pmatrix}$$

$$T^8 = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

$$T^{100} = \begin{pmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{pmatrix}$$

Modelli compartimentali ciclici possono essere utilizzati nella descrizione di tessuti se si assume che in una sequenza di stadi di differen-

ziazione l'uscita dall'ultimo stadio provoca un evento di retroazione stimolando così la nascita di nuove cellule nello stadio iniziale.

Esempio 7.8 (modello di Ehrenfest). Sia $S := \{1, \dots, q\}$ con $q \geq 2$. $m := q - 1$ particelle siano distribuite in modo arbitrario tra due compartimenti A e B di un contenitore, separati da una membrana attraverso la quale le particelle possono diffondere. In ogni stato di tempo una molecola diffonde da A a B oppure da B ad A con probabilità rispettivamente pari a $\frac{j-1}{m}$ e $1 - \frac{j-1}{m}$.

Nell'interpretazione si ha $X_n = j$, se il comparto A contiene $j - 1$ particelle. Abbiamo quindi

$$p(X_{n+1} = i | X_n = j) = \begin{cases} 0 & \text{se } i \neq j \pm 1 \\ 1 - \frac{j-1}{m} & \text{se } i = j + 1 \\ \frac{j-1}{m} & \text{se } i = j - 1 \end{cases}$$

Ciò porta alla matrice di transizione

$$T = \begin{pmatrix} 0 & 1/m & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 2/m & 0 & \dots & 0 & 0 \\ 0 & 1 - 1/m & 0 & 3/m & \dots & 0 & 0 \\ 0 & 0 & 1 - 2/m & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 - 3/m & \dots & 1/m & 1 \end{pmatrix}$$

Per $m = 4$ abbiamo quindi

$$T = \begin{pmatrix} 0 & 0.25 & 0 & 0 & 0 \\ 1 & 0 & 0.50 & 0 & 0 \\ 0 & 0.75 & 0 & 0.75 & 0 \\ 0 & 0 & 0.50 & 0 & 0 \\ 0 & 0 & 0 & 0.25 & 1 \end{pmatrix}$$

Esempio 7.9. Riportiamo da Kemeny/Snell, pagg. 39-40, il seguente esempio.

Tre carri armati combattono tra di loro. Il carro armato A colpisce i suoi avversari con probabilità $2/3$, B con probabilità $1/2$, C con probabilità $1/3$. I colpi sono sparati simultaneamente e se un carro armato è colpito non è più in grado di combattere. Assumiamo che ciascuno dei carri armati scelga di colpire ogni volta il nemico più forte tra quelli rimasti. Allora otteniamo la seguente matrice di transizione:

$$T = \begin{pmatrix} & E & A & B & C & AC & BC & ABC \\ E & 1 & 0 & 0 & 0 & 2/9 & 1/6 & 0 \\ A & 0 & 1 & 0 & 0 & 4/9 & 0 & 0 \\ B & 0 & 0 & 1 & 0 & 0 & 1/3 & 0 \\ C & 0 & 0 & 0 & 1 & 1/9 & 1/6 & 4/9 \\ AC & 0 & 0 & 0 & 0 & 2/9 & 0 & 2/9 \\ BC & 0 & 0 & 0 & 0 & 0 & 1/3 & 2/9 \\ ABC & 0 & 0 & 0 & 0 & 0 & 0 & 1/9 \end{pmatrix}$$

In essa con E indichiamo l'evento in cui nessuno dei carri armati so-

pravvive, con A (risp. B, C) l'evento in cui solo il carro armato A (risp. B, C) sopravvive, con AC (risp. BC) l'evento in cui sopravvivono solo i carri A e C (risp. B e C); con ABC l'evento in cui sopravvivono tutti e tre i carri armati. Non è presente l'evento AB poiché il carro armato C può essere eliminato solo dopo la distruzione di uno degli altri due.

$$\text{Si trova che } \lim_{n \rightarrow \infty} T^n \simeq T^{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0.29 & 0.25 & 0.13 \\ 0 & 1 & 0 & 0 & 0.57 & 0 & 0.14 \\ 0 & 0 & 1 & 0 & 0 & 0.50 & 0.12 \\ 0 & 0 & 0 & 1 & 0.14 & 0.25 & 0.60 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Guardando l'ultima colonna, che corrisponde alla situazione iniziale ABC , un po' a sorpresa si scopre che il carro armato C sopravvive nel 60% dei casi. Ciò è naturalmente dovuto al fatto che fino a quando siamo nella situazione ABC gli altri due non gli sparano.

Esempio 7.10. L'esempio 7.5 con $r = 0$ e $q = 5$ può essere usato per descrivere le situazioni in caso di parità alla fine di un set di tennis-tavolo (20:20 nel conteggio classico, 10:10 in quello introdotto pochi anni fa) o nella pallavolo; cfr. Ferschl, pag. 27. Dal punto di vista del giocatore A abbiamo gli stati

1	2 punti di vantaggio
2	1 punto di vantaggio
3	parità
4	1 punto di svantaggio
5	2 punti di svantaggio

Se a rappresenta la probabilità che uno scambio venga vinto da A, b la probabilità che venga vinto da B, allora

$$T = \begin{pmatrix} 1 & a & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 \\ 0 & b & 0 & a & 0 \\ 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & b & 1 \end{pmatrix}$$

Per $a = 0.7, b = 0.3$ la terza colonna, che corrisponde allo stato iniziale

di parità, converge a $\begin{pmatrix} 0.845 \\ 0 \\ 0 \\ 0 \\ 0.155 \end{pmatrix}$.

Mentre ciò da un lato conferma che la superiorità del primo giocatore nel singolo scambio (70%) si traduce in una probabilità del 84.5% a vincere il set, anche le possibilità per il secondo giocatore con il 15.5% sono maggiori del 9% che si otterrebbe calcolando solo la probabilità di vincere direttamente 12:10.

8. Un programma di simulazione

Nota 8.1. Creiamo un programma di simulazione:

```
def simula (T,volte,p=None,j=None):
    S=xrange(len(T))
    bloccanti=filter(lambda j: T[j][j]==1,S)
    if p:
        a=0; u=random.random()
        for j in S:
            b=a+p[j]
            if a<=u<b: break
            a=b
    else: j-=1
    v=[j+1]
    for t in xrange(volte):
        if j in bloccanti: break
        a=0; u=random.random()
        for i in S:
            b=a+T[i][j]
            if a<=u<b: v.append(i+1); j=i; break
            a=b
    return v
```

- (1) Gli argomenti sono:
 - T , che indica la matrice di transizione;
 - $volte$, che indica il numero di simulazioni;
 - p , che rappresenta il vettore di probabilità da cui partire.
- (2) S è il vettore degli indici usati.
- (3) $bloccanti$ contiene gli indici j tali che gli elementi diagonali $T_{(j)}^{(j)}$ coincidano con 1.
- (4) Se p è definito, lo usiamo per calcolare uno stato iniziale j .
- (5) Se invece p non è definito, allora deve essere definito j che, per il conteggio che inizia da zero, viene sostituito con $j - 1$.
- (6) Nel doppio ciclo finale vengono calcolati gli stati successivi per il numero prescritto di volte oppure fino a quando non venga raggiunto uno stato bloccante.
- (7) Il programma restituisce v .

Esempio 8.2. Consideriamo la matrice $T = \begin{pmatrix} 0.5 & 0.3 \\ 0.5 & 0.7 \end{pmatrix}$:

```
T=[[0.5,0.3],[0.5,0.7]]; p=[0.6,0.4]
v=simula(T,p,200)
for i in xrange(4): print ''.join(map(str,v[50*i:50*(i+1)]))
```

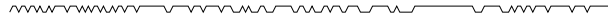
Otteniamo ad esempio l'output

```

1221221221211212222122212121221211212212221222222
2211222221222122222122211121211122111222222112211
211122122211222211112222221121111122222222222222
222211122222211121221221222212222221222212222222

```

La figura rappresenta graficamente l'output indicato.



Si vede abbastanza bene che le probabilità nella seconda colonna di T si rispecchiano in una tendenza a rimanere nello stato 2 una volta raggiunto.

Esempio 8.3. Consideriamo ancora la simulazione della parte finale di un set di tennistavolo vista nell'es. 7.10. Effettuiamo alcune prove.

```

a=0.7; b=0.3
T=[[1,a,0,0,0],[0,0,a,0,0],[0,b,0,a,0],[0,0,b,0,0],[0,0,0,b,1]]
for k in xrange(5):
    v=simula(T,200,j=3)
    print ''.join(map(str,v))

```

Otteniamo (ad esempio) i seguenti risultati, che traduciamo nei punteggi corrispondenti:

```

3234321
11:10 11:11 11:12 12:12 13:12 14:12

```

```

34321
10:11 11:11 12:11 13:11

```

```

321
11:10 12:10

```

```

34345
10:11 11:11 11:12 11:13

```

```

323234321
11:10 11:11 12:11 12:12 12:13 13:13 14:13 15:13

```

9. Una formula per le potenze

Situazione 9.1. Siano $A \in \mathbb{C}_q^q$ ed $\mathcal{M}_A := (x - \lambda_1)^{m_1} \cdots (x - \lambda_s)^{m_s}$ il polinomio minimale di A con i λ_k tutti distinti ed $m_k \geq 1$ per ogni k . Sia $m := m_1 + \dots + m_s$.

Osservazione 9.2. In Sage il polinomio minimale di A si ottiene con `minpoly(A)`, la sua fattorizzazione con `factor(minpoly(A))`.

Esempio 9.3. Proviamo queste istruzioni con il seguente esempio, che si trova in Gantmacher, pagg. 112-120:

```
A=matrix(CDF,[[3,-3,2],[-1,5,-2],[-1,3,0]]
print factor(minpoly(A))
# (1.0*x-4.0)*(1.0*x-2.0)}
```

Teorema 9.4. Per ogni $k = 1, \dots, s$ ed ogni $j = 0, \dots, m_k - 1$ sia dato un $v_{kj} \in \mathbb{C}$. Allora esiste esattamente un polinomio $H \in \mathbb{C}[x]$ di grado minore di m tale che per ogni $k = 1, \dots, s$ si abbia $H^{(j)}(\lambda_k) = v_{kj}$ per ogni $j = 0, \dots, m_k - 1$.

Dimostrazione. Giamtesio, pag. 46.

Definizione 9.5. Il polinomio H nel teorema 9.4 si chiama *polinomio di interpolazione di Hermite* rispetto al problema di interpolazione dato. Per indicare i parametri del problema di interpolazione denotiamo H con

$$H[\lambda_1 : (v_{10}, \dots, v_{1m_1-1}), \dots, \lambda_s : (v_{s0}, \dots, v_{sm_s-1})]$$

Definizione 9.6. Per $k = 1, \dots, s$ e $j = 0, \dots, m_k - 1$ definiamo H_{kj} come la soluzione del problema di interpolazione

$$H_{kj}^{(\alpha)}(\lambda\beta) = \delta_{k\beta}\delta_{j\alpha}$$

per $\beta = 1, \dots, s$ ed $\alpha = 0, \dots, m_\beta - 1$. Abbiamo quindi

$$\begin{aligned} H_{10} &= H[\lambda_1 : (1, 0, \dots, 0), \lambda_2 : (0, \dots, 0), \dots] \\ H_{11} &= H[\lambda_1 : (0, 1, 0, \dots, 0), \lambda_2 : (0, \dots, 0), \dots] \\ H_{12} &= H[\lambda_1 : (0, 0, 1, \dots, 0), \lambda_2 : (0, \dots, 0), \dots] \\ &\vdots \\ H_{20} &= H[\lambda_1 : (0, \dots, 0), \lambda_2 : (1, 0, \dots, 0), \dots] \\ H_{21} &= H[\lambda_1 : (0, \dots, 0), \lambda_2 : (0, 1, \dots, 0), \dots] \end{aligned}$$

I polinomi H_{kj} sono detti *polinomi di interpolazione fondamentali di Hermite*.

Teorema 9.7 (formula spettrale di Sylvester-Buheim). Siano Ω un dominio di \mathbb{C} che contiene lo spettro di A ed $f : \Omega \rightarrow \mathbb{C}$ una funzione analitica. Allora

$$f(A) = \sum_{k=1}^s \sum_{j=0}^{m_k-1} f^{(j)}(\lambda_k) H_{kj}(A)$$

Dimostrazione. Giantesio, pagg. 57-58.

Definizione 9.8. Per $a \in \mathbb{C}$ e $j \in \mathbb{N} + 1$ sia $a_{[j]} := a(a-1) \cdots (a-j+1)$.
Poniamo $a_{[0]} := 1$.

Osservazione 9.9. Per $n \in \mathbb{N}$ fissato sia $f := \bigcirc_x x^n : \mathbb{C} \rightarrow \mathbb{C}$.

Per ogni $j \in \mathbb{N}$ allora $f^{(j)}(z) = n_{[j]} z^{n-j}$.

Corollario 9.10. $A^n = \sum_{k=1}^s \sum_{j=0}^{m_k-1} n_{[j]} \lambda_k^{n-j} H_{kj}(A)$.

Esempio 9.11. Sia

$$A = \begin{pmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{pmatrix}$$

Calcoliamo prima gli H_{kj} , poi A^n . Come nell'esempio 9.3 gli autovalori sono $\lambda = 2, \mu = 4$. Dobbiamo calcolare

$$H_{10} = H[2 : (1), 4 : (0)]$$

$$H_{20} = H[2 : (0), 4 : (1)]$$

Utilizziamo lo schema alle differenze

$$\begin{array}{l} 2: (1), (0) \\ \quad \quad \quad 1, 0 \\ 4: (0), (1) \quad [2, 4] \quad -1/2, 1/2 \\ \quad \quad \quad \quad \quad 0, 1 \end{array}$$

Otteniamo quindi

$$H_{10} = 1 - \frac{1}{2}(x-2) = 2 - \frac{x}{2}$$

$$H_{20} = \frac{1}{2}(x-2) = -1 + \frac{x}{2}$$

Per calcolare A^n utilizziamo la formula del corollario 9.10:

$$\begin{aligned} A^n &= \lambda^n \left(2\delta - \frac{A}{2} \right) + \mu^n \left(-\delta + \frac{A}{2} \right) = 2^n \left(2 - \frac{A}{2} \right) + 4^n \left(-\delta + \frac{A}{2} \right) \\ &= (2^{n+1} - 2^{2n})\delta + (2^{2n-1} - 2^{n-1})A \\ &= 2^{n-1} [(4 - 2^{n+1})\delta + (2^n - 1)A] \end{aligned}$$

Esempio 9.12. Sia

$$A = \begin{pmatrix} 1 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0 \\ 0 & 0.3 & 0 & 0.7 & 0 \\ 0 & 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 1 \end{pmatrix}$$

Troviamo $\mathcal{M}_A = x(x-1)^2(x-\lambda)(x+\lambda)$ con $\lambda \cong 0.648$. Per calcolare A^n dobbiamo prima determinare

$$H_{10} = H[0 : (1), 1 : (0, 0), \lambda : (0), -\lambda : (0)]$$

$$H_{20} = H[0 : (0), 1 : (1, 0), \lambda : (0), -\lambda : (0)]$$

$$H_{21} = H[0 : (0), 1 : (0, 1), \lambda : (0), -\lambda : (0)]$$

$$H_{30} = H[0 : (0), 1 : (0, 0), \lambda : (1), -\lambda : (0)]$$

$$H_{40} = H[0 : (0), 1 : (0, 0), \lambda : (0), -\lambda : (1)]$$

0: (1);(0);(0);(0);(0)

1;0;0;0;0

1: (0, 0);(1, 0);(0, 1);(0, 0);(0, 0)

[0, 1]

-1;1;0;0;0

0;1;0;0;0

[0, 1]

1;-1;1;0;0

[1]

0;0;1;0;0

[0, λ]

0;1;0;0;0

[1, λ]

0;-8.071;2.841;8.074;0

λ : (0);(0);(0);(1);(0)

[1, λ]

0;2.841;0;-2.842;0

[1, $-\lambda$]

0;0;0;1;0

[1, $-\lambda$]

0;1.724;0;1.724;0.468

$-\lambda$: (0);(0);(0);(0);(1)

[λ , $-\lambda$]

0;0;0;0.772;-0.772

0;0;0;0;1

-1.543;-10.912;2.841;12.460;0

[0, $-\lambda$]

-2.381;-7.667;1.724;9.614;0.438

0;-5.944;1.724;6.230;-0.284

Otteniamo quindi

$$\begin{aligned} H_{10} &= 1 - x + (x-1)x - 1.543x(x-1)^2 - 2.381x(x-1)^2(x-\lambda) \\ &= 1 + (2.381\lambda - 3.543)x + (1.705 - 4.762\lambda)x^2 + (3.219 + 2.381\lambda)x^3 - 2.381x^4 \\ &= 1 - 2x - 1.381x^2 + 4.762x^3 - 2.381x^4 \end{aligned}$$

$$\begin{aligned} H_{20} &= x - x(x-1) - 10.912x(x-1)^2 - 7.667x(x-1)^2(x-\lambda) \\ &= (-8.912 + 7.667\lambda)x + (13.157 - 15.334\lambda)x^2 + (4.442 + 7.667\lambda)x^3 - 7.667x^4 \\ &= -3.944x + 3.221x^2 + 9.410x^3 - 7.667x^4 \end{aligned}$$

$$\begin{aligned} H_{21} &= x(x-1) + 2.841x(x-1)^2 + 1.724x(x-1)^2(x-\lambda) \\ &= (1.841 - 1.724\lambda)x + (-2.958 + 3.448\lambda)x^2 + (-0.607 - 1.724\lambda)x^3 + 1.724x^4 \\ &= 0.724x - 0.724x^2 - 1.724x^3 + 1.724x^4 \end{aligned}$$

$$\begin{aligned}
H_{30} &= 12.460x(x-1)^2 + 9.614x(x-1)^2(x-\lambda) \\
&= (12.460 - 9.614\lambda)x + (19.225\lambda - 15.306)x^2 - (6.768 + 9.614\lambda)x^3 + 9.614x^4 \\
&= 6.23x - 2.848x^2 - 12.998x^3 + 9.614x^4 \\
H_{40} &= 0.438x(x-1)^2(x-\lambda) \\
&= -0.438\lambda x + (0.438 + 0.876\lambda)x^2 - (0.876 + 0.438\lambda)x^3 + 0.438x^4 \\
&= -0.284x + 1.006x^2 - 1.16x^3 + 0.438x^4
\end{aligned}$$

Ponendo $\lambda_1 := 0$, $\lambda_2 := 1$, $\lambda_3 := \lambda$, $\lambda_4 := -\lambda$ abbiamo per $n > 0$ quindi

$$\begin{aligned}
A^n &= \lambda_1^n H_{10}(A) + \lambda_2^n H_{20}(A) + n\lambda_2^{n-1} H_{21}(A) + \lambda_3^n H_{30}(A) + \lambda_4^n H_{40}(A) \\
&= H_{20}(A) + nH_{21}(A) + \lambda^n H_{30}(A) + (-\lambda)^n H_{40}(A) \\
&= -3.944A + 3.221A^2 + 9.410A^3 - 7.667A^4 \\
&\quad + n(0.724A - 0.724A^2 - 1.724A^3 + 1.724A^4) \\
&\quad + \lambda^n(6.23A - 2.848A^2 - 12.998A^3 + 9.614A^4) \\
&\quad + (-\lambda)^n(-0.284A + 1.006A^2 - 1.16A^3 + 0.438A^4) \\
&= (-3.944 + 0.724n + 6.23p - 0.284q)A \\
&\quad + (3.221 - 0.724n - 2.848p + 1.006q)A^2 \\
&\quad + (9.410 - 1.724n - 12.998p - 1.160q)A^3 \\
&\quad + (-7.667 + 1.724n + 9.614p + 0.438q)A^4
\end{aligned}$$

dove abbiamo posto $p := \lambda^n$, $q := (-\lambda)^n$. Abbiamo quindi

$$\begin{aligned}
A^n &= (-3.944 + 0.724n + 5.946\lambda^n)A \\
&\quad + (3.221 - 0.724n - 1.842\lambda^n)A^2 \\
&\quad + (9.410 - 1.724n - 14.158\lambda^n)A^3 \\
&\quad + (-7.667 + 1.724n + 10.052\lambda^n)A^4
\end{aligned}$$

se n è pari e > 0 ,

$$\begin{aligned}
A^n &= (-3.944 + 0.724n + 6.514\lambda^n)A \\
&\quad + (3.221 - 0.724n - 3.854\lambda^n)A^2 \\
&\quad + (9.410 - 1.724n - 11.838\lambda^n)A^3 \\
&\quad + (-7.667 + 1.724n + 9.176\lambda^n)A^4
\end{aligned}$$

se n è dispari.

Dovremmo adesso calcolare le prime quattro potenze di A e sostituirle nell'espressione che risulta piuttosto complicata e numericamente instabile. Vedremo nel prossimo capitolo che in molti casi (compreso quello di quest'esempio) il limite delle potenze di una matrice stocastica può essere calcolato con una semplice procedura di algebra lineare.

10. Il teorema ergodico matriciale

Osservazione 10.1. V sia uno spazio vettoriale di dimensione finita e $\varphi : V \rightarrow V$ un'applicazione lineare tale che $\ker \varphi \cap \operatorname{Im} \varphi = 0$. Allora $V = \ker \varphi \oplus \operatorname{Im} \varphi$.

Dimostrazione. Dobbiamo dimostrare che $V = \ker \varphi + \operatorname{Im} \varphi$. Abbiamo però, usando due risultati noti dell'algebra lineare,

$$\begin{aligned} \dim(\ker \varphi + \operatorname{Im} \varphi) &= \dim \ker \varphi + \dim \operatorname{Im} \varphi - \dim(\ker \varphi \cap \operatorname{Im} \varphi) \\ &= \dim \ker \varphi + \dim \operatorname{Im} \varphi = \dim V \end{aligned}$$

Siccome V ha dimensione finita, da ciò segue che $V = \ker \varphi + \operatorname{Im} \varphi$.

Lemma 10.2. T sia una matrice stocastica $q \times q$ ed $x \in \mathbb{R}^q$. Allora $|Tx| \leq \sqrt{q}|x|$.

Dimostrazione. Scriviamo \sum_i invece di $\sum_{i=1}^q$. Usando la disuguaglianza di Cauchy-Schwartz (CS) e l'ipotesi che la matrice T sia ≥ 0 (1) e stocastica (2) abbiamo

$$\begin{aligned} |Tx|^2 &= \sum_i |T^{(i)}x|^2 \stackrel{CS}{\leq} \sum_i |T^{(i)}|^2 |x|^2 = |x|^2 \sum_i |T^{(i)}|^2 \\ &\stackrel{(1)}{=} |x|^2 \sum_i \sum_j (T_{(j)}^{(i)})^2 \stackrel{(2)}{\leq} |x|^2 \sum_i \sum_j T_{(j)}^{(i)} \\ &= |x|^2 \sum_j \sum_i T_{(j)}^{(i)} = |x|^2 \sum_j 1 = q|x|^2 \end{aligned}$$

Ciò implica l'enunciato.

Definizione 10.3. V sia uno spazio vettoriale e $\varphi : V \rightarrow V$ un'applicazione lineare. Poniamo $\operatorname{Fix} \varphi := \{x \in V \mid \varphi x = x\}$.

È chiaro che $\operatorname{Fix} \varphi$ è un sottospazio vettoriale di V .

Osservazione 10.4. Sia T una matrice stocastica. Allora:

- (1) 1 è autovalore di T .
- (2) $\operatorname{Fix} T = \ker(T - \delta) \neq 0$

Dimostrazione. (1) Il vettore $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$ è autovettore della matrice tras-

posta T^t , come segue immediatamente dalla condizione che ogni colonna di T abbia somma uguale a 1. Ciò mostra che 1 è autovalore di T^t e quindi anche di T .

(2) Segue direttamente da (1).

Definizione 10.5. Per $A \in \mathbb{C}_q^q$ poniamo $\|A\|_F := \sqrt{\sum_{i=1}^q \sum_{j=1}^q |A_{(j)}^{(i)}|^2}$.

$\|A\|_F$ si chiama la *norma di Frobenius* di A .

Si osservi che $\|A\|_F^2 = \sum_{i=1}^q |A^{(i)}|^2 = \sum_{j=1}^q |A_{(j)}|^2$.

Osservazione 10.6. $\bigcirc_n A_n$ sia una successione di matrici $q \times q$ e

$B \in \mathbb{R}_q^q$. Allora

$$\lim_{n \rightarrow \infty} A_n = B \iff \lim_{n \rightarrow \infty} \|A_n - B\|_F = 0$$

Dimostrazione. Ciò è chiaro perché la norma di Frobenius di una matrice coincide con la lunghezza del vettore complessivo dei suoi coefficienti.

Osservazione 10.7. Siano $A, B \in \mathbb{C}_q^q$ ed $x \in \mathbb{C}^q$. Allora:

- (1) $|Ax| \leq \|A\|_F |x|$.
- (2) $\|AB\|_F \leq \|A\|_F \|B\|_F$.

Dimostrazione. (1) Utilizzando la disuguaglianza di Cauchy-Schwartz abbiamo:

$$|Ax|^2 = \sum_i |A^{(i)}x|^2 \stackrel{CS}{\leq} \sum_i |A^{(i)}|^2 |x|^2 = |x|^2 \|A\|_F^2$$

$$(2) \|AB\|_F^2 = \sum_j |AB_{(j)}|^2 \stackrel{(1)}{\leq} \|A\|_F^2 \sum_j |B_{(j)}|^2 = \|A\|_F^2 \|B\|_F^2.$$

Osservazione 10.8. $\bigcirc_n A_n$ sia una successione di matrici $q \times q$ e

$B \in \mathbb{R}_q^q$. Allora le seguenti affermazioni sono equivalenti:

- (1) $\lim_{n \rightarrow \infty} A_n = B$.
- (2) $\lim_{n \rightarrow \infty} A_n x = Bx$ per ogni $x \in \mathbb{R}^q$.

Dimostrazione. Possiamo assumere che $B = 0$.

(1) \implies (2): Sia $x \in \mathbb{R}^q$. Dobbiamo dimostrare che $\lim_{n \rightarrow \infty} A_n x = 0$.

Per l'oss. 10.7 abbiamo $|A_n x| \leq \|A_n\|_F$, inoltre $\lim_{n \rightarrow \infty} \|A_n\|_F = 0$ per l'oss. 10.6.

(2) \implies (1): L'ipotesi implica in particolare che $\lim_{n \rightarrow \infty} (A_n)_{(j)} = \lim_{n \rightarrow \infty} A_n \delta_j = 0$ per ogni j e ciò mostra l'enunciato.

Teorema 10.9. Sia T una matrice stocastica $q \times q$. Allora:

- (1) $\mathbb{R}^q = \text{Fix } T \oplus \text{Im}(T - \delta)$.
- (2) $P : \mathbb{R}^q \longrightarrow \mathbb{R}^q$ sia la proiezione su $\text{Fix } T$ nella decomposizione (1).

$$\text{Allora } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n = P$$

Dimostrazione. Per $N \in \mathbb{N} + 1$ poniamo $S_N := \frac{1}{N} \sum_{n=0}^{N-1} T^n$.

- (A) Se $x \in \text{Fix } T$, allora $S_N x = x$ per ogni N .
 (B) Sia $x \in \text{Im}(T - \delta)$. Allora esiste $y \in \mathbb{R}^q$ tale che $x = Ty - y$.
 Per il lemma 10.2 e usando il fatto che le matrici T^n sono stocastiche, abbiamo

$$\begin{aligned} |S_N x| &= |S_N(Ty - y)| = \left| \frac{1}{N} \sum_{n=1}^N T^n y - \frac{1}{N} \sum_{n=0}^{N-1} T^n y \right| \\ &= \frac{1}{N} |T^N y - y| \leq \frac{1}{N} |T^N y| + \frac{1}{N} |y| \\ &\leq \frac{1}{N} (\sqrt{q}|y| + |y|) = \frac{\sqrt{q} + 1}{N} |y| \end{aligned}$$

Ciò mostra che $\lim_{n \rightarrow \infty} S_N x = 0$ per ogni $x \in \text{Im}(T - \delta)$.

- (C) Sia $x \in \text{Fix } T \cap \text{Im}(T - \delta)$. Allora per il punto (B) $\lim_{N \rightarrow \infty} S_N x = 0$, mentre per il punto (A) $S_N x = x$ per ogni N , e ciò è possibile solo se $x = 0$. Vediamo così che $\text{Fix } T \cap \text{Im}(T - \delta) = 0$. Però $\text{Fix } T = \ker(T - \delta)$ e dall'oss. 10.1 segue $\mathbb{R}^q = \text{Fix } T \oplus \text{Im}(T - \delta)$.
 (D) La proiezione P è quindi ben definita e per $x \in \mathbb{R}^q$ si ha $x = u + v$, con $u = Px \in \text{Fix } T$ e $v = x - Px \in \text{Im}(T - \delta)$. Inoltre

$$\lim_{N \rightarrow \infty} S_N x = \lim_{N \rightarrow \infty} S_N(u + v) \stackrel{(A)}{=} u + \lim_{N \rightarrow \infty} S_N v = u = Px$$

per ogni $N \in \mathbb{N} + 1$. L'enunciato segue dall'oss. 10.8.

Osservazione 10.10. Nelle ipotesi del teorema 10.9 si hanno le relazioni:

$$P^2 = P = PT = TP$$

Dimostrazione. Sia $x \in \mathbb{R}^q$. Per il teorema 10.9 $x = u + Tv - v$ con $u, v \in \mathbb{R}^q$ e $Tu = u$. È chiaro che $PPx = Pu = u = Px$. Inoltre $Tx = Tu + TTv - Tv = u + TTv - Tv$, per cui $PTx = u$, mentre $TPx = Tu = u$.

Osservazione 10.11. $\bigcirc_n A_n$ sia una successione di matrici stocastiche $q \times q$ e $B \in \mathbb{R}_q^q$ tale che $\bigcirc_n A_n \rightarrow B$. Allora anche la matrice B è stocastica.

Dimostrazione. Chiaro.

Osservazione 10.12. La matrice P nel teorema 10.9 è stocastica.

Definizione 10.13. Sia $A \in \mathbb{C}_q^q$. Il raggio spettrale di A è definito come $\rho(A) := \max\{|\lambda| \mid \lambda \text{ è autovalore di } A\}$.

Proposizione 10.14. Sia $A \in \mathbb{C}_q^q$. Allora $\rho(A) \leq \|A\|_F$.

Dimostrazione. Siano v un autovettore di A e $\lambda \in \mathbb{C}$ tali che $Av = \lambda v$. Per l'oss. 10.7 allora $|\lambda||v| = |\lambda v| = |Av| \leq \|A\|_F |v|$ e poiché $v \neq 0$ ciò implica $|\lambda| \leq \|A\|_F$ e quindi l'enunciato.

Definizione 10.15. Per $\lambda \in \mathbb{C}$ ed $m \in \mathbb{N} + 1$ definiamo la *casella di Jordan* $J_m(\lambda) \in \mathbb{C}_m^m$ nel modo seguente:

$$\begin{aligned}
 J_1(\lambda) &:= \lambda \\
 J_2(\lambda) &:= \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \\
 J_3(\lambda) &:= \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \\
 &\vdots \\
 J_m(\lambda) &:= \begin{pmatrix} \lambda & 1 & 0 & \dots & \dots & 0 \\ 0 & \lambda & 1 & \dots & \dots & 0 \\ 0 & 0 & \lambda & \dots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \dots & 1 & 0 \\ 0 & 0 & \dots & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}
 \end{aligned}$$

Teorema 10.16 (teorema della forma normale di Jordan).

Per ogni $A \in \mathbb{C}_q^q$ esiste $E \in \text{GL}(q, \mathbb{C})$ tale che $E^{-1}AE$ sia della forma

$$\begin{pmatrix} M_1 & 0 & \dots & 0 \\ 0 & M_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & M_p \end{pmatrix},$$

dove ogni M_j è una casella di Jordan.

Gli autovalori di A coincidono con gli elementi della diagonale principale di $E^{-1}AE$ e vi appaiono con la stessa molteplicità con cui appaiono nel polinomio caratteristico di A .

Dimostrazione. Ad es. Huppert, pagg. 23-26.

Osservazione 10.17. Siano $A \in \mathbb{C}_q^q$ ed $E \in \text{GL}(q, \mathbb{C})$. Per ogni $n \in \mathbb{N}$ allora $(E^{-1}AE)^n = E^{-1}A^nE$.

Lemma 10.18. Siano $\lambda \in \mathbb{C}$, $m \in \mathbb{N} + 1$ ed $n \in \mathbb{N}$. Allora

$$(J_m(\lambda))^n = \begin{pmatrix} \binom{n}{0}\lambda^n & \binom{n}{1}\lambda^{n-1} & \binom{n}{2}\lambda^{n-2} & \dots \\ 0 & \binom{n}{0}\lambda^n & \binom{n}{1}\lambda^{n-1} & \dots \\ & & \dots & \\ 0 & 0 & 0 & \binom{n}{0}\lambda^n \end{pmatrix}$$

Dimostrazione. Induzione su n .

L'enunciato è banale per $n = 0$.

$n \rightarrow n + 1$: Poniamo $A := J_m(\lambda)$. Dobbiamo dimostrare che

$(A^n)_{(j)}^{(i)} = \binom{n}{j-1-(i+1)} \lambda^{n-[j-1-(i+1)]} = \binom{n}{j-i} \lambda^{n-j+i}$ per $j \geq i$ perché è chiaro che $(A^n)_{(j)}^{(i)} = 0$ per $j < i$. Anche il caso $j = i$ è ovvio. Assumiamo quindi che $j > i$. Allora

$$\begin{aligned} (A^{n+1})_{(j)}^{(i)} &= \sum_{k=1}^m A_{(k)}^{(i)} (A^n)_{(j)}^{(k)} = A_{(i)}^{(i)} (A^n)_{(j)}^{(i)} + A_{(i+1)}^{(i)} (A^n)_{(j)}^{(i+1)} \\ &= \lambda (A^n)_{(j)}^{(i)} + (A^n)_{(j)}^{(i+1)} \\ &\stackrel{ind}{=} \lambda \binom{n}{j-i} \lambda^{n-j+i} + \binom{n}{j-i-1} \lambda^{n-j+i+1} \\ &= \lambda^{n+1-j+i} \left[\binom{n}{j-i} + \binom{n}{j-i-1} \right] \\ &= \lambda^{n+1-j+i} \binom{n+1}{j-i} \end{aligned}$$

Proposizione 10.19. Sia $A \in \mathbb{C}_q^q$ tale che $\rho(A) < 1$. Allora $\lim_{n \rightarrow \infty} A^n = 0$.

Dimostrazione. Per il teorema 10.16 esiste $E \in GL(q, \mathbb{C})$ tale che $E^{-1}AE$ sia in forma normale di Jordan e per l'oss. 10.17 $\lim_{n \rightarrow \infty} A^n = 0$ se e solo se $\lim_{n \rightarrow \infty} (E^{-1}AE)^n = 0$. È quindi sufficiente dimostrare che per ogni $m \in \mathbb{N} + 1$ ed ogni $\lambda \in \mathbb{C}$ con $|\lambda| \leq 1$ vale $\lim_{n \rightarrow \infty} (J_m(\lambda))^n = 0$. Ma ciò è evidente dal lemma 10.18.

Osservazione 10.20. Siano $A \in \mathbb{C}_q^q$ ed $n \in \mathbb{N}$.

- (1) Se $\lambda_1, \dots, \lambda_q$ sono gli autovalori di A , allora $\lambda_1^n, \dots, \lambda_q^n$ sono gli autovalori di A^n .
- (2) Perciò $\rho(A^n) = (\rho(A))^n$.

Dimostrazione. Ciò è una conseguenza immediata del teorema 10.16 e del lemma 10.18.

Teorema 10.21. Sia $A \in \mathbb{C}_q^q$. Allora $\rho(A) = \lim_{n \rightarrow \infty} \sqrt[n]{\|A^n\|_F}$.

Dimostrazione. Dimostriamo che per ogni $\varepsilon > 0$ esiste $n_0 \in \mathbb{N}$ tale che $\rho(A) \leq \sqrt[n]{\|A^n\|_F} < \rho(A) + \varepsilon$ per ogni $n \geq n_0$.

(1) Per l'oss. 10.20 e la prop. 10.14 abbiamo $(\rho(A))^n = \rho(A^n) \leq \|A^n\|_F$ e quindi $\rho(A) \leq \sqrt[n]{\|A^n\|_F}$ per ogni $n \in \mathbb{N} + 1$.

(2) Siano $\varepsilon > 0$ ed $A_\varepsilon := \frac{1}{\rho(A) + \varepsilon}$. Allora $\rho(A_\varepsilon) = \frac{\rho(A)}{\rho(A) + \varepsilon} < 1$. Per la prop. 10.19 quindi $\lim_{n \rightarrow \infty} A_\varepsilon^n = 0$. In particolare esiste $n_0 \in \mathbb{N}$ tale che $\|A_\varepsilon^n\|_F < 1$ per ogni $n \geq n_0$. Ciò significa però $\frac{\|A^n\|_F}{(\rho(A) + \varepsilon)^n} < 1$, ovvero $\sqrt[n]{\|A^n\|_F} < \rho(A) + \varepsilon$ per ogni $n \geq n_0$.

Definizione 10.22. Per $A, B \in \mathbb{R}_q^q$ scriviamo $A \geq B$ se $A_{(j)}^{(i)} \geq B_{(j)}^{(i)}$ per ogni i, j . La matrice A si dice *non negativa* se $A \geq 0$.

Lemma 10.23. Siano $A, B \in \mathbb{R}_q^q$ tali che $A \geq B \geq 0$. Allora $\rho(A) \geq \rho(B)$.

Dimostrazione. L'ipotesi implica che anche $A^n \geq B^n \geq 0$ e quindi $\|A^n\|_F \geq \|B^n\|_F$ per ogni n . Usando il teorema 10.21 abbiamo

$$\rho(A) = \lim_{n \rightarrow \infty} \sqrt[n]{\|A^n\|_F} \geq \lim_{n \rightarrow \infty} \sqrt[n]{\|B^n\|_F} = \rho(B)$$

Definizione 10.24. Per $v \in \mathbb{C}^q$ ed $A \in \mathbb{C}_q^q$ siano

$$\|v\|_1 := \sum_{i=1}^q |v^{(i)}|$$

$$\|A\|_1 := \max\{\|A_{(j)}\|_1 \mid j = 1, \dots, q\}$$

Per $A \geq 0$ abbiamo quindi in particolare

$$\|A\|_1 = \max\left\{\sum_{i=1}^q A_{(j)}^{(i)} \mid j = 1, \dots, q\right\}$$

Osservazione 10.25. Sia T una matrice stocastica. Allora $\|T\|_1 = 1$.

Dimostrazione. Infatti $\|T_{(j)}\|_1 = 1$ per ogni j , poiché T è stocastica.

Lemma 10.26. Sia $A \in \mathbb{C}_q^q$. Allora:

- (1) $\|A\|_1 = \sup\left\{\frac{\|Av\|_1}{\|v\|_1} \mid v \in \mathbb{C}^q \setminus 0\right\}$.
- (2) In particolare si ha $\|Av\|_1 \leq \|A\|_1 \|v\|_1$ per ogni $v \in \mathbb{C}^q$.

Dimostrazione. Dimostriamo prima il punto (2), scrivendo \sum_i invece di $\sum_{i=1}^q$. Sia $v \in \mathbb{C}^q$. Allora, usando la disuguaglianza triangolare, abbiamo

$$\begin{aligned} \|Av\|_1 &= \sum_i |A^{(i)}v| = \sum_i \left| \sum_j A_{(j)}^{(i)} v^{(j)} \right| \leq \sum_i \sum_j |A_{(j)}^{(i)}| |v^{(j)}| \\ &= \sum_j |v^{(j)}| \sum_i |A_{(j)}^{(i)}| = \sum_j |v^{(j)}| \|A_{(j)}\|_1 \\ &\leq \sum_j |v^{(j)}| \|A\|_1 = \|A\|_1 \sum_j |v^{(j)}| = \|A\|_1 \|v\|_1 \end{aligned}$$

Viceversa scegliamo j in modo tale che $\|A_{(j)}\|_1 = \|A\|_1$. Allora $\|A\delta_j\|_1 = \|A_{(j)}\|_1 = \|A\|_1$, e siccome $\|\delta_j\|_1 = 1$, abbiamo

$$\|A\|_1 = \frac{\|A\delta_j\|_1}{\|\delta_j\|_1}.$$

Lemma 10.27. Siano $A, B \in \mathbb{C}_q^q$. Allora $\|AB\|_1 \leq \|A\|_1 \|B\|_1$.

Dimostrazione.

$$\begin{aligned}\|AB\|_1 &= \max\{\|(AB)_{(j)}\|_1 \mid j = 1, \dots, q\} \\ &= \max\{\|AB_{(j)}\|_1 \mid j = 1, \dots, q\} \\ &\stackrel{10.26}{\leq} \max\{\|A\|_1 \|B_{(j)}\|_1 \mid j = 1, \dots, q\} \\ &= \|A\|_1 \max\{\|B_{(j)}\|_1 \mid j = 1, \dots, q\} \\ &= \|A\|_1 \|B\|_1\end{aligned}$$

Proposizione 10.28. Sia $A \in \mathbb{C}_q^q$. Allora $\rho(A) \leq \|A\|_1$.

Dimostrazione. Uguaile alla dimostrazione della prop. 10.14, utilizzando il lemma 10.26.

Teorema 10.29. Sia T una matrice stocastica. Allora $\rho(T) = 1$.

Dimostrazione. Per la prop. 10.27 si ha che $\rho(T) \leq \|T\|_1 \stackrel{10.25}{=} 1$. Ma per l'oss. 10.4 sappiamo che 1 è autovalore di T , quindi necessariamente $\rho(T) = 1$.

Lemma 10.30. Sia $\lambda \in \mathbb{C}$. Allora

- (1) Se $|\lambda| > 1$, allora la successione $\bigcirc_n \lambda^n$ non converge.
- (2) Se $|\lambda| < 1$, allora $\lim_{n \rightarrow \infty} \lambda^n = 0$.
- (3) Se $|\lambda| = 1$ e $\lambda \neq 1$, allora la successione $\bigcirc_n \lambda^n$ non converge.
- (4) Per $\lambda = 1$ naturalmente $\lim_{n \rightarrow \infty} \lambda^n = 1$.

Dimostrazione. È sufficiente dimostrare il punto (3). Sia $\lim_{n \rightarrow \infty} \lambda^n = a$. Allora da un lato $\lim_{n \rightarrow \infty} \lambda^{n+1} = \lambda a$, ma anche $\lim_{n \rightarrow \infty} \lambda^{n+1} = a$, per cui $\lambda a = a$. Per la continuità del valore assoluto si ha però $|a| = 1$, cosicché $a \neq 0$ e ciò implica $\lambda = 1$, in contraddizione all'ipotesi.

Osservazione 10.31. Siano $\lambda \in \mathbb{C}$ ed $m \in \mathbb{N}$. Assumiamo che $m \geq 2$ e $|\lambda| \geq 1$. Allora la successione $\bigcirc_n (J_m(\lambda))^n$ non converge.

Dimostrazione. Utilizziamo il lemma 10.18. Dal lemma 10.30 segue che la successione non converge per $\lambda \neq 1$. Siano $\lambda = 1$ ed $m \geq 2$. Allora $((J_m(\lambda))^n)_{(2)}^{(1)} = n$ e si vede che anche in questo caso la successione non può convergere.

Corollario 10.32. Siano $\lambda \in \mathbb{C}$ ed $m \in \mathbb{N}+1$. Sia $A := J_m(\lambda)$. Allora:

- (1) Se $|\lambda| < 1$, allora $\lim_{n \rightarrow \infty} A^n = 0$.
- (2) Se $|\lambda| > 1$, allora la successione $\bigcirc_n A^n$ non converge.
- (3) Se $|\lambda| = 1$ ed $m \geq 2$, allora la successione $\bigcirc_n A^n$ non converge.

(4) Se $|\lambda| = 1$ ed $m = 1$, allora la successione $\bigcirc_n A^n$ converge se e solo se $\lambda = 1$.

Dimostrazione. (1) e (2): Seguono dal lemma 10.18.

(3): Segue dal corollario 10.31.

(4): Segue dal lemma 10.30.

Proposizione 10.33. *T sia una matrice stocastica $q \times q$ ed $E \in GL(q, \mathbb{C})$ tale che $S := E^{-1}TE$ sia in forma normale di Jordan. λ sia un autovalore di T con $|\lambda| = 1$. Allora per ogni casella di Jordan $J_m(\lambda)$ che appare in S si ha $m = 1$.*

Dimostrazione. Per ogni $n \in \mathbb{N}$ anche la matrice T^n è stocastica per cui $\|T^n\|_1 = 1$. Per il lemma 10.27 abbiamo

$$\|S^n\|_1 = \|E^{-1}T^nE\|_1 \leq \|E^{-1}\|_1 \|T^n\|_1 \|E\|_1 = \|E^{-1}\|_1 \|E\|_1 =: \alpha$$

Teorema 10.34. *Sia T una matrice stocastica. Allora sono equivalenti:*

(1) La successione $\bigcirc_n T^n$ converge.

(2) T non possiede un autovalore λ con $\lambda \neq 1$ e $|\lambda| = 1$.

Dimostrazione. Possiamo trovare $E \in GL(q, \mathbb{C})$ tale che $S := E^{-1}TE$ sia in forma normale di Jordan. La successione $\bigcirc_n T^n$ converge se e solo se converge la successione $\bigcirc_n S^n$ e ciò accade se e solo se per ogni casella di Jordan $A := J_m(\lambda)$ converge la successione $\bigcirc_n A^n$. Per il corollario 10.32 è sufficiente considerare il caso $|\lambda| = 1$. Per la prop. 10.33 in quel caso però $m = 1$ e l'enunciato segue dal corollario 10.32.

Lemma 10.35 (primo teorema di Gershgorin). *Siano*

$A \in \mathbb{C}_q^q$ e λ un autovalore di A . Allora esiste $j \in \{1, \dots, q\}$ tale che $|\lambda - A_{(j)}^{(j)}| \leq \|A_{(j)}\|_1 - |A_{(j)}^{(j)}|$.

Dimostrazione. λ è anche autovalore di A^t , perciò esiste un vettore riga $f \in \mathbb{C}_q \setminus 0$ tale che $fA = \lambda f$. Possiamo assumere che $|f_{(1)}| \geq |f_{(k)}|$ per ogni $k \geq 2$. In particolare $f_{(1)} \neq 0$. Per ipotesi abbiamo

$$\lambda f_{(1)} = (fA)_{(1)} = fA_{(1)} = f_{(1)}A_{(1)}^{(1)} + \sum_{k=2}^q f_{(k)}A_{(1)}^{(k)}$$

e quindi

$$|\lambda - A_{(1)}^{(1)}| = \left| \sum_{k=2}^q \frac{f_{(k)}}{f_{(1)}} A_{(1)}^{(k)} \right| \leq \sum_{k=2}^q \left| \frac{f_{(k)}}{f_{(1)}} \right| |A_{(1)}^{(k)}| \leq \sum_{k=2}^q |A_{(1)}^{(k)}| = \|A_{(1)}\|_1 - |A_{(1)}^{(1)}|$$

Corollario 10.36. *Siano T una matrice stocastica $q \times q$ e λ un autovalore di T . Allora esiste $j \in \{1, \dots, q\}$ tale che*

$$|\lambda - A_{(j)}^{(j)}| \leq 1 - A_{(j)}^{(j)}.$$

Definizione 10.37. Una matrice $A \in \mathbb{C}_q^q$ si dice *a diagonale positiva* se $A_{(j)}^{(j)} > 0$ per ogni $j \in \{1, \dots, q\}$.

Corollario 10.38. Siano T una matrice stocastica a diagonale positiva e λ un autovalore di T con $|\lambda| = 1$. Allora $\lambda = 1$.

Dimostrazione. Scegliamo j come nel corollario 10.36 e poniamo $a := A_{(j)}^{(j)}$. Per ipotesi, inoltre $|\lambda - a| \leq 1 - a$. Siano $x, y \in \mathbb{R}$ con $\lambda = a + x + iy$. Siccome $|\lambda| = 1$, necessariamente $a + x \leq 1$. Da $|\lambda| = 1$ segue però anche $a^2 + 2ax + x^2 + y^2 = 1$, mentre $|\lambda - a|^2 \leq (1 - a)^2$ implica $x^2 + y^2 \leq 1 - 2a + a^2$ cosicché $1 - a^2 - 2ax \leq 1 - 2a + a^2$ ovvero $2a \leq 2a^2 + 2ax$. Siccome $a > 0$ otteniamo $1 \leq a + x$. Cosicché $a + x = 1$. Ciò implica $\lambda = 1 + iy$. Ma $|\lambda| = 1$, per cui $y = 0$ e $\lambda = 1$.

Corollario 10.39. Sia T una matrice stocastica a diagonale positiva. Allora la successione $\bigcirc_n T^n$ converge.

Dimostrazione. Teorema 10.34 e corollario 10.38.

Definizione 10.40. Per $U \in \mathbb{R}^2$ sia $\text{conv}(U)$ l'involuppo convesso di U . Denotiamo inoltre con $S^1, \mathbb{D}, \overline{\mathbb{D}}$ la circonferenza unitaria, il disco unitario aperto e il disco unitario chiuso in \mathbb{R}^2 .

$\lambda \in \mathbb{C}$ si chiama una *radice dell'unità*, se esiste $k \in \mathbb{N} + 1$ tale che $\lambda^k = 1$.

Osservazione 10.41. Sia $U \in \overline{\mathbb{D}}$. Allora $\text{conv}(U) \cap S^1 \subset U$.

Dimostrazione. Geometricamente evidente.

Proposizione 10.42. Siano T una matrice stocastica e λ un autovalore di T con $|\lambda| = 1$. Allora $|\lambda|$ è una radice dell'unità.

Dimostrazione. Sia $T \in \mathbb{R}_q^q$. λ è autovalore anche di T^t , perciò esiste un vettore riga $f \in \mathbb{C}_q \setminus 0$ tale che $fT = \lambda f$. Possiamo assumere che $\max\{|f_{(j)}| \mid j = 1, \dots, q\} = 1$. Sia $U := \{f_{(1)}, \dots, f_{(q)}\}$. Allora $U \subset \overline{\mathbb{D}}$ e $U \cap S^1 \neq \emptyset$. Possiamo quindi scegliere un elemento $z \in U \cap S^1$. Sia ad esempio $z = f_{(1)}$. Allora $\lambda z = \lambda f_{(1)} = fT_{(1)} = \sum_{k=1}^q f_{(k)}T_{(1)}^{(k)} \in \text{conv}(U)$, perché T è stocastica.

D'altra parte $|\lambda z| = |\lambda||z| = 1$, per cui $\lambda z \in S^1 \cap \text{conv}(U)$ e dall'oss. 10.41 segue che $\lambda z \in S^1 \cap U$. Ciò mostra $\lambda(U \cap S^1) \subset U \cap S^1$.

Qui abbiamo usato soltanto che $z \in U \cap S^1$, perciò vediamo che $\lambda(U \cap S^1) \subset U \cap S^1$ e quindi anche $\lambda^n z \in U \cap S^1$ per ogni n . L'insieme $U \cap S^1$ è finito, per cui devono esistere $m, n \in \mathbb{N} + 1$ con $m > n$ tali che $\lambda^m z = \lambda^n z$. Però $z \neq 0$ e ciò implica $\lambda^m = \lambda^n$ ovvero $\lambda^{m-n} = 1$.

Lemma 10.43. V sia uno spazio vettoriale (reale o complesso) normato e $\bigcirc_n x_n$ una successione di elementi di V ed $x \in V$ con $\bigcirc_n x_n \rightarrow x$.

Allora $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} x_n = x$.

Dimostrazione. Sia $\varepsilon > 0$. Per ipotesi esiste N_1 tale che $\|x_n - x\| < \varepsilon/2$ per ogni $n \geq N_1$. Scegliamo $N_0 \geq N_1$ in modo tale che

$$\frac{1}{N_0} \left\| \sum_{n=0}^{N_1-1} (x_n - x) \right\| < \varepsilon/2. \text{ Per } N \geq N_0 \text{ allora}$$

$$\begin{aligned} \left\| \frac{1}{N} \sum_{n=0}^{N-1} x_n - x \right\| &= \left\| \frac{1}{N} \sum_{n=0}^{N-1} (x_n - x) \right\| \\ &\leq \frac{1}{N} \left\| \sum_{n=0}^{N_1-1} (x_n - x) \right\| + \frac{1}{N} \left\| \sum_{n=N_1}^{N-1} (x_n - x) \right\| \\ &< \frac{\varepsilon}{2} + \frac{1}{N} \left\| \sum_{n=N_1}^{N-1} (x_n - x) \right\| \leq \frac{\varepsilon}{2} + \frac{1}{N} \sum_{n=N_1}^{N-1} \|x_n - x\| \\ &< \frac{\varepsilon}{2} + \frac{N - N_1}{N} \frac{\varepsilon}{2} < \varepsilon \end{aligned}$$

Corollario 10.44. T sia una matrice stocastica $q \times q$ e la successione $\bigcirc_n T^n$ sia convergente. Allora:

- (1) $\mathbb{R}^q = \text{Fix } T \oplus \text{Im}(T - \delta)$.
- (2) Se $P : \mathbb{R}^q \rightarrow \mathbb{R}^q$ è la proiezione su $\text{Fix } T$ nella decomposizione (1), allora $\lim_{n \rightarrow \infty} T^n = P$.
- (3) $P^2 = P = PT = TP$.

Dimostrazione. Lemma 10.43, teorema 10.9 e oss. 10.10.

Corollario 10.45. T sia una matrice stocastica, la successione $\bigcirc_n T^n$ sia convergente e $\dim \text{Fix } T = 1$.

Allora $\text{Fix } T$ contiene un unico vettore stocastico v e si ha $\lim_{n \rightarrow \infty} T^n = (v, \dots, v)$.

Dimostrazione. Come nel corollario 10.44 sia P la proiezione su $\text{Fix } T$. Allora $TP = P$. Sia v una qualsiasi colonna di P . Allora $Tv = v$, quindi $v \in \text{Fix } T$. Per l'oss. 10.12 la matrice P è stocastica e quindi v è un vettore stocastico. In particolare si ha $\sum_{i=1}^q v^{(i)} = 1$ e ciò determina v univocamente essendo per ipotesi $\dim \text{Fix } T = 1$. Ciò implica però che $P = (v, \dots, v)$.

Esempio 10.46. Come nella nota 7.2 sia $T = \begin{pmatrix} a & b \\ 1-a & 1-b \end{pmatrix}$ con $|a - b| < 1$. Allora $b \neq 0$ e risolvendo $T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} := v$ troviamo $ax + by = x$, per cui $y = \frac{1-a}{b}x$. Ciò mostra che $\dim \text{Fix } T = 1$. Affinché v sia un vettore stocastico dobbiamo avere $x + y = 1$, cioè $x + x \frac{1-a}{b} = 1$ ovvero $x = \frac{b}{1-a+b}$ e $y = \frac{1-a}{b} \frac{b}{1-a+b} = \frac{1-a}{1-a+b}$, cosicché

$P = (v, v) = \frac{1}{1-a+b} \begin{pmatrix} b & b \\ 1-a & 1-a \end{pmatrix}$ in accordo con quanto trovato nella nota 7.2.

Esempio 10.47. Consideriamo ancora la matrice $T = \begin{pmatrix} 1 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0.7 & 0 & 0 \\ 0 & 0.3 & 0 & 0.7 & 0 \\ 0 & 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 1 \end{pmatrix}$.

Nell'esempio 9.12 abbiamo trovato che $\mathcal{M}_T = x(x-1)^2(x-\lambda)(x+\lambda)$ con $\lambda \cong 0.648$ per cui dal teorema 10.34 segue che la successione $\bigcirc_n T^n$ converge. Possiamo applicare il corollario 10.44. Risolvendo l'equazione $Tx = x$ troviamo che $\text{Fix } T = SV(\delta_1, \delta_5)$.

Abbiamo inoltre $T - \delta = \begin{pmatrix} 0 & 0.7 & 0 & 0 & 0 \\ 0 & -1 & 0.7 & 0 & 0 \\ 0 & 0.3 & -1 & 0.7 & 0 \\ 0 & 0 & 0.3 & -1 & 0 \\ 0 & 0 & 0 & 0.3 & 0 \end{pmatrix}$.

Poniamo $e_3 = \begin{pmatrix} 0.7 \\ -1 \\ 0.3 \\ 0 \\ 0 \end{pmatrix}$, $e_4 = \begin{pmatrix} 0 \\ 0.7 \\ -1 \\ 0.3 \\ 0 \end{pmatrix}$, $e_5 = \begin{pmatrix} 0 \\ 0 \\ 0.7 \\ -1 \\ 0.3 \end{pmatrix}$.

Allora e_3, e_4, e_5 è una base di $\text{Im}(T - \delta)$, mentre $\delta_1, \delta_5, e_3, e_4, e_5$ è una base di \mathbb{R}^5 . Sia $E := (\delta_1, \delta_5, e_3, e_4, e_5)$. Sia $x \in \mathbb{R}^5$. Allora x possiede una rappresentazione $x = Ea = a^{(1)}\delta_1 + a^{(2)}\delta_5 + a^{(3)}e_3 + a^{(4)}e_4 + a^{(5)}e_5$ con $a = E^{-1}x$. Con Sage troviamo

$$E^{-1} = \begin{pmatrix} 1 & 0.95 & 0.84 & 0.59 & 0 \\ 0 & 0.05 & 0.16 & 0.41 & 1 \\ & & \dots & & \end{pmatrix}$$

per cui

$$\begin{aligned} a^{(1)} &= x^{(1)} + 0.95x^{(2)} + 0.84x^{(3)} + 0.59x^{(4)} \\ a^{(2)} &= 0.05x^{(2)} + 0.16x^{(3)} + 0.41x^{(4)} + x^{(5)} \end{aligned}$$

$$Px = a^{(1)}\delta_1 + a^{(2)}\delta_5 = \begin{pmatrix} x^{(1)} + 0.95x^{(2)} + 0.84x^{(3)} + 0.59x^{(4)} \\ 0 \\ 0 \\ 0 \\ 0.05x^{(2)} + 0.16x^{(3)} + 0.41x^{(4)} + x^{(5)} \end{pmatrix}$$

e quindi
$$P = \begin{pmatrix} 1 & 0.95 & 0.84 & 0.59 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05 & 0.16 & 0.41 & 1 \end{pmatrix}$$

Lemma 10.48. *K sia un campo e V uno spazio vettoriale su K . Siano $v_1, \dots, v_s \in V$ e $W := SV(v_1, \dots, v_s)$. Sia $1 \leq m \leq s$ tale che i vettori v_1, \dots, v_m siano linearmente indipendenti, mentre per ogni $j > m$ i vettori v_1, \dots, v_m, v_j siano linearmente dipendenti. Allora $SV(v_1, \dots, v_m) = W$. I vettori v_1, \dots, v_m costituiscono quindi una base di W .*

Dimostrazione. Sia $j > m$. È sufficiente dimostrare che $v_j \in SV(v_1, \dots, v_m)$. Poiché per ipotesi v_1, \dots, v_m, v_j sono linearmente dipendenti, esistono $\beta_1, \dots, \beta_m, \beta_j \in K$, non tutti nulli, tali che $\beta_1 v_1 + \dots + \beta_m v_m + \beta_j v_j = 0$. Necessariamente $\beta_j \neq 0$, perché altrimenti v_1, \dots, v_m sarebbero linearmente dipendenti, perciò possiamo scrivere

$$v_j = -\frac{\beta_1}{\beta_j} v_1 - \dots - \frac{\beta_m}{\beta_j} v_m \in SV(v_1, \dots, v_m)$$

Osservazione 10.49. *U e W siano due sottospazi vettoriali di \mathbb{R}^q tali che $\mathbb{R}^q = U \oplus W$. e_1, \dots, e_r sia una base di U , e_{r+1}, \dots, e_q una base di W . Siano $E := (e_1, \dots, e_q)$ e $Q := (\delta_1, \dots, \delta_r, 0, \dots, 0)$. Allora $P := EQE^{-1}$ è la proiezione di \mathbb{R}^q su U (più precisamente, come in precedenza, la matrice di questa proiezione nella base standard di \mathbb{R}^q).*

Dimostrazione. Sia $x \in \mathbb{R}^q$. Allora $x = Ea$ con $a \in \mathbb{R}^q$. Allora

$$Px = E \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(r)} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = EQa = EQE^{-1}x, \text{ e ciò mostra l'enunciato.}$$

Nota 10.50. Utilizzando l'oss. 10.48, dal corollario 10.44 otteniamo un semplice metodo generale per calcolare $\lim_{n \rightarrow \infty} T^n =: P$ per una matrice stocastica $T \in \mathbb{R}^q$ per la quale questo limite esiste:

- (1) Risolvendo l'equazione $Tx = x$ troviamo una base di $\text{Fix } T$.
- (2) Estraendo un numero massimo di colonne linearmente indipendenti di $T - \delta$ troviamo una base di $\text{Im}(T - \delta)$.
- (3) Adesso possiamo calcolare P come proiezione su $\text{Fix } T$ nella decomposizione $\mathbb{R}^q = \text{Fix } T \oplus \text{Im}(T - \delta)$, usando il metodo dell'oss. 10.49.

Esempio 10.51. Sia $T = \begin{pmatrix} 0.8 & 0.5 & 0.1 & 0 \\ 0.1 & 0.2 & 0.7 & 0.1 \\ 0.1 & 0.2 & 0.1 & 0.3 \\ 0 & 0.1 & 0.1 & 0.6 \end{pmatrix}$. Con Sage troviamo

il polinomio minimale $\mathcal{M}_T = (x - 0.901)(x - 0.655)(x^2 - 0.144x + 0.006)$, da cui vediamo che T non possiede autovalori λ con $\lambda \neq 1$ e $|\lambda| = 1$. Per il teorema 10.34 esiste il limite $P := \lim_{n \rightarrow \infty} T^n$. Risolvendo l'equazione $Tx = x$ troviamo

$$\text{Fix } T = SV(e_1) \text{ con } e_1 = \begin{pmatrix} 88 \\ 31 \\ 21 \\ 13 \end{pmatrix}.$$

Con il vettore stocastico $v := \frac{1}{153} \begin{pmatrix} 88 \\ 31 \\ 21 \\ 13 \end{pmatrix}$ per il corollario 10.45 abbiamo quindi $P = (v, \dots, v)$.

Esempio 10.52. Sia $T = \begin{pmatrix} 1 & 1/2 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1 \end{pmatrix}$.

Il polinomio caratteristico di T è uguale a $x(x-1)^2(x-1/2)^2$, cosicché dal teorema 10.34 segue che $\lim_{n \rightarrow \infty} T^n =: P$ esiste.

Risolvendo $Tx = x$ otteniamo che $\text{Fix } T = SV(\delta_1, \delta_5)$. Essendo

$$T - \delta = \begin{pmatrix} 0 & 1/2 & 1/4 & 0 & 0 \\ 0 & -3/4 & 0 & 1/4 & 0 \\ 0 & 0 & -1/2 & 0 & 0 \\ 0 & 1/4 & 0 & -3/4 & 0 \\ 0 & 0 & 1/4 & 1/2 & 0 \end{pmatrix}$$

poniamo

$$e_2 = \begin{pmatrix} 1/2 \\ -3/4 \\ 0 \\ 1/4 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 1/4 \\ 0 \\ -1/2 \\ 0 \\ 1/4 \end{pmatrix}, e_4 = \begin{pmatrix} 0 \\ 1/4 \\ 0 \\ -3/4 \\ 1/2 \end{pmatrix}$$

Allora e_2, e_3, e_4 è una base di $\text{Im}(T - \delta)$. Siano $E := (\delta_1, \delta_5, e_2, e_3, e_4)$ e $Q := (\delta_1, \delta_2, 0, 0, 0)$. Calcoliamo E^{-1} con Sage:

$$E^{-1} = \begin{pmatrix} 1 & 0.75 & 0.5 & 0.25 & 0 \\ 0 & 0.25 & 0.5 & 0.75 & 1 \\ 0 & -1.5 & 0 & -0.5 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & -0.5 & 0 & -1.5 & 0 \end{pmatrix}$$

Per la nota 10.50 abbiamo

$$\begin{aligned}
P = EQE^{-1} &= \begin{pmatrix} 1 & 0 & 0.5 & 0.25 & 0 \\ 0 & 0 & -0.75 & 0 & 0.25 \\ 0 & 0 & 0 & -0.5 & 0 \\ 0 & 0 & 0.25 & 0 & -0.75 \\ 0 & 1 & 0 & 0.25 & 0.5 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0.75 & 0.5 & 0.25 & 0 \\ 0 & 0.25 & 0.5 & 0.75 & 1 \\ 0 & -1.5 & 0 & -0.5 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & -0.5 & 0 & -1.5 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0.75 & 0.5 & 0.25 & 0 \\ 0 & 0.25 & 0.5 & 0.75 & 1 \\ 0 & -1.5 & 0 & -0.5 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & -0.5 & 0 & -1.5 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0.75 & 0.5 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.5 & 0.75 & 0 \end{pmatrix}
\end{aligned}$$

II. MATRICI NON NEGATIVE

11. Sistemi dinamici stocastici finiti

Definizione 11.1. Un *sistema dinamico stocastico finito* (SDSF) è una tripla $(S, F, \bigcirc_j r_j) = (S, F, \bigcirc_j r_j)$ con le seguenti componenti:

- (1) $S = \{1, \dots, q\}$ con $q \in \mathbb{N} + 1$.
- (2) F è un sottoinsieme non vuoto di S^S , cioè un insieme non vuoto di applicazioni $S \rightarrow S$.
- (3) Per ogni $j \in S$ è dato un vettore stocastico $r_j \in [0, 1]^F$.

Gli elementi di S si chiamano *stati* del sistema dinamico, gli elementi di F *applicazioni* (o operatori), i vettori r_j sono detti *regole* del sistema. Per $i, j \in S$ poniamo $F_{ij} := \{f \in F \mid f(j) = i\}$.

Definizione 11.2. $(S, F, \bigcirc_j r_j)$ sia un SDSF con $|S| = q$. Per ogni $i, j \in S$

sia $T_{(j)}^{(i)} := \sum_{f \in F_{ij}} r_j(f)$. Naturalmente la somma è uguale a zero quando $F_{ij} = \emptyset$. Otteniamo così una matrice $T \in [0, 1]_q^q$ che chiameremo la *matrice di transizione* associata al SDSF dato. Siccome, come vedremo nella prop. 11.3, questa matrice è stocastica, ad ogni SDSF è associata una catena di Markov finita.

Proposizione 11.3. La matrice T definita nella def. 11.2 è stocastica.

Dimostrazione. È chiaro che $T \geq 0$. Sia $j \in S$. Siccome per ogni $f \in F$ esiste esattamente un $i \in S$ tale che $f(j) = i$ abbiamo

$F = F_{1j} \dot{\cup} \dots \dot{\cup} F_{qj}$. Perciò $\sum_{i=1}^q T_{(j)}^{(i)} = \sum_{i=1}^q \sum_{f \in F_{ij}} r_j(f) = \sum_{f \in F} r_j(f) = 1$, utilizzando l'ipotesi che il vettore r_j sia stocastico.

Nota 11.4. Usiamo l'abbreviazione $[i_1, \dots, i_q]$ per l'applicazione $f : \{1, \dots, q\} \rightarrow \{1, \dots, q\}$ definita da $f(j) = i_j$.

Esempio 11.5. Siano $S = \{1, 2, 3, 4\}$ ed $F = \{f, g, h\}$ con $f = [3132]$, $g = [1324]$, $h = [2334]$. Le regole siano date dai vettori stocastici

$$r_1 = \begin{pmatrix} 0.2 \\ 0.3 \\ 0.5 \end{pmatrix}, r_2 = \begin{pmatrix} 0.5 \\ 0.1 \\ 0.4 \end{pmatrix}, r_3 = \begin{pmatrix} 0.8 \\ 0.1 \\ 0.1 \end{pmatrix}, r_4 = \begin{pmatrix} 0.4 \\ 0.2 \\ 0.4 \end{pmatrix}$$

e si riferiscono all'ordine (f, g, h) su F .

Per ottenere la matrice di transizione T , costruiamo prima la matrice degli insiemi F_{ij} :

$$\begin{pmatrix} F_{11} & F_{12} & F_{13} & F_{14} \\ F_{21} & F_{22} & F_{23} & F_{24} \\ F_{31} & F_{32} & F_{33} & F_{34} \\ F_{41} & F_{42} & F_{43} & F_{44} \end{pmatrix} = \begin{pmatrix} \{g\} & \{f\} & \emptyset & \emptyset \\ \{h\} & \emptyset & \{g\} & \{f\} \\ \{f\} & \{g, h\} & \{f, h\} & \emptyset \\ \emptyset & \emptyset & \emptyset & \{g, h\} \end{pmatrix}$$

Perciò

$$\begin{aligned}
 T &= \begin{pmatrix} r_1(g) & r_2(f) & 0 & 0 \\ r_1(h) & 0 & r_3(g) & r_4(f) \\ r_1(f) & r_2(g) + r_2(h) & r_3(f) + r_3(h) & 0 \\ 0 & 0 & 0 & r_4(g) + r_4(h) \end{pmatrix} \\
 &= \begin{pmatrix} 0.3 & 0.5 & 0.2 & 0 \\ 0.5 & 0 & 0 & 0.4 \\ 0.2 & 0.1 + 0.4 & 0.8 + 0.1 & 0 \\ 0 & 0 & 0 & 0.2 + 0.4 \end{pmatrix} \\
 &= \begin{pmatrix} 0.3 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.1 & 0.4 \\ 0.2 & 0.5 & 0.9 & 0 \\ 0 & 0 & 0 & 0.6 \end{pmatrix}
 \end{aligned}$$

Con Sage si verifica che 1 è l'unico autovalore λ di T con $|\lambda| = 1$, cosicché dal teorema 10.34 segue che $\lim_{n \rightarrow \infty} T^n =: P$ esiste. Risolvendo il sistema $Tx = x$ si trova $\text{Fix} T = SV(v)$, dove v è il vettore stocastico

dato da $v = \begin{pmatrix} 5/57 \\ 7/57 \\ 45/57 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.09 \\ 0.123 \\ 0.790 \\ 0 \end{pmatrix}$, per cui dal corollario 10.45

otteniamo $P = (v, v, v, v) = \begin{pmatrix} 0.09 & 0.09 & 0.09 & 0.09 \\ 0.123 & 0.123 & 0.123 & 0.123 \\ 0.790 & 0.790 & 0.790 & 0.790 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Osservazione 11.6. La matrice di transizione di un SDSF $(S, F, \bigcirc_j r_j)$

può essere ottenuta in modo più meccanico nella maniera seguente.

Identifichiamo ogni $f \in F$ con la matrice che corrisponde all'applicazione lineare $\mathbb{R}^q \rightarrow \mathbb{R}^q$ definita da $\delta_j \mapsto \delta_{f(j)}$. Sopra a questa matrice scriviamo i valori $r_j(f)$ per $j = 1, \dots, q$. Nell'es. 11.5 otteniamo

$$\begin{array}{cccccccccccc}
 0.2 & 0.5 & 0.8 & 0.4 & & 0.3 & 0.1 & 0.1 & 0.2 & & 0.5 & 0.4 & 0.1 & 0.4 \\
 f = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, & g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & h = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
 \end{array}$$

Adesso moltiplichiamo ogni colonna con il valore soprastante; la som-

ma delle matrici che così si ottengono è uguale a T :

$$\begin{aligned}
 T &= \begin{pmatrix} 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0.4 \\ 0.2 & 0 & 0.8 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0.3 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.2 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.4 & 0.1 & 0 \\ 0 & 0 & 0 & 0.4 \end{pmatrix} \\
 &= \begin{pmatrix} 0.3 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.1 & 0.4 \\ 0.2 & 0.5 & 0.9 & 0 \\ 0 & 0 & 0 & 0.6 \end{pmatrix}
 \end{aligned}$$

Esempio 11.7. La matrice di transizione dell'es. 11.5 si può ottenere anche nel seguente modo:

$$\delta_1 \mapsto 0.2f(\delta_1) + 0.3g(\delta_1) + 0.5h(\delta_1) = 0.2\delta_3 + 0.3\delta_1 + 0.5\delta_2 = \begin{pmatrix} 0.3 \\ 0.5 \\ 0.2 \\ 0 \end{pmatrix}$$

$$\delta_2 \mapsto 0.5f(\delta_2) + 0.1g(\delta_2) + 0.4h(\delta_2) = 0.5\delta_1 + 0.1\delta_3 + 0.4\delta_3 = \begin{pmatrix} 0.5 \\ 0 \\ 0.5 \\ 0 \end{pmatrix}$$

$$\delta_3 \mapsto 0.8f(\delta_3) + 0.1g(\delta_3) + 0.1h(\delta_3) = 0.8\delta_3 + 0.1\delta_2 + 0.1\delta_3 = \begin{pmatrix} 0 \\ 0.1 \\ 0.9 \\ 0 \end{pmatrix}$$

$$\delta_4 \mapsto 0.4f(\delta_4) + 0.2g(\delta_4) + 0.4h(\delta_4) = 0.4\delta_2 + 0.2\delta_4 + 0.4\delta_4 = \begin{pmatrix} 0 \\ 0.4 \\ 0 \\ 0.6 \end{pmatrix}$$

Nota 11.8. Nella modellazione, ad esempio tramite reti di Petri (cfr. Leardini), di sistemi biochimici o metabolici si considerano spesso sistemi della forma (P, K, F) , dove P (insieme delle *posizioni*) e K (insieme dei *valori*) sono insiemi finiti ed F è un insieme non vuoto di applicazioni $K^P \rightarrow K^P$. Possiamo allora definire l'insieme degli *stati* $S := K^P$. Nella simulazione del sistema è naturale associare ad ogni stato j un vettore stocastico $r_j \in [0, 1]^F$, cosicché ci troviamo nella situazione della def. 11.1.

Nota 11.9. Sia dato un SDSF $(S, F, \bigcirc_j r_j)$. Assumiamo che esso corrisponda al comportamento di un insieme di cellule che possono passare da uno stato all'altro tramite i passaggi corrispondenti alle transizioni $f \in F$. Supponiamo inoltre che il raggiungimento dello stato j induca una proliferazione (o diminuzione) della popolazione in quello stato di

un fattore λ_j prima che venga effettuata un'altra transizione. Allora per descrivere la dinamica del sistema dobbiamo sostituire la matrice T della def. 11.2 con la matrice TD , dove D è una matrice diagonale

della forma $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_q \end{pmatrix}$, per cui $TD = (\lambda_1 T_{(1)}, \dots, \lambda_q T_{(q)})$.

Questa matrice è non negativa ed evidentemente ogni matrice non negativa A può essere ottenuta in questo modo da una matrice stocastica, di cui la j -esima colonna è univocamente determinata se $A_{(j)} \neq 0$.

12. Una stima per il raggio spettrale

Situazione 12.1. Siano $q \in \mathbb{N} + 1$ e $A \in \mathbb{R}_q^q$.

Scriviamo \sum_i invece di $\sum_{i=1}^q$ e $\min\{\dots\}$ invece di $\min\{\dots \mid i = 1, \dots, q\}$.

Definizione 12.2. Per $B \in \mathbb{R}_q^q$ scriviamo $A > B$ se $A_{(j)}^{(i)} > B_{(j)}^{(i)}$ per ogni i, j . A si dice *positiva* se $A_{(j)}^{(i)} > 0$ per ogni i, j .

Similmente, per $v, w \in \mathbb{R}^q$ scriviamo $v \geq w$ se $v^{(i)} \geq w^{(i)}$ per ogni i e $v > w$ se $v^{(i)} > w^{(i)}$ per ogni i . Il vettore v si dice *non negativo* se $v \geq 0$ e v *positivo* se $v > 0$.

Matrici non negative sono state definite nella def. 10.22.

Definizione 12.3. Sia $1 \leq k_1 < \dots < k_r \leq q$. Denotiamo allora con $A[k_1, \dots, k_r]$ la matrice quadratica che si ottiene da A togliendo tutte le righe $A_{(i)}^{(i)}$ e tutte le colonne $A_{(j)}$ per le quali i risp. j non appartengono a $\{k_1, \dots, k_r\}$. Ogni matrice che si ottiene in questo modo si chiama una *sottomatrice principale* di ordine r di A .

Esempio 12.4. Sia $A := \begin{pmatrix} 2 & 5 & 8 & 1 & 6 \\ 3 & 9 & 2 & 7 & 4 \\ 1 & 4 & 6 & 2 & 3 \\ 5 & 1 & 8 & 9 & 2 \\ 6 & 8 & 2 & 1 & 4 \end{pmatrix}$.

Allora $A[1, 3, 4] = \begin{pmatrix} 2 & 8 & 1 \\ 1 & 6 & 2 \\ 5 & 8 & 9 \end{pmatrix}$.

Lemma 12.5. A sia non negativa e B una sottomatrice principale di A . Allora $\rho(B) \leq \rho(A)$.

Dimostrazione. Sia \bar{B} la matrice ottenuta da A sostituendo i coefficienti tolti per ottenere B con coefficienti nulli. Con una matrice di permutazione P si ottiene $P\bar{B}P^{-1} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$. Il polinomio caratteristico di \bar{B} è uguale a quello di B moltiplicato con una potenza di x , per cui $\rho(B) = \rho(\bar{B})$. Poiché $\bar{B} \leq A$, per il lemma 10.23 si ha $0 \leq \rho(B) \leq \rho(A)$.

Osservazione 12.6. Nell'esempio 12.4 abbiamo $\bar{B} = \begin{pmatrix} 2 & 0 & 8 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 6 & 2 & 0 \\ 5 & 0 & 8 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Con

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

abbiamo allora

$$P\bar{B} = \begin{pmatrix} 2 & 0 & 8 & 1 & 0 \\ 1 & 0 & 6 & 2 & 0 \\ 5 & 0 & 8 & 9 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, P^{-1} = P^t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$P\bar{B}P^{-1} = \begin{pmatrix} 2 & 8 & 1 & 0 & 0 \\ 1 & 6 & 2 & 0 & 0 \\ 5 & 8 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Corollario 12.7. *A sia non negativa. Allora $\rho(A) \geq A_{(i)}^{(i)}$ per ogni i .*

Osservazione 12.8. *A sia non negativa.*

- (1) Se le somme di colonna di A sono tutte uguali, allora $\rho(A) = \|A\|_1$.
- (2) Se le somme di riga di A sono tutte uguali, allora $\rho(A) = \|(A^t)\|_1$.

Dimostrazione. (1) L'enunciato è banale per $A = 0$. Sia $A \neq 0$ ed s il valore comune delle somme di colonna. Allora $s \neq 0$ e la matrice $\frac{1}{s}A$ è stocastica. Per il teorema 10.29 si ha $\rho(\frac{1}{s}A) = 1$ e quindi $\rho(A) = s = \|A\|_1$.

(2) Segue dal punto (1) considerando la matrice trasposta A^t , utilizzando il fatto che $\rho(A^t) = \rho(A)$.

Definizione 12.9. *A sia non negativa. Allora poniamo*

$$\begin{aligned} \max_{\text{sdc}}(A) &:= \max_j \|A_{(j)}\|_1 = \|A\|_1 \\ \max_{\text{sdr}}(A) &:= \max_i \|A^{(i)}\|_1 = \|A^t\|_1 \\ \min_{\text{sdc}}(A) &:= \min_j \|A_{(j)}\|_1 \\ \min_{\text{sdr}}(A) &:= \min_i \|A^{(i)}\|_1 \end{aligned}$$

Osservazione 12.10. *A sia non negativa ed $m := \min_{\text{sdc}}(A)$.*

Definiamo la matrice $B \in \mathbb{R}_q^q$ ponendo $B_{(j)}^{(i)} := m \frac{A_{(j)}^{(i)}}{\|A_{(j)}\|_1}$.

Allora $0 \leq B \leq A$ e le somme di colonna di B sono tutte uguali ad m . Per l'oss. 12.8 $\rho(B) = m$.

Dimostrazione. (1) È chiaro che $B \geq 0$.

(2) Siccome per definizione $m \leq \|A_{(j)}\|_1$, abbiamo $B_{(j)}^{(i)} \leq A_{(j)}^{(i)}$ per ogni i, j .

3) Infine per ogni j abbiamo $\sum_i B_{(j)}^{(i)} = m \frac{\sum_i A_{(j)}^{(i)}}{\|A_{(j)}\|_1} = m \frac{\|A_{(j)}\|_1}{\|A_{(j)}\|_1} = m$.

Proposizione 12.11. *A sia non negativa. Allora:*

- (1) $\text{minsdc}(A) \leq \rho(A) \leq \text{maxsdc}(A)$.
- (2) $\text{minsdr}(A) \leq \rho(A) \leq \text{maxsdr}(A)$.

Dimostrazione. (1) Sappiamo dalla prop. 10.28 che $\rho(A) \leq \|A\|_1 = \text{maxsdc}(A)$. Sia $m := \text{minsdc}(A)$. Per l'oss. 12.10 possiamo trovare una matrice B con $0 \leq B \leq A$ e $\rho(B) = m$. Usando il lemma 10.23 abbiamo $m \leq \rho(A)$.

(2) Si ottiene dal punto (1) considerando la matrice trasposta A^t .

Esempio 12.12. Le somme di colonna della matrice A dell'esempio 12.4 sono 17, 27, 26, 20, 19, le somme di riga 22, 25, 16, 25, 21. Dalla prop. 12.11 otteniamo le stime

$$17 \leq \rho(A) \leq 27 \quad \text{e} \quad 16 \leq \rho(A) \leq 25,$$

perciò $17 \leq \rho(A) \leq 25$.

Con Sage troviamo che gli autovalori di A , arrotondati a due cifre decimali, sono $-2.67, 2.67, 21.80, 3.99 \pm 4.16i$, per cui $\rho(A) = 21.80$.

Corollario 12.13. *A sia non negativa e tutte le colonne di A oppure tutte le righe di A non nulle. Allora $\rho(A) > 0$. In particolare si ha che $\rho(A) > 0$ se $A > 0$.*

Osservazione 12.14. *A sia non negativa. Le disuguaglianze della prop. 12.11 possono essere riscritte in modo più esplicito:*

- (1) $\min_j \sum_i A_{(j)}^{(i)} \leq \rho(A) \leq \max_j \sum_i A_{(j)}^{(i)}$.
- (2) $\min_i \sum_j A_j^{(i)} \leq \rho(A) \leq \max_i \sum_j A_j^{(i)}$.

Osservazione 12.15. Siano $\varepsilon_1, \dots, \varepsilon_q \neq 0$ e $D := \begin{pmatrix} \varepsilon_1 & & 0 \\ & \ddots & \\ 0 & & \varepsilon_q \end{pmatrix}$.

Allora $(D^{-1}AD)_{(j)}^{(i)} = \frac{\varepsilon_j}{\varepsilon_i} A_{(j)}^{(i)}$.

Dimostrazione. Infatti AD si ottiene da A moltiplicando la j -esima colonna con ε_j , e $D^{-1}AD$ da AD moltiplicando la i -esima riga con $1/\varepsilon_i$.

Corollario 12.16. *A sia non negativa ed $\varepsilon_1, \dots, \varepsilon_q > 0$. Allora:*

- (1) $\min_j \frac{1}{\varepsilon_j} \sum_i A_{(j)}^{(i)} \varepsilon_i \leq \rho(A) \leq \max_j \frac{1}{\varepsilon_j} \sum_i A_{(j)}^{(i)} \varepsilon_i$.
- (2) $\min_i \frac{1}{\varepsilon_j} \sum_j A_j^{(i)} \varepsilon_i \leq \rho(A) \leq \max_i \frac{1}{\varepsilon_j} \sum_j A_j^{(i)} \varepsilon_i$.

Dimostrazione. Sia $D := \begin{pmatrix} \varepsilon_1 & & 0 \\ & \ddots & \\ 0 & & \varepsilon_q \end{pmatrix}$.

Le matrici DAD^{-1} e $D^{-1}AD$ sono anch'esse non negative e possiedono lo stesso raggio spettrale di A .

Sostituendo nel punto (1) dell'oss. 12.14 A con DAD^{-1} e nel punto (2) A con $D^{-1}AD$, otteniamo direttamente l'enunciato, tenendo conto dell'oss. 12.15.

Lemma 12.17. *A sia non negativa e $v \in \mathbb{R}^q$ con $v > 0$. Siano $\alpha, \beta \in \mathbb{R}$ con $\alpha, \beta > 0$. Allora:*

(1) $\alpha v \leq Av \implies \alpha \leq \rho(A)$.

(2) $\alpha v < Av \implies \alpha < \rho(A)$.

(3) $Av \leq \beta v \implies \rho(A) \leq \beta$.

(4) $Av < \beta v \implies \rho(A) < \beta$.

Dimostrazione. (1) $\alpha v \leq Av$ significa $\alpha v^{(i)} \leq A^{(i)}v = \sum_j A_{(j)}^{(i)} v^{(j)}$ per ogni i , ovvero $\alpha \leq \min_i \sum_j A_{(j)}^{(i)} \frac{v^{(j)}}{v^{(i)}} \stackrel{12.16}{\leq} \rho(A)$.

(2)-(4) Si ottengono nello stesso modo.

Proposizione 12.18. *A sia non negativa e λ un autovalore di A per il quale esiste un autovettore positivo. Allora $\rho(A) = \lambda$.*

Dimostrazione. Sia $v \in \mathbb{R}^q$ con $v > 0$ ed $Av = \lambda v$. Ciò implica in particolare $A^{(1)}v = \lambda v^{(1)}$, per cui $\lambda = \frac{1}{v^{(1)}} A^{(1)}v \geq 0$. Siccome banalmente $\lambda v \leq Av \leq \lambda v$, dal lemma 12.17 otteniamo $\lambda \leq \rho(A) \leq \lambda$, quindi $\rho(A) = \lambda$.

13. Il teorema di Perron

Situazione 13.1. Siano $q \in \mathbb{N} + 1$ e $A \in \mathbb{R}_q^q$. Scriviamo \sum_i invece di $\sum_{i=1}^q$.

Definizione 13.2. Per $v = \begin{pmatrix} v^{(1)} \\ \vdots \\ v^{(q)} \end{pmatrix} \in \mathbb{C}^q$ poniamo $[v] := \begin{pmatrix} |v^{(1)}| \\ \vdots \\ |v^{(q)}| \end{pmatrix}$.

Definizione 13.3. Denotiamo con $\mathcal{M}(A)$ l'insieme degli autovalori di A di modulo massimo, cioè l'insieme degli autovalori λ di A per i quali $|\lambda| = \rho(A)$. Vogliamo adesso dimostrare che, se A è positiva, allora $\mathcal{M}(A) = \{\rho(A)\}$.

Osservazione 13.4. A sia positiva ed $u \in \mathbb{R}^q$ con $u \neq 0$ e $u \geq 0$. Allora $Au > 0$.

Osservazione 13.5. A sia non negativa e $v \in \mathbb{C}^q$. Allora $[Av] \leq A[v]$.

Dimostrazione. Usando la disuguaglianza triangolare e l'ipotesi $A \geq 0$, per ogni i abbiamo

$$[Av]^{(i)} = |A^{(i)}v| = \left| \sum_j A_{(j)}^{(i)} v^{(j)} \right| \leq \sum_j A_{(j)}^{(i)} |v^{(j)}| = (A[v])^{(i)}$$

Lemma 13.6. A sia positiva, $\lambda \in \mathcal{M}(A)$ e $v \in \mathbb{C}^q$ un autovettore di A rispetto a λ . Allora:

- (1) $A[v] = \rho(A)[v] = [Av]$.
- (2) $[v] > 0$ e quindi i coefficienti di v sono tutti non nulli.

Dimostrazione. (1) Siccome $v \neq 0$ ed A è positiva, l'oss. 13.4 implica $A[v] > 0$. Sia $x := A[v] - \rho(A)[v] = A[v] - |\lambda|v$. Dobbiamo dimostrare che $x = 0$.

Abbiamo $\rho(A)[v] = |\lambda|[v] = [\lambda v] = [Av] \leq A[v]$ per l'oss. 13.5, per cui $\rho(A)[v] = [Av]$ ed $x \geq 0$.

Assumiamo adesso, per assurdo, che $x \neq 0$. Allora per l'oss. 13.4 $Ax > 0$, e ciò significa $\rho(A)A[v] < AA[v]$. Però $A[v] > 0$, cosicché il lemma 12.17 implica $\rho(A) < \rho(A)$, una contraddizione.

(2) Dal cor. 12.13 sappiamo che $\rho(A) > 0$. Perciò dal punto (1) otteniamo $[v] = \frac{1}{\rho(A)}[Av] > 0$.

Corollario 13.7. A sia positiva. Allora $\rho(A)$ è un autovalore di A e possiede un autovettore positivo.

Proposizione 13.8. A sia positiva e v, w due autovettori di A rispetto all'autovalore $\rho(A)$. Allora esiste $t \in \mathbb{C}$ tale che $v = tw$.

Dimostrazione. Sia ad esempio $w^{(1)} \neq 0$. Allora con $t := \frac{v^{(1)}}{w^{(1)}}$ formiamo il vettore $u := v - tw$. Dimostriamo che $u = 0$.

Se così non fosse, u sarebbe ancora un autovettore di A rispetto a $\rho(A)$ e $[u] > 0$ per il lemma 13.6. Però $u^{(1)} = 0$ per la definizione di t .

Corollario 13.9. *A sia positiva e v, w due autovettori non negativi di A rispetto all'autovalore $\rho(A)$. Allora esiste $t > 0$ tale che $v = tw$.*

Osservazione 13.10. *A sia positiva. La prop. 13.8 afferma che l'autovalore $\rho(A)$ possiede molteplicità geometrica 1. Dimosteremo fra poco che anche la molteplicità algebrica di $\rho(A)$ è uguale a 1.*

Definizione 13.11. *A sia positiva. Dai corollari 13.7 e 13.9 segue che A possiede un unico autovettore positivo e stocastico rispetto all'autovalore $\rho(A)$.*

Questo autovettore si chiama il *vettore di Perron* di A .

Proposizione 13.12. *A sia positiva. Allora ogni autovettore non negativo di A è un multiplo (reale positivo) del vettore di Perron di A e quindi in particolare autovettore rispetto a $\rho(A)$.*

Dimostrazione. Siano $v \in \mathbb{R}^q \setminus \{0\}$, $v \geq 0$ e $\lambda \in \mathbb{C}$ tali che $Av = \lambda v$. Per il corollario 13.9 è sufficiente dimostrare che $\lambda = \rho(A)$.

La matrice trasposta A^t è anch'essa positiva e possiede lo stesso raggio spettrale di A . Per il cor. 13.7 esiste un vettore positivo u tale che $A^t u = \rho(A)u$. Abbiamo allora

$$\rho(A)\|u, v\| = \|\rho(A)u, v\| = \|A^t u, v\| = \|u, Av\| = \|u, \lambda v\| = \lambda\|u, v\|$$

Le ipotesi $u > 0$, $v \geq 0$ e $u \neq 0$ implicano adesso che $\|u, v\| > 0$, e da ciò segue $\rho(A) = \lambda$.

Osservazione 13.13. Siano $m, n \in \mathbb{N} + 1$ e $z_1, \dots, z_m \in \mathbb{R}^n$ tali che $|z_1 + \dots + z_m| = |z_1| + \dots + |z_m|$. Allora i vettori z_1, \dots, z_m si trovano tutti su uno stesso raggio passante per l'origine, esistono cioè $u \in \mathbb{R}^n$ e $t_1, \dots, t_m \in [0, \infty)$ tali che $z_1 = t_1 u, \dots, z_m = t_m u$. Per $n = 2$ ciò significa che esiste $\alpha \in \mathbb{R}$ tale che $z_1 = e^{i\alpha}|z_1|, \dots, z_m = e^{i\alpha}|z_m|$.

Dimostrazione. Il caso $m = 2$ si trova ad esempio in Beckenbach/Bellman, pag. 70, o Collatz, pag. 28. Nel caso generale l'ipotesi implica

$$\begin{aligned} |z_1 + \dots + z_{m-1}| + |z_m| &\leq |z_1| + \dots + |z_{m-1}| + |z_m| = |z_1 + \dots + z_m| \\ &\leq |z_1 + \dots + z_{m-1}| + |z_m| \end{aligned}$$

per cui $|z_1| + \dots + |z_{m-1}| = |z_1 + \dots + z_{m-1}|$ e si conclude per induzione.

Lemma 13.14. *A sia positiva, $\lambda \in \mathcal{M}(A)$ e $v \in \mathbb{C}^q$ un autovettore di A rispetto a λ . Allora esiste $\alpha \in \mathbb{R}$ tale che $v = e^{i\alpha}[v]$.*

Dimostrazione. Per il lemma 13.6 abbiamo $[Av] = A[v]$. Quindi in particolare $[Av]^{(1)} = (A[v])^{(1)}$ e ciò significa

$$\sum_j \left| A_{(j)}^{(1)} v^{(j)} \right| = \sum_j A_{(j)}^{(1)} |v^{(j)}| = \left| \sum_j A_{(j)}^{(1)} v^{(j)} \right|$$

e per l'oss. 13.13 esiste $\alpha \in \mathbb{R}$ tale che $A_{(j)}^{(1)} v^{(j)} = e^{i\alpha} |A_{(j)}^{(1)} v^{(j)}| = e^{i\alpha} A_{(j)}^{(1)} |v^{(j)}|$ per ogni j . Siccome per ipotesi $A_{(j)}^{(1)} > 0$, vediamo che $v^{(j)} = e^{i\alpha} |v^{(j)}|$ per ogni j e quindi $v = e^{i\alpha} [v]$.

Proposizione 13.15. *A sia positiva. Allora $\mathcal{M}(A) = \{\rho(A)\}$.*

Dimostrazione. Per il cor. 13.7 abbiamo $\rho(A) \in \mathcal{M}(A)$.

Sia $\lambda \in \mathcal{M}(A)$ e v un autovettore rispetto a λ . Per il lemma 13.14 esiste $\alpha \in \mathbb{R}$ con $v = e^{i\alpha} [v]$, per cui

$$\lambda v = Av = e^{i\alpha} [Av] = e^{i\alpha} \rho(A) [v] = e^{i\alpha} \rho(A) e^{-i\alpha} v = \rho(A) v$$

Siccome $v \neq 0$, ciò implica $\lambda = \rho(A)$.

Osservazione 13.16. *A sia non negativa. Allora $\rho(A)$ è un autovalore di A . A possiede un autovettore non negativo rispetto a $\rho(A)$.*

Dimostrazione. Ponendo $(A_n)_{(j)}^{(i)} := A_{(j)}^{(i)} + 1/n$ per $n \in \mathbb{N} + 1$, troviamo una successione $\bigcirc_n A_n$ di matrici positive che converge ad A . Per ogni n sia v_n il vettore di Perron di A_n . Siccome autovalori e autovettori di una matrice dipendono in modo continuo dai suoi coefficienti, $\lambda := \lim_{n \rightarrow \infty} \rho(A_n)$ è un autovalore di A e $\lim_{n \rightarrow \infty} v_n =: v$ esiste ed è ancora un autovettore ≥ 0 di A rispetto a λ (questo limite è non nullo perché ogni $|v_n| = 1$). Siccome gli autovalori di A si ottengono come limiti degli autovalori degli A_n , per ogni n maggiorati da $\rho(A_n)$, si deve avere $\lambda \geq \rho(A)$ e quindi $\lambda = \rho(A)$.

Lemma 13.17. *$v \in \mathbb{C}^q$ sia un vettore fisso di A , $w \in \mathbb{C}^q$ un vettore fisso di A^t . Assumiamo inoltre che $w^t v = 1$ e poniamo $P := vw^t$. Allora:*

(1) $PA = AP = P = P^2$.

(2) $Pv = v$.

v è quindi un vettore non nullo fisso comune di A e P .

Dimostrazione. (1) Per ipotesi $A^t w = w$ e quindi $w^t = w^t A$. Abbiamo perciò:

$$PA = vw^t A = vw^t = P.$$

$$AP = Avw^t = vw^t = P.$$

$$P^2 = vw^t vw^t = vw^t = P.$$

(2) $Pv = vw^t v = v$.

Infine $v \neq 0$ perché $w^t v = 1$.

Lemma 13.18. *Sia $P \in \mathbb{C}_q^q$ tale che $PA = AP = P = P^2$. Allora:*

(1) $P(A - P) = 0$.

(2) $A^n = P + (A - P)^n$ per ogni $n \geq 1$.

(3) *Se λ è autovalore $\neq 0$ di $A - P$, allora ogni autovettore di $A - P$ rispetto a λ è anche autovettore di A rispetto a λ .*

Il punto (3) implica in particolare che ogni autovalore $\neq 0$ di $A - P$ è anche autovalore di A .

Dimostrazione. (1) $P(A - P) = PA - P^2 = 0$.

(2) Induzione su n .

$$n = 1: A = P + (A - P).$$

$n \longrightarrow n + 1$:

$$\begin{aligned} A^{n+1} &= AA^n \stackrel{ind}{=} A(P + (A - P)^n) = AP + A(A - P)^n \\ &= P + A(A - P)^n \stackrel{(1)}{=} P + (A - P)(A - P)^n = P + (A - P)^{n+1}. \end{aligned}$$

(3) Sia $v \in \mathbb{C}^q \setminus 0$ e $(A - P)v = \lambda v$. Per il punto (1) allora $\lambda Pv = P(A - P)v = 0$, e siccome $\lambda \neq 0$ ciò implica $Pv = 0$. Ma allora $\lambda v = (A - P)v = Av$.

Osservazione 13.19. Sia $P \in \mathbb{C}_q^q$ tale che $PA = AP = P = P^2$. λ sia un autovalore $\neq 0$ di molteplicità geometrica 1 di A . A e P abbiano un autovettore in comune rispetto a λ . Allora λ non è autovalore di $A - P$.

Dimostrazione. Siano $v \in \mathbb{C}^q \setminus 0$ tale che $Av = \lambda v = Pv$ e $w \in \mathbb{C}^q \setminus 0$ tali che $(A - P)w = \lambda w$. Per il punto (3) del lemma 13.18 abbiamo anche $Aw = \lambda w$, per cui $Pw = 0$. Per ipotesi esiste $t \in \mathbb{C}$ tale che $w = tv$ e quindi $0 = Pw = tPv = \lambda tv$. Siccome necessariamente $t \neq 0$, arriviamo ad una contraddizione.

Lemma 13.20. Sia $P \in \mathbb{C}_q^q$ e siano soddisfatte le seguenti condizioni:

- (1) $PA = AP = P = P^2$.
- (2) A e P hanno un vettore fisso non nullo in comune.
- (3) $\mathcal{M}(A) = \{1\}$.
- (4) 1 abbia molteplicità geometrica 1 come autovalore di A .

Allora:

- (I) $\rho(A - P) < 1$.
- (II) $\lim_{n \rightarrow \infty} A^n = P$.

Dimostrazione. (I) Il punto (3) implica che $\rho(A) = 1$ e ogni autovalore diverso da 1 ha modulo minore di 1. Per l'oss. 13.19 dalla condizione (4) segue invece che 1 non è autovalore di $A - P$.

Per il lemma 13.18 però ogni autovalore $\neq 0$ di $A - P$ è anche autovalore di A . Ciò implica $\rho(A - P) < 1$.

(II) Per l'oss. 10.19 abbiamo $\lim_{n \rightarrow \infty} (A - P)^n = 0$. Dal punto (2) del lemma 13.18 otteniamo $\lim_{n \rightarrow \infty} A^n = P$.

Teorema 13.21. *A sia positiva, v il vettore di Perron di A e w il vettore di Perron di A^t . Sia $P := \frac{vw^t}{v^t w}$. Allora:*

$$(1) \lim_{n \rightarrow \infty} \left(\frac{1}{\rho(A)} A \right)^n = P.$$

(2) P è una matrice positiva di rango 1.

Dimostrazione. Siccome i vettori v e w sono positivi, il prodotto scalare $v^t w$ è positivo.

(1) Sia $B := \frac{1}{\rho(A)} A$. Allora v è un vettore fisso di B e $\tilde{w} := \frac{w}{v^t w}$ un vettore fisso di B^t e si ha $v^t \tilde{w} = 1$. Inoltre $P = v \tilde{w}^t$. Dal lemma 13.17 segue che allora $PB = BP = P = P^2$ e che P e B hanno in v un vettore fisso $\neq 0$ in comune. È chiaro che $\rho(B) = 1$. Per la prop. 13.15 $\mathcal{M}(B) = \{1\}$.

Per l'oss. 13.10 inoltre 1 possiede molteplicità geometrica 1 come autovalore di B . Applicando il lemma 13.20 con B al posto di A otteniamo $\lim_{n \rightarrow \infty} B^n = P$.

(2) È chiaro che $P > 0$, essendo v e w vettori positivi. Ogni colonna di P è un multiplo di v . Essendo $P \neq 0$, ciò implica che P è di rango 1.

Corollario 13.22. *A sia positiva e $\rho(A) = 1$. La matrice P sia costruita come nel teorema 13.21. Allora $\lim_{n \rightarrow \infty} A^n = P$.*

Osservazione 13.23. Sia $1 \leq p < q$. I vettori $e_1, \dots, e_p \in \mathbb{C}^q$ siano ortogonali tra loro e di lunghezza 1. Allora esistono vettori e_{p+1}, \dots, e_q tali che la matrice $E := (e_1, \dots, e_q)$ sia unitaria.

Dimostrazione. È una conseguenza immediata del procedimento di ortogonalizzazione di Gram-Schmidt.

Teorema 13.24 (teorema di Schur). *Esiste una matrice unitaria $E \in U(q)$ tale che $E^{-1}AE$ sia una matrice triangolare superiore.*

Dimostrazione. Induzione su q .

$q = 1$: Banale.

$q - 1 \rightarrow q$: Siano λ un autovalore di A ed e_1 un autovettore di A corrispondente a λ e di lunghezza 1. Per l'oss. 13.23 esiste una matrice unitaria $F \in U(q)$, la cui prima colonna è e_1 . Siccome $Ae_1 = \lambda e_1$, abbiamo $F^{-1}AF = \begin{pmatrix} \lambda & a \\ 0 & B \end{pmatrix}$ con $a \in \mathbb{C}_{q-1}$, $B \in \mathbb{C}_{q-1}^{q-1}$.

Per ipotesi di induzione esiste una matrice unitaria $G \in U(q - 1)$ tale che $T := G^{-1}BG$ sia triangolare superiore. Allora

$E := F \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix} \in U(q)$ e la matrice

$$\begin{aligned} E^{-1}AE &= \begin{pmatrix} 1 & 0 \\ 0 & G^{-1} \end{pmatrix} F^{-1}AF \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} \lambda & a \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & G \end{pmatrix} = \begin{pmatrix} \lambda & aG \\ 0 & T \end{pmatrix} \end{aligned}$$

è ancora triangolare superiore.

Proposizione 13.25. *A sia positiva. Allora la molteplicità algebrica dell'autovalore $\rho(A)$ è uguale a 1.*

Dimostrazione. $\rho := \rho(A)$ abbia molteplicità algebrica m . Gli altri autovalori di A siano $\lambda_{m+1}, \dots, \lambda_q$. Per la prop. 13.15 $|\lambda_j| < \rho$ per ogni j . Per il teorema 13.24 esiste $E \in U(q)$ tale che

$$A = E \begin{pmatrix} \rho & & & & \\ & \ddots & & & \\ & & \rho & & * \\ & & & \lambda_{m+1} & \\ 0 & & & & \ddots \\ & & & & & \lambda_q \end{pmatrix} E^{-1}$$

Con la matrice P definita come nel teorema 13.21 abbiamo allora

$$\begin{aligned} \lim_{n \rightarrow \infty} E(A/\rho)^n E^{-1} &= \lim_{n \rightarrow \infty} \begin{pmatrix} 1 & & & & \\ & \ddots & & & * \\ & & 1 & & \\ & & & \frac{\lambda_{m+1}}{\rho} & \\ 0 & & & & \ddots \\ & & & & & \frac{\lambda_q}{\rho} \end{pmatrix}^n \\ &= E \begin{pmatrix} 1 & & & & \\ & \ddots & & & * \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix} E^{-1} = P \end{aligned}$$

essendo $|\lambda_j/\rho| < 1$ per ogni j .

La matrice P ha perciò rango $\geq m$. Per il teorema 13.21 però P possiede rango 1 e ciò è possibile solo se $m = 1$.

Osservazione 13.26. I risultati di questo capitolo, in particolare la prop. 13.15, il teorema 13.21 e la prop. 13.25, vengono spesso raccolti sotto il nome di *teorema di Perron*.

14. Matrici irriducibili

Situazione 14.1. Siano $q \in \mathbb{N} + 1$ e $A \in \mathbb{R}_q^q$.

Con S_q denotiamo il gruppo delle permutazioni di $\{1, \dots, q\}$.

Osservazione 14.2. La matrice identica mostra che i risultati del capitolo precedente non rimangono validi se l'ipotesi che A sia positiva viene sostituita con la condizione che A sia non negativa. È stato Frobenius ad individuare la classe di matrici a cui le conclusioni del teorema di Perron possono essere generalizzate, le matrici irriducibili (cfr. Salce, pag. 137).

Definizione 14.3. Un δ -sottospazio di \mathbb{R}^q è un sottospazio vettoriale V di \mathbb{R}^q della forma $V = SV(\delta_{i_1}, \dots, \delta_{i_k})$ con $k \in \mathbb{N}$ ed $i_1, \dots, i_k \in \{1, \dots, q\}$.

Definizione 14.4. La matrice A si dice *riducibile*, se esiste un δ -sottospazio A -invariante V di \mathbb{R}^q con $0 < \dim V < q$.

Altrimenti A si dice *irriducibile*.

Osservazione 14.5. Sia $E = (e_1, \dots, e_q) \in GL(q, \mathbb{R})$ una base ordinata di \mathbb{R}^q . Allora la matrice dell'applicazione $\bigcirc_x Ax : \mathbb{R}^q \rightarrow \mathbb{R}^q$ rispetto alla base E è uguale a $E^{-1}AE$.

Dimostrazione. Corso di Geometria I.

Utilizzando la notazione matriciale si può vedere direttamente nel modo seguente. Sia B la matrice dell'applicazione $\bigcirc_x Ax : \mathbb{R}^q \rightarrow \mathbb{R}^q$ rispetto alla base E . Per ogni j allora $B_{(j)}$ è uguale al vettore delle coordinate di Ae_j rispetto alla base E . Ciò significa $AE_{(j)} = EB_{(j)}$ per ogni j , quindi $AE = EB$ e $B = E^{-1}AE$.

Osservazione 14.6. $E = (f_1, \dots, f_k, g_1, \dots, g_{q-k})$ sia una base di \mathbb{R}^q . Allora sono equivalenti:

- (1) Il sottospazio $SV(f_1, \dots, f_k)$ è A -invariante.
- (2) Esistono matrici $X \in \mathbb{R}_k^k, Y \in \mathbb{R}_{q-k}^k, Z \in \mathbb{R}_{q-k}^{q-k}$ tali che

$$E^{-1}AE = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}.$$

Dimostrazione. Ciò è una conseguenza immediata dell'oss. 14.5.

Osservazione 14.7. Sia $1 \leq k \leq q$. Allora sono equivalenti:

- (1) Esiste un δ -sottospazio A -invariante di dimensione k di \mathbb{R}^q .
- (2) Esistono matrici $X \in \mathbb{R}_k^k, Y \in \mathbb{R}_{q-k}^k, Z \in \mathbb{R}_{q-k}^{q-k}$ e una matrice di

$$\text{permutazione } P \in \mathbb{R}_q^q \text{ tali che } P^{-1}AP = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}.$$

Dimostrazione. Ciò segue dall'oss. 14.6.

Corollario 14.8. Sono equivalenti:

(1) A è riducibile.

(2) Esistono $k \in \mathbb{N}$ con $0 < k < q$ e una matrice di

$$\text{permutazione } P \in \mathbb{R}_q^q \text{ tali che } P^{-1}AP = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$$

$$\text{con } X \in \mathbb{R}_k^k, Y \in \mathbb{R}_{q-k}^k, Z \in \mathbb{R}_{q-k}^{q-k}.$$

Definizione 14.9. Per $\sigma \in S_q$ siano $\delta_\sigma := (\delta_{\sigma(1)}, \dots, \delta_{\sigma(q)})$ e

$\delta^\sigma := \delta_\sigma^t = \delta_\sigma^{-1}$. Abbiamo quindi $(\delta_\sigma)_{(j)}^{(i)} = \delta_{\sigma(j)}^i$ e

$(\delta^\sigma)_{(j)}^{(i)} = (\delta_\sigma)_{(j)}^{(i)} = \delta_{\sigma(i)}^j = \delta_j^{\sigma(i)}$ per ogni i, j .

Osservazione 14.10. Siano $\sigma \in S_q, v \in \mathbb{R}^q, f \in \mathbb{R}_q, A \in \mathbb{R}_q^q$. Allora

$$\delta^\sigma v = \begin{pmatrix} v^{\sigma(1)} \\ \vdots \\ v^{\sigma(q)} \end{pmatrix}$$

$$f \delta_\sigma = (f_{\sigma(1)} \quad \dots \quad f_{\sigma(q)})$$

$$\delta^\sigma A = \begin{pmatrix} A_1^{\sigma(1)} & \dots & A_q^{\sigma(1)} \\ \vdots & & \vdots \\ A_1^{\sigma(q)} & \dots & A_q^{\sigma(q)} \end{pmatrix}$$

$$A \delta_\sigma = \begin{pmatrix} A_{\sigma(1)}^1 & \dots & A_{\sigma(q)}^1 \\ \vdots & & \vdots \\ A_{\sigma(1)}^q & \dots & A_{\sigma(q)}^q \end{pmatrix}$$

$$\delta^\sigma A \delta_\sigma = \begin{pmatrix} A_{\sigma(1)}^{\sigma(1)} & \dots & A_{\sigma(q)}^{\sigma(1)} \\ \vdots & & \vdots \\ A_{\sigma(1)}^{\sigma(q)} & \dots & A_{\sigma(q)}^{\sigma(q)} \end{pmatrix}$$

dove abbiamo tralasciato le parentesi tonde negli indici.

Dimostrazione. Siano $i, j \in \{1, \dots, q\}$. Allora:

(1) $(\delta^\sigma v)^i = \delta^{\sigma(i)} v = v^{\sigma(i)}$.

(2) $(f \delta_\sigma)_j = f \delta_{\sigma(j)} = f_{\sigma(j)}$.

(3) $(\delta^\sigma A)_j^i = \delta^{\sigma(i)} A_j = A_j^{\sigma(i)}$.

(4) $(A \delta_\sigma)_j^i = A^i \delta_{\sigma(j)} = A_{\sigma(j)}^i$.

(5) $(\delta^\sigma A \delta_\sigma)_j^i = \delta^{\sigma(i)} (A \delta_\sigma)_j = \delta^{\sigma(i)} A_{\sigma(j)} = A_{\sigma(j)}^{\sigma(i)}$.

Osservazione 14.11. A è riducibile se e solo se $A + \delta$ è riducibile.

Dimostrazione. A sia riducibile. Siano X, Y, Z, P come nel cor. 14.8. Allora

$$P^{-1}(A + \delta)P = P^{-1}AP + \delta = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} + \delta = \begin{pmatrix} X + \delta & Y \\ 0 & Z + \delta \end{pmatrix}$$

Le matrici $X + \delta$ e $Z + \delta$ sono ancora quadratiche. Quindi $A + \delta$ è riducibile.

Nello stesso modo si dimostra la direzione opposta.

Osservazione 14.12. A sia riducibile ed $m \in \mathbb{N}$. Allora A^m è riducibile.

Dimostrazione. Siano X, Y, Z, P come nel cor. 14.8. Allora

$$P^{-1}A^mP = (P^{-1}AP)^m = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}^m = \begin{pmatrix} X^m & R \\ 0 & Z^m \end{pmatrix}$$

con $R \in \mathbb{R}_{q-k}^k$, per cui A^m è riducibile.

Osservazione 14.13. Una matrice positiva è irriducibile.

Dimostrazione. Ciò segue direttamente dal corollario 14.8.

Osservazione 14.14. Se esiste $m \in \mathbb{N}$ tale che $A^m > 0$ oppure $(A + \delta)^m > 0$, allora A è irriducibile.

Dimostrazione. (1) Se $A^m > 0$, allora per l'oss. 14.13 A^m è irriducibile. Ciò implica per l'oss. 14.12 che anche A è irriducibile.

(2) Sia $(A + \delta)^m > 0$. Allora per il punto (1) $A + \delta$ è irriducibile e l'oss. 14.11 implica che anche A è irriducibile.

Lemma 14.15. A sia non negativa ed irriducibile e $v \in \mathbb{R}^q$ un vettore non negativo. Sia $1 \leq k < n$ tale che esattamente k delle coordinate di v siano positive.

Allora il vettore $(A + \delta)v$ possiede almeno $k + 1$ coordinate positive.

Dimostrazione. Seguiamo Salce, pag. 138. Sia $w := (A + \delta)v = Av + v$. È chiaro che $w^{(i)} > 0$ se $v^{(i)} > 0$. w ha perciò almeno k coordinate positive. Assumiamo per assurdo che tutte le altre coordinate di w siano nulle. Allora ogni volta che $v^{(i)} = 0$ anche $(Av)^{(i)} = 0$.

Sia ora $\sigma \in S_q$ tale che le prime k componenti di $u := \delta^\sigma v$ siano positive e quindi le rimanenti nulle. u è quindi della forma $u = \begin{pmatrix} u^+ \\ 0 \end{pmatrix}$ con $u^+ \in \mathbb{R}^k$ positivo.

Per quanto osservato prima, anche le ultime $q - k$ coordinate di $\delta^\sigma Av = \delta^\sigma A\delta_\sigma u$ sono nulle.

$$\text{Possiamo scrivere } B := \delta^\sigma A\delta_\sigma \text{ nella forma } B = \begin{pmatrix} X & Y \\ R & Z \end{pmatrix},$$

con $X \in \mathbb{R}_k^k$, $Y \in \mathbb{R}_{q-k}^k$, $R \in \mathbb{R}_k^{q-k}$, $Z \in \mathbb{R}_{q-k}^{q-k}$. Allora

$$Bu = \begin{pmatrix} X & Y \\ R & Z \end{pmatrix} \begin{pmatrix} u^+ \\ 0 \end{pmatrix} = \begin{pmatrix} Xu^+ \\ Ru^+ \end{pmatrix} = \begin{pmatrix} Xu^+ \\ 0 \end{pmatrix}$$

Ma u^+ è positivo ed $R \geq 0$, e ciò è possibile solo se $R = 0$. Ciò implica che A è riducibile, in contrasto con l'ipotesi.

Proposizione 14.16. A sia non negativa. Allora sono equivalenti:

(1) A è irriducibile.

(2) $(A + \delta)^{q-1}$ è positiva.

(3) Esiste $m \in \mathbb{N}$ tale che $(A + \delta)^m$ è positiva.

Dimostrazione. (1) \implies (2): A sia irriducibile. Sia $j \in \{1, \dots, q\}$. Dimostriamo che la j -esima colonna $((A + \delta)^{q-1})_{(j)}$ è positiva. Questa colonna però è uguale a $(A + \delta)^{q-1} \delta_j$ e per il lemma 14.15 questo vettore possiede q coordinate positive ed è quindi positivo.

(2) \implies (3): Banale.

(3) \implies (1): Oss. 14.14.

Esempio 14.17. Sia $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Allora $A + \delta = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ è positiva e dalla prop. 14.16 segue che A è irriducibile.

Però $A^m = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ se m è pari e $A^m = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ se m è dispari, e quindi la matrice A^m non è mai positiva.

Osservazione 14.18. A è riducibile se e solo se A^t è riducibile.

Dimostrazione. Se A è non negativa, ciò segue dalla prop. 14.16.

Il caso generale, non difficile, si trova in Huppert, pag. 353.

Osservazione 14.19. Sono equivalenti:

(1) A è riducibile.

(2) Esistono $k \in \mathbb{N}$ con $0 < k < q$ e una matrice di permutazione

$$P \in \mathbb{R}_q^q \text{ tale che } P^{-1}AP = \begin{pmatrix} X' & 0 \\ Y' & Z' \end{pmatrix} \text{ con } X' \in \mathbb{R}_k^k, Y' \in \mathbb{R}_k^{q-k}, \\ Z' \in \mathbb{R}_{q-k}^{q-k}.$$

Dimostrazione. (1) \implies (2): A sia riducibile. Per l'oss. 14.18 anche A^t è riducibile, per cui esiste una matrice di permutazione $P \in \mathbb{R}_q^q$ tale che $P^{-1}A^tP = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$ con $X \in \mathbb{R}_k^k, Y \in \mathbb{R}_{q-k}^k, Z \in \mathbb{R}_{q-k}^{q-k}$.

$$\text{Ciò implica } P^{-1}AP = \begin{pmatrix} X^t & 0 \\ Y^t & Z^t \end{pmatrix}.$$

(2) \implies (1): Per simmetria, usando ancora l'oss. 14.18.

15. Il teorema di Frobenius

Situazione 15.1. Siano $q \in \mathbb{N} + 1$ ed $A \in \mathbb{R}_q^q$.

Lemma 15.2.

- (1) $\rho(A + \delta) \leq 1 + \rho(A)$.
- (2) Se A è non negativa, allora $\rho(A + \delta) = 1 + \rho(A)$.

Dimostrazione. (1) Gli autovalori di A siano $\lambda_1, \dots, \lambda_q$ con le rispettive molteplicità. Dalla forma normale di Jordan è immediato che gli autovalori di $A + \delta$ con le rispettive molteplicità sono $1 + \lambda_1, \dots, 1 + \lambda_q$. Perciò

$$\begin{aligned}\rho(A + \delta) &= \max\{|1 + \lambda_1|, \dots, |1 + \lambda_q|\} \leq \max\{1 + |\lambda_1|, \dots, 1 + |\lambda_q|\} \\ &= 1 + \max\{|\lambda_1|, \dots, |\lambda_q|\} = 1 + \rho(A)\end{aligned}$$

(2) Se A è non negativa, anche $A + \delta$ è non negativa e per l'oss. 13.16 $1 + \rho(A)$ è autovalore di $A + \delta$. Perciò $1 + \rho(A) \leq \rho(A + \delta)$ e per il punto (1) ciò implica $1 + \rho(A) = \rho(A + \delta)$.

Osservazione 15.3. Una matrice irriducibile non può avere righe o colonne nulle.

Dimostrazione. Supponiamo $A_{(j)} = A\delta_j = 0$. Allora $SV(\delta_j)$ è un δ -sottospazio A -invariante. Quindi per definizione A è riducibile, in contrasto con l'ipotesi.

Teorema 15.4 (teorema di Frobenius). *La matrice A sia non negativa e irriducibile. Allora:*

- (1) $\rho(A)$ è autovalore di A .
- (2) $\rho(A) > 0$.
- (3) Esiste un autovettore positivo rispetto all'autovalore $\rho(A)$.
- (4) La molteplicità algebrica dell'autovalore $\rho(A)$ è uguale a 1.

Dimostrazione. (1) Oss. 13.16.

(2) Per l'oss. 15.3 A non possiede colonne nulle. Dal corollario 12.13 segue che $\rho(A) > 0$.

(3) Per l'oss. 13.16 esiste un autovettore $v \geq 0$ rispetto a $\rho(A)$. Allora $(A + \delta)v = Av + v = \rho(A)v + v = (1 + \rho(A))v$ e perciò anche $(A + \delta)^{q-1}v = (1 + \rho(A))^{q-1}v$. Per la prop. 14.16 però $(A + \delta)^{q-1} > 0$, quindi per l'oss. 13.4 anche $(A + \delta)^{q-1}v > 0$. Abbiamo così $(1 + \rho(A))^{q-1}v > 0$, e ciò implica $v > 0$.

(4) La molteplicità algebrica di $\rho(A)$ come autovalore di A è uguale alla molteplicità algebrica di $1 + \rho(A)$ come autovalore di $A + \delta$. Ma la molteplicità algebrica di $1 + \rho(A)$ come autovalore di $A + \delta$ è minore o uguale alla molteplicità algebrica di $(1 + \rho(A))^{q-1}$ come autovalore di $(A + \delta)^{q-1}$, come si vede ad esempio dall'oss. 10.20. Per il lemma 15.2 il raggio spettrale di $A + \delta$ è $1 + \rho(A)$, per cui, ancora per l'oss. 10.20, $(1 + \rho(A))^{q-1}$ è il raggio spettrale di $(A + \delta)^{q-1}$. Quest'ultima matrice

è però positiva, per cui dalla prop. 13.25 vediamo che la molteplicità algebrica di $(1 + \rho(A))^{q-1}$ come autovalore di $(A + \delta)^{q-1}$ è uguale a 1.

Osservazione 15.5. Sia $A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Questa matrice è non negativa ed irriducibile. Gli autovalori di A sono 1 e -1 , quindi $\mathcal{M}(A) = \{1, -1\}$.

Ciò mostra che la prop. 13.15 non rimane più valida se si sostituisce l'ipotesi che A sia positiva con la condizione che A sia non negativa.

Definizione 15.6. A sia non negativa ed irriducibile. Dal teorema 15.4 segue che A possiede un unico autovettore positivo e stocastico rispetto all'autovalore $\rho(A)$. Questo vettore si chiama il *vettore di Perron* di A .

Il vettore di Perron di una matrice positiva è stato introdotto nella def. 13.11.

Proposizione 15.7. A sia non negativa ed irriducibile. Allora ogni autovettore non negativo di A è un multiplo (reale positivo) del vettore di Perron di A e quindi in particolare un autovettore rispetto a $\rho(A)$.

Dimostrazione. Usando i risultati di questo capitolo, è sufficiente trascrivere la dimostrazione della prop. 13.12.

16. Il teorema di Wielandt

Situazione 16.1. Siano $q \in \mathbb{N} + 1$ ed $A \in \mathbb{R}_q^q$.

Definizione 16.2. Per $B \in \mathbb{C}_q^q$ poniamo $[B] := \begin{pmatrix} |B_{(1)}^{(1)}| & \dots & |B_{(q)}^{(1)}| \\ \vdots & & \vdots \\ |B_{(1)}^{(q)}| & \dots & |B_{(q)}^{(q)}| \end{pmatrix}$.

La notazione analoga per vettori è stata introdotta nella def. 13.2.

Osservazione 16.3. Sia $B \in \mathbb{C}_q^q$. Allora $\rho(B) \leq \rho([B])$.

Dimostrazione. Per ogni $n \in \mathbb{N}$ la disuguaglianza triangolare implica $[B^n] \leq [B]^n$. Dalla definizione della norma di Frobenius (introdotta nella def. 10.5) segue immediatamente che $\|B^n\|_F = \|[B^n]\|_F \leq \|[B]^n\|_F$. Utilizzando il teorema 10.21 abbiamo

$$\rho(B) = \lim_{n \rightarrow \infty} \sqrt[n]{\|B^n\|_F} \leq \lim_{n \rightarrow \infty} \sqrt[n]{\|[B]^n\|_F} = \rho([B])$$

Osservazione 16.4. Siano $B \in \mathbb{C}_q^q$ e $v \in \mathbb{C}^q$. Allora $[Bv] \leq [B][v]$.

Dimostrazione. Come nell'oss. 13.5 usando la disuguaglianza triangolare per ogni i abbiamo

$$[Bv]^{(i)} = |Bv|^{(i)} = |B^{(i)}v| = \left| \sum_{j=1}^q B_{(j)}^{(i)} v^{(j)} \right| \leq \sum_{j=1}^q |B_{(j)}^{(i)}| |v^{(j)}| = ([B][v])^{(i)}$$

Lemma 16.5. Sia $B \in \mathbb{C}_q^q$. Supponiamo che siano soddisfatte le seguenti condizioni:

- (1) A è non negativa e irriducibile.
- (2) $[B] \leq A$.
- (3) $\rho(A) = \rho(B)$.

Sia $\lambda \in \mathcal{M}(B)$ e sia $v \in \mathbb{C}^q$ un autovettore di B rispetto a λ . Allora $[v]$ è un autovettore di A rispetto a $\rho(A)$ e quindi positivo per 15.7. Inoltre $A = [B]$.

Dimostrazione. Poniamo $\rho := \rho(A)$. Per ipotesi ρ coincide anche con il raggio spettrale di B . Per il teorema 15.4 $\rho > 0$.

Siccome $\lambda \in \mathcal{M}(B)$, si ha $\left| \frac{\rho}{\lambda} \right| = 1$, perciò

$$\rho[v] = [\rho v] = \left[\frac{\rho}{\lambda} \lambda v \right] = [\lambda v] = [Bv] \stackrel{16.4}{\leq} [B][v] \leq A[v] \quad (\star)$$

Per l'oss. 14.18 anche A^t è irriducibile. Per il teorema 15.4 esiste $f \in \mathbb{R}_q$ con $f > 0$ e tale che $fA = \rho f$. Ciò implica $fA[v] = \rho f[v]$, ovvero $f(A[v] - \rho[v]) = 0$.

Per (\star) $A[v] - \rho[v] \geq 0$, e siccome $f > 0$, dobbiamo avere $A[v] = \rho[v]$. Dalla prop. 15.7 segue $[v] > 0$, mentre in (\star) abbiamo adesso $[B][v] = A[v]$, cioè $(A - [B])[v] = 0$. Per ipotesi $A - [B] \geq 0$ e ciò implica $A = [B]$.

Corollario 16.6. Sia $B \in \mathbb{R}_q^q$ tale che $0 \leq B \leq A$. A sia irriducibile e $A \neq B$. Allora $\rho(B) < \rho(A)$.

Dimostrazione. Altrimenti il lemma 16.5, essendo per ipotesi $B = [B]$, implicherebbe $A = B$.

Osservazione 16.7. $w \in \mathbb{R}^q$ sia un vettore positivo e $C \in \mathbb{C}_q^q$ tale che $Cw = [C]w$. Allora $C = [C]$.

Dimostrazione. Denotiamo con $\operatorname{Re} C$ la parte reale di C e con $\operatorname{Im} C$ la parte immaginaria.

(1) Le ipotesi implicano in primo luogo che

$$\operatorname{Re}(Cw) = \operatorname{Re}([C]w) = [C]w = Cw$$

Però $Cw = (\operatorname{Re} C)w + i(\operatorname{Im} C)w$. Siccome $w > 0$, ciò implica $C = \operatorname{Re} C$. La matrice C è quindi reale.

(2) A questo punto però da $w > 0$ e $Cw = [C]w$ segue direttamente $C = [C]$.

Lemma 16.8. Sia $B \in \mathbb{C}_q^q$. $[B]$ sia irriducibile e $\rho(B) = \rho([B])$. Sia $\lambda \in \mathcal{M}(B)$. Allora esistono $\alpha_1, \dots, \alpha_q \in \mathbb{R}$ tali che con

$$D := \begin{pmatrix} e^{i\alpha_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\alpha_q} \end{pmatrix} \text{ si ha } D^{-1}BD = \frac{\lambda}{\rho}[B].$$

Dimostrazione. Poniamo $\rho := \rho(B)$. Per il teorema 15.4 $\rho > 0$.

Sia $v \in \mathbb{C}^q$ un autovettore di B rispetto a λ . Per il lemma 16.5 (applicato ad $A := [B]$) abbiamo $[v] > 0$ e $[B][v] = \rho[v]$. Possiamo scrivere $v^{(k)} = e^{i\alpha_k}|v^{(k)}|$, con $\alpha_1, \dots, \alpha_q \in \mathbb{R}$.

$$\text{Con } D := \begin{pmatrix} e^{i\alpha_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\alpha_q} \end{pmatrix} \text{ abbiamo allora } D[v] = v.$$

Ciò implica $BD[v] = Bv = \lambda v = \frac{\lambda}{\rho}\rho v = \frac{\lambda}{\rho}\rho D[v]$ e quindi $\frac{\rho}{\lambda}D^{-1}BD[v] = \rho[v] = [B][v]$. Sia $C := \frac{\rho}{\lambda}D^{-1}BD$. Allora $C[v] = [B][v]$.

Come nell'oss. 12.15 abbiamo $C_{(j)}^{(k)} = \frac{\rho}{\lambda}e^{-i\alpha_k}e^{i\alpha_j}B_{(j)}^{(k)}$, e quindi $|C_{(j)}^{(k)}| = |B_{(j)}^{(k)}|$ per ogni k, j , per cui $[C] = [B]$, cosicché $C[v] = [C][v]$.

L'oss. 17.6 implica adesso $C = [C]$ e quindi $C = [B]$.

Ciò significa $\frac{\rho}{\lambda}D^{-1}BD = [B]$, ovvero $D^{-1}BD = \frac{\lambda}{\rho}[B]$.

Definizione 16.9. Due matrici $B, C \in \mathbb{C}_q^q$ si dicono *simili*, se esiste $E \in \operatorname{GL}(n, \mathbb{C})$ tali che $C = E^{-1}BE$. Scriviamo allora $B \sim C$.

Osservazione 16.10. La relazione di somiglianza introdotta nella def. 16.9 è una relazione di equivalenza su \mathbb{C}_q^q .

Inoltre, se $B \sim C$, allora $tB \sim tC$ per ogni $t \in \mathbb{C}$.

Matrici simili hanno lo stesso polinomio caratteristico.

Corollario 16.11. *A sia non negativa ed irriducibile. Siano $\rho := \rho(A)$ e $\lambda \in \mathcal{M}(B)$. Allora $A \sim \frac{\lambda}{\rho}A$. Poiché $\lambda \in \mathcal{M}(B)$, $\frac{\lambda}{\rho} \neq 0$.*

Osservazione 16.12. Siano $B \in \mathbb{C}_q^q$ e $t \in \mathbb{C}$ con $t \neq 0$. λ sia un autovalore di B di molteplicità algebrica m . Allora $t\lambda$ è autovalore di tB della stessa molteplicità algebrica m .

Dimostrazione. Ciò è chiaro dalla forma normale di Jordan.

Teorema 16.13 (teorema di Wielandt). *A sia non negativa ed irriducibile. Sia $h := |\mathcal{M}(A)|$. Allora $\mathcal{M}(A) = \{e^{\frac{2\pi i}{h}j} \rho(A) \mid j = 0, \dots, h-1\}$.*

Ogni elemento di $\mathcal{M}(A)$ ha molteplicità algebrica 1 come autovalore di A .

Dimostrazione. Poniamo di nuovo $\rho := \rho(A)$. Dal teorema 15.4 sappiamo che $\rho > 0$.

(1) Sia $\lambda \in \mathcal{M}(A)$. Per il cor. 16.11 $A \sim \frac{\lambda}{\rho}A$ e dall'oss. 16.12 e dal teorema 15.4 segue che $\lambda = \frac{\lambda}{\rho}\rho$ è autovalore di $\frac{\lambda}{\rho}A$ di molteplicità algebrica 1. Però A e $\frac{\lambda}{\rho}A$ hanno lo stesso polinomio caratteristico, di cui λ è quindi radice semplice.

(2) Sia $G := \{\frac{\lambda}{\rho} \mid \lambda \in \mathcal{M}(A)\}$. Per ipotesi $|G| = h$ ed è chiaro che $G \subset S^1$. È sufficiente dimostrare che G è sottogruppo di S^1 , perchè ciò implica che $\varepsilon^h = 1$ per ogni $\varepsilon \in G$ ed è ben noto che $\{\varepsilon \in S^1 \mid \varepsilon^h = 1\} = \{e^{\frac{2\pi i}{h}j} \mid j = 0, \dots, h-1\}$.

Siano $\lambda, \mu \in \mathcal{M}(A)$. Per il cor. 16.11 abbiamo $A \sim \frac{\lambda}{\rho}A$ e $A \sim \frac{\mu}{\rho}A$, quindi l'oss. 16.10 implica $A \sim \frac{\lambda}{\mu}A$. Per l'oss. 16.12 ciò implica che $\frac{\lambda}{\mu}\rho \in \mathcal{M}(A)$, cosicché $\frac{\lambda}{\mu} \in G$.

Proposizione 16.14. *A sia non negativa ed irriducibile ed $h := |\mathcal{M}(A)|$. Allora lo spettro di A è invariante rispetto alla rotazione di angolo $\frac{2\pi}{h}$.*

Dimostrazione. Siano $\rho := \rho(A)$ ed $\varepsilon := e^{\frac{2\pi i}{h}}$. Per il teorema 16.13 abbiamo $\lambda := \varepsilon\rho \in \mathcal{M}(A)$. Per il cor. 16.11 $A \sim \frac{\lambda}{\rho}A = \varepsilon A$. Per l'oss. 16.10 A ed εA hanno lo stesso polinomio caratteristico, ma d'altra parte per l'oss. 16.12 lo spettro di εA consiste degli autovalori di A moltiplicati per ε .

17. Matrici primitive

Situazione 17.1. Siano $q \in \mathbb{N} + 1$ ed $A \in \mathbb{R}_q^q$.

Definizione 17.2. A si dice *primitiva*, se sono soddisfatte le seguenti condizioni:

- (1) A è non negativa e irriducibile.
- (2) $|\mathcal{M}(A)| = 1$.

Lemma 17.3. A sia non negativa e irriducibile. Siano $h := |\mathcal{M}(A)|$ e $\rho := \rho(A)$. Sia $m \in \mathbb{N} + 1$. Allora:

- (1) $\mathcal{M}(A^{mh}) = \{\rho^{mh}\}$.
- (2) ρ^{mh} ha molteplicità algebrica h come autovalore di A^{mh} .

Dimostrazione. Per il teorema 16.13 il polinomio caratteristico di A è della forma $(x - \rho)(x - e^{\frac{2\pi i}{h}}\rho) \dots (x - e^{\frac{2\pi i}{h}(h-1)}\rho)\psi$, dove ψ è il prodotto dei fattori $(x - \lambda)$, dove λ sono gli autovalori di A di modulo minore di ρ . È chiaro a questo punto l'enunciato.

Proposizione 17.4. A sia non negativa. Se esiste $m \in \mathbb{N}$ tale che $A^m > 0$, allora A è primitiva.

Dimostrazione. Per l'oss. 14.14 A è irriducibile. Siano $h := |\mathcal{M}(A)|$ e $\rho := \rho(A)$. Il caso $m = 0$ è banale (e implica $q = 1$). Sia quindi $m \geq 1$. Allora anche $mh \geq 1$, e chiaramente $A^{mh} > 0$. Per il lemma 17.3 $\mathcal{M}(A^{mh}) = \{\rho^{mh}\}$ e ρ^{mh} possiede molteplicità algebrica h come autovalore di A^{mh} . Per la prop. 13.25 questa molteplicità algebrica è però uguale a 1.

Teorema 17.5. A sia primitiva e $\rho := \rho(A)$. Usando l'oss. 14.18 e il teorema 15.4 possiamo trovare vettori positivi $v, w \in \mathbb{R}^q$ tali che $Av = \rho v$ e $A^t w = \rho w$. Possiamo assumere che $w^t v = 1$.

$$\text{Allora } \lim_{n \rightarrow \infty} \left(\frac{1}{\rho} A\right)^n = vw^t.$$

Dimostrazione. Possiamo ricalcare la dimostrazione del teorema 13.21. Si osservi che $\rho > 0$ per il teorema 15.4.

Siano $B := \frac{1}{\rho} A$ e $P := vw^t$. Allora v è un vettore fisso di B e w un vettore fisso di B^t . Dal lemma 13.17 segue che $PB = BP = P^2 = P$ e che P e B hanno in v un vettore fisso positivo in comune. È chiaro che $\rho(B) = 1$. Naturalmente anche B è primitiva e quindi $\mathcal{M}(B) = \{1\}$. Per il teorema 16.13 l'autovalore 1 di B possiede molteplicità algebrica e quindi anche geometrica 1. Applicando il lemma 13.20 con B al posto di A otteniamo $\lim_{n \rightarrow \infty} B^n = P$.

Teorema 17.6. A sia non negativa. Allora sono equivalenti:

- (1) A è primitiva.
- (2) Esiste $m \in \mathbb{N}$ tale che $A^m > 0$.

Dimostrazione. (1) \implies (2): Con la notazione del teorema 17.5 abbiamo $\lim_{n \rightarrow \infty} \left(\frac{1}{\rho} A\right)^n = vw^t$. I vettori v e w sono positivi, per cui anche la matrice vw^t è positiva. Ciò implica che deve esistere $m \in \mathbb{N}$ tale che $\left(\frac{1}{\rho} A\right)^m > 0$, e quindi anche $A^m > 0$.

(2) \implies (1): Prop. 17.4.

Corollario 17.7. *A sia non negativa. Se esiste $n \in \mathbb{N}$ tale che A^n sia primitiva, allora anche A è primitiva.*

Dimostrazione. Per il teorema 17.6 esiste $m \in \mathbb{N}$ tale che $A^{mn} > 0$. Ma sempre per il teorema 17.6 segue che A è primitiva.

Corollario 17.8. *A sia non negativa ed $m \in \mathbb{N}$ tale che $A^m > 0$. Allora $A^n > 0$ per ogni $n \geq m$.*

Dimostrazione. Per l'oss. 14.14 A è irriducibile e quindi non può avere righe nulle per l'oss. 15.3. Ciò implica $A^{m+1} > 0$. Ripetendo il ragionamento si trova $A^{m+2} > 0$, ecc.

Corollario 17.9. *A sia primitiva. Allora A^n è primitiva per ogni $n \in \mathbb{N} + 1$.*

Proposizione 17.10. *A sia non negativa ed irriducibile. Se esiste un indice i tale che $A_{(j)}^{(i)} > 0$, allora A è primitiva.*

Dimostrazione. Per la dimostrazione (non troppo difficile) di questo risultato rimandiamo a Salce, pag. 149.

Proposizione 17.11. *A sia primitiva. Allora $A^{q^2-2q+2} > 0$.*

Dimostrazione. Per la dimostrazione, piuttosto complicata, rimandiamo a Horn/Johnson, pag. 520, Gantmacher, pagg. 424-425, oppure Cull/Flahive/Robson, pagg. 198-199.

Teorema 17.12. *A sia primitiva e $\rho := \rho(A)$. x sia un vettore non negativo arbitrario, ma non nullo. Allora il limite $\lim_{n \rightarrow \infty} \rho^{-n} A^n x =: u$ esiste ed è un multiplo positivo del vettore di Perron di A .*

Se scegliamo v e w come nel teorema 17.5, abbiamo:

(1) $u = vw^t x = w^t x \cdot v$.

(2) $\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n x)^{(k)}} = \frac{v^{(i)}}{v^{(k)}}$ per ogni i, k .

Questo limite quindi non dipende da x .

(3) Se y è un altro vettore non negativo e non nullo, allora

$\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n y)^{(i)}} = \frac{w^t x}{w^t y}$ per ogni i .

Questo limite quindi non dipende da i .

(4) Abbiamo perciò

$\lim_{n \rightarrow \infty} \rho^{-n} A^n x = \alpha v$

$$\lim_{n \rightarrow \infty} \rho^{-n} A^n y = \beta v$$

con $\alpha, \beta > 0$ e $\frac{\alpha}{\beta} = \frac{w^t x}{w^t y}$.

Dimostrazione. (1) Per il teorema 17.5 esiste il limite $u := \lim_{n \rightarrow \infty} \rho^{-n} A^n x = v w^t x$. Notiamo che $u > 0$.

Allora $Au = Av w^t x = \rho v w^t x = \rho u$. Per la prop. 15.7 u è un multiplo positivo del vettore di Perron di A .

(2) Osserviamo che $A^n x > 0$ per $n \gg 0$. Per il punto (1) abbiamo

$$\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n x)^{(k)}} = \lim_{n \rightarrow \infty} \frac{(\rho^{-n} A^n x)^{(i)}}{(\rho^{-n} A^n x)^{(k)}} = \frac{(v w^t x)^{(i)}}{(v w^t x)^{(k)}} = \frac{v^{(i)} w^t x}{v^{(k)} w^t x} = \frac{v^{(i)}}{v^{(k)}}$$

(3) Di nuovo $A^n y > 0$ per ogni $n \gg 0$. Per il punto (1) abbiamo

$$\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n y)^{(i)}} = \lim_{n \rightarrow \infty} \frac{(\rho^{-n} A^n x)^{(i)}}{(\rho^{-n} A^n y)^{(i)}} = \frac{(v w^t x)^{(i)}}{(v w^t y)^{(i)}} = \frac{v^{(i)} w^t x}{v^{(i)} w^t y} = \frac{w^t x}{w^t y}$$

(4) Chiaro.

Nota 17.13. A sia primitiva. Se A descrive la crescita di una popolazione suddivisa in sottopopolazioni $1, \dots, q$, i punti (2) e (3) del teorema 17.12 diventano molto importanti. Il punto (2) afferma infatti che i rapporti tra le sottopopolazioni convergono a valori che non dipendono dal vettore di partenza x (che naturalmente deve essere non negativo e non nullo). Il punto (3) descrive, per ogni sottopopolazione i , il modo in cui il numero degli individui appartenenti ad i dipende dai vettori di partenza.

Esempio 17.14. Sia $A = \begin{pmatrix} 4 & 1 & 0 & 1 \\ 1 & 5 & 0 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 5 \end{pmatrix}$.

$$\text{Allora } A^3 = \begin{pmatrix} 77 & 64 & 22 & 64 \\ 62 & 139 & 2 & 14 \\ 11 & 42 & 26 & 42 \\ 2 & 24 & 82 & 149 \end{pmatrix}.$$

La matrice A è quindi primitiva.

Il polinomio minimale è $(x-6)(x-5)(x-\lambda)(x-\mu)$, con $\lambda, \mu = \frac{5 \pm \sqrt{5}}{2}$.

Perciò $\rho(A) = 6$. Calcolando gli autovettori di A e A^t rispetto all'autovalore 6 vediamo che nei teoremi 17.5 e 17.12 possiamo porre

$$v = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 2 \end{pmatrix}, w = \frac{1}{11} \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}. \text{ Perciò}$$

$$\lim_{n \rightarrow \infty} 6^{-n} A^n = \frac{1}{11} \begin{pmatrix} 2 & 4 & 2 & 4 \\ 2 & 4 & 2 & 4 \\ 1 & 2 & 1 & 2 \\ 1 & 4 & 2 & 4 \end{pmatrix} \simeq \begin{pmatrix} 0.18 & 0.36 & 0.18 & 0.36 \\ 0.18 & 0.36 & 0.18 & 0.36 \\ 0.09 & 0.18 & 0.09 & 0.18 \\ 0.18 & 0.36 & 0.18 & 0.36 \end{pmatrix}$$

Infatti

$$6^{-10}A^{10} \simeq \begin{pmatrix} 0.19 & 0.36 & 0.18 & 0.36 \\ 0.24 & 0.43 & 0.12 & 0.27 \\ 0.09 & 0.18 & 0.09 & 0.18 \\ 0.12 & 0.30 & 0.25 & 0.46 \end{pmatrix}$$

$$6^{-20}A^{20} \simeq \begin{pmatrix} 0.18 & 0.36 & 0.18 & 0.36 \\ 0.19 & 0.37 & 0.17 & 0.35 \\ 0.09 & 0.18 & 0.09 & 0.18 \\ 0.17 & 0.35 & 0.19 & 0.38 \end{pmatrix}$$

Se x ed y sono vettori non negativi e non nulli, allora si avrà

$$\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n y)^{(i)}} = \frac{x^{(1)} + 2x^{(2)} + x^{(3)} + 2x^{(4)}}{y^{(1)} + 2y^{(2)} + y^{(3)} + 2y^{(4)}} \text{ indipendentemente da } i.$$

Nota 17.15. A sia primitiva e $\rho := \rho(A)$. Se nella notazione della situazione 9.1 poniamo $\lambda_1 = \rho$, il cor. 9.10 diventa

$$A^n = \rho^n H_{10}(A) + \sum_{k=2}^s \sum_{j=0}^{m_k-1} n_{[j]} \lambda_k^{n-j} H_{kj}(A)$$

Siccome $\rho > |\lambda_k|$ per $k = 2, \dots, s$ troviamo che $\lim_{n \rightarrow \infty} \rho^{-n} A^n = H_{10}(A)$.

Esempio 17.16. Sia A la matrice dell'esempio 17.14. Allora

$$H_{10} = H[6 : (1), 5 : (0), \lambda : (0), \mu : (0)].$$

Calcolando H_{10} (in questo caso un polinomio di interpolazione di Lagrange) troviamo $H_{10} \simeq 0.09091x^3 - 0.0909x^2 + 2.72727x - 2.2727$ e

$$H(A) \simeq \begin{pmatrix} 0.18 & 0.36 & 0.18 & 0.36 \\ 0.18 & 0.36 & 0.18 & 0.36 \\ 0.09 & 0.18 & 0.09 & 0.18 \\ 0.18 & 0.36 & 0.18 & 0.36 \end{pmatrix}$$

in accordo con il risultato dell'esempio 17.14.

Il metodo del teorema 17.5 è comunque più semplice e numericamente più stabile.

Osservazione 17.17. Nella teoria delle popolazioni e nelle ipotesi del teorema 17.5 il vettore stocastico $\frac{v}{\|v\|_1}$ (il vettore di Perron di A) è detto anche *vettore strutturale* della popolazione, il vettore stocastico $\frac{w}{\|w\|_1}$ (il vettore di Perron di A^t) si chiama il *vettore riproduttivo* della popolazione.

Definizione 17.18. A sia non negativa. La *matrice di incidenza* di A è la matrice $\langle A \rangle$ che si ottiene da A sostituendo ogni coefficiente non nullo con 1.

Se B è un'altra matrice non negativa appartenente a \mathbb{R}_q^q , scriviamo $A \approx B : \iff \langle A \rangle = \langle B \rangle$. È chiaro che in questo modo otteniamo una relazione di equivalenza sull'insieme delle matrici non negative in \mathbb{R}_q^q .

Osservazione 17.19. Le matrici $A, B, C, D \in \mathbb{R}_q^q$ siano non negative e tali che $A \approx B, C \approx D$. Allora $AC \approx BD$.

Dimostrazione. Tutti i coefficienti che appaiono in $\sum_{k=1}^q A_{(k)}^{(i)} C_{(j)}^{(k)}$ sono ≥ 0 , perciò affinché $(AC)_{(j)}^{(i)}$ sia positivo è necessario e sufficiente che esista almeno un k tale che $A_{(k)}^{(i)}$ e $C_{(j)}^{(k)}$ siano entrambi positivi. Per ipotesi ciò accade se e solo se $B_{(k)}^{(i)}$ e $D_{(j)}^{(k)}$ sono entrambi positivi.

Osservazione 17.20. A sia non negativa. Dall'oss. 17.19 vediamo che per verificare che $A^m > 0$, è sufficiente che sia positiva la matrice che si ottiene dal seguente algoritmo:

- (1) $P = \delta; X = \langle A \rangle$
- (2) se $m == 0$: termina
- (3) se m è pari: $X = \langle X^2 \rangle; m = m/2$
- (4) altrimenti: $P = \langle XP \rangle; m = m - 1$
- (5) riprendi da (2)

Corollario 17.21. A sia non negativa. Allora

- (1) A è irriducibile se e solo se $(\langle A \rangle + \delta)^{q-1} > 0$.
- (2) A è primitiva se e solo se $\langle A \rangle^{q^2-2q+2} > 0$.

Dimostrazione. Ciò segue dalle prop. 14.16 e 17.11.

Nota 17.22. Daremo adesso, per $q = 2$ e $q = 3$, elenchi completi delle matrici di incidenza $q \times q$ che sono irriducibili o primitive. Usando il cor. 17.21, questi elenchi possono essere utilizzati per verificare, per una qualsiasi matrice non negativa in \mathbb{R}_q^q , se essa è irriducibile o primitiva.

(1) Per generare tutte le matrici in $\{0, 1\}_q^q$, associamo ad ogni numero naturale n con $0 \leq n < 2^q$ il vettore v della sua rappresentazione binaria (completata a q cifre) utilizzando la funzione

```
def rapp2 (n,cifre=0):
    if n<=1: v=[n]
    else:
        v=rapp2(n//2)
        if n%2==0: v.append(0)
        else: v.append(1)
    d=cifre-len(v)
    if d>0: v=[0]*d+v
    return v
```

In Sage otteniamo poi la matrice A corrispondente a v con l'istruzione $A=matrix(q,v)$. La matrice corrispondente ad n è quindi $matrix(q, rapp2(n,q2))$ dove $q2$ è uguale a q^2 .

(2) Per determinare se una matrice A è positiva, dobbiamo prima trasformare A in una lista di numeri con

```
def lista(A):
    v=[]
    for x in A: v.extend(x)
    return v
```

Possiamo perciò usare la funzione

```
def positiva(A):
    for x in lista(A):
        if x<=0: return 0
    return 1
```

(3) Per verificare se la matrice A è irriducibile o primitiva, applichiamo il cor. 17.21:

```
def irriducibile(A):
    q=A.ncols()
    return positiva(pow(A+identity_matrix(q),q-1))

def primitiva(A):
    q=A.ncols()
    return positiva(pow(A,q*q-2*q+2))
```

(4) Otteniamo l'elenco delle matrici di incidenza primitive con

```
def primitive(q):
    q2=q*q
    return filter(primitiva,
        [matrix(q, rapp2(n,q2)) for n in range(pow(2,q2))])
```

e l'elenco delle matrici irriducibili, ma non primitive, con

```
def irriducibilinonprimitive(q):
    q2=q*q
    return filter(lambda A:
        irriducibile(A) and not primitiva(A),
        [matrix(q,rapp2(n,q2)) for n in range(pow(2,q2))])
```

Osservazione 17.23. Sage contiene una funzione `latex` che fornisce la rappresentazione in Latex per una matrice. Per generare il nostro elenco possiamo quindi usare la funzione

```
def elencolatex(u,colonne=6):
    k=1
    for A in u:
        print '$',latex(A),'$'
        if not k%colonne: print '\\smallskip'; print
        k+=1
```


$$\begin{aligned}
& \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\
& \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}
\end{aligned}$$

Solo 5 matrici sono irriducibili, ma non primitive:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Osservazione 17.26. Usando le stesse funzioni, si trova che 25575 delle $2^{16} = 65536$ matrici di incidenza 4×4 sono primitive, mentre solo 121 sono irriducibili, ma non primitive.

Nota 17.27. Siccome una matrice di incidenza A è sicuramente primitiva, se esiste una matrice di incidenza primitiva $B \leq A$, possiamo limitarci ad un elenco delle matrici di incidenza primitive minimali che otteniamo con la seguente funzione:

```

def maggiorante(A,B):
    u=lista(A); v=lista(B)
    for a,b in zip(u,v):
        if a<b: return 0
    return 1

```

```

def primitiveminimali(q):
    u=[]; q2=q*q
    for A in primitive(q):
        for B in u:
            if maggiorante(A,B): break;
        else: u.append(A)
    return u

```

Troviamo così che esistono 21 matrici di incidenza primitive minimali 3×3 :

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}
 \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}
 \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}
 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

18. Un modello demografico

Nota 18.1. Nel modello di Leslie le fasce di età in cui tradizionalmente è suddivisa la popolazione possono anche essere interpretate come stadi di differenziazione o compartimenti, ad esempio compartimenti funzionali nelle cellule di un tessuto.

Definizione 18.2. La matrice di transizione che determina il modificarsi della distribuzione in fasce d'età (o stadi di vita) da uno stadio a quello successivo prende il nome di *matrice di Leslie* ed è della forma

$$\begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} & a_n \\ b_1 & 0 & \dots & 0 & 0 \\ 0 & b_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & b_{n-1} & 0 \end{pmatrix}$$

con $0 \leq b_j < 1$ per ogni j e $a_j \geq 0$ per ogni j .

Gli a_i rappresentano i *tassi di natalità*, cioè il numero medio di figli nati da un singolo individuo che si trova nella i -esima fascia d'età; i b_j , invece, rappresentano gli *indici di sopravvivenza*, cioè la frazione degli individui della j -esima fascia d'età che sopravvivono nella fascia successiva, e sono supposti costanti nel tempo.

Osservazione 18.3. Dalla tabella nel capitolo precedente si vede che una matrice di Leslie 3×3 è primitiva se e solo se

$$\langle A \rangle \geq \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ oppure } \langle A \rangle \geq \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Esempio 18.4. Consideriamo la matrice di Leslie (Salce, pagg. 123-124)

$$A = \begin{pmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{pmatrix}$$

Con Sage troviamo che $\rho(A) = 3/2 = 1.5$. Calcolando gli autovettori di A e A^t rispetto all'autovalore 1.5 vediamo che nei teoremi 17.5 e 17.12 possiamo porre

$$v = \begin{pmatrix} 18 \\ 6 \\ 1 \end{pmatrix}, w = \frac{1}{38} \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

Per il punto (2) del teorema 17.12 ciò significa che la popolazione tende a stabilizzarsi nelle tre fasce d'età secondo il rapporto $18 : 6 : 1$.

Per due vettori non negativi e non nulli x ed y dal punto (3) del teorema 17.12 segue che

$$\lim_{n \rightarrow \infty} \frac{(A^n x)^{(i)}}{(A^n y)^{(i)}} = \frac{x^1 + 3x^2 + 2x^3}{y^1 + 3y^2 + 2y^3} \quad \text{indipendentemente da } i.$$

Bibliografia

- E. Beckenbach/R. Bellman:** An introduction to inequalities. MAA 1961.
- P. Brémaud:** Markov chains. Springer 1999.
- M. Brin/G. Stuck:** Introduction to dynamical systems. Cambridge UP 2002.
- A. Castaldi:** Catene di Markov finite e matrici stocastiche.
Tesi, Ferrara 1995.
- L. Collatz:** Funktionalanalysis und numerische Mathematik. Springer 1968.
- P. Cull/M. Flahive/R. Robson:** Difference equations. Springer 2005.
- J. Elstrodt:** Maß- und Integrationstheorie. Springer 2005.
- K. Engel/R. Nagel:** A short course on operator semigroups. Springer 2006.
- L. Fahrmeir/H. Kaufmann/F. Ost:** Stochastische Prozesse. Hanser 1981.
- F. Ferschl:** Markovketten. Springer 1970.
- G. Frobenius:** Über Matrizen aus nicht negativen Elementen.
Sber. Preu. Akad. Wiss. Berlin (1912), 456-477.
- F. Gantmacher:** Matrizentheorie. Springer 1986.
- G. Gantesio:** Modelli matematici per la crescita tumorale.
Tesi LT, Ferrara 2007.
- G. Golub/C. Van Loan:** Matrix computations. Johns Hopkins UP 1987.
- G. Greiner:** Zur Perron-Frobenius-Theorie stark stetiger Halbgruppen.
Math. Zeitschr. 177 (1981), 401-423.
- O. Häggström:** Finite Markov chains and algorithmic applications.
Cambridge UP 2007.
- B. Huppert:** Angewandte lineare Algebra. De Gruyter 1990.
- M. Iosifescu/C. Kraaikamp:** Metrical theory of continued fractions.
Kluwer 2002.
- M. Iosifescu/P. Tautu:** Stochastic processes and applications in biology
and medicine I. Springer 1973.
- V. Keicher/R. Nagel:** Positive semigroups behave asymptotically as rotation groups. Positivity 12 (2008), 93-103.
- J. Kemeny/J. Snell:** Finite Markov chains. Springer 1976.
- N. Keyfitz/H. Caswell:** Applied mathematical demography. Springer 2005.
- C. Leardini:** Reti di Petri e analisi formale dei concetti.
Tesi, Ferrara 2009.
- C. Li/H. Schneider:** Applications of Perron-Frobenius theory to population dynamics. J. Math. Biol. 44 (2002), 450-462.
- D. Lind/B. Marcus:** An introduction to symbolic dynamics and coding.
Cambridge UP 1995.
- D. Logofet/I. Belova:** Nonnegative matrices as a tool to model population dynamics - classical models and contemporary expansions.
J. Math. Sci. 155/6 (2008), 894-907.
- M. Lothaire:** Applied combinatorics on words. Cambridge UP 2005.
- H. Lotz:** Über das Spektrum positiver Operatoren.
Math. Zeitschr. 108 (1968), 15-32.
- C. Lugli:** Equazioni alle differenze e problemi inversi in sistemi dinamici finiti. Tesi LT, Ferrara 2009.
- R. Nagel (ed.):** One-parameter semigroups of positive operators.
Springer LN Math. 1184 (1986).
- O. Perron:** Zur Theorie der Matrizen. Math. Ann. 64 (1907), 248-263.

- M. Queffélec:** Substitution dynamical systems - spectral analysis.
Springer LN Math. 1294 (1987).
- M. Rabin:** Probabilistic automata. Inf. Control 6 (1963), 230-245.
- L. Salce:** Lezioni sulle matrici. Zanichelli 1993.
- H. Schaefer:** Some spectral properties of positive linear operators.
Pacific J. Math. 10 (1960), 1009-1019.
- H. Schaefer:** Banach lattices and positive operators. Springer 1974.
- E. Seneta:** Non-negative matrices and Markov chains. Springer 1981.
- H. Wielandt:** Unzerlegbare, nicht negative Matrizen.
Math. Zeitschr. 52 (1950), 642-648.
- W. Woess:** Catene di Markov e teoria del potenziale nel discreto.
Pitagora 1996.