PROCESSI STOCASTICI E ANALISI DI SERIE TEMPORALI

Laurea magistrale in Ingegneria Informatica

Università di Roma, "Tor Vergata", a.a. 2020-21

Mario Abundo

II PARTE: Processi di Markov

(versione aggiornata a Ottobre 2020)

INDICE

1. Catene di Markov a tempo discreto

Si chiama processo stocastico una famiglia $\{X(t)\}$ di v.a. definite su uno spazio di probabilità Ω, dove $t \in T \subset \mathbb{R}^+$. I processi stocastici sono modelli matematici di fenomeni aleatori che si evolvono nel tempo.

Un processo stocastico viene detto *senza memoria* se il suo stato al tempo t dipende solo dallo stato all'istante immediatamente precedente. Dunque, la *storia* del sistema non è rilevante per la conoscenza dello stato attuale.
Contrapposti ai processi senza memoria sono quelli *con memoria*, per i quali lo stato attuale del sistema dipende da tutti gli stati precedentemente assunti. Esempi sono dati dal processo di isteresi magnetica per sostanze ferromagnetiche, e dal processo di deformazione irreversibile di un materiale sottoposto ad una sollecitazione oltre il limite di elasticità.

Per ora, ci limiteremo a considerare il caso in cui T è discreto e $X(t) \in E \subset \mathbb{R}$, dove E è pure un insieme discreto. Sia dunque dato uno spazio degli stati discreto e finito $E = \{x_1, x_2, \dots, x_N\}$; siccome E può essere messo in corrispondenza biunivoca con l'insieme $\{1, 2, \dots, N\}$, supporremo per semplicità che $E = \{1, 2, \dots, N\}$.

Definizione 1.1 (Catena di Markov a tempo discreto con spazio degli stati discreto finito) *Si chiama Catena di Markov (CM) discreta con spazio degli stati discreto $E = \{1, 2, \dots, N\}$ una famiglia di v.a. $\{X_k\}$ indiciate dal tempo discreto $k \in \mathbb{N}$, per cui vale la seguente relazione per le probabilità condizionali:*

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1} \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) = p_{ij}(n) \quad (1.1)$$

per $i_0, i_1, \dots, i_{n-1}, i, j \in E$.
La quantità $p_{ij}(n)$ $(i, j \in E)$ si chiama probabilità di transizione della CM dallo stato i allo stato j.

Una tale CM si dice temporalmente *omogenea* se $\forall i, j \in E$ risulta $p_{ij}(n) = p_{ij}$, ovvero se le probabilità di transizione non dipendono da n.
Nel seguito supporremo che $\{X_k\}$ sia una CM discreta con spazio degli stati finito, temporalmente omogenea.

Ad esempio, per $n = 1$ si ha:

$$P(X_2 = i_2 | X_1 = i_1, X_0 = i_0) = \frac{P(X_2 = i_2, X_1 = i_1, X_0 = i_0)}{P(X_1 = i_1, X_0 = i_0)} =$$

$$= \frac{P(X_2 = i_2, X_1 = i_1, X_0 = i_0)}{P(X_1 = i_1 | X_0 = i_0) P(X_0 = i_0)} \quad (1.2)$$

da cui segue che

$$P(X_2 = i_2, X_1 = i_1, X_0 = i_0) =$$

MA: appunti relativi alla seconda parte del corso: *Complementi di Probabilità e Statistica*, Laurea magistrale in Ingegneria Informatica, Università di Roma "Tor Vergata", versione aggiornata, a.a. 2014-15

$$= P(X_2 = i_2|X_1 = i_1, X_0 = i_0)P(X_1 = i_1|X_0 = i_0)P(X_0 = i_0) \tag{1.3}$$

In generale:

$$P(X_{n+1} = i_{n+1}, X_n = i_n, \dots, X_0 = i_0) =$$
$$= P(X_{n+1} = i_{n+1}|X_n = i_n, \dots, X_0 = i_0)P(X_n = i_n|X_{n-1} = i_{n-1}, \dots, X_0 = i_0)\cdot$$
$$\cdots P(X_1 = i_1|X_0 = i_o)P(X_0 = i_0) \tag{1.4}$$

Se $\{X_k\}$ è una CM omogenea, per la poprietà di Markov (1.1), la relazione (1.4) diventa:

$$P(X_{n+1} = i_{n+1}, X_n = i_n, \dots, X_0 = i_0) = P(X_0 = i_0)p_{i_0 i_1}p_{i_1 i_2}\cdots p_{i_n i_{n+1}} \tag{1.5}$$

La matrice $P = (p_{ij})$ $(i, j \in E)$ è detta matrice delle probabilità di transizione ad un passo; $p_{ij} = P\{i \to j\}$ rappresenta la probabilità di passare dallo stato $i \in E$ allo stato $j \in E$ in un passo.

Definizione 1.2 (matrice stocastica)
Una matrice $A = (a_{ij})$, $i, j = 1, 2, \dots, N$, *si dice stocastica se* $a_{ij} \geq 0$ *per ogni* $i, j = 1, 2, \dots, N$ *e risulta* $\sum_{j=1}^{N} a_{ij} = 1$.

La matrice P delle probabilità di transizione di una CM omogenea, definita da $p_{ij} = P(X_{n+1} = j|X_n = i)$, è una *matrice stocastica.* Infatti:

$$\sum_{j=1}^{N} p_{ij} = P\left(\bigcup_{j=1}^{N}\{X_{n+1} = j|X_n = i\}\right) = P(X_{n+1} \in E|X_n = i) = 1 \tag{1.6}$$

Ad ogni CM omogenea resta associata la matrice quadrata $P = (p_{ij})$, $i, j = 1, 2, \dots, N$. Viceversa, data una matrice stocastica P, esiste sempre una CM omogenea che abbia P come matrice delle probabilità di transizione.

Esercizio 1.1 Mostrare che se P e Q sono matrici stocastiche $N \times N$, allora $P \cdot Q$ è anch'essa una matrice stocastica, dove $\cdot$ denota il prodotto riga per colonna tra matrici.

Soluzione:

$$(PQ)_{ij} = \sum_{h=1}^{N} p_{ih}q_{hj} \tag{1.7}$$

Dunque

$$\sum_{j=1}^{N}(PQ)_{ij} = \sum_{j=1}^{N}\sum_{h=1}^{N} p_{ih}q_{hj} = \sum_{h=1}^{N} p_{ih}\sum_{j=1}^{N} q_{hj}$$

Siccome Q è stocastica, $\sum_{j=1}^{N} q_{hj} = 1$, e quindi $\sum_{j=1}^{N}(PQ)_{ij} = \sum_{h=1}^{N} p_{ih} = 1$, per il fatto che anche P è stocastica.
Conseguenza: se P è una matrice stocastica, anche il prodotto $P^n = P \cdot P \cdots P$ (n volte) è una matrice stocastica.

Esempio 1.1 (Successione di prove indipendenti, schema di Bernoulli)
Sia $E = \{1, 2\}$ lo spazio degli stati. Poniamo $p_{12} = p_{22} = p \in (0, 1)$ e $p_{11} = p_{21} = q = (1 - p)$. Otteniamo una CM omogenea con matrice delle probabilità di transizione:

$$P = \begin{pmatrix} q & p \\ q & p \end{pmatrix}$$

Si osservi che le righe di P sono uguali.

Esempio 1.2 Sia X_k lo stato di un' apparecchiatura al tempo k. Lo spazio degli stati è $E = \{0, 1\}$ ($0 =$ guasto, $1 =$ funzionante). Supponiamo che dallo stato 0 si possa transire in 1 con probabilità $1 - p \in (0, 1)$ o rimanere in 0 con probabilità p; analogamente, dallo stato 1 si può rimanere in 1 con probabilità $q \in (0, 1)$, oppure andare nello stato 0 con probabilità $1 - q$. X_k è una CM omogenea con matrice delle probabilità di transizione:

$$P = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$$

Si può anche descrivere la CM tramite il grafo di Fig. 1.

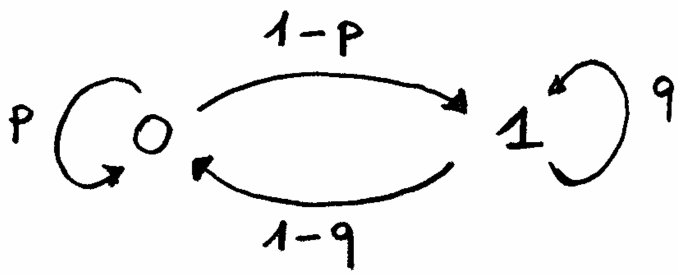

Fig. 1: Grafo di transizione della CM dell'Esempio 1.2

Esempio 1.3 (Stato energetico di un elettrone)
Sia X_k lo stato energetico di un elettrone al tempo k; nel modello di Bohr dell' atomo di idrogeno, $P(X_{k+1} = j | X_k = i) = p_{ij}$ dipende solo dalla variazione di energia $\Delta_{ij} = E_j - E_i$, durante il passaggio da un' orbita di energia E_i ad una di energia E_j, avendosi $p_{ij} = K \exp(-K\Delta_{ij})$.

Esempio 1.4 (Automa stocastico)
Un automa finito (E, A, f) può leggere una sequenza di lettere di un alfabeto A scritte su un nastro infinito. Esso può stare in uno degli stati dell'insieme finito E, e la sua evoluzione è descritta da una funzione $f : E \times A \longrightarrow E$ nel modo seguente: quando l'automa è nello

stato $i \in E$ e legge la lettera $a \in A$, esso passa dallo stato i nello stato $j = f(i,a)$ e quindi legge sul nastro la successiva lettera alla destra.
L'automa può essere rappresentato dal suo grafo di transizione G avente per nodi gli stati di E. Esiste un cammino orientato dal nodo (stato) i al nodo j se e solo se esiste $a \in A$ tale che $j = f(i,a)$, e a questo cammino viene data l'etichetta (label) a. Se $j = f(i,a_1) = f(i,a_2)$ con $a_1 \neq a_2$, vi sono due cammini orientati da i a j con labels a_1 e a_2 o, più brevemente, un cammino con label (a_1, a_2). Più in generale, un dato cammino orientato può avere labels multiple di qualunque ordine.

Ad esempio, consideriamo l'automa con alfabeto $A = \{0,1\}$ corrispondente al grafo di transizione della figura 2 e supponiamo che l'automa parta dallo stato iniziale 0.

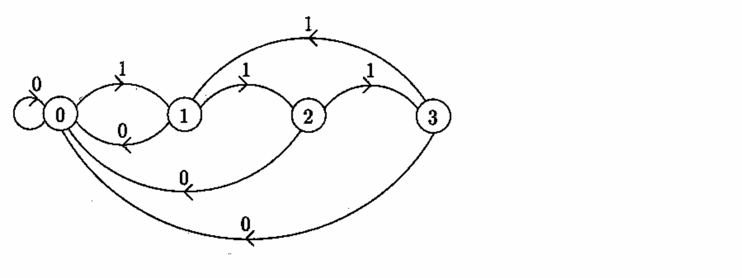

Fig. 2: Grafo di transizione dell'automa dell' Esempio 1.4

Quando l'automa legge la seguente sequenza di 0 e 1 da sinistra a destra:

$$1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 1\ 0$$

esso passa successivamete attraverso gli stati (incluso lo stato iniziale 0):

$$0\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 2\ 3\ 0\ 1\ 0$$

Infatti, se indichiamo con X_k lo stato dell'automa al tempo k, in base al grafo di figura 2 e alla stringa letta, abbiamo:

$X_0 = 0$ e, siccome il primo carattere letto è 1, il grafo impone di passare nello stato 1, cioé $X_1 = f(0,1) = 1$;

$X_1 = 1$, quindi, siccome il secondo carattere letto è 0, il grafo impone di passare nello stato 0, ovvero $X_2 = f(1,0) = 0$;

seguitando, si ottiene $X_3 = f(0,0) = 0$, $X_4 = f(0,1) = 1$, $X_5 = f(1,1) = 2$, $X_6 = f(2,1) = 3$, $X_7 = f(3,1) = 1$, $X_8 = f(1,0) = 0$, etc. Pertanto, l'automa visita successivamente gli stati (partendo da 0:)

$$0\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 2\ 3\ 0\ 1\ 0$$

Riscriviamo ora la sequenza di stati sotto la stringa di lettere:

$$1\ 0\ 0\ \underline{1\ 1\ 1}\ 1\ 0\ 0\ \underline{1\ 1\ 1}\ \underline{1\ 1\ 1}\ 0\ 1\ 0$$

$$0\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 0\ 0\ 1\ 2\ 3\ 1\ 2\ 3\ 0\ 1\ 0$$

possiamo vedere che l'automa è nello stato 3 dopo che ha letto tre volte consecutivamente il carattere 1. Tale automa è allora in grado di riconoscere e contare i blocchi di 1 (contrassegnati dalla sottolineatura), anche se esso non sa tener conto dei blocchi di 1 che si sovrappongono.
Se la sequenza di lettere lette dall'automa è $\{Z_n\}_{n\geq 1}$, la sequenza di stati $\{X_n\}_{n\geq 0}$ è data dall' equazione ricorsiva $X_{n+1} = f(X_n, Z_n)$. Supponiamo ora che le $Z_n \in \{0,1\}$ siano variabili aleatorie, indipendenti tra loro e dallo stato iniziale X_0 e identicamente distribuite, in modo che $P(Z_n = 0) = P(Z_n = 1) = \frac{1}{2}$; allora si ottiene il grafo di figura 3, che rappresenta ua CM omogenea X_n con spazio degli stati $E = \{0,1,2,3\}$ e matrice delle probabilità di transizione:

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

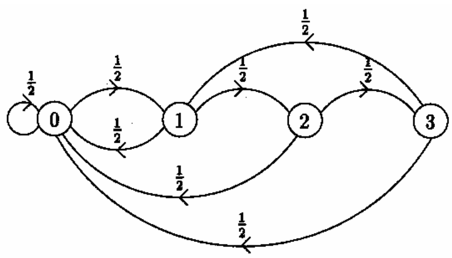

Fig. 3: Grafo di transizione della CM associata all'automa dell'Esempio 1.4

Esempio 1.5 (Random walk con barriere assorbenti)
Nella passeggiata a caso sul reticolo $E = \{0, 1, \ldots, a\}$ con barriere 0 ed a assorbenti, si considera una particella che si può spostare di una quantità unitaria, verso destra con probabilità $p \in (0,1)$, o verso sinistra con probabilità $q = 1-p$, ma non può mai rimanere ferma in una posizione (stato); gli stati 0 e a (barriere) vengono detti *assorbenti* poiché si richiede che $p\{0 \to 0\} = p_{00} = 1$ e $P\{0 \to j\} = p_{0j} = 0\ \forall j \neq 0$; in modo analogo $p\{a \to a\} = p_{aa} = 1$ e $P\{a \to j\} = p_{aj} = 0\ \forall j \neq a$. Dunque, una volta finita nello stato 0 oppure a, la particella non può più allontanarvisi, ma resta lì per sempre. In maniera colorita, gli stati assorbenti 0 e a si comportano come "buchi neri". La CM omogenea

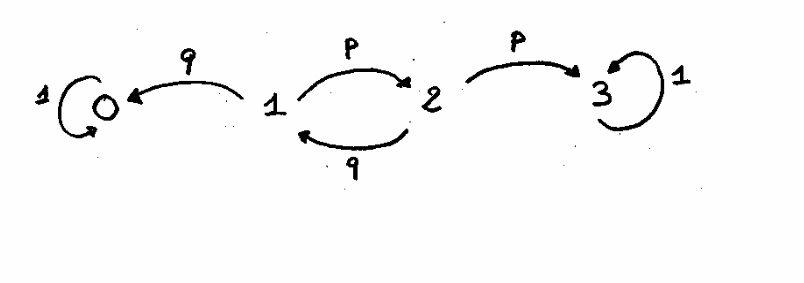

Fig. 4: Grafo di transizione della CM dell'Esempio 1.4

corrispondente ha per spazio degli stati $E = \{0, 1, \ldots, a\}$; ad esempio, per $a = 3$, la matrice delle probabilità di transizione è:

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

ed il relativo grafo è illustrato in Fig. 4.

Esempio 1.6 (Random walk con barriere riflettenti)
Nella passeggiata a caso sul reticolo $E = \{0, 1, \ldots, a\}$ con barriere 0 ed a riflettenti, si considera una particella che si può muovere di una quantità unitaria, verso destra con probabilità $p \in (0, 1)$, o verso sinistra con probabilità $q = 1 - p$, ma non può mai rimanere ferma in uno stato; Gli stati 0 e a vengono detti *riflettenti* poiché $P\{0 \to 0\} = p_{00} = 0$ e $P\{0 \to 1\} = p_{01} = 1$; in modo analogo $P\{a \to a\} = p_{aa} = 0$ e $P\{a \to a-1\} = p_{a\ a-1} = 1$. Dunque, una volta finita nello stato 0 oppure a, la particella viene riflessa all'interno di E, spostandosi un passo a destra se parte da 0, un passo a sinistra se parte da a. La CM omogenea corrispondente ha per spazio degli stati $E = \{0, 1, \ldots, a\}$; ad esempio, per $a = 3$, la matrice delle probabilità di transizione è:

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

ed il relativo grafo è illustrato in Fig. 5.

Esempio 1.7 (Random walk con barriere parzialmente riflettenti)
Si tratta di una CM simile a quella dell Esempio 1.5. La differenza è che ora le barriere 0 e a sono in parte assorbenti, in parte riflettenti. $p\{0 \to 0\} = p_{00} = r \in (0, 1)$ e $P\{0 \to 1\} = p_{01} = 1 - r$; in modo analogo $p\{a \to a\} = p_{aa} = r$ e $P\{a \to a-1\} = p_{a\ a-1} = 1 - r$.

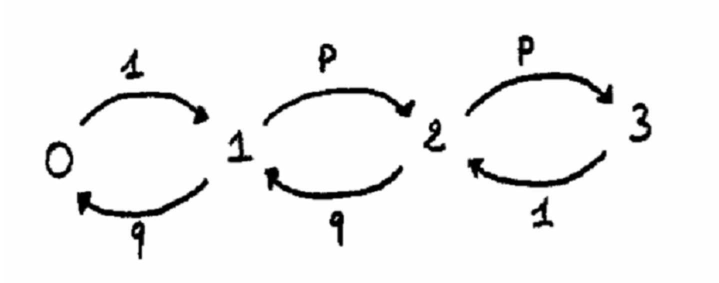

Fig. 5: Grafo di transizione della CM dell'Esempio 1.6

Dunque, una volta finita nello stato 0 oppure a, la particella viene riflessa all'interno di E, con probabilità $1-r$, oppure rimane ferma nello stato occupato con probabilità r. La CM omogenea corrispondente ha per spazio degli stati $E = \{0, 1, \ldots, a\}$; ad esempio, per $a = 3$, la matrice delle probabilità di transizione è:

$$P = \begin{pmatrix} r & 1-r & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 1-r & r \end{pmatrix}$$

ed il relativo grafo è illustrato in Fig. 6.

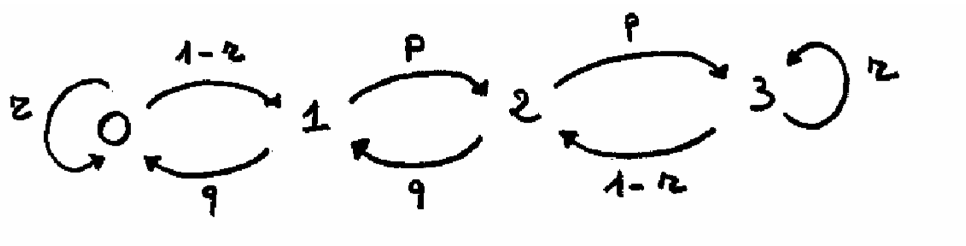

Fig. 6: Grafo di transizione della CM dell'Esempio 1.7

Esempio 1.8 (Il problema della rovina del giocatore)
Si consideri un gioco condotto tra due concorrenti. Ad ogni partita, un giocatore vince 1 euro con probabilità $p \in (0, 1)$ o perde 1 euro con probabilità $q = 1 - p$; non è prevista parità. Sia i il capitale iniziale (in euro) di uno dei due giocatori (ad esempio il primo) e $a - i$ il capitale (in euro) del suo avversario. Si suppone che il capitale complessivo dei due concorrenti (in euro) durante tutte le partite sia costante ed uguale ad a, dove a è un intero positivo. Il gioco termina quando il capitale di un giocatore si riduce a 0 o ad a, cioé uno di essi vince tutto, e l'altro è rovinato. Si può modellizzare il gioco mediante un random walk con barriere assorbenti, 0 ed a : ad ogni partita, il capitale, diciamo del primo giocatore, è rappresentato dalla posizione di una particella che si muove sul reticolo $0, 1, \ldots, a$ dalla

posizione iniziale $X_0 = i$ $(0 \leq i \leq a)$, verso destra con probabilità $p \in (0,1)$ e verso sinistra con probabilità $q = 1-p$. La posizione della particella dopo n passi è il capitale del primo giocatore dopo l'ennesima partita. Il gioco termina quando la particella raggiunge per la prima volta lo stato 0 o lo stato a. Sia q_i la probabilità che il primo giocatore perda il suo capitale iniziale i, ovvero la sua probabilità di rovina, cioé la probabilità che la particella, partendo dalla posizione iniziale i, sia definitivamente assorbita in 0. La CM omogenea associata a questo random walk con barriere assorbenti, con insieme degli stati $E = \{0, 1, \ldots, a\}$, ha matrice delle probabilità di transizione:

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

Deve aversi:

$$\begin{cases} q_i = p \cdot q_{i+1} + q \cdot q_{i-1} \\ q_0 = 1, \ q_a = 0 \end{cases} \tag{1.8}$$

Infatti, se la particella parte dallo stato i al tempo 0, l'evento che essa sia assorbita in 0 è l'unione disgiunta dei due eventi: $A_1 = \{$la particella fa un passo avanti raggiungendo la posizione $i+1$, e da qui viene assorbita in $0\}$; $A_2 = \{$la particella fa un passo indietro raggiungendo la posizione $i-1$, e da qui viene assorbita in $0\}$. Risulta $P(A_1) = p \ q_{i+1}$, $P(A_2) = q \ q_{i-1}$. Inoltre, se la particella parte da $i = 0$, essa viene assorbita in 0 con probabilità 1; se essa parte da a, viene assorbita in 0 con probabilità 0, cioè viene assorbita in a. La (1.8) è un' equazione alle differenze, lineare del II ordine, omogenea, a coefficienti costanti e con condizioni al bordo; essa si risolve in maniera simile alle equazioni differenziali lineari, omogenee, a coefficienti costanti.

1. Sia $p \neq q$. Si cerca una soluzione della forma $q_i = \alpha^i$; sostituendo in (1.8) si trova:

$$\alpha^i = p\alpha^{i+1} + q\alpha^{i-1}$$

e, mettendo in evidenza α^{i-1}, si ottiene:

$$\alpha^{i-1}\left(p\alpha^2 - \alpha + q\right) = 0$$

Pertanto, si ottiene l'equazione di II grado $p\alpha^2 - \alpha + q = 0$ con discriminante $1 - 4pq > 0$; risolvendola, si trovano le due soluzioni indipendenti $\alpha_1 = 1$ e $\alpha_2 = (1-p)/p = q/p$. Si dice che $q_i = 1$ e $q_i = \left(\frac{q}{p}\right)^i$ sono le soluzioni fondamentali di (1.8). La soluzione generale si ottiene come combinazione lineare delle due soluzioni fondamentali:

$$q_i = A + B\left(\frac{q}{p}\right)^i \tag{1.9}$$

dove A e B sono due costanti positive da determinare. Le condizioni al bordo impongono che

$$\begin{cases} A+B=1 \\ A+B(\frac{q}{p})^a=0 \end{cases}$$

Risolvendo tale sistema, si trova:

$$A=\frac{\left(\frac{q}{p}\right)^a}{\left(\frac{q}{p}\right)^a-1}\ ,\ B=-\frac{1}{\left(\frac{q}{p}\right)^a-1} \tag{1.10}$$

Sostituendo in (1.9), si trova allora:

$$q_i=\frac{\left(\frac{q}{p}\right)^a-\left(\frac{q}{p}\right)^i}{\left(\frac{q}{p}\right)^a-1} \tag{1.11}$$

2. Sia $p=q=\frac{1}{2}$. Procedendo come nel caso $p\neq q$, si trova ora l'equazione di II grado (con discriminante uguale a zero): $\frac{1}{2}\alpha^2-\alpha+\frac{1}{2}$ che ha due soluzioni coincidenti $\alpha_1=\alpha_2=1$; in tal caso un sistema di soluzioni fondamentale è dato da $1^i=1$ e $i\cdot 1^i=i$. Dunque, la soluzione generale dell'equazione alle differenze è $q_i=A\cdot 1+B\cdot i$; imponendo le condizioni al bordo, si trova $A=1$ e $B=-1/a$. Pertanto, la soluzione cercata è

$$q_i=1-\frac{i}{a} \tag{1.12}$$

Lo stesso risultato si può ottenere passando al limite per $p/q\to 1$ in (1.11); usando la regola di L'Hospital *:

$$q_i=\lim_{p/q\to 1}\frac{\left(\frac{q}{p}\right)^a-\left(\frac{q}{p}\right)^i}{\left(\frac{q}{p}\right)^a-1}=1-\frac{i}{a}$$

Indichiamo con $p_i=1-q_i$ la probabilità che il giocatore vinca tutto (cioé che la particella sia definitivamente assorbita in a). La probabilità p_i di vittoria è la probabilità di rovina dell'avversario, pertanto $p_i=q_{a-i}$, da cui si ottiene $p_i+q_i=1$ e $q_i+q_{a-i}=1$.

Introduciamo ora la v.a. G = guadagno; essa assume valori $a-i$ e $-i$ con probabilità $1-q_i$ e q_i; pertanto la sua media è:

$$E(G)=(a-i)\cdot(1-q_i)+q_i\cdot(-i)=a(1-q_i)-i \tag{1.13}$$

Nel caso particolare $p=q=\frac{1}{2}$, tenendo conto della (1.12), si trova $E(G)=0$ (gioco equo).

* Posto $x=p/q$, si ha $\lim_{x\to 1}\frac{x^a-x^i}{x^a-1}=\lim_{x\to 1}\frac{ax^{a-1}-ix^{i-1}}{ax^{a-1}}=1-i\cdot\lim_{x\to 1}\frac{x^{i-1-a+1}}{a}=$ $=1-\frac{i}{a}$.

Per $a \to \infty$ si ha un gioco contro un avversario infinitamente ricco (il banco); si ottiene

$$\lim_{a\to\infty} q_i = \begin{cases} 1 & \text{se } p \le q \\ \left(\frac{q}{p}\right)^i & \text{se } p > q \end{cases} \tag{1.14}$$

Durata media del gioco
Sia τ_i il tempo (random) di durata del gioco, con la condizione che il capitale iniziale (del primo giocatore) sia i, ovvero il tempo di assorbimento in 0 oppure in a. Si può dimostrare che la durata media del gioco $D_i \doteq E(\tau_i)$ è finita e soddisfa la seguente equazione alle differenze:

$$D_i = p \cdot D_{i+1} + q \cdot D_{i-1} + 1,\ \ 0 < i < a \tag{1.15}$$

con le condizioni al bordo

$$D_0 = 0,\ \ D_a = 0$$

Procedendo in maniera analoga a quanto fatto per l'equazione (1.8), si trova che la soluzione di (1.15), per $p \neq q$ è:

$$D_i = \frac{i}{q-p} - \frac{a}{q-p}\,\frac{1-(q/p)^i}{1-(q/p)^a} \tag{1.16}$$

Se $p = q = 1/2$, si trova:

$$D_i = i(a-i) \tag{1.17}$$

Esercizio 1.2 Si consideri la seguente generalizzazione della CM relativa alla rovina del giocatore. Si tratta di una passeggiata aleatoria sul reticolo $\{0, 1, \ldots, a\}$ modificata, in cui la particella si può spostare verso sinistra con probabilità q, verso destra con probabilità p, e può rimanere ferma con probabilità r; naturalmente ora $p+q+r = 1$ (il caso di prima si ottiene per $r = 0$). Supposto che gli stati 0 e a siano assorbenti, si scriva la matrice delle probabilità di transizione relativa a questa CM e si calcoli la probabilità di assorbimento in 0, partendo dallo stato i.

1.2 Probabilità di transizione ad n passi

Sia $\{X_k\}$ una CM omogenea con spazio degli stati $E = \{1, 2, \ldots, N\}$. Vogliamo calcolare la probabilità di transizione ad n passi, ovvero

$$p_{ij}^{(n)} \doteq P(X_{k+n} = j | X_k = i) \tag{1.18}$$

che, per l'omogeneità, è anche uguale a $P(X_n = j|X_0 = i)$. Si ha:

$$P(X_n = j|X_0 = i) = \frac{P(X_n = j, X_0 = i)}{P(X_0 = i)} = \tag{1.19}$$

$$= \sum_h \left(\frac{P(X_n = j, X_{n-1} = h, X_0 = i)}{P(X_{n-1} = h, X_0 = i)} \cdot \frac{P(X_{n-1} = h, X_0 = i)}{P(X_0 = i)} \right) =$$

$$= \sum_h P(X_n = j|X_{n-1} = h, X_0 = i) P(X_{n-1} = h|X_0 = i) =$$

(per la proprietà di Markovianetà)

$$= \sum_h P(X_n = j|X_{n-1} = h) P(X_{n-1} = h|X_0 = i) = \sum_h p_{hj}\, p_{ih}^{(n-1)}$$

Abbiamo pertanto ottenuto la formula di *Chapmann-Kolmogorov*:

$$p_{ij}^{(n)} = \sum_{h \in E} p_{hj} p_{ih}^{(n-1)} \tag{1.20}$$

Se indichiamo con $P^{(n)}$ la matrice delle probabilità di transizione ad n passi, cioé la matrice di elementi $p_{ij}^{(n)}$, dalla formula di Chapmann-Kolmogorov si ottiene $P^{(n)} = P^{(n-1)}P$, in particolare $P^{(2)} = P^{(1)}P = P^2$. Procedendo per induzione, si trova allora $P^{(n)} = P^n$, dove P^n è la matrice ottenuta moltiplicando (riga per colonna) n volte la matrice P per se stessa.

Ricordando l'Esercizio 1.1, si ottiene che la matrice $P^{(n)} = P^n$ delle probabilità di transizione ad n passi è ancora una matrice stocastica.

1.3 Legge di $\mathbf{X}_n$ e distribuzione invariante

Sia $\{X_k\}$ una CM omogenea con spazio degli stati $E = \{1, 2, \ldots, N\}$ e matrice delle probabilità di transizione P. La distribuzione della CM al tempo n è individuata dai numeri $w_i = P(X_n = i)$, $i = 1, 2, \ldots, N$. Il vettore $w = (w_1, w_2, \ldots, w_N)^T$ è un vettore riga con componenti $w_i \geq 0$ e tale che $\sum_{i=1}^{N} w_i = 1$. Se X_0 ha legge (o distribuzione) $u = (u_1, u_2, \ldots, u_N)^T = (P(X_0 = 1), P(X_0 = 2), \ldots, P(X_0 = N))^T$, si ha:

$$w_i = P(X_n = i) = \sum_{h \in E} P(X_n = i | X_0 = h) P(X_0 = h) = \sum_{h \in E} p_{hi}^{(n)} u_h \tag{1.21}$$

Pertanto, risulta:

$$w = uP^n \tag{1.22}$$

(notare che P^n è moltiplicata **a sinistra** per u).
Nel caso particolare in cui la CM parte quasi certamente dallo stato k, ovvero

$$u_h = \delta_{hk} = \begin{cases} 1 & \text{se } h = k \\ 0 & \text{se } h \neq k \end{cases} \tag{1.23}$$

la relazione di sopra diventa:

$$w_i = P(X_n = i) = p_{ki}^{(n)} \tag{1.24}$$

Definizione 1.3 (Distribuzione invariante)
Una distribuzione di probabilità v su E si dice invariante per la CM se risulta $v = vP$.
Se la distribuzione iniziale u è invariante (ovvero $u = uP$) da (1.22) si ottiene:

$$w = uP^n = (uP)P^{n-1} = uP^{n-1} = (uP)P^{n-2} = uP^{n-2} = \ldots = u \tag{1.25}$$

ovvero $w_i = P(X_n = i) = u_i = P(X_0 = i)$, $i \in E$ e dunque X_n ha la stessa distribuzione, per ogni n.

1.4 Classificazione degli stati di una CM

Sia $\{X_k\}$ una CM omogenea con spazio degli stati $E = \{1, 2, \ldots, N\}$ e matrice delle probabilità di transizione P.
1) Si dice che lo stato $i \in E$ *comunica* con $j \in E$, e si scrive $i \rightharpoonup j$, se esiste un intero $n > 0$ tale che $p_{ij}^{(n)} > 0$.
2) Un s.i. $C \subset E$ si dice *classe chiusa* se gli stati di C non comunicano con quelli del complementare di C.
3) Una classe chiusa si dice *irriducibile* se tutti i suoi stati comunicano tra loro.
4) Se uno stato costituisce da solo una classe irriducibile, esso si dice *assorbente*.
5) Una CM si dice *irriducibile* se tutti i suoi stati comunicano, cioé E è formata da un'unica classe irriducibile.
6) Dato uno stato $j \in E$, sia $\tau_j = \min\{n > 0 : X_n = j\}$ il primo istante (random) in cui la CM "visita" lo stato j; τ_j può essere anche infinito, se lo stato j non viene mai visitato. Si osservi che τ_j è una v.a. Poniamo

$$\rho_{ij} = P(\tau_j < \infty | X_0 = i)$$

cioé la probabilità che la CM, partendo dallo stato i al tempo zero, visiti prima o poi lo stato j.
7) Uno stato $i \in E$ si dice *transitorio* se $\rho_{ii} < 1$, *ricorrente* se $\rho_{ii} = 1$. Dunque, se i è uno stato transitorio, l'evento che la CM parta da i e non vi faccia più ritorno, ha probabilità positiva (in altre parole, se lo stato i è transitorio, esiste un percorso della CM che parte da i e non può più ritornare indietro su i). Invece, se la CM parte da uno stato ricorrente, essa può farvi ritorno infinite volte, con probabilità 1. Ovviamente, uno stato assorbente h è ricorrente, poiché $P(\tau_h = 1 | X_0 = h) = 1$, visto che la CM non può mai allontanarsi da h (cioé $p_{hh} = 1, \; p_{hk} = 0, \; k \neq h$).

Da quanto detto sopra, segue che lo stato i è ricorrente se e solo se, partendo dallo stato i, il numero medio di volte che la CM visita lo stato i è infinito. Allora, posto

$$I_n = \begin{cases} 1 & \text{se } X_n = i \\ 0 & \text{se } X_n \neq i \end{cases},$$

si ha che $\sum_{n=0}^{\infty} I_n$ rappresenta il numero di volte che la CM visita lo stato i. Dunque:

$$E\left[\sum_{n=0}^{\infty} I_n \Big| X_0 = i\right] = \sum_{n=0}^{\infty} E[I_n | X_0 = i]$$

$$= \sum_{n=0}^{\infty} P(X_n = i | X_0 = i) = \sum_{n=0}^{\infty} p_{ii}^{(n)}.$$

Abbiamo pertanto ottenuto:

lo stato i è:

ricorrente se $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$,

transitorio se $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$.

Osservazione 1.1
Se $i \rightharpoonup j$ e $j \rightharpoonup h$, allora $i \rightharpoonup h$.
Infatti, esistono $n, m > 0$ tali che $p_{ij}^{(n)} > 0$ e $p_{jh}^{(m)} > 0$. Allora

$$p_{ih}^{(n+m)} = \sum_{k \in E} p_{ik}^{(n)} p_{kh}^{(m)} \geq p_{ij}^{(n)} \cdot p_{jh}^{(m)} > 0$$

e dunque $i \rightharpoonup h$.

Esercizio 1.3 Provare che, se lo stato i comunica con lo stato transitorio j, allora anche i è transitorio.

Esercizio 1.4 Provare che, se lo stato ricorrente i comunica con j e j comunica con i, allora anche j è ricorrente.

Esempio 1.8 Consideriamo la CM discreta omogenea, con spazio degli stati $E = \{1, 2, \ldots, 10\}$ e matrice delle probabilità di transizione

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & * & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 & * & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & * & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.26}$$

dove $*$ indica una quantità positiva e minore di 1; ricordiamo che P è una matrice stocastica, quindi la somma degli elementi su ciascuna riga deve valere 1. Dall'esame di P si trova:
1) lo stato 1 comunica con 7 e 9; 7 comunica con 1 e 9; 9 comunica con 7 e 9. Pertanto gli stati $1, 7, 9$ comunicano tra loro e $C_1 = \{1, 7, 9\}$ è una classe irriducibile.
2) 2 comunica con 2 e con 4; 4 comunica con 2. $C_2 = \{2, 4\}$ è una classe irriducibile.
3) $3 \rightharpoonup 5$; $5 \rightharpoonup 2$; dunque da 3 si può andare in 2 in due passi, ma 2 non comunica nè con 3 nè con $5:3$ e 5 sono stati transitori.
4) $6 \rightharpoonup 1$ ma 1 non comunica con 6: lo stato 6 è transitorio.
5) $8 \rightharpoonup 3$, $3 \rightharpoonup 5$, $5 \rightharpoonup 2$; quindi 8 comunica con 2 che non comunica con $8:$ lo stato 8 è transitorio.
6) 10 comunica solo con se stesso: esso è uno stato assorbente.

L'insieme degli stati E si decompone in 3 classi irriducibili: $C_1 = \{1, 7, 9\}, C_2 = \{2, 4\}$, $C_3 = \{10\}$ e $T = \{3, 5, 6, 8\}$. Quest'ultima classe contiene gli stati transitori, gli altri stati sono ricorrenti. In generale, per una CM vale la decomposizione in modo unico: $E = T \cup C_1 \cup C_2 \cup \ldots$, ove le C_i sono classi irrudicibili di stati comunicanti tra loro e T è l'insieme degli stati transitori.

Osservazione. Con riferimento all' Esempio precedente, potrebbe sembrare che, per esempio, lo stato 1 è transitorio, poiché se la CM effettua una transizione da 1 a 9 e poi rimane in 9 per sempre, si viene a realizzare un percorso per il quale è impossibile ritornare nello stato 1. Ciò non implica che 1 è transitorio; infatti, per esserlo, in base alla definizione dovrebbe accadere che la probabilità di non ritorno nello stato 1 è strettamente positiva, mentre la probabilità di realizzare il percorso di cui sopra (che consiste nel soggiornare definitivamente, ovvero un numero infinito di volte, nello stato 9 è minore o uguale del $\lim_{n\to\infty}(p_{99})^n = 0$, pertanto tale probabilità è zero, ovvero $\rho_{11} = 1$ e lo stato 1 è ricorrente. Analoghi ragionamenti valgono per gli altri stati di C_1.

Esempio 1.9 (rovina del giocatore)
Nella CM della rovina del giocatore (v. Esempio 1.8) con $0 < p < 1$, gli stati 0 e a sono assorbenti, gli stati $\{1, 2, \ldots, a-1\}$ sono transitori; infatti, se $i \in \{1, 2, \ldots, a-1\}$, $i \rightharpoonup 0$, ma 0 non comunica con i.

Ritorniamo a parlare del tempo τ_j di primo passaggio della CM per lo stato j. Se j è uno stato transitorio, allora $P(\tau_j = +\infty | X_0 = j) > 0$; in tal caso si ha

$$E(\tau_j) = 1 \cdot P(\tau_j = 1) + 2 \cdot P(\tau_j = 2) + \ldots + \ldots + \infty \cdot P(\tau_j = +\infty) \tag{1.27}$$

Se $P(\tau_j = +\infty) > 0$, porremo per convenzione $E(\tau_j) = +\infty$. Se invece $P(\tau_j = +\infty) = 0$, allora τ_j è una v.a propria e $E(\tau_j)$ risulta finita se la serie

$$\sum_{k=1}^{\infty} k \cdot P(\tau_j = k) \tag{1.28}$$

è convergente.

1.5 Problemi di assorbimento

Sia C una classe chiusa, non necessariamente composta da tutti gli stati ricorrenti; se la CM raggiunge C, allora resterà per sempre in C. Ci proponiamo di calcolare la probabilità di giungere in C, partendo dallo stato i, ovvero

$$\lambda_i = P(X_n \in C \text{ per qualche } n > 0 | X_0 = i) \tag{1.29}$$

Naturalmente, se $i \in C$, allora risulta $\lambda_i = 1$; se $i \in C'$ con C' altra classe chiusa diversa da C, allora risulta $\lambda_i = 0$, poiché X_n resta sempre in C' per ogni n. Dunque, λ_i è da calcolare solo se i è uno stato transitorio che non sta in C. I numeri λ_i si chiamano le *probabilità di assorbimento* in C. Si può dimostare la seguente:

Proposizione 1.1 *Sia T l'insieme degli stati transitori della CM, che non fanno parte di C; allora per $i \in T$ le probabilità λ_i di assorbimento in C sono le soluzioni (uniche se E è finito) del sistema lineare:*

$$\lambda_i = \sum_{h \in C} p_{ih} + \sum_{j \in T} p_{ij} \lambda_j, \ i \in T \tag{1.30}$$

■

Da un punto di vista intuitivo, la (1.30) non è sorprendente, poiché la prima somma fornisce la probabilità di fare una transizione direttamente da i in C al primo passo, la seconda è la probabilità di andare al primo passo in uno stato $j \in T$ e poi essere assorbiti in C partendo da j.

Esercizio 1.5 Verificare che, per la CM della rovina del giocatore, le probabilità di assorbimento in $C = \{0\}$, partendo dallo stato transitorio $i \in \{1, 2, \ldots, a-1\}$, sono date da (cfr. con (1.11), (1.12)):

$$\lambda_i = \begin{cases} \frac{(q/p)^a - (q/p)^i}{(q/p)^a - 1} & \text{se } q \neq p \\ 1 - \frac{i}{a} & \text{se } q = p = \frac{1}{2} \end{cases} \tag{1.31}$$

(per a fissato è sufficiente verificare che le (1.31) risolvono (1.30))

Tempo medio trascorso nella classe degli stati transitori

Sia T la classe degli stati transitori della CM e siano $i, j \in T$ due stati transitori. Indichiamo con s_{ij} il numero medio di volte che la CM è nello stato j, essendo partita dallo stato i. Si può dimostrare che vale la seguente:

Proposizione 1.2 *I numeri $s_{i,j}$, $i, j \in T$ sono soluzioni dell' equazione:*

$$s_{ij} = \delta_{i,j} + \sum_{k \in T} p_{ik} s_{kj} \ , \ i, j \in T \tag{1.32}$$

Dim. Se conteggiamo separatamente il primo passo, questo dà un contributo δ_{ij}. La media della permanenza successiva è s_{ir} se la CM passa dallo stato transitorio i allo stato transitorio r, mentre è nulla se passa in uno stato ricorrente. In conclusione, vale la (1.32).

■

Sia ora S la matrice dei numeri s_{ij}, $i, j = 1, \ldots, t$ dove t indica il numero degli stati transitori; denotiamo inoltre con P_T la matrice delle probabilità di transizione ristretta all'insieme T degli stati transitori. Allora, in notazione matriciale l'equazione (1.32) si scrive:

$$S = Id + P_T S$$

dove Id è la matrice identità di ordine t. Questa equazione è equivalente a $(Id - P_T)S = Id$, ovvero:

$$S = (Id - P_T)^{-1}.$$

Dunque, le quantità $s_{ij},\ i, j \in T$, si possono ottenere, invertendo la matrice $Id - P_T$. L'unica cosa da controllare è che la matrice $Id - P_T$ sia non singolare, ovvero $det(Id - P_T) \neq 0$; ciò equivale a dire che l'equazione $(Id - P_T)X = 0$ ammette l'unica soluzione $X = 0$. Tale equazione si può scrivere $X = P_T X$; se X è soluzione, allora:

$$X = P_T X = P_T(P_T X) = (P_T)^2 X = \ldots = (P_T)^n X \to 0, \quad \text{per } n \to \infty,$$

poiché $(P_T)^n_{ij} = p^{(n)}_{ij}$ è la probabilità di andare in n passi dallo stato transitorio i ad un altro transitorio j, e quindi tende a zero, per $n \to \infty$, poiché prima o poi la CM esce dall'insieme dcegli stati transitori T, e finisce nella classe degli stati ricorrenti. Ne segue che $X = 0$, dunque $Id - P_T$ è invertibile.

Esempio 1.10 Si consideri il problema della rovina del giocatore con $p = 0.4$ e $a = 7$. Supponiamo che un giocatore parta con un capitale iniziale di 3 euro, determinare:
(a) il numero medio di volte che il giocatore ha 5 euro;
(b) il numero medio di volte che il giocatore ha 2 euro.

Soluzione. La matrice P_T, che specifica $P_{ij},\ i, j \in \{1, 2, 3, 4, 5, 6\}$ è:

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

Invertendo $Id - P_T$ (per esempio con Scilab), si ottiene:

$$S = (Id - P_T)^{-1} = \begin{pmatrix} 1.6149 & 1.0248 & 0.6314 & 0.3691 & 0.1943 & 0.0777 \\ 1.5372 & 2.5619 & 1.5784 & 0.9228 & 0.4857 & 0.1943 \\ 1.4206 & 2.3677 & 2.9990 & 1.7533 & 0.9228 & 0.3691 \\ 1.2458 & 2.0763 & 2.6299 & 2.9990 & 1.5784 & 0.6314 \\ 0.9835 & 1.6391 & 2.0763 & 2.3677 & 2.5619 & 1.0248 \\ 0.5901 & 0.9835 & 1.2458 & 1.4206 & 1.5372 & 1.6149 \end{pmatrix}$$

Quindi, $s_{35} = 0.9228, \quad s_{32} = 2.3677$.

Se $i \in T$, indichiamo ora con τ_i il tempo che la CM trascorre nell' insieme degli stati transitori T, prima di finire nella classe degli stati ricorrenti, con la condizione che $X_0 = i$

(ovvero il tempo di assorbimento nella classe degli stati ricorrenti, con la condizione che la CM è partita da $i \in T$ al tempo zero). Allora, evidentemente $\lambda_i = 1$ e quindi $\tau_i = \inf\{n > 0 : X_n \in C | X_0 = i\}$ è finito. In tal caso vale la seguente:

Proposizione 1.2' *Se lo stato di partenza i è uno stato transitorio, allora i tempi medi $E(\tau_i) \doteq \eta_i$ di assorbimento nella classe chiusa formata dagli stati ricorrenti sono le soluzioni (uniche se E è finito) del sistema:*

$$\eta_i = 1 + \sum_{r \in T} p_{ir} \eta_r, \ \ i \in T \tag{1.32'}$$

ove T è l'insieme degli stati transitori; notare che $E(\tau_i) \geq 1, \ i \in T$, mentre $E(\tau_i) = 0$ se $i \in C$.

Prova. Si ha ovviamente $\eta_i = E(\tau_i) = \sum_{j \in T} s_{ij}$. Allora, sommando sull'indice $j \in T$ in entrambi i membri dell'equazione (1.32), si ottiene:

$$E(\tau_i) = \sum_{j \in T} \delta_{ij} + \sum_{k \in T} p_{ik} \sum_{j \in T} s_{kj},$$

ovvero

$$\eta_i = 1 + \sum_{k \in T} p_{ik} \eta_k,$$

visto che $\sum_{j \in T} \delta_{ij} = 1$. La prova è conclusa. ∎

Esercizio 1.6 Verificare che, per la CM della rovina del giocatore, i tempi medi di assorbimento in $C = \{0, a\}$, partendo dallo stato transitorio $i \in \{1, 2, \ldots, a-1\}$, sono soluzioni del sistema:

$$\begin{cases} \eta_1 = 1 + p\eta_2 \\ \eta_2 = 1 + q\eta_1 + p\eta_3 \\ \ \cdots\cdots \\ . \\ . \\ . \\ \eta_{a-1} = 1 + q\eta_{a-2} \end{cases} \tag{1.33}$$

Se $p = q = \frac{1}{2}$, la soluzione del sistema di sopra è $\eta_i = i(a-i)$, come già trovato in precedenza (cfr. (1.17)). Si noti che $\eta_0 = \eta_a = 0$ e $\eta_{a/2} = \frac{a}{2}(a - \frac{a}{2}) = \frac{a^2}{4}$.

Esercizio 1.7 E' data la CM con spazio degli stati $E = \{1, 2, 3, 4\}$ e matrice delle probabilità di transizione:

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

(i) Classificare gli stati della CM.
(ii) Dopo aver verificato che $C = \{1, 2\}$ è una classe chiusa, calcolare i tempi medi di assorbimento in C, partendo dallo stato $i \in \{3, 4\}$ e trovare le probabilità di assorbimento in C, partendo da $i \in \{3, 4\}$.

Soluzione (i) Come si verifica facilmente $T = \{3, 4\}$ è l'insieme degli stati transitori mentre $C = \{1, 2\}$ è una classe chiusa costituita da stati ricorrenti.
(ii) I tempi medi di assorbimento in C, partendo da $i \in T$ sono soluzioni del sistema:

$$\begin{cases} \eta_3 = 1 + p_{33}\eta_3 + p_{34}\eta_4 \\ \eta_4 = 1 + p_{43}\eta_3 + p_{44}\eta_4 \end{cases}$$

Inserendo i valori di p_{ij}, e risolvendo si trova $\eta_3 = \eta_4 = 2$. Le probabilità di assorbimento in C partendo da $i \in T$ sono soluzioni del sistema:

$$\begin{cases} \lambda_3 = p_{31} + p_{32} + p_{33}\lambda_3 + p_{34}\lambda_4 \\ \lambda_4 = p_{41} + p_{42} + p_{43}\lambda_3 + p_{44}\lambda_4 \end{cases}$$

Inserendo i valori di p_{ij}, e risolvendo si trova $\lambda_3 = \lambda_4 = 1$, come ci si aspettava.

1.6 Distribuzione stazionaria di una CM

Può accadere che, per $n \to \infty$, la probabilità di transizione in n passi dallo stato i allo stato j, $p_{ij}^{(n)}$, tenda ad un limite π_j, indipendente dallo stato iniziale i, ovvero

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)} \tag{1.34}$$

Passando al limite, per $n \to \infty$ nella formula di Chapman-Kolmogorov (1.20), si ottiene allora

$$\pi_j = \sum_{h\in E} \pi_h p_{hj} \tag{1.35}$$

oppure, in notazione matriciale

$$\pi = \pi P \tag{1.36}$$

il che significa che $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ è autovettore sinistro di P relativo all'autovalore 1 (oppure π^T è autovettore di P^T relativo all'autovalore 1). Il vettore π sopra trovato è il vettore delle probabilità stazionarie. Il numero π_i rappresenta la probabilità che la CM sia nello stato i all'equilibrio, cioé ad un tempo infinito. Ricordando la definizione di distribuzione invariante, si ottiene che la distribuzione stazionaria (se esiste, è unica) di una CM è anche invariante, ma non è vero il viceversa: una distribuzione può essere invariante, ma non stazionaria, nel senso che può non esistere $\lim_{n\to\infty} p_{ij}^{(n)}$.

Un sistema dinamico discreto si dice *ergodico* se la media temporale di un osservabile coincide con la media di stato (all' equilibrio), ovvero, per ogni funzione f continua:

$$\lim_{M\to+\infty} \frac{1}{M} \sum_{n=1}^{M} f(X_n) \equiv E(f(X_n)) = \sum_{i=1}^{N} f(i)\pi_i$$

prendendo allora $f(x) = \mathbf{1}_A(x)$, si ottiene, per $M \to \infty$:

$$\frac{\#\text{ visite ad A in un tempo M}}{M} \longrightarrow \sum_{i\in A} \pi_i = P(X(\infty) \in A)$$

Dunque, se $A = \{i\}$, si ha che la frequenza di visita della CM allo stato i tende a π_i .

Supponiamo ora che $\{X_k\}$ sia una CM stazionaria (ovvero esiste la distribuzione stazionaria π). Allora vale la seguente:

Proposizione 1.3 Sia $\{X_k\}$ una CM discreta con spazio degli stati $\{1, \ldots, N\}$ e supponiamo che esista la distribuzione stazionaria $\pi = (\pi_1, \ldots, \pi_N)$ tale che $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$, $\forall i, j \in \{1, \ldots, N\}$; allora la frequenza di visita della CM allo stato $j \in \{1, \ldots, N\}$ tende in probabilità a π_j.

■

Per questo motivo, con riferimento alla proprietà di ergodicità, se esiste la distribuzione stazionaria, la CM si dice anche ergodica.

Prova della Prop. 1.3 Per ipotesi:

$$P(X_n = j|X_0 = i) = p_{ij}^{(n)} \to \pi_j, \text{ per } n \to \infty$$

Da ciò segue anche:

$$\frac{1}{n}\sum_{k=0}^{n-1} P(X_k = j|X_0 = i) \to \pi_j, \text{ per } n \to \infty$$

Infatti, è facile provare (convergenza in media di Cesaro) che se $a_k \to l$, per $k \to \infty$, allora $\frac{1}{n}\sum_{k=0}^{n-1} a_k \to l$, per $n \to \infty$. Sia $X_0 = i \in \{1, \dots, N\}$; introduciamo, per $j \in \{1, \dots, N\}$ fissata la v.a Y_k che vale 1 se $X_k = j$, e 0 altrimenti. Allora $P(X_k = j|X_0 = i) = E(Y_k)$ e quindi, per quanto visto sopra, per $n \to \infty$:

$$\frac{1}{n}\sum_{k=0}^{n-1} E(Y_k) = E\left(\frac{1}{n}\sum_{k=0}^{n-1} Y_k\right) = E(\text{freq. di visita allo stato j}) \to \pi_j \ .$$

Inoltre:

$$E\left[\left(\frac{1}{n}\sum_{k=0}^{n-1} Y_k\right)^2\right] = \frac{1}{n^2}\sum_{k=0}^{n-1} E(Y_k)^2 + \frac{2}{n^2}\sum_{h<k} E(Y_h Y_k) =$$

$$= \frac{1}{n^2}\sum_{k=0}^{n-1} E(Y_k) + \frac{2}{n^2}\sum_{h<k} P(X_h = j, X_k = j) =$$

$$= \frac{1}{n}\cdot\frac{1}{n}\sum_{k=0}^{n-1} E(Y_k) + \frac{2}{n^2}\sum_{h<k} P(X_h = j) p_{jj}^{(k-h)} \sim$$

$$\sim \frac{1}{n}\pi_j + \frac{2}{n^2}\ \frac{n(n-1)}{2}\pi_j\pi_j \to \pi_j^2$$

Quindi, per $n \to \infty$:

$$E\left(\frac{1}{n}\sum_{k=0}^{n-1} Y_k\right) \to \pi_j \ \text{ e } \ Var\left(\frac{1}{n}\sum_{k=0}^{n-1} Y_k\right) \to 0.$$

Allora, per la disuguaglianza di Chebicev, $\forall \epsilon > 0$:

$$P\left(\left|\frac{1}{n}\sum_{k=0}^{n-1} Y_k - \pi_j\right| > \epsilon\right) \le \frac{1}{\epsilon^2} Var\left(\frac{1}{n}\sum_{k=0}^{n-1} Y_k\right) \to 0,$$

ovvero la frequenza di visita allo stato j converge in probabilità a π_j.

$\blacksquare$

Definizione 1.4 *Una CM discreta con spazio degli stati E, si dice regolare se, detta $P = (P)_{ij}$ la matrice delle probabilità di transizione associata, esistono un intero $s > 0$ ed un numero $\alpha \in (0,1)$ tali che $p_{ij}^{(s)} \geq \alpha \ \forall i,j \in E$ (ovvero P^s ha tutti gli elementi strettamente positivi)*

Criterio di regolarità *Se la CM è irriducibile (gli stati formano un'unica classe di elementi tra loro comunicanti), condizione sufficiente per la regolarità è che esista un elemento strettamente positivo sulla diagonale principale di P, ovvero esista $h \in E$ per cui $p_{hh} > 0$.*

Dim. Per ipotesi, per ogni $i,j \in E$ esiste $n = n(i,j)$ tale che $p_{ij}^{(n(i,j))} > 0$. Sia $m \doteq \max_{i,j\in E} n(i,j)$. Si ha allora:

$$p_{ij}^{(2m)} = \sum_{h_1\in E}\sum_{h_2\in E}\cdots\sum_{h_{2m}\in E} p_{ih_1}p_{h_1h_2}\cdots p_{h_{2m-1}h_{2m}}\cdot p_{h_{2m}j} \geq$$

$$\geq p_{ih}^{(n(i,h))}\cdot \underbrace{p_{hh}\cdots\ p_{hh}}_{2m-n(i,h)-n(h,j)\ \text{volte}} \cdot p_{hj}^{(n(h,j))} > 0 \tag{1.37}$$

e quindi P^{2m} ha tutti gli elementi positivi, cioè P è regolare. $\blacksquare$

Teorema 1.1 (ergodico I) *Sia X_k una catena di Markov regolare con spazio degli stati $E = \{1,2\ldots,N\}$, ovvero esistono un intero $s > 0$ ed un numero $\alpha \in (0,1)$ tali che $p_{ij}^{(s)} \geq \alpha \ \forall i,j \in E$. Allora esiste un'unica distribuzione invariante $\pi = (\pi_1,\pi_2,\ldots,\pi_n)$ (cioé tale che $\pi P = \pi$), con $\pi_i > 0 \ \forall i \in E$. Inoltre $\forall j \in E$, esiste ed è indipendente dallo stato $i \in E$, il limite $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$ (π è anche distribuzione stazionaria) e la velocità di convergenza è esponenziale, ovvero risulta:*

$$|p_{ij}^{(n)} - \pi_j| \leq 2\beta^{n-s} \tag{1.38}$$

dove si è posto $\beta = (1-\alpha)^{1/s} < 1$. $\blacksquare$

Se esiste la distibuzione stazionaria, cioé la CM è stazionaria, il limite per $n \to \infty$ di P^n è una matrice con le righe tutte uguali. Gli elementi di ciascuna riga formano il vettore delle probabilità stazionarie π. Naturalmente, se la CM non è regolare, può accadere che esista la distribuzione stazionaria π, ma il vettore π non ha componenti tutte positive.

Esempio 1.11 Si consideri la CM con spazio degli stati $E = \{1,2,3,4\}$ e matrice delle probabilità di transizione

$$P = \begin{pmatrix} 1/3 & 2/3 & 0 & 0 \\ 3/4 & 1/8 & 1/8 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/3 & 2/3 \end{pmatrix}$$

E' facile verificare con un computer (ad es. col software Scilab, Matlab, o con un programma scritto in un qualunque linguaggio) che P^n conterrà sempre qualche zero, qualunque sia n. Dunque la CM non è regolare e non si può applicare il teorema ergodico. Tuttavia, la CM è stazionaria; infatti, si verifica numericamente (ancora col software Scilab o altro) che, per n grande P^n approssima una matrice con le righe tutte uguali al vettore $\pi = (0, 0, \frac{2}{5}, \frac{3}{5})$, che è evidentemente il vettore delle probabilità stazionarie. Le componenti non sono tutte positive, poiché la CM non è regolare.

Esercizio 1.8
In ambiente Scilab (un software freeware analogo a Matlab che si scarica gratuitamente da Internet), l'istruzione di input, ad esempio della matrice $P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, è la seguente:

$$P = [0\ 1\ 0; 0\ 0\ 1; 1\ 0\ 0]$$

Per calcolare P^n, si digita $P\ \hat{}\ n$.
L'istruzione $spec(P)$ fornisce gli autovalori di P, mentre l'istruzione
$[m, q] = eigenmarkov(P)$ fornisce il vettore delle probabilità invarianti.
L'istruzione $[perm, rec, tr, insRec, indsT] = classmarkov(P)$ fornisce la classificazione degli stati della CM associata a P; $perm$ è un vettore di permutazione di interi; rec, tr sono due interi contenenti, rispettivamente, il numero degli stati ricorrenti e il numero degli stati transitori; $insRec$ e $indsT$ sono vettori interi contenenti, rispettivamente, gli stati ricorrenti e quelli transitori.
Usando Scilab, classificare gli stati, studiare la regolarità, la stazionarietà, e trovare autovalori e distribuzione/i invariante/i e/o stazionaria/e per le CM associate alle seguenti matrici:

$$P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2/3 & 0 & 1/3 \end{pmatrix}, \ P_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$P_3 = \begin{pmatrix} 1/3 & 2/3 & 0 & 0 \\ 3/4 & 1/8 & 1/8 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/3 & 2/3 \end{pmatrix}, \ P_4 = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

Verificare che la CM associata a P_2 non è stazionaria e $(P_2)^4 = P_2$ (CM periodica di periodo 3); la CM associata a P_3 non è regolare (qualunque potenza di P_3 contiene degli zeri), ma è stazionaria, ed esiste la distribuzione stazionaria $\pi = (0, 0, 2/5, 3/5)$. La matrice P_4 è bistocastica (v. Definizione 1.5) e risulta $(P_4)^{2k} = (P_4)^2$, $(P_4)^{2k+1} = (P_4)^3$; la CM associata è periodica di periodo 2, non esiste la distribuzione stazionaria, ma c'è un'unica distribuzione invariante $\pi = (1/4, 1/4, 1/4, 1/4)$.

Esercizio 1.9
Scrivere un programma in Java, C, oppure C^{++}, per calcolare autovalori, distribuzione invariante, potenze della matrice delle probabilità di transizione associata a una CM.

Osservazione 1.2 Se esistono (almeno) due stati assorbenti, diciamo 1 e 2, la CM non può essere stazionaria. Infatti, risulta $p_{1j}^{(n)} = 0 \ \forall j \neq 1, \ p_{2h}^{(n)} = 0 \ \forall h \neq 2$. Allora, per ogni n risulta $p_{12}^{(n)} = 0$ e $p_{22}^{(n)} = 1$; se esistesse la distribuzione stazionaria π, dovrebbe aversi contemporaneamente:

$$\pi_2 = \lim_{n\to\infty} p_{12}^{(n)} = 0 \text{ e } \pi_2 = \lim_{n\to\infty} p_{22}^{(n)} = 1$$

ma ciò è impossibile, in quanto il limite che definisce π, se esiste, è unico.

Osservazione 1.3 Supponiamo che esistano due distinte classi chiuse C_1 e C_2, tali che la CM ristretta a C_2 sia regolare. Allora, non può esistere la distribuzione stazionaria per la CM. Infatti, siano $C_1 = \{i_1, i_2, \ldots, i_\ell\}$ e $C_2 = \{j_1, j_2, \ldots, j_m\}$. Ovviamente, risulta $\forall n$ e $\forall k = 1, 2, \ldots, \ell : \ \ p_{i_k j_2}^{(n)} = 0$ (poiché è impossibile andare da C_1 a C_2). Inoltre, siccome la sottocatena ristretta a C_2 è regolare, deve esistere, ad es.

$$\lim_{n\to\infty} p_{j_1 j_2}^{(n)} = \nu_{j_2} > 0$$

Ma, se esistesse la distribuzione stazionaria π per la CM intera, cioé tale che $\pi_k = \lim_{n\to\infty} p_{hk}^{(n)}$, dovrebbe aversi contemporaneamente:

$$\begin{cases} 0 = \pi_{j_2} = \lim_{n\to\infty} p_{i_k j_2}^{(n)} \\ \pi_{j_2} = \nu_{j_2} = \lim_{n\to\infty} p_{j_1 j_2}^{(n)} > 0 \end{cases}$$

ovvero π_{j_2} dovrebbe essere contemporaneamente 0 e > 0, il che è assurdo.

Osservazione 1.4 Una CM può essere irriducibile, senza essere regolare. Vale però la seguente:

Proposizione 1.4 *Se una CM con spazio degli stati E finito ha matrice di transizione P irriducibile, allora esiste un'unica distribuzione invariante per P (non è detto che sia anche stazionaria).*

Dim. Consideriamo la matrice $Q \doteq \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} P^n$. Essa è stocastica e irriducibile, cioé $q_{ij} > 0 \ \forall i, j$. Infatti:

$$\sum_j q_{ij} = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \underbrace{\sum_j p_{ij}^{(n)}}_{= 1} = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} = 1$$

e quindi Q è stocastica.
Inoltre, visto che $q_{ij} = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} p_{ij}^{(n)}$, e P è irriducibile, nell'ultima somma compare almeno un termine $p_{ij}^{(n(i,j))} > 0$, per cui si conclude che $q_{ij} > 0 \ \forall i, j \in E$ e dunque Q è regolare.

Ora, se v è distribuzione invariante per P, essa lo è anche per $Q \doteq \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} P^n$. Infatti:

$$vQ = \sum_n \frac{1}{2^{n+1}} vP^n = \sum_n \frac{1}{2^{n+1}} v = v$$

dal momento che la serie geometrica $\sum_n \frac{1}{2^{n+1}}$ vale 1.
Siccome, come abbiamo visto, Q è regolare, per il Teorema ergodico 1.1, esiste ed è unica la distribuzione stazionaria ed invariante per Q. Pertanto, non possono esistere due diverse distribuzioni invarianti per P, poiché in tal caso esse lo sarebbero anche per Q che di tali distribuzioni ne ha una sola. ∎

Tempo medio di primo ritorno per una CM ergodica

Sia $T_i = \inf\{n \geq 1 : X_n = i\}$; allora, $E(T_i|X_0 = i) := m_i$ si chiama *il tempo medio di primo ritorno nello stato i* . Vale la seguente:

Proposizione 1.5 Sia X_k una CM con spazio degli stati $E = \{1, 2, \ldots, N\}$ ed ergodica (ovvero esiste la distribuzione stazionaria); allora il tempo medio di ritorno, m_i, nello stato i è finito $\forall i \in E$, e risulta $m_i = 1/\pi_i$, dove $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$ è la distribuzione stazionaria.

Diamo di seguito una dimostrazione non rigorosa di questa Proposizione.
Per $i \in E$ fissato e $k = 0, 1, \ldots, n$ consideriamo la v.a. Y_k definita da:

$$Y_k = \begin{cases} 1 & \text{se } X_k = i \\ 0 & \text{se } X_k \neq i \end{cases}$$

Per la Proposizione 1.3, $\frac{1}{n}(Y_0 + Y_1 + \ldots Y_{n-1})$ converge in probabilità, per $n \to \infty$, a $P(X_\infty = i) = \pi_i$, ovvero $\forall \epsilon > 0$, risulta:

$$\lim_{n\to\infty} P\left(\left|\frac{Y_0 + \ldots + Y_{n-1}}{n} - \pi_i\right| > \epsilon\right) = 0.$$

Poiché m_i è il numero medio di passi per ritornare nello stato i, la frazione $(n-1)/m_i$ è uguale approssimativamente al numero di visite allo stato i nei primi $n-1$ passi, partendo da $X_0 = i$. Quindi:

$$\frac{n-1}{m_i} \approx Y_0 + Y_1 + \ldots + Y_{n-1}$$

ovvero

$$\frac{1}{m_i} \approx \frac{Y_0 + Y_1 + \ldots + Y_{n-1}}{n-1} = \frac{Y_0 + Y_1 + \ldots + Y_{n-1}}{n} \cdot \frac{n}{n-1} \approx \pi_i \text{ , per } n \to \infty.$$

∎

Esempio 1.12. Si lancia ripetutamente un dado perfetto. Mostriamo, usando la Teoria delle CM, che il tempo medio tra due uscite successive di una data faccia del dado è 6. In verità, questo risultato segue subito, se si considera che, per l'indipendenza dei lanci e la proprietà di mancanza di memoria, il tempo tra due uscite successive di una faccia è una v.a. geometrica modificata di parametro $p = 1/6$, e quindi il tempo medio tra due uscite successive di una data faccia del dado è $1/p = 6$. Se però modellizziamo i lanci del dado con una CM, possiamo considerare la CM $X_k \in \{1, 2, 3, 4, 5, 6\}$ con matrice delle probabilità di transizione:

$$P = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}$$

Si tratta di una matrice bistocastica (v. più avanti), e la CM è regolare, per cui esiste la distribuzione stazionaria $\pi = (1/6, 1/6, 1/6, 1/6, 1/6, 1/6)$. Allora, per la Proposizione 1.5 si ottiene che il tempo medio tra due uscite successive di una data faccia del dado è $1/\pi_i = 6$.

Esempio 1.13. Consideriamo la CM a due stati con matrice delle probabilità di transizione:

$$P = \begin{pmatrix} 1/2 & 1/2 \\ p & 1-p \end{pmatrix}$$

dove il parametro $p \in (0, 1]$ è incognito. Supponiamo che, se si osserva la CM per un numero di passi abbastanza grande, essa finisce nello stato 1 approssimativamente il 20% delle volte e nello stato 2 l'80% delle volte. Serviamoci di questi dati per stimare il valore del parametro p.

Siccome la CM è regolare, esiste la distribuzione stazionaria π; inoltre, dalle ipotesi si può ritenere che $\pi_1 = 0.2$ e $\pi_2 = 0.8$, quindi il tempo medio di ritorno nello stato 1 è $1/\pi_1 = 5$, mentre quello nello stato 2 è $1/\pi_2 = 5/4$. Siccome il vettore π deve soddisfare all'equazione $\pi = \pi P$ con la condizione $\pi_1 + \pi_2 = 1$, imponendo

$$\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ p & 1-p \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$$

si trova $1/2 \cdot 0.2 + p \cdot 0.8 = 0.2$ che implica $p = 1/8$.

Definizione 1.5 *Uno stato i di una CM si dice periodico di periodo $d_i > 1$ (indicato anche con per(i)) se*

$$MCD\{n > 0 : p_{ii}^{(n)} > 0\} = d_i \tag{1.39}$$

Se tale MCD vale 1, *lo stato i si dice aperiodico.*
Se risulta $p_{ii}^{(n)} > 0$ e $p_{ii}^{(m)} > 0$ con n e m primi tra loro, ovviamente lo stato i è aperiodico. Inoltre, come si vede facilmente, se una CM ha stati periodici (di periodo maggiore di 1), essa non può essere stazionaria, cioé non può esistere $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$.

Gli stati di una stessa classe hanno lo stesso periodo, come evince dalla seguente:

Proposizione 1.6 *Se i e j appartengono alla stessa classe di elementi comunicanti, allora $per(i) = per(j)$.*
Dim. Siccome i comunica con j e j comunica con i, esistono interi m, n tali che $p_{ij}^{(m)} > 0$ e $p_{ji}^{(n)} > 0$; dunque $p_{ii}^{(m+n)} \geq p_{ij}^{(m)} \cdot p_{ji}^{(n)} > 0$ il che implica che $m + n$ è divisibile per $d_i = per(i)$.
Consideriamo lo stato j, di periodo $d_j = per(j)$; esistono infiniti interi r abbastanza grandi per cui $p_{jj}^{(rd_j)} > 0$. Ma

$$p_{ii}^{(m+rd_j+n)} \geq p_{ij}^{(m)} \cdot p_{jj}^{(rd_j)} \cdot p_{ji}^{(n)} > 0$$

da cui segue che $m+rd_j+n$ è divisibile per d_i. Siccome d_i divide $m+n$, allora d_i divide rd_j, e vista l'arbitrarietà di r (poiché i valori di tali r sono infiniti), d_i deve necessariamente dividere d_j, cioé esiste k intero positivo tale che $d_j = kd_i$. D'altra parte, scambiando i con j, si ottiene con ragionamento analogo che d_j divide d_i, cioé esiste un intero $h > 0$ tale che $d_i = hd_j$; pertanto $d_j = k \cdot hd_j$, da cui segue $kh = 1$, ovvero $k = h = 1$ e concludiamo che $d_i = d_j$. ∎

Un caso particolare di CM con stati periodici si ha quando la CM è *ciclica*, ovvero tutti gli stati comunicano tra loro, ed esiste un intero $d > 0$ tale che $P^{d+1} = P$. Si osservi che ciò non implica $P^d = I$ (matrice identità) a meno che P sia non degenere. Se la CM è ciclica, tutti gli stati hanno stesso periodo d. Naturalmente, una CM può avere stati periodici, senza che essa sia ciclica. Inoltre, può accadere che esista un intero $d > 0$ tale che $P^{d+1} = P$, senza che la CM sia ciclica (cfr. Esempio 1.16, per il quale gli stati della CM non appartengono tutti alla stessa classe).

Esempio 1.14 La CM con 3 stati e matrice delle probabilità di transizione:

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

è ciclica e periodica di periodo 3, infatti $P^4 = P$ (oppure $P^3 = I$, essendo $detP = 1 \neq 0$). Partendo da uno stato i si ritorna in i esattamente dopo 3 passi (vedi Figura 7.)

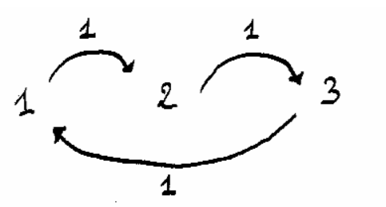

Fig. 7: Grafo delle probabilità di transizione della CM dell'Esempio 1.11

Questa CM non è stazionaria (la distribuzione stazionaria non esiste, poiché non esiste $\lim_{n\to\infty} P^n$), ma la distribuzione invariante, come si verifica facilmente, calcolando l'autovettore sinistro di P relativo all'autovalore 1, è $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, cioé la distribuzione uniforme sugli stati.

Esempio 1.15 (random walk con barriere riflettenti)
Consideriamo la CM con 3 stati e matrice delle probabilità di transizione:

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

con $p, q > 0$ e $p + q = 1$. Si ha:

$$P^2 = \begin{pmatrix} q & 0 & p \\ 0 & 1 & 0 \\ q & 0 & p \end{pmatrix}$$

e

$$P^3 = P$$

per cui $P^{2n} = P^2$, mentre $P^{2n+1} = P$, avendosi $p_{ii}^{(2n+1)} = p_{ii}^{(3)} = 0 \; \forall n$. Quindi gli stati sono tutti periodici di periodo 2, e la distribuzione stazionaria non esiste, poiché non esiste $\lim_{n\to\infty} P^n$. Infatti, ad esempio:

$$p_{ii}^{(n)} = \begin{cases} 0 & \text{se n è dispari} \\ > 0 & \text{se n è pari} \end{cases}$$

La distribuzione invariante è $\pi = (q/2, 1/2, p/2)$.

Esempio 1.16
Consideriamo la CM con 3 stati e matrice delle probabilità di transizione:

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

Lo stato 1 è transitorio, gli stati 2 e 3 sono ricorrenti ed hanno periodo 2. Infatti:

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

$$P^3 = P$$

per cui $P^{2n} = P^2$ e $P^{2n+1} = P$, e la distribuzione stazionaria non esiste. La distribuzione invariante è $\pi = (0, 1/2, 1/2)$.

Definizione 1.6 *Una matrice stocastica P si dice bistocastica se $\sum_{i\in E} p_{ij} = 1$ ovvero se anche P^T è stocastica. Per la CM associata ad una tale P, la distribuzione uniforme $v = (\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N})$ è invariante (non è detto sia unica e che sia stazionaria!). Infatti, $v_i = 1/N$ soddisfa banalmente il sistema $v_i = \sum_j v_j p_{ji}$, visto che, essendo P bistocastica, risulta $\sum_j p_{ji} = 1$.*

Esempio 1.17
Consideriamo la CM con 4 stati e matrice delle probabilità di transizione:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}$$

$C_1 = \{1, 2\}$ e $C_2 = \{3, 4\}$ sono classi chiuse di stati ricorrenti; gli stati 3 e 4 sono aperiodici, mentre gli stati 1 e 2 sono periodici di periodo 2. Infatti:

$$P^2 = \begin{pmatrix} Id & 0 \\ 0 & P_2 \end{pmatrix},\ P^3 = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = P,\ P^4 = \begin{pmatrix} Id & 0 \\ 0 & P_2 \end{pmatrix}, \dots$$

ovvero, per ogni intero n :

$$P^{2n+1} = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} = P,\ P^{2n} = \begin{pmatrix} Id & 0 \\ 0 & P_2 \end{pmatrix}$$

La distribuzione stazionaria non esiste; quella invariante, essendo P bistocastica, è $\pi = (1/4, 1/4, 1/4, 1/4)$.

Definizione 1.7 *Una distribuzione π sugli stati si dice reversibile se è verificata l'equazione del bilancio dettagliato, ovvero:*

$$\pi_i p_{ij} = \pi_j p_{ji}\ \forall i, j \in E \tag{1.40}$$

In tal caso, π è anche invariante, avendosi:

$$\sum_{i\in E} \pi_i p_{ij} = \sum_i \pi_j p_{ji} = \pi_j \left(\sum_i p_{ji} \right) = \pi_j$$

visto che la somma in ( ) vale 1.

Vediamo ora il significato di distribuzione reversibile. Sia X_k una CM ergodica con probabilità stazionarie π_i.
Partendo dall'istante n, consideriamo la sequenza di stati $X_n,\ X_{n-1},\ X_{n-2}, \dots$; questa sequenza di stati individua un'altra CM *invertita* X'_k con probabilità di transizione

$$q_{ij} = P(X + m = j | X_{m+1} = i) = \frac{P(X_m = j, X_{m+1} = i)}{P(X_{m+1} = i)}$$

$$= \frac{P(X_m = j)P(X_{m+1} = i|X_m = j)}{P(X_{m+1} = i)} \approx \frac{\pi_j p_{ji}}{\pi_i},$$

per n grande.
Se $q_{ij} = p_{ij}$ per ogni i, j allora la CM X_k si dice reversibile, e vale dunque l'equazione del bilancio dettagliato (1.40). Questo significa che il tasso con il quale la CM X_k effettua in regime stazionario una transizione da i a j, ovvero $\pi_i p_{ij}$, uguaglia il tasso col quale essa va da j a i, ovvero $\pi_j p_{ji}$. Ciò è ovviamente una condizione necessaria per la reversibilità temporale della CM X_k.

1.7 Matrice di transizione ad n passi

Ci occuperemo ora del calcolo esplicito della matrice delle probabilità di transizione ad n passi di una CM, ovvero della matrice $P^n = (p_{ij}^{(n)})$ il cui elemento di posto i, j rappresenta $P(X_n = j|X_0 = i)$. E' facile provare che la matrice (stocastica) P ha sempre un autovalore $\lambda_1 = 1$ (corrispondente all'autovettore avente tutte le componenti uguali ad 1) e tutti gli altri autovalori hanno modulo ≤ 1. In altre parole, una matrice stocastica ammette sempre l'autovalore 1, e comunque tutti i suoi autovalori sono situati all'interno del cerchio unitario nel piano complesso.
Per dimostrare ciò, osserviamo che se λ è un qualunque autovalore di P relativo all' autovettore u, si ha:

$$\lambda u_i = \sum_{j=1}^{N} p_{ij} u_j \ , \forall i = 1, 2, \ldots, N$$

Sia ora $|u_h| = \max_{j=1,2,\ldots,N} |u_j|$, per un certo $h \in \{1, 2, \ldots, N\}$. Allora:

$$|\lambda||u_h| = |\sum_j p_{hj} u_j| \leq \sum_j p_{hj} |u_j| \leq |u_h| \sum_j p_{hj} = |u_h|$$

da cui, semplificando $|u_h| > 0$, si ottiene $|\lambda| \leq 1$.

Ricordiamo dall'algebra lineare la nozione di matrice diagonalizzabile. Sia A una matrice quadrata di ordine N con autovalori $\lambda_1, \lambda_2, ..., \lambda_N$ e supponiamo che per ciascuno dei λ_i, $i = 1, .., N$, la molteciplità geometrica uguagli la molteplicità algebrica; allora A è diagonalizzabile, cioé esiste una matrice U di ordine N, tale che $U^{-1}AU$ è una matrice diagonale e risulta

$$U^{-1}AU = \Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_N) \tag{1.41}$$

La matrice U ha per colonne gli N autovettori indipendenti di A.

Consideriamo ora, al posto di A, la matrice di transizione P di una CM con spazio degli stati $E = \{1, 2, ..., N\}$, e supponiamo che gli autovalori di P siano tutti distinti. Allora, P è diagonalizzabile e risulta $P = U\Lambda U^{-1}$; pertanto, come è immediato verificare, si ha $P^n = U\Lambda^n U^{-1}$.
Effettuiamo, ad esempio il calcolo esplicito di P^n, nel caso in cui P è una matrice stocastica di ordine 2, con elementi strettamente positivi, cioé essa è la matrice delle probabilità di

transizione di una CM regolare a due stati (nel caso di matrici di ordine > 2 il calcolo è più complicato e lo tratteremo in qualche esercizio); P può essere posta nella forma:

$$P = \begin{pmatrix} 1-p & p \\ a & 1-a \end{pmatrix} \tag{1.42}$$

dove $a, p \in (0,1)$. Gli autovalori di P sono: $\lambda_1 = 1$ e $\lambda_2 = 1 - a - p$ con $|\lambda_2| \leq 1$ (v. osservazioni precedenti), ed essi sono distinti; dunque:

$$\Lambda = \begin{pmatrix} 1 & 0 \\ 0 & 1-a-p \end{pmatrix} \tag{1.43}$$

I due autovettori relativi a λ_1 e λ_2 sono, a meno di un fattore di proporzionalità:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ -\frac{a}{p} \end{pmatrix}$$

Si ha allora:

$$U = \begin{pmatrix} 1 & 1 \\ 1 & -\frac{a}{p} \end{pmatrix}$$

e

$$U^{-1} = \frac{p}{a+p} \begin{pmatrix} \frac{a}{p} & 1 \\ 1 & -1 \end{pmatrix}$$

Dunque:

$$P^n = U\Lambda^n U^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & -\frac{a}{p} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1-a-p)^n \end{pmatrix} \begin{pmatrix} \frac{a}{a+p} & \frac{p}{a+p} \\ \frac{p}{a+p} & -\frac{p}{a+p} \end{pmatrix}$$

ed effettuando i calcoli si ottiene infine:

$$P^n = \begin{pmatrix} \frac{a}{a+p} & \frac{p}{a+p} \\ \frac{a}{a+p} & \frac{p}{a+p} \end{pmatrix} + \frac{(1-a-p)^n}{a+p} \begin{pmatrix} p & -p \\ -a & a \end{pmatrix} \tag{1.44}$$

Le due righe (uguali) della prima matrice sono costituite dalle componenti del vettore π delle probabilità stazionarie (π è l' autovettore sinistro di P relativo all'autovalore 1), che è:

$$\pi = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} \frac{a}{a+p} \\ \frac{p}{a+p} \end{pmatrix} \tag{1.45}$$

Abbiamo allora ottenuto:

$$P^n = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix} + (\lambda_2)^n \cdot R \tag{1.46}$$

ove R è la matrice (non stocastica) data da:

$$R = \frac{1}{a+p} \begin{pmatrix} p & -p \\ -a & a \end{pmatrix} \tag{1.47}$$

La velocità di convergenza nel teorema ergodico è fornita allora dalla quantità ${\lambda_2}^n$ ove, come abbiamo visto, $|\lambda_2| < 1$: più vicino a zero è il secondo autovalore di P, maggiore è la velocità di convergenza di $p_{ij}^{(n)}$ a π_j per $n \to \infty$. La velocità di convergenza alla stazionarietà può essere valutata anche tramite la (1.38).

Concludiamo col seguente teorema, valevole anche nel caso in cui la CM non è regolare.

Teorema 1.2 (ergodico II)
Se X_k è una CM con insieme degli stati E finito, irriducibile e i cui stati sono tutti aperiodici, allora esiste la distribuzione stazionaria per la CM. ∎

Osservazione 1.5 (Relazione tra spettro della matrice di transizione e sua regolarità)
Lo spettro di una matrice è costituito dai suoi autovalori; come abbiamo visto, la matrice di transizione P di una CM ammette sempre un autovalore $\lambda_1 = 1$, mentre gli altri suoi autovalori hanno valore assoluto minore o uguale di 1, ovvero il *raggio spettrale* di P è 1. Viene da chiedersi se vi sia una relazione tra lo spettro di P e la sua regolarità. In effetti, si può dimostrare il seguente risultato:

Se P è regolare, allora 1 *è autovalore semplice di P e tutti gli altri autovalori di P sono in modulo* < 1.

Il risultato di sopra non si può invertire, ovvero esistono matrici di transizione P che hanno 1 come autovalore semplice, gli altri autovalori hanno modulo strettamente < 1, ma *non* sono regolari. Ad esempio, la matrice di transizione:

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

ha autovalori $\lambda_1 = 1,\ \lambda_2 = 3/4,\ \lambda_3 = 0$ (quindi 1 è autovalore semplice e tutti gli altri autovalori sono in modulo < 1), ma come si verifica facilmente P non è regolare, poiché qualunque potenza di P avrà sempre almeno un elemento nullo. Tuttavia, la CM che ha P come matrice di transizione ammette distribuzione stazionaria $\pi = (1, 0, 0)$, visto che, per $n \to \infty$, P^n tende ad una matrice avente per righe $(1, 0, 0)$.
Si osservi che la matrice di transizone P dell'esempio di sopra ha zero come autovalore, quindi è degenere (cioé $det(P) = 0$).

Vediamo ora come calcolare esplicitamente P^n per una CM a tre stati, nel caso di autovalori distinti $1, \lambda_2, \lambda_3$.
Siccome P è diagonalizzabile, esiste una matrice invertibile Γ per cui $P = \Gamma\Lambda\Gamma^{-1}$, dove $\Lambda = diag(1, \lambda_2, \lambda_3)$. Allora:

$$P^n = \Gamma \begin{pmatrix} 1 & 0 & 0 \\ 0 & (\lambda_2)^n & 0 \\ 0 & 0 & (\lambda_3)^n \end{pmatrix} \Gamma^{-1}$$

Quindi:

$$p_{ij}^{(n)} = (P^n)_{ij} = \sum_h \sum_k \Gamma_{ik} (\Lambda)^n_{kh} (\Gamma^{-1})_{hj}$$

Siccome $(\Lambda^n)_{kh} = \delta_{kh}(\lambda_h)^n$, si ottiene:

$$p_{ij}^{(n)} = \sum_{h=1}^{3} \Gamma_{ih}(\lambda_h)^n (\Gamma^{-1})_{hj} =$$

$$= \Gamma_{i1} \cdot 1 \cdot (\Gamma^{-1})_{1j} + \Gamma_{i2}(\lambda_2)^n (\Gamma^{-1})_{2j} + \Gamma_{i3}(\lambda_3)^n (\Gamma^{-1})_{3j} = A_{ij} + B_{ij}(\lambda_2)^n + C_{ij}(\lambda_3)^n$$

ove

$$\begin{cases} A_{ij} = \Gamma_{i1}(\Gamma^{-1})_{1j} \\ B_{ij} = \Gamma_{i2}(\Gamma^{-1})_{2j} \\ C_{ij} = \Gamma_{i3}(\Gamma^{-1})_{3j} \end{cases}$$

Ad esempio, sia:

$$P = \begin{pmatrix} 1/3 & 0 & 2/3 \\ 1/3 & 2/3 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

Gli autovalori sono $\lambda_1 = 1, \lambda_2 = 1/3, \lambda_3 = 0$. Allora

$$p_{ij}^{(n)} = A_{ij} + B_{ij} \left(\frac{1}{3}\right)^n + C_{ij} \cdot 0^n = A_{ij} + B_{ij} \left(\frac{1}{3}\right)^n \qquad (*)$$

Supponiamo di voler calcolare, per es., $p_{11}^{(n)}$.
Si ha $p_{11} = 1/3$ e, dalla formula di Chapman-Kolmogorov (o calcolando la matrice P^2), risulta:

$$p_{11}^{(2)} = \sum_{h=1}^{3} p_{1h}p_{h1} = p_{11}^2 + p_{12}p_{21} + p_{13}p_{31} =$$

$$\frac{1}{9} + 0 + \frac{2}{3} \cdot \frac{1}{3} = \frac{1}{3}$$

Sostituendo in (*) $n = 1$ e $i = j = 1$, si ha $1/3 = p_{11} = A + B/3 = 1/3$; sostituendo $n = 2$, si ha $1/3 = p_{11}^{(2)} = A + B/9$. Dunque, si ottiene il sistema (si è posto per semplicità $A_{ij} = A$ e $B_{ij} = B$) :

$$\begin{cases} A + B/3 = 1/3 \\ A + B/9 = 1/3 \end{cases}$$

che ha soluzione $A = 1/3,\ B = 0$, per cui $p_{11}^{(n)} = 1/3$.

Se vogliamo trovare $p_{12}^{(n)}$, dobbiamo prima calcolare $p_{12}^{(2)} = p_{11}p_{12} + p_{12}p_{22} + p_{13}p_{32} = \frac{1}{3} \cdot 0 + 0 + \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$; sostituendo in (*) $n = 1$ e $i = 1, j = 2$, si ha $0 = p_{12} = A + B \cdot \frac{1}{3}$; sostituendo $n = 2$, si ha $\frac{2}{9} = p_{12}^{(2)} = A + \frac{B}{9}$. Dunque, si ottiene il sistema:

$$\begin{cases} A + B/3 = 0 \\ A + B/9 = 2/9 \end{cases}$$

che ha soluzione $A = 1/3,\ B = -1$, per cui $p_{12}^{(n)} = 1/3 - \left(\frac{1}{3}\right)^n$.

Infine $p_{13}^{(n)} = 1 - p_{11}^{(n)} - p_{12}^{(n)} = 1 - 1/3 - 1/3 + (1/3)^n = 1/3 + (1/3)^n$.
Pertanto, per $n \to \infty$, $p_{1j}^{(n)} \to 1/3 \equiv \pi_j$, $\forall i$. Ciò è in accordo col Teorema ergodico (la distribuzione stazionaria π esiste, poiché P è regolare, essendo $P^2 > 0$). Inoltre, siccome P è bistocastica, la distribuzione stazionaria è quella uniforme.

Osservazione 1.6
Se esistono due o più distribuzioni invarianti per una CM, allora non può esistere la distribuzione stazionaria.
Infatti, siano π e μ due diverse distribuzioni invarianti; allora si ha:

$$\pi = \pi P^n \text{ e } \mu = \mu P^n \ \forall n$$

Se esistesse la distibuzione stazionaria, diciamo π, si avrebbe $P^n \to \Pi$, ove Π è la matrice che ha tutte le righe uguali al vettore π. Allora, dovrebbe aversi $\mu = \mu\Pi$, ovvero:

$$\mu_i = \sum_j \mu_j \Pi_{ji} = \sum_j \mu_j \pi_i = \pi_i \underbrace{\sum_j \mu_j}_{=1} = \pi_i,\ i = 1, \ldots, N$$

il che implica $\mu = \pi$.

1.8 L'algoritmo di Metropolis e il Simulated Annealing

L'algoritmo del simulated annealing (SAA) è uno strumento utilissimo in molti problemi di ottimizzazione in cui occorre trovare il minimo o il massimo globale di una funzione f definita su uno spazio finito Ω (lo spazio degli stati o delle configurazioni). Il SAA è una strategia ottenuta per mezzo di un metodo di rilassamento stocastico; l'interpretazione meccanico-statistica è quella di un sistema fisico sul quale si considera una funzione f ed una distribuzione di Gibbs a temperatura T. La distribuzione di Gibbs fornisce la probabilità di trovare il sistema in una particolare configurazione (stato) ad una data temperatura.

L'algoritmo consiste nel far evolvere il sistema in base alla distribuzione di Gibbs e contemporaneamente *raffreddare* il sistema abbastanza lentamente, in modo che (per $T \to 0$) la distribuzione limite coincida con la distribuzione uniforme sugli stati di minima energia. Dunque, se il sistema non è raffreddato troppo rapidamente, si può costruire un'opportuna catena di Markov X_n le cui probabilità di transizione soddisfano l'equazione del bilancio dettagliato (1.40) e tale che, quando $T \to 0$ e $n \to \infty$, la CM X_n converge verso lo stato di minima energia (si veda, ad es. S. Geman and D. Geman "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images" IEEE Trans. Pattern and Mach. intell. 6 (1984), 721–741). Si può dimostrare (E. Aarts, J. Korst "Simulated annealing and Boltzmann machines" John Wiley and Sons, 1989) che X_n ammette distribuzione stazionaria ed essa coincide con la distribuzione uniforme sui mimimi globali, mentre la convergenza dell'algoritmo è garantita dalla proprietà di Markov della catena.

Supponiamo che lo spazio degli stati Ω sia un insieme finito i cui elementi sono le configurazioni σ, e sia $f : \Omega \longrightarrow \mathbb{R}$ una funzione reale sullo spazio degli stati. Il SAA risolve il problema di trovare un elemento $\sigma_{opt} \in \Omega$ tale che

$$\min_{\sigma \in \Omega} f(\sigma) = f(\sigma_{opt}) \tag{1.48}$$

In generale, tale configurazione σ_{opt} non è unica. L'insieme delle configurazioni ottimali sarà denotato con $\Omega_{opt} = \{\sigma_{opt} \in \Omega\}$.

Per $T > 0$, consideriamo la distribuzione di Gibbs Π_T su Ω definita da:

$$\Pi_T(\sigma) = \exp\left(-\frac{f(\sigma)}{T}\right)\left[\sum_{\sigma \in \Omega} \Pi_T(\sigma)\right]^{-1} \tag{1.49}$$

(si osservi che questa distribuzione di probabilità (o densità discreta) su Ω è ben definita, risultando $\sum_{\sigma \in \Omega} \Pi_T(\sigma) = 1$, grazie alla costante di normalizzazione).
Osserviamo che, per grandi valori della temperatura T la distribuzione tende ad essere uniforme su Ω (approssimando il valore $\frac{1}{\#\Omega} = \left[\sum_{\sigma \in \Omega} \Pi_T(\sigma)\right]^{-1}$); per piccoli valori di T gli stati favorevoli di Ω, cioé quei $\sigma \in \Omega$ per i quali $f(\sigma)$ è piccola, vengono pesati con maggiore probabilità (infatti, per $T \approx 0$ la frazione $q = \frac{f(\sigma)}{T}$ diventa in generale grande, e quindi $\exp(-q)$ diviene piccolo; ma se il valore $f(\sigma)$ è confrontabile con T, per esempio $f(\sigma) \approx \frac{T}{m}$, con m grande, si ottiene che $\exp(-q) = \exp(-1/m)$ è poco meno di 1). Quindi, una soluzione probabilistica al problema di trovare il minimo globale di f si ottiene campionando con la distribuzione Π_T per piccoli valori di $T > 0$. Ciò si può realizzare simulando una CM X_n con spazio degli stati Ω, che abbia Π_T come distribuzione stazionaria, e lasciando che la CM raggiunga l'equilibrio.

In primo luogo, è necessario determinare per ogni $\sigma \in \Omega$ un insieme di stati "vicini" $\Omega_\sigma \subset \Omega$ in cui è possibile effettuare una transizione, a partire da σ, in modo tale che la CM sia irriducibile e aperiodica (ciò è sufficiente - Teorema 1.2 - per garantire che esiste la distribuzione stazionaria). Supponiamo che $\sigma \in \Omega_{\sigma'} \Leftrightarrow \sigma' \in \Omega_\sigma$ e che per ogni $\sigma \in \Omega$ sia $\#\Omega_\sigma = \Theta = costante$.

Ora, definiamo la probabilità di transizione dallo stato $\sigma \in \Omega$ ad un altro stato $\sigma' \in \Omega$;

$$\begin{aligned} p_T(\sigma, \sigma') &= P_T(X_{n+1} = \sigma' | X_n = \sigma) = \\ &= \begin{cases} 0 & \text{se } \sigma' \notin \Omega_\sigma \\ cost \cdot \frac{1}{T} \exp\{-(f(\sigma') - f(\sigma))^+\} & \text{se } \sigma' \in \Omega_\sigma \end{cases} \end{aligned} \tag{1.50}$$

dove la costante *cost* si ricava imponendo la condizione che $\sum_{\sigma' \in \Omega_\sigma} p_T(\sigma, \sigma') = 1$. Qui g^+ denota la parte positiva della funzione g, cioé

$$g^+(\sigma) = \begin{cases} g(\sigma) & \text{se } g(\sigma) \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Si può dimostrare che Π_T definita da (1.49) soddisfa l'equazione del bilancio dettagliato (v. E. Aarts, J. Korst) :

$$p_T(\sigma, \sigma')\Pi_T(\sigma) = p_T(\sigma', \sigma)\Pi_T(\sigma') \tag{1.51}$$

ciò è sufficiente a garantire che Π_T è una distribuzione invariante per la CM e coincide con la distribuzione stazionaria, che esiste, unica, per il Teorema Ergodico II.

In pratica, la CM è simulata nel seguente modo: quando il sistema è nello stato σ un vicino σ' viene selezionato uniformemente in Ω_σ e il valore $f(\sigma')$ calcolato. Quindi, la transizione da σ a σ' viene accettata con probabilità $p = \frac{1}{T}\exp\{-(f(\sigma') - f(\sigma))^+\}$ e il nuovo stato della CM è σ', altrimenti la transizione viene rigettata ed il sistema rimane nello stato σ. Geman and Geman hanno dimostrato che, se il sistema è raffreddato abbastanza lentamente, diminuendo la temperatura $T = T_n$ proporzionalmente a $1/\log n$, ovvero se $T_n = c/\log n,\ n \to \infty$, e se al tempo discreto n si utilizzano le probabilità di transizione (1.50) con $T = T_n$, allora la distribuzione limite $\Pi_0 = \lim_{T\to 0^+}\Pi_T$ è la distribuzione stazionaria (e invariante) della CM X_n ed essa è la distribuzione uniforme sugli stati di minima energia f. L' algoritmo del simulated annealing (tradotto letteralmente "algoritmo del raffreddamento simulato") viene così chiamato, proprio per i motivi sopra descritti.

Il SAA è un metodo generale per risolvere problemi combinatori di ottimizzazione; osserviamo che, sia la scelta della struttura dei vicini in Ω, che della funzione di energia f, sono di peculiare importanza per riuscire a implementare l'algoritmo in maniera efficiente.

Per un insieme degli stati Ω discreto, una possibile buona scelta della struttura dei vicini è quella utilizzata nel problema del commesso viaggiatore: esso consiste nel trovare il giro ottimale (cioé quello di minima lunghezza) affinché il viaggiatore possa visitare ognuna di n città $C_1, C_2, \ldots, C_n$ ed infine tornare a casa. In questo caso Ω è l'insieme delle permutazioni degli indici $\{1, 2, \ldots, n\}$; ogni permutazione $\sigma = \{i_1, i_2, \ldots, i_n\} \in \Omega$ è in corrispondenza biunivoca con il giro effettuato secondo l'ordine impartito da σ.
Per il problema del commesso viaggiatore, sono state proposte e studiate varie strutture di vicini (si veda ad es. E. Bonomi, J. L. Lutton "The N-city travelling salesman problem: statistical mechanics and the Metropolis algorithm" SIAM Rev. 26 (1984), 551–568; S. Lin "Computer solutions of the travelling salesman problem" Bell System Tech. J. 44 (1965), 2245–2269).

Un altro problema combinatorio di ottimizzazione è il "multiple digest problem" che trova applicazione, oltre che in Biologia molecolare (v. e.g. M. Abundo, "Sequencing DNA fragments by using a simulated Annealing algorithm" Technical Report Centro Matematico V. Volterra N. 46, 1990), in vari altri contesti. Ad esempio, consideriamo il seguente problema. Siano $Q_1, Q_2, \ldots, Q_n$ n distinti apparecchi telefonici e supponiamo che i tempi di arrivo delle chiamate a ciascun apparecchio siano regolati da un processo di Poisson. Supponiamo di osservare gli n apparecchi nell'intervallo $[0, T]$ e denotiamo con $T_1^i, T_2^i, \ldots, T_r^i$ i tempi di arrivo delle chiamate all'apparecchio $Q_i,\ i = 1, \ldots, n$; indichiamo inoltre con $\delta_j^i = T_{j+1}^i - T_j^i,\ j = 1, \ldots, r-1$ gli intertempi di arrivo tra due succesive chiamate all'apparecchio Q_i, con $\sum_{j=1}^{r-1}\delta_j^i = T$. Per fissare le idee, supponiamo che $n = 2$. Dunque, osserviamo le chiamate che arrivano a Q_1 e registriamo i tempi δ_j^1 che intercorrono tra due succesive chiamate verso Q_1 (conserviamo traccia solo degli intertempi, ma non dei tempi di arrivo); facciamo la stessa cosa per Q_2. Analogamente, osserviamo le chiamate verso il cluster (Q_1, Q_2) e registriamo i tempi che intercorrono tra due successive chiamate all' uno o all'altro dei due apparecchi Q_1 e Q_2, senza distinguere di quale apparecchio si tratti. Il problema del doppio digest consiste allora nel ricostruire esattamente i tempi di arrivo delle chiamate verso Q_1 e verso Q_2.

2. CM a tempo continuo

2.1 Il processo di Poisson

Supponiamo di voler studiare i tempi di arrivo di successive richieste di collegamento ad un server remoto da parte di utenti autorizzati. Se l'intervallo di tempo tra due successivi arrivi è una variabile esponenziale di parametro $\lambda > 0$, per T fissato, cosa possiamo dire riguardo al numero di richieste di collegamento giunte nell'intervallo $[0, T]$? Precisamente, siamo interessati a conoscere la probabilità che giungano esattamente k richieste in $[0, T]$. Indichiamo con $N(t)$ il numero delle richieste di collegamento giunte fino al tempo $t \le T$, e supponiamo che $N(0) = 0$. Sia τ_i il tempo di arrivo della $i-$esima rihiesta e poniamo $T_1 = \tau_1,\ T_2 = \tau_2 - \tau_1$, e in generale $T_i = \tau_i - \tau_{i-1}$ gli intertempi di arrivo delle richieste, ovvero $\tau_1 \equiv T_1 > 0$ è il tempo di arrivo della prima richiesta, $\tau_2 = T_1 + T_2$ è il tempo di arrivo della seconda richiesta, ..., $\tau_k = T_1 + T_2 + \ldots + T_k$ è il tempo di arrivo della $k-$esima richiesta. Facciamo l'ipotesi che le v.a. T_i $(i = 1, 2, \ldots)$ siano indipendenti ed esponenziali di parametro $\lambda > 0$, ovvero abbiano densità

$$f(t) = \begin{cases} 0 & \text{se } t < 0 \\ \lambda e^{-\lambda t} & \text{se } t \ge 0 \end{cases} \tag{2.1}$$

L'evento $\{N(T) \le k\} \equiv \{N(T) < k+1\} \equiv \{\tau_{k+1} > T\} \equiv \{T_1 + T_2 + \ldots + T_{k+1} > T\}$, ovvero il tempo del $(k+1)-$ esimo arrivo è $> T$. Siccome $T_i \sim Esp(\lambda)$ (ovvero ha legge Gamma di parametri $(1, \lambda)$), per il teorema di addizione di v.a. indipendenti con legge Gamma, risulta per il tempo del $k-$esimo arrivo

$$\tau_k = T_1 + T_2 + \ldots + T_k \sim \Gamma(k, \lambda)$$

ovvero τ_k ha distribuzione Gamma di parametri k e λ, cioé ha densità:

$$g(t) = \begin{cases} 0 & \text{se } t < 0 \\ \frac{\lambda^k}{\Gamma(k)} t^{k-1} e^{-\lambda t} & \text{se } t \ge 0 \end{cases} \tag{2.2}$$

ove

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx \tag{2.3}$$

è la funzione Gamma di Eulero (si ricordi che $\Gamma(k) = (k-1)!$).
Allora:

$$P(N(T) \le k) = P(T_1 + \ldots + T_{k+1} > T) = P(\tau_{k+1} > T) =$$

(ricordando l'espressione della funzione di distribuzione cumulativa di una Gamma con primo parametro k intero)

$$= 1 - P(\tau_{k+1} \le T) = \sum_{i=0}^{k} \frac{(\lambda T)^i}{i!} e^{-\lambda T} \tag{2.4}$$

Questa è la funzione di distribuzione cumulativa di una v.a. di Poisson di parametro λT. Ne segue che $N(T)$ è una v.a. di Poisson di parametro λT, oppure, sostituendo t a T, possiamo scrivere $N(t) \sim$ Poisson (λt).

La famiglia di v.a. $N(t),\ t \geq 0$ è un facile esempio di processo stocastico a tempo continuo: esso viene chiamato processo di *Poisson*. La v.a. $N(t)$ "conta" il numero di volte che uno specifico evento (nel nostro esempio l'arrivo di una richiesta di collegamento al server) si è verificato nell'intervallo di tempo $[0,t]$. Tale processo è anche detto processo *di conteggio*, proprio perché si tratta di una famiglia di v.a. a valori interi ≥ 0, che "conta" il numero degli "arrivi" in $[0,t]$. Una traiettoria (aleatoria) di $N(t)$ è rappresentata da una funzione non decrescente a scalino. Ad ogni istante di salto, τ_k, $N(t)$ si incrementa di una unità. Naturalmente, $N(0) = 0$ (v. Figura 8).

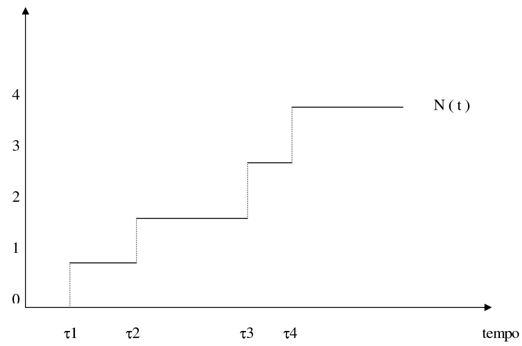

Fig. 8: Traiettoria del processo di Poisson $N(t)$.

Altri esempi di fenomeni aleatori modellizzabili con un processo di Poisson sono: 1) il numero di particelle emesse in $[0,t]$ da una sostanza che va incontro a decadimento radioattivo; 2) il numero di guasti che subisce un'apparecchiatura nell'intervallo $[0,t]$; 3) il numero di arrivi di clienti ad un servizio nell'intervallo $[0,t]$.

Definizione 2.1 *Si chiama* **processo di Poisson di intensità** $\lambda > 0$ *un processo stocastico* $N(t)$ *a valori interi* ≥ 0 *per cui risulti:*

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \ k \geq 0 \tag{2.5}$$

Diamo ora un'altra costruzione/interpretazione del processo di Poisson.

Proposizione 2.1 *Sia* $X(t) \in \mathbb{N}$ *un processo stocastico che "conta" il numero di eventi occorsi fino al tempo* t*, e supponiamo che valgano le seguenti condizioni:*
1) il numero di eventi occorsi in due intervalli disgiunti di tempo sono stocasticamente indipendenti: se $s_1 < t_1 < s_2 < t_2$*, allora* $X(t_1) - X(s_1)$ *è indipendente da* $X(t_2) - X(s_2)$*;*
2) $X(s+t) - X(s)$ *dipende solo da* t *e non da* s *e da* $X(s)$ *(omogeneità);*
3) $X(0) = 0$ *(oppure* $P(X(0) = 0) = 1$*) ;*
4) dallo stato $j \in \mathbb{N}$ *si può passare direttamente soltanto nello stato* $j+1$*;*
5) la probabilità di effettuare un salto nell'intervallo di tempo $[t, t+h]$ *è uguale a* $\lambda h + o(h)$ *(*$\lambda > 0$ *è l'intensità del processo);*
6) la probabilità di effettuare due o più salti in $[t, t+h]$ *è uguale a* $o(h)$*.*

Allora, $X(t)$ risulta essere un processo di Poisson di intensità λ.

Dim. Per $n \geq 1$, poniamo $p_n(t) = P(X(t) = n)$; consideriamo l'evento che al tempo $t + h$ risulti $X(t + h) = n$; la sua probabilità è $p_n(t + h)$. L'evento in questione, $\{X(t + h) = n\}$ si può scrivere come unione disgiunta di tre eventi:
(i) al tempo t è $X(t) = n$ e nessun salto avviene tra t e $t + h$; la probabilità di questo evento è $p_n(t)p_0(h) = p_n(t)(1 - \lambda h) + o(h)$;
(ii) al tempo t è $X(t) = n - 1$ ed esattamente un salto avviene tra t e $t + h$; la probabilità di questo evento è $p_{n-1}(t)\lambda h + o(h)$;
(iii) al tempo t è $X(t) = j$ con $|n - j| > 1$ e nell'intervallo $[t, t + h]$ è avvenuto più di un salto; la probabilità di questo evento è $o(h)$.
Sommando i contributi (i), (ii), (iii), si ottiene:

$$p_n(t + h) = p_n(t)(1 - \lambda h) + p_{n-1}(t)\lambda h + o(h) \tag{2.6}$$

che si può riscrivere:

$$\frac{p_n(t + h) - p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t) + \frac{o(h)}{h} \tag{2.7}$$

Calcolando il limite per $h \to 0$, si ottiene che $p_n(t)$ è derivabile e per $n \geq 1$ * :

$$p_n'(t) = -\lambda p_n(t) + \lambda p_{n-1}(t) \tag{2.8}$$

E' ora facile verificare che le equazioni (2.8), con la condizione $p_0(0) = P(X(0) = 0) = 1$, sono verificate da:

$$p_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \tag{2.9}$$

Per esempio, $p_0(t) = e^{-\lambda t}$ (con la condizione $p_{-1}(t) = 0$) soddisfa (2.8), infatti $\frac{d}{dt}e^{-\lambda t} = -\lambda e^{-\lambda t}$; $p_1(t) = e^{-\lambda t}\lambda t$ soddisfa (2.8), infatti $\frac{d}{dt}(e^{-\lambda t}\lambda t) = -\lambda e^{-\lambda t}\lambda t + \lambda e^{-\lambda t}$, e così via.

∎

Osservazione 2.1 Per la (2.9) possiamo dire che $X(t) = N(t)$ è un processo di Poisson di intensità λ, ovvero $X(t)$ ha distribuzione di Poisson di parametro λt, e si ha $E(X(t)) = Var(X(t)) = \lambda t$.

Osservazione 2.2 Se $X(t)$ è un processo di Poisson (in base alla II formulazione), è facile riottenere che gli intertempi di arrivo $T_k,\ k = 1, \ldots,$ sono v.a. indipendenti ed esponenziali di parametro λ.

---

* in realtà, passando al limite per $h \to 0^+$ si mostra che esiste la derivata destra di $p_n(t)$; poi, sostituendo $t-h$ al posto di t, si ottiene $\frac{p_n(t)-p_n(t-h)}{h} = -\lambda p_n(t-h)+\lambda p_{n-1}(t-h)+\frac{o(h)}{h}$ ovvero $\frac{p_n(t-h)-p_n(t)}{-h} = -\lambda p_n(t-h)+\lambda p_{n-1}(t-h)+\frac{o(h)}{h}$ e per $h \to 0^+$ si ottiene che anche la derivata sinistra di $p_n(t)$ esiste e soddisfa la stessa equazione della derivata destra, dunque $p_n(t)$ è derivabile e soddisfa (2.8)

Infatti, per mostrare che T_k ha distribuzione esponenziale, basta osservare che:

$$P(T_k > t) = P(\tau_{k+1} - \tau_k > t) = P(X(\tau_k + t) = X(\tau_k)) =$$

$$= P(X(\tau_k + t) - X(\tau_k) = 0) =$$

(per l'omogeneità)

$$= P(X(t) = 0) = e^{-\lambda t}.$$

Dunque, $P(T_k \leq t) = 1 - e^{-\lambda t}$.
Mostriamo ora che le v.a. T_k sono indipendenti; proviamo ad esempio che T_1 e T_2 sono indipendenti. Fissiamo i tempi $0 < t_1 < t_1 + h < t_2 < t_2 + k$ con $h, k > 0$ opportunamente piccoli, in modo che $t_2 + k < \tau_3$. Allora:

$$P(t_1 < \tau_1 \leq t_1 + h, t_2 < \tau_2 \leq t_2 + k) =$$

$$= P(X(t_1) = 0, X(t_1 + h) - X(t_1) = 1, X(t_2) - X(t_1 + h) = 0, X(t_2 + k) - X(t_2) = 1) =$$

(per l'indipendenza degli incrementi e per l'omogeneità)

$$= P(X(t_1) = 0) \cdot P(X(h) = 1) \cdot P(X(t_2 - t_1 - h) = 0) \cdot P(X(k) = 1) =$$

$$= e^{-\lambda t_1} \cdot \lambda h e^{-\lambda h} \cdot e^{-\lambda(t_2 - t_1 - h)} \cdot \lambda k e^{-\lambda k} =$$

$$= e^{-\lambda t_1} \cdot \lambda^2 e^{-\lambda(t_2 - t_1)} \cdot e^{-\lambda k} hk$$

Dividendo per hk e passando al limite per $h, k \to 0$, si ottiene che la densità congiunta di (τ_1, τ_2) è:

$$f_{\tau_1, \tau_2}(t_1, t_2) = \lambda^2 e^{-\lambda t_2} \mathbf{1}_{\{0 < t_1 < t_2\}}(t_1, t_2)$$

Effettuiamo ora il cambio di variabili $T_1 = \tau_1,\ T_2 = \tau_2 - \tau_1$. Il determinante della matrice Jacobiana è uguale ad 1, per cui otteniamo che la densità congiunta di (T_1, T_2) è:

$$f_{T_1, T_2}(s_1, s_2) = f_{\tau_1, \tau_2}(s_1, s_1 + s_2) = \lambda^2 e^{-\lambda(s_1 + s_2)} \cdot \mathbf{1}_{\{s_1 > 0, s_2 > 0\}}(s_1, s_2) =$$

$$= \lambda e^{-\lambda s_1} \mathbf{1}_{\{s_1 > 0\}}(s_1) \cdot \lambda e^{-\lambda s_2} \mathbf{1}_{\{s_2 > 0\}}(s_2) = f_{T_1}(s_1) \cdot f_{T_2}(s_2),$$

pertanto T_1 e T_2 sono indipendenti.

Applicazione: se in $[0, T]$ si verifica un solo arrivo, allora il tempo di arrivo ha distribuzione uniforme in $[0, T]$.

Sia $T > 0$ fissato e supponiamo che si sia verificato in $[0, T]$ un solo arrivo, cioé $N(T) = 1$. Vogliamo studiare la distribuzione condizionata del tempo di questo arrivo, diciamo τ, ovvero $P(\tau \leq t | N(T) = 1),\ t \in [0, T]$. Si ha, per $0 \leq t \leq T$:

$$P(\tau \leq t | N(T) = 1) = \frac{P(\{\tau \leq t\} \cap \{N(T) = 1\})}{P(N(T) = 1)} =$$

$$\frac{P(\{N(t)=1\}\cap\{N(T)-N(t)=0\})}{P(N(T)=1)}=$$

(per l'indipendenza degli incrementi)

$$\frac{P(\{N(t)=1\}\cdot P(\{N(T)-N(t)=0\})}{P(N(T)=1)}=$$

(per l'omogeneità)

$$\frac{P(\{N(t)=1\}\cdot P(\{N(T-t)=0\})}{P(N(T)=1)}=\frac{\lambda t e^{-\lambda t}e^{-\lambda(T-t)}}{\lambda T e^{-\lambda T}}=\frac{t}{T},$$

il che vuol dire che τ è uniformemente distribuita in $[0,T]$.

2.2 Processi di nascita e morte

Mediante i *processi di nascita e morte* si può descrivere la dinamica di popolazioni, la diffusione di epidemie, etc.
Supponiamo che un dato sistema S possa ad ogni istante trovarsi in uno stato intero j e che l'insieme di tali stati sia finito o numerabile. Supponiamo ancora che ad ogni istante il sistema possa cambiare stato e che la probabibilità di passare dallo stato n al tempo t nello stato $n+1$ al tempo $t+h$ sia uguale a $\lambda_n h+o(h)$, mentre la probabilità che nello stesso intervallo di tempo passi nello stato $n-1$ sia $\mu_n h+o(h)$. Supponiamo inoltre che la probabilità che durante l'intervallo $(t,t+h)$ il sistema passi dallo stato n ad uno stato j con $|n-j|>1$ sia $o(h)$, cioé infinitesima di ordine superiore ad h. In queste ipotesi, segue che la probabilità che il sistema rimanga nello stato n nell'intervallo $(t,t+h)$ è $1-\lambda_n h-\mu_n h+o(h)$. Le quantità λ_n e μ_n dipendono da n ma non dal tempo t, nè dalla storia passata del sistema (markovianetà). Un proceso stocastico che rientra nello schema descritto prende il nome di *processo di nascita e morte*. Se il processo descrive l'evoluzione di una popolazione, la transizione dallo stato n allo stato $n+1$ significa che la popolazione è aumentata di una unità (nascita di un nuovo membro); la transizione da n a $n-1$ significa che la popolazione è diminuita di una unità (decesso di un membro). Se per ogni $n\geq 1$ si ha $\mu_n=0$, ovvero se sono proibite le transizioni $n\to n-1$, il processo prende il nome di processo di *pura nascita*, mentre se le transizioni proibite sono quelle $n\to n+1$, ossia se $\lambda_n=0, n=0,1,\dots,$ allora il processo si dice *processo di morte.* Il processo di Poisson è un *processo di pura nascita* con $\lambda_n=\lambda$ per ogni $n\geq 0$. Analogamente a quanto fatto nel caso del processo di Poisson, si possono trovare delle equazioni differenziali per $p_k(t)=P(X(t)=k)$, cioé la probabilità che il sistema si trovi nello stato k al tempo t. Ricaviamo ora tali equazioni; l'evento $\{X(t+h)=n\}$ si può scrivere come unione disgiunta di quattro eventi $E_1,\ E_2,\ E_3,\ E_4$:
$E_1:=\{X(t)=n$ e nessun salto è avvenuto nell'intervallo $[t,t+h]\}$, $E_2:=\{X(t)=n-1$ e nell'intervallo $[t,t+h]$ è avvenuto un salto in avanti$\}$, $E_3:=\{X(t)=n+1$ e nell'intervallo $[t,t+h]$ è avvenuto un salto indietro$\}$, $E_4:=\{X(t)=j$, con $|j-n|>1\}$.
Si ha:

$$P(E_1)=p_n(t)(1-\lambda_n h-\mu_n h)+o(h),\ P(E_2)=p_{n-1}(t)\lambda_{n-1}h+o(h),$$

$$P(E_3) = p_{n+1}(t)\mu_{n+1}h + o(h),\ \ P(E_4) = o(h).$$

Allora, per n≥ 1 :

$$p_n(t+h) = P(X(t+h) = n) = p_n(t)(1-\lambda_n h-\mu_n h)+p_{n-1}(t)\lambda_{n-1}h+p_{n+1}(t)\mu_{n+1}h+o(h),$$

da cui segue:

$$\frac{p_n(t+h) - p_n(t)}{h} = -(\lambda_n + \mu_n)p_n(t) + p_{n-1}(t)\lambda_{n-1} + p_{n+1}(t)\mu_n + \frac{o(h)}{h}$$

e quindi, passando al limite per $h \to 0$, si ottiene infine:

$$p'_n(t) = -(\lambda_n + \mu_n)p_n(t) + \lambda_{n-1}p_{n-1}(t) + \mu_{n+1}p_{n+1}(t).$$

Se $n = 0$, manca il termine $p_{n-1}(t)\lambda_{n-1}$, e si ha:

$$p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t).$$

In definitiva, si ottiene:

$$\begin{cases} p'_0(t) = -\lambda_0 p_0(t) + \mu_1 p_1(t) \\ p'_k(t) = -(\lambda_k + \mu_k)p_k(t) + \lambda_{k-1}p_{k-1}(t) + \mu_{k+1}p_{k+1}(t),\ \ k \geq 1 \end{cases} \tag{2.10}$$

In realtà, noi siamo interessati a trovare soluzioni stazionarie $p_k(t) = \pi_k$; allo scopo, basterà risolvere il sistema ottenuto uguagliando a zero le derivate $p'_k(t)$ e ponendo $p_k(t) = \pi_k$, cioé:

$$\begin{cases} -\lambda_0\pi_0 + \mu_1\pi_1 = 0 \\ -(\lambda_k + \mu_k)\pi_k + \lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1} = 0,\ \ k \geq 1 \end{cases} \tag{2.11}$$

L'equazione (2.11) è suscettibile di una rappresentazione delle transizioni di stato:

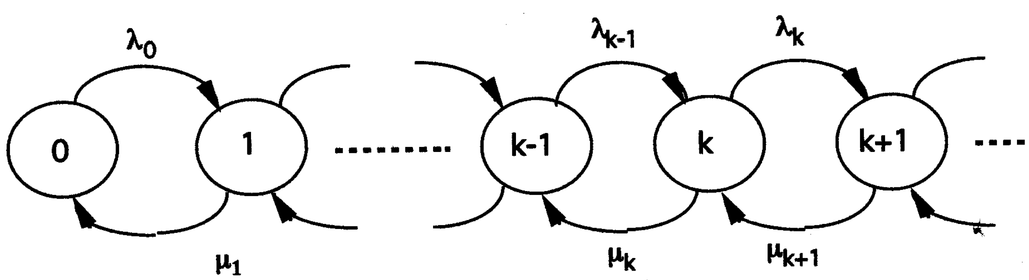

Fig. 9: Rappresentazione delle transizioni di stato di un processo di nascita e morte

Dalla (2.11) si può ricavare la relazione di bilancio globale:

$$(\lambda_k + \mu_k)\pi_k = \lambda_{k-1}\pi_{k-1} + \mu_{k+1}\pi_{k+1} \tag{2.12}$$

che si può interpretare, relativamente alla figura sopra, dicendo che in equilibrio statistico il *flusso di probabilità* che esce dallo stato k equivale al *flusso di probabilità* che entra nello stato k. Dal diagramma si ottiene:

$$\lambda_0 \pi_0 = \mu_1 \pi_1, \; \lambda_k \pi_k = \mu_{k+1} \pi_{k+1} \tag{2.13}$$

che sono le equazioni di bilancio locale, riguardani coppie di stati adiacenti.
Le equazioni (2.13) si possono risolvere esplicitamente; infatti, dalla prima equazione, si ricava $\pi_1 = \frac{\lambda_0}{\mu_1}\pi_0$; sostituendo $k = 1$ nella seconda delle (2.13), si ottiene $\pi_2 = \frac{\lambda_1}{\mu_2}\pi_1 = \frac{\lambda_1}{\mu_2}\,\frac{\lambda_0}{\mu_1}\pi_0$; sostituendo $k = 2$, si ricava $\pi_3 = \frac{\lambda_2}{\mu_3}\pi_2 = \frac{\lambda_2}{\mu_3}\,\frac{\lambda_1}{\mu_2}\,\frac{\lambda_0}{\mu_1}\pi_0$.
Dunque, in generale si ha, per $k = 1, 2, \dots$:

$$\pi_k = \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\;\pi_0.$$

Imponendo la condizione $\sum_{k=0}^{\infty} \pi_k = 1$, si ottiene:

$$\pi_0 = \left(1 + \sum_{k=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{k-1}}{\mu_1 \mu_2 \cdots \mu_k}\right)^{-1},$$

che, insieme alle equazioni per π_k determina esplicitamente π_k, per $k = 0, 1, \dots$.
Se λ_{k-1}/μ_k è costante, ovvero $\lambda_{k-1}/\mu_k = \rho$, si ottiene che $\pi_0 = \left(\sum_{k=0}^{\infty} \rho^k\right)^{-1}$; la serie (geometrica) è convergente solo se $\rho < 1$ e quindi questa è la condizione per l'esistenza della distribuzione stazionaria. Ad esempio, risulta $\lambda_{k-1}/\mu_k = \rho$, se $\lambda_i = \lambda$ e $\mu_i = \mu, \forall i \geq 0$. Per $\rho < 1$, calcolando la somma della serie, si trova $\pi_0 = (1/(1-\rho))^{-1} = 1 - \rho$, e quindi la distribuzione stazionaria è:

$$\pi_k = \rho^k (1-\rho), \; k = 0, 1, \dots;$$

si tratta di una distribuzione Geometrica di parametro $1 - \rho$.
Se λ_{k-1}/μ_k non è costante, ma $\lambda_{k-1}/\mu_k \leq c$, ove $c < 1$ è una costante, allora la serie che definisce il reciproco di π_0 converge per il criterio del confronto, ed esiste la distribuzione stazionaria π, anche se il calcolo di π_0, e quindi di π_k, è più complicato.

Le equazioni di Kolmogorov (2.10) diventano particolarmente semplici nel caso di un processo di pura nascita. Tale processo si ottiene dal processo di Poisson imponendo che il tasso di crescita dipenda dallo stato n, ovvero si sostituisce λ con λ_n, oppure dal processo di nascita e morte, ponendo il tasso di morte μ_n uguale a zero. Si ottengono in tal modo le equazioni:

$$\begin{cases} p_0'(t) = -\lambda_0 p_0(t) \\ p_k'(t) = -\lambda_k p_k(t) + \lambda_{k-1} p_{k-1}(t), \; k \geq 1 \end{cases} \tag{2.14}$$

Osservazione 2.3 Nel caso di un processo di nascita e morte in cui $\lambda_n = n\lambda$ e $\mu_n = n\mu$, se denotiamo con $X(t)$ lo stato del sistema al tempo t (popolazione in senso generalizzato) e $\Delta X(t) = X(t+h) - X(t)$, si ha:

$$\begin{cases} P(\Delta X(t) = -1 | X(t)) = \mu X(t) h + o(h) \\ P(\Delta X(t) = +1 | X(t)) = \lambda X(t) h + o(h) \end{cases} \tag{2.15}$$

Quindi:

$$E[X(t+h)|X(t)] = X(t) + (\lambda - \mu)hX(t) + o(h)$$

Passando alla media:

$$E(X(t+h)) = E[E(X(t+h)|X(t))] = E(X(t)) + (\lambda - \mu)E(X(t))h + o(h)$$

e dunque:

$$E(X(t+h)) - E(X(t)) = (\lambda - \mu)E(X(t))h + o(h) \tag{2.16}$$

Posto $M(t) = E(X(t))$, dividendo entrambi i membri di (2.16) per h e passando al limite per $h \to 0$, si ottiene:

$$M'(t) = (\lambda - \mu)M(t) \tag{2.17}$$

che ha soluzione $M(t) = X(0)e^{(\lambda-\mu)t}$. Se $\lambda = \mu$ la popolazione rimane a media costante per ogni t; se $\lambda > \mu$ la popolazione cresce in media, al variare di t; infine, se $\lambda < \mu$, la popolazione si estingue per $t \to \infty$.

2.3 Q-matrici e matrici di transizione

Definizione 2.2 *Una $Q-$matrice su uno spazio degli stati I è una matrice reale $Q = (q_{ij})$ tale che:*

$$\begin{cases} (i)\ q_{ii} \leq 0,\ \forall i \in I \\ (ii)\ q_{ij} \geq 0,\ \forall i \neq j,\ i,j \in I \\ (iii)\ \sum_j q_{ij} = 0,\ \forall i \in I \end{cases} \tag{2.18}$$

Per $i \neq j$ il valore q_{ij} rappresenta il *tasso di transizione dallo stato i allo stato j.*
Da (iii) segue che

$$-q_{ii} = \sum_{j \neq i} q_{ij} \tag{2.19}$$

Il valore $-q_{ii} \doteq q_i$ rappresenta il *tasso di uscita dallo stato i.*

Esempio 2.1 La matrice nulla:

$$Q = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \cdot & \cdot & & \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

è una $Q-$matrice.

Esempio 2.2 Se $I = \{1, 2\}$ (due stati), la generica $Q-$matrice ha la forma:

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

con $\alpha, \beta \geq 0$.

Definizione 2.3 (esponenziale di una matrice)
Si pone:

$$e^{tQ} = Id + \sum_{k=1}^{\infty} \frac{(tQ)^k}{k!} \tag{2.20}$$

ovvero

$$(e^{tQ})_{ij} = \sum_{k=0}^{\infty} \frac{t^k (Q^k)_{ij}}{k!},\ t \geq 0$$

ove $Q^0 = Id$ è la matrice identità.

La serie di sopra converge, infatti, se $\parallel A \parallel$ denota una norma della matrice A, si ha:

$$\parallel e^{tQ} \parallel = \parallel \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!} \parallel \leq \sum_{k=0}^{\infty} \frac{|t^k| \parallel Q \parallel^k}{k!} = e^{|t| ||Q||} < +\infty$$

Proprietà delle $Q-$ matrici

(I)

$$e^{tQ} \cdot e^{sQ} = e^{sQ} \cdot e^{tQ} = e^{(t+s)Q}, \; s, t \in \mathbb{R}^+ \tag{2.21}$$

In particolare $e^{-tQ} = \left(e^{tQ}\right)^{-1}$.

(II)

$$\frac{d}{dt} e^{tQ} = Q e^{tQ} = e^{tQ} Q, \; t \geq 0 \tag{2.22}$$

ovvero

$$\frac{d}{dt} \left(e^{tQ}\right)_{ij} = \left(Q e^{tQ}\right)_{ij} = \left(e^{tQ} Q\right)_{ij}$$

Inoltre:

$$\frac{d^n}{dt^n} e^{tQ} = Q^n e^{tQ} = e^{tQ} Q^n, \; n \in \mathbb{N}, \; \mathrm{t} \geq 0 \tag{2.23}$$

(III)

$$det\left(e^{tQ}\right) = e^{t \cdot TrQ} \tag{2.24}$$

dove det denota il *determinante* di una matrice e $TrQ = \sum_i q_{ii}$ denota la *traccia* di Q.

La (I) e (II) sono conseguenza del seguente:

Teorema 2.1 *Sia Q una $Q-$matrice finita, allora per la famiglia di matrici $P(t) = e^{tQ}, \; t \geq 0$, valgono le seguenti proprietà:*

(a)

$$P(t+s) = P(s)P(t), \; s, t \geq 0 \tag{2.25}$$

(proprietà di semigruppo)

(b) $P(t)$ è l'unica soluzione di:

$$\frac{d}{dt} P(t) = P(t)Q, \; t \geq 0 \text{ (forward equation)} \tag{2.26}$$

$$\frac{d}{dt} P(t) = QP(t), \; t \geq 0 \text{ (backward equation)} \tag{2.27}$$

con $P(0) = Id =$ matrice identità .
(c) per ogni intero $n \in \mathbb{N}$, si ha:

$$\frac{d^n}{dt^n} P(t) = P(t)Q^n = Q^n P(t) \tag{2.28}$$

In particolare:

$$\frac{d^n}{dt^n} P(t) \mid_{t=0} = Q^n \tag{2.29}$$

che, per $n = 1$, fornisce $P'(0) = Q$.

Dim.
(a) Dalla formula del binomio di Newton:

$$((t+s)Q)^k = (t+s)^k Q^k = \sum_{i=0}^{k} \binom{k}{i} t^i s^{k-i} Q^k = \sum_{i=0}^{k} \binom{k}{i} (tQ)^i (sQ)^{k-i}$$

Allora:

$$P(t+s) = e^{(t+s)Q} = \sum_{k=0}^{\infty} \frac{(t+s)^k Q^k}{k!} =$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{i=0}^{k} \binom{k}{i} (tQ)^i (sQ)^{k-i} \right) =$$

$$= \sum_{k=0}^{\infty} \sum_{i=0}^{k} \frac{1}{i!(k-i)!} (tQ)^i (sQ)^{k-i}$$

Posto $k - i = h$, l'ultima sommatoria diventa:

$$\sum_{h=0}^{\infty} \sum_{i=0}^{\infty} \frac{1}{i!h!} (tQ)^i (sQ)^h =$$

$$= \left(\sum_{h=0}^{\infty} \frac{(sQ)^h}{h!} \right) \left(\sum_{i=0}^{\infty} \frac{(tQ)^i}{i!} \right) = e^{sQ} e^{tQ}$$

Siccome $P(t+s) = P(s+t)$, si ottiene:

$$e^{(t+s)Q} = e^{sQ} e^{tQ} = P(s+t) = e^{tQ} e^{sQ}.$$

(b) Si ha:

$$\frac{d}{dt} P(t) = \frac{d}{dt} \left(\sum_{k=0}^{\infty} \frac{t^k Q^k}{k!} \right) = Q \sum_{k=1}^{\infty} \frac{t^{k-1} Q^{k-1}}{(k-1)!} = QP(t),$$

ma è anche uguale a

$$\left(\sum_{k=1}^{\infty} \frac{t^{k-1} Q^{k-1}}{(k-1)!} \right) Q = P(t)Q.$$

Osservazione 2.4 L'equazione (b) vale anche con $-t$ al posto di t; pertanto:

$$\frac{d}{dt} e^{-tQ} = -e^{-tQ} Q = -Q e^{-tQ}$$

La condizione iniziale $P(0) = Id$ è banalmente verificata:

$$P(0) = e^{0 \cdot Q} = Id + \sum_{k=1}^{\infty} \frac{(0 \cdot Q)^k}{k!} = Id$$

Per mostrare l'unicità della soluzione dell'equazione differenziale, sia $Z(t)$ una qualunque soluzione di:

$$\begin{cases} \frac{d}{dt}Z(t) = Z(t)Q \\ Z(0) = Id \end{cases}$$

Allora:

$$\frac{d}{dt}\left(Z(t)e^{-tQ}\right) = \left(\frac{d}{dt}Z(t)\right)e^{-tQ} + Z(t)\frac{d}{dt}e^{-tQ} =$$

$$= Z(t)Qe^{-tQ} + Z(t)\left(-Qe^{-tQ}\right) = 0$$

Dunque $Z(t)e^{-tQ}$ non dipende da t, è costante, ed è uguale a $Z(0)e^{-0\cdot Q} = Z(0) = Id$. Abbiamo quindi ottenuto $Z(t)e^{-tQ} = Id$, $\forall t \geq 0$, ovvero $Z(t) = e^{tQ}$. Lo stesso argomento si può appicare per l'equazione differenziale:

$$\frac{d}{dt}Z(t) = QZ(t).$$

∎

Osservazione 2.5 Se Q è finita, la proprietà di semigruppo

$$P(t+s) = P(s)P(t) = P(t)P(s)$$

vale $\forall s, t \in \mathbb{R}$.

Teorema 2.2 *Se Q è una matrice finita, allora $P(t) = e^{tQ}$ è una matrice stocastica $\forall t \geq 0$ se e solo se Q è una $Q-$matrice.*

Dim. ($\Rightarrow$)
Sia $P(t) = e^{tQ}$ una matrice stocastica $\forall t \geq 0$. Osserviamo che:

$$\lim_{t\to 0^+} \frac{P(t) - P(0)}{t} = \lim_{t\to 0^+} \frac{e^{tQ} - Id}{t} =$$

$$= \lim_{t\to 0^+} \left(\sum_{i=1}^{\infty} \frac{t^i Q^i}{i!}\right) \cdot \frac{1}{t} = Q$$

Dunque, se $i \neq j$, $p_{ij}(t) = q_{ij}t + o(t)$, per $t \to 0^+$, mentre $p_{ii}(t) - 1 = q_{ii}t + o(t)$.
Allora, siccome $\sum_j p_{ij}(t) = 1$ (P è stocastica), si ha:

$$0 = \frac{d}{dt}\sum_j p_{ij}(t) = \sum_j \frac{d}{dt}p_{ij}(t) = \sum_j (QP(t))_{ij}$$

Pertanto, per $t = 0$, visto che $P(0) = Id$, si ottiene $\sum_j q_{ij} = 0$. Inoltre, dalla relazione $0 \leq p_{ij}(t) = q_{ij}t + o(t)$, per $i \neq j$, segue che $q_{ij} \geq 0$, $i \neq j$; da $0 \geq p_{ii}(t) - 1 = q_{ii}t + o(t)$ segue che $q_{ii} \leq 0$. Dumque, concludiamo che Q è una $Q-$matrice.

($\Leftarrow$)
Sia Q una $Q-$matrice; sappiamo che la somma degli elementi su ogni riga deve valere zero, ovvero $\sum_j q_{ij} = 0$. Osserviamo che questa proprietà si conserva per le potenze di Q, ovvero, per ogni intero n, anche la somma degli elementi di ciascuna riga della matrice Q^n vale zero. Per esempio, per $n = 2$, $(Q^2)_{ij} = \sum_k q_{ik}q_{kj}$, per cui:

$$\sum_j (Q^2)_{ij} = \sum_j \sum_k q_{ik}q_{kj} = \sum_k q_{ik} \sum_j q_{kj} = 0 \cdot 0 = 0$$

Procedendo per induzione, la cosa si dimostra per ogni intero n.
Cominciamo col dimostrare la tesi per $P(h)$, dove $h > 0$ è sufficientemente piccolo, ovvero che $P(h) = e^{hQ}$ è una matrice stocastica, se h è abbastanza piccolo. Tenendo presenti le proprietà di Q, dal fatto che $P(h) = Id + hQ + o(h)$, per $h \to 0^+$, segue che, per $h > 0$ sufficientemente piccolo:
(i) $p_{ii}(h) = 1 + hq_{ii} + o(h) > 0$
(ii) per $i \neq j$, $p_{ij}(h) = hq_{ij} + o(h) \geq 0$
Infatti, ciò è immediato, se $q_{ij} > 0$; se invece $q_{ij} = 0$, allora risulta $p_{ij}(h) = \frac{1}{2}h^2(Q^2)_{ij} + o(h^2)$. Osserviamo che $(Q^2)_{ij} = \sum_k q_{ik}q_{kj} \geq 0$; se $(Q^2)_{ij} > 0$ si vede subito che (ii) vale. Nel caso in cui fosse $(Q^2)_{ij} = 0$, si passerebbe a considerare Q^3, fino a trovare una potenza n di Q per cui $(Q^n)_{ij} > 0$, concludendo che vale (ii), visto che risulterebbe $p_{ij}(h) = \frac{1}{n!}h^n(Q^n)_{ij} + o(h^n) > 0$. Infine, se dovesse accadere che $(Q^n)_{ij} = 0$ per ogni n, si avrebbe $p_{ij}(h) = 0$. In conclusione (ii) è verificata in tutti i casi.

Si ha noltre:

$$P(h) = e^{hQ} = Id + \sum_{k=1}^{\infty} \frac{h^k}{k!} Q^k$$

e quindi:

(iii) $$\sum_j p_{ij}(h) = \sum_j \delta_{ij} + \sum_k \frac{h^k}{k!} \sum_j (Q^k)_{ij} = 1,$$

visto che $\sum \delta_{ij} = 1$ e $\sum_j (Q^k)_{ij} = 0$.
Allora, da (i), (ii) e (iii) segue che $P(h)$ è una matrice stocastica.
Per concludere la dimostrazione, occorre provare che $P(t) = e^{tQ}$ è stocastica per ogni $t > 0$. Posto $h = t/n$, risulta che h si può rendere piccolo come si vuole, prendendo n opportunamente grande. Per tale n, risulta allora che $P(h) = P(t/n)$ è stocastica. Ma, per la proprietà di semigruppo:

$$P(t) = P\left(\frac{t}{n} + \frac{t}{n} + \ldots + \frac{t}{n}\right) = P\left(\frac{t}{n}\right) P\left(\frac{t}{n}\right) \cdots P\left(\frac{t}{n}\right) = \left(P\left(\frac{t}{n}\right)\right)^n = (P(h))^n$$

che è una matrice stocastica, essendo la potenza ennesima della matrice stocastica $P(h)$.

$\blacksquare$

Osservazione 2.6
Dalla dimostrazione fatta segue che $p_{ii}(h)$ è sempre positivo, mentre per $i \neq j$, $p_{ij}(h)$ è positivo, qualora esista un intero $n \geq 1$ per cui risulti $(Q^n)_{ij} > 0$. Dunque, se è soddisfatta quest'ultima condizione, la matrice $P(h) = e^{hQ}$ è regolare (addirittura risulta $P(h) > 0$).

Osservazione 2.7
$P(t) = e^{tQ}$ è bistocastica se e solo se la somma degli elementi su ciascuna colonna di Q è zero (la somma degli elementi su ciascuna riga è già zero per le proprietà di Q).

Dim. ($\Rightarrow$)
Se $\sum_i p_{ij}(t) = 1$, allora:

$$0 = \frac{d}{dt}\sum_i p_{ij}(t) = \sum_i \frac{d}{dt} p_{ij}(t) = \sum_i (QP(t))_{ij}$$

e quindi, per $t = 0$, visto che $P(0) = Id$, si ottiene $\sum_i Q_{ij} = 0$.

($\Leftarrow$)
Se $\sum_i Q_{ij} = 0$, allora per ogni intero n risulta $\sum_i (Q^n)_{ij} = 0$. Ad esempio, per $n = 2$:

$$\sum_i (Q^2)_{ij} = \sum_i \sum_k q_{ik} q_{kj} = \sum_k q_{kj} \left(\sum_i q_{ik} \right) = 0$$

(per $n > 2$, si procede per induzione).
Allora, poiché $P(t) = e^{tQ} = Id + \sum_{k=1}^{\infty} \frac{t^k}{k!} Q^k$, si ha:

$$\sum_i p_{ij}(t) = \sum_i \delta_{ij} + \sum_{k=1}^{\infty} \frac{t^k}{k!} \sum_i (Q^k)_{ij} = 1,$$

visto che $\sum_i (Q^k)_{ij} = 0$.

Osservazione 2.8
Se γ è autovalore di Q, allora $e^{t\gamma}$ è autovalore di $P(t)$. In particolare, e^{γ} è autovalore di $P(1)$.
Infatti, se γ è autovalore di Q esiste un vettore $u \neq 0$ tale che $Qu = \gamma u$ (ricordiamo che Q ha sempre un autovalore $\gamma_1 = 0$). Allora:

$$P(t)u = e^{tQ} u = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k u = \sum_{k=0}^{\infty} \frac{t^k}{k!} \gamma^k u = e^{t\gamma} u$$

e quindi $e^{t\gamma}$ è autovalore di $P(t)$.
Dunque, gli autovalori di $P(1)$ (eccetto $\lambda_1 = 1$) hanno modulo minore di 1, se gli autovalori di Q (eccetto $\gamma_1 = 0$) hanno parte reale negativa. Inoltre, visto che $detP(t) = e^{tTrQ} > 0$, $P(t)$ non è degenere.

Esempio 2.3 ($Q-$ matrice 2×2)
Sia, per $\alpha,\beta\geq 0$:

$$Q=\begin{pmatrix}-\alpha & \alpha\\ \beta & -\beta\end{pmatrix}$$

Gli autovalori di Q sono $\lambda_1=0$ e $\lambda_2=-(\alpha+\beta)$. Q è diagonalizzabile e risulta $Q=U\Lambda U^{-1}$, ove

$$\Lambda=\begin{pmatrix}0 & 0\\ 0 & -(\alpha+\beta)\end{pmatrix}.$$

Dunque:

$$\begin{aligned}P(t)&=\sum_{k=0}^{\infty}\frac{t^k}{k!}Q^k=\sum_{k=0}^{\infty}\frac{t^k}{k!}U\Lambda^kU^{-1}=\\&=U\sum_{k=0}^{\infty}\frac{t^k}{k!}\Lambda^kU^{-1}=Ue^{t\Lambda}U^{-1}=\\&=U\begin{pmatrix}1 & 0\\ 0 & e^{-t(\alpha+\beta)}\end{pmatrix}U^{-1}\end{aligned}$$

Se $\alpha=\beta=0$, allora $P(t)=Id$. Altrimenti $p_{ij}(t)=A_{ij}+B_{ij}e^{-(\alpha+\beta)t}$ con:

$$\begin{cases}p_{ij}(0)=A_{ij}+B_{ij}=\delta_{ij}\\ \frac{d}{dt}p_{ij}(t)\mid_{t=0}=q_{ij}=-(\alpha+\beta)B_{ij}\end{cases}.$$

Per esempio, per $i=j=1$, si ha:

$$\begin{cases}A_{11}+B_{11}=1\\ -\alpha=q_{11}=-(\alpha+\beta)B_{11}\end{cases}$$

da cui segue:

$$A_{11}=\frac{\beta}{\alpha+\beta},\ B_{11}=\frac{\alpha}{\alpha+\beta}$$

Effettuando i calcoli per ogni i e j, si trova:

$$P(t)=\begin{pmatrix}\frac{\beta}{\alpha+\beta}+\frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta}-\frac{\alpha}{\alpha+\beta}e^{-(\alpha+\beta)t}\\ \frac{\beta}{\alpha+\beta}-\frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t} & \frac{\alpha}{\alpha+\beta}+\frac{\beta}{\alpha+\beta}e^{-(\alpha+\beta)t}\end{pmatrix}$$

Si noti che, per $t\to+\infty$, $P(t)$ tende a:

$$\begin{pmatrix}\frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta}\end{pmatrix}$$

Esempio 2.4 (autovalori multipli)

$$Q=\begin{pmatrix}-2 & 1 & 1 & 0\\ 1 & -2 & 1 & 0\\ 0 & 1 & -2 & 1\\ 1 & 0 & 1 & -2\end{pmatrix}$$

Gli autovalori sono $\lambda_1 = 0,\ \lambda_2 = -2,\ \lambda_3 = \lambda_4 = -3$. La molteplicità geometrica di ciascun autovalore è 1. Se $P(t) = e^{tQ}$, si ha per $i, j = 1, 2, 3, 4$:

$$(P(t))_{ij} = p_{ij}(t) = A_{ij} + B_{ij}e^{-2t} + (C_{ij} + D_{ij}t)e^{-3t}$$

Osservazione 2.9

1. Una $Q-$ matrice è sempre degenere, cioé ha determinante zero, e quindi almeno un autovalore zero. Ciò è dovuto al fatto che una riga di Q è banalmente somma di altre due righe, visto che la somma degli elementi su ciascuna riga è zero.

2. Il quadrato Q^2 di una $Q-$ matrice non fornisce necessariamente una $Q-$ matrice, a meno che Q non sia la matrice nulla. Infatti, se per $\alpha, \beta \geq 0$:

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

si ha:

$$Q^2 = \begin{pmatrix} \alpha^2 + \alpha\beta & -\alpha^2 - \alpha\beta \\ -\alpha\beta - \beta^2 & \alpha\beta + \beta^2 \end{pmatrix}$$

La somma sulle righe è zero (ciò è vero anche per Q^n); però, affinché gli elementi diagonali siano ≤ 0, deve essere $\alpha = \beta = 0$, ovvero $Q = 0$.

3. Per il Teorema 2.2, data una $Q-$ matrice, risulta che $P(t) = e^{tQ}$ è una matrice stocastica, il che vuol dire che esiste una CM discreta che ha $P(t)$ come matrice delle probabilità di transizione.

4. Non è vero che ogni matrice di transizione P può essere scritta come e^{tQ} per qualche $Q-$ matrice e qualche $t \geq 0$. Ad esempio,

$$P_1 = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}, \ P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

hanno determinante ≤ 0, mentre $e^{t \cdot trQ} > 0$, il che contraddice la proprietà (III), ovvero $det(e^{tQ}) = e^{t \cdot trQ}$.

A proposito, la proprietà (III) (2.24) è facilmente dimostrabile nel caso in cui Q è diagonalizzabile. Infatti, in tal caso $Q = U\Lambda U^{-1}$, dove Λ è la matrice diagonale degli autovalori di Q. Allora $e^{tQ} = Ue^{t\Lambda}U^{-1}$ e quindi

$$dete^{tQ} = detU \cdot dete^{t\Lambda} \cdot detU^{-1} = dete^{t\Lambda} =$$

$$= det(diag(1, e^{t\lambda_2}, \ldots, e^{t\lambda_n}) = e^{t\sum_i \lambda_i} = e^{t \cdot trQ},$$

visto che la somma degli autovalori di una matrice è sempre uguale alla sua traccia.

2.4 Catene di Markov con spazio degli stati discreto, tempo continuo ed omogenee

Definizione *Una CM a stati discreti, tempo continuo ed omogenea, con* $Q-$ *matrice* Q *e distribuzione iniziale* ν *è una famiglia di v.a.* $\{X_t,\ t \geq 0\}$ *a valori in un insieme discreto* I *tale che:*
(a) $P(X_0 = i) = \nu_i,\ i \in I$
(b) $\forall\ 0 < t_1 < t_2 < \ldots, t_n$ *e stati* $i_0, i_1, \ldots, i_n \in I$, *si ha:*

$$P\left(X_0 = i_0, X_{t_1} = i_1, \ldots, X_{t_n} = i_n\right) = \nu_{i_0} p_{i_0 i_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1})$$

dove $p_{ij}(t)$ *sono gli elementi della matrice* $P(t) = e^{tQ}$.

La matrice Q è chiamata *generatore* della CM a tempo continuo X_t. Più precisamente, X_t è detta $(\nu, Q)-$ CM a tempo continuo.

Proprietà 1 (omogeneità)
La matrice $P(t) = (p_{ij}(t))$ è detta matrice delle probabilità di transizione al tempo t :

$$p_{ij}(t) = P(X_t = j | X_0 = i) = P(X_{t+s} = j | X_s = i) \tag{2.30}$$

Proprietà 2 (perdita di memoria)

$$P(X(t_n) = j | X(t_{n-1}) = i, X(t_{n-2}) = i_{n-2}, \ldots, X_0 = i_0) =$$

$$= P(X(t_n) = j | X(t_{n-1}) = i) = p_{ij}(t_n - t_{n-1}) \tag{2.31}$$

Proprietà 3

$$P(X_t = j) = \sum_i \nu_i p_{ij}(t) = (\nu P(t))_j \tag{2.32}$$

Definizione 2.4
Diciamo che $\nu = \pi$ è una *distribuzione invariante* se $P(X_t = j) = \pi_j,\ \forall t \geq 0$, i.e. $\pi P(t) = \pi$. Allora:

$$0 = \frac{d}{dt}(\pi P(t)) = \pi \frac{d}{dt} P(t) = \pi P(t) Q$$

che, per $t = 0$, fornisce:

$$\pi Q = 0 \tag{2.33}$$

ovvero, se la distribuzione π è invariante, π è autovettore sinistro di Q relativo all'autovalore 0.
D'altra parte, se $\pi Q = 0$, allora $\pi P(t) = \pi,\ \forall t > 0$; infatti:

$$\pi P(t) = \pi e^{tQ} = \pi \left[Id + \sum_{k=1}^{\infty} \frac{t^k Q^k}{k!} \right] = \pi + \sum_{k=1}^{\infty} \frac{t^k}{k!} (\pi Q^k) = \pi,$$

poiché $\pi Q^k = 0$, per ogni $k \geq 1$.
In conclusione, π è invariante se e solo se $\pi Q = 0$.

Proprietà 4
Se esiste $\lim_{t\to\infty} P(t) = \Pi$, allora, utilizzando il fatto che $P(t+s) = P(s)P(t)$:

$$\Pi = \lim_{s\to\infty} P(t+s) = \left(\lim_{s\to\infty} P(s)\right) P(t) = \Pi P(t) \tag{2.34}$$

In tal caso esiste la *distribuzione stazionaria* che è anche invariante; la matrice Π ha le righe tutte uguali e in ciascuna riga vi è il vettore delle probabilità stazionarie $\pi = (\pi_1, \ldots, \pi_n)$, che è anche soluzione dell'equazione $\pi Q = 0$. Si noti che risulta $\pi = \pi P(t)$ $\forall t \geq 0$.

Esempio 2.5 Sia:

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

Se $\alpha + \beta > 0$, allora la distribuzione stazionaria ed invariante è

$$\pi = \left(\tfrac{\beta}{\alpha+\beta} \ , \quad \tfrac{\alpha}{\alpha+\beta} \right)$$

Invece, se $\alpha = \beta = 0$, si ha $Q = 0$, $P(t) = Id$, ed ogni distribuzione è invariante.
Vale il seguente:

Teorema 2.3 *Sia X_t una CM discreta a tempo continuo con generatore Q, allora, se $X_t = i$, posto $q_i = -q_{ii} > 0$, il tempo di attesa residuo R_i nello stato i ha distribuzione esponenziale di parametro q_i, ovvero:*

$$P(R_i \leq \tau | X_t = i) = 1 - e^{-q_i \tau}$$

o anche:

$$P(R_i \geq \tau | X_t = i) = P(X_{t+s} = i, \ \forall\ 0 \leq s < \tau | X_t = i) = e^{-q_i \tau}$$

Inoltre, al tempo $t + R_i$ la CM passa nello stato $j \neq i$ con probabilità $\hat{p}_{ij} = q_{ij}/q_i$, cioé:

$$P(X_{t+R_i} = j | X_t = i) = \frac{q_{ij}}{q_i}$$

Posto $\hat{p}_{ii} = 0$ e $\hat{p}_{ij} = q_{ij}/q_i$, è immediato verificare che la matrice $\hat{P}$ tale che $(\hat{P})_{ij} = \hat{p}_{ij}$ è stocastica. Infatti, risulta $\hat{p}_{ij} \geq 0$, $\hat{p}_{ii} = 0$; inoltre, per la (2.19):

$$\sum_j \hat{p}_{ij} = \sum_{j\neq i} \frac{q_{ij}}{-q_{ii}} = -\frac{1}{q_{ii}} \sum_{j \neq i} q_{ij} = -\frac{1}{q_{ii}} \cdot (-q_{ii}) = 1$$

Osservazione 2.10 Se per uno stato i risulta $q_i = 0$, allora $\forall \tau$ si ha $P(R_i \geq \tau | X_t = i) = 1$, il che vuol dire che la CM, una volta raggiunto lo stato i, vi rimane per sempre, i è uno stato *assorbente* (naturalmente, in tal caso anche q_{ij} deve essere zero).

Esempio 2.6 Un virus si presenta sotto la forma di $N+1$ diverse specie mutanti (varianti) $0, 1, \dots, N$. Esso permane nella stessa specie per un tempo aleatorio con distribuzione esponenziale di parametro $\lambda > 0$, dopodiché avviene una mutazione in una delle rimanenti specie, con uguale probabilità. Calcoliamo la probabilità che la specie al tempo t sia la stessa di quella al tempo 0.
Intanto, si ha $q_i = -q_{ii} = \lambda$, $\hat{p}_{ij} = q_{ij}/q_i = q_{ij}/\lambda$ per cui $q_{ij} = \lambda \hat{p}_{ij}$. Siccome una delle altre specie viene scelta uniformemente con la stessa probabilità, deve essere $\hat{p}_{ij} = 1/N$ e quindi $q_{ij} = \lambda/N,\ 1 \le i, j \le N+1,\ i \ne j$. Dunque, si ottiene:

$$Q = \begin{pmatrix} -\lambda & \lambda/N & \dots & \lambda/N \\ \lambda/N & -\lambda & \dots & \lambda/N \\ \dots & \dots & \dots & \dots \\ \lambda/N & \lambda/N & \dots & -\lambda \end{pmatrix}$$

Supponiamo ora, per semplicità, che $N = 1$ (2 specie mutanti); allora

$$Q = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}$$

Ricordando l'Esempio 2.3, si ottiene:

$$P(t) = e^{tQ} = \begin{pmatrix} \frac{1}{2} + \frac{1}{2} e^{-2\lambda t} & \frac{1}{2} - \frac{1}{2} e^{-2\lambda t} \\ & \\ \frac{1}{2} - \frac{1}{2} e^{-2\lambda t} & \frac{1}{2} + \frac{1}{2} e^{-2\lambda t} \end{pmatrix} \to \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ & \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \text{ (per } t \to \infty\text{)}$$

ovvero $p_{ij}(t) \to 1/2,\ i, j = 0, 1$ per $t \to \infty$, avendosi distribuzione stazionaria $\pi = (1/2, 1/2)$.

Ritorniamo ora a considerare il processo di Poisson/nascita e quello di nascita e morte, già visti precedentemente. Vogliamo illustrare come essi possano essere visti come CM a tempo continuo e trovarne esplicitamente il generatore Q. Iniziamo dal processo di nascita; ricordiamo che deve aversi:

$$\begin{gathered} P(X(t+h) = k+1 | X(t) = k) = \lambda_k h + o(h) \\ P(X(t+h) = k | X(t) = k) = 1 - \lambda_k h + o(h) \\ P(X(t+h) = m | X(t) = k) = o(h) \text{ se } |m-k| \ge 2 \end{gathered} \tag{2.35}$$

Calcoliamo ora il generatore:

$$Q = \lim_{t \to 0^+} \frac{P(t) - Id}{t} = \lim_{t \to 0^+} \frac{e^{tQ} - Id}{t} = \lim_{t \to 0^+} \left(\sum_{i=1}^{\infty} t^i Q^i \right) \cdot \frac{1}{t} \tag{2.36}$$

Si ha:

(i)
$$\lim_{h \to 0^+} \frac{p_{kk+1}(h) - \delta_{kk+1}}{h} = \lim_{h \to 0^+} \frac{\lambda_k h}{h} = \lambda_k$$

(ii)
$$\lim_{h\to 0^+} \frac{p_{kk}(h)-\delta_{kk}}{h} = \lim_{h\to 0^+} \frac{1-\lambda_k h - 1}{h} = -\lambda_k$$

(iii)
$$\lim_{h\to 0^+} \frac{p_{kk-1}(h)-\delta_{kk-1}}{h} = 0$$

(iv)
$$\lim_{h\to 0^+} \frac{p_{ki}(h)-\delta_{ki}}{h} = 0 \text{ se } |k-i|>1$$

Riasssumendo, per $i, j = 0, 1, \ldots,$:

$$(Q)_{ij} = q_{ij} = \begin{cases} \lambda_i & \text{se } j = i+1 \\ -\lambda_i & \text{se } i = j \\ 0 & \text{altrimenti} \end{cases} \tag{2.37}$$

e, per esteso:

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & -\lambda_1 & \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & -\lambda_2 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \tag{2.38}$$

È immediato verificare che $\sum_j q_{ij} = 0$ e $q_{ii} = -\lambda_i$, dove λ_i rappresenta il reciproco del tempo medio di permanenza nello stato i, prima di andare in $i+1$. Inoltre, da $p_{ij}(h) = (Id + hQ)_{ij} + o(h)$, si ottiene, ad esempio, come verifica:

$$p_{11}(h) = 1 + hq_{11} + o(h) = 1 - \lambda_1 h + o(h)$$

$$p_{12}(h) = 0 + h\lambda_1 + o(h) = h\lambda_1 + o(h)$$

$$p_{13}(h) = o(h), \; p_{21}(h) = hq_{21} + o(h) = o(h)$$

$$p_{22}(h) = 1 - \lambda_2 h + o(h), \; p_{23}(h) = 0 + \lambda_2 h + o(h), \quad \text{etc.}$$

Per il processo di Poisson, si può ripetere il discorso precedente, sostituendo per ogni k, λ_k con λ.

Si può verificare che, in tal caso, risulta:

$$P(t) = e^{tQ} = \begin{pmatrix} e^{-\lambda t} & \lambda t e^{-\lambda t} & \frac{(\lambda t)^2}{2} e^{-\lambda t} & \ldots & \ldots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \ldots & \ldots \\ 0 & 0 & e^{-\lambda t} & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} =$$

$$= \begin{pmatrix} \textit{Distribuzione di Poisson}(\lambda t) \\ 0 \;\; \textit{Distribuzione di Poisson}(\lambda t) \\ 0\; 0 \;\; \textit{Distrib. di Poisson}(\lambda t) \\ \ldots\ldots \end{pmatrix} \tag{2.39}$$

Per quanto riguarda il processo di nascita e morte, è facile vedere che il generatore Q è dato da $(i, j = 0, 1, \ldots)$:

$$(Q)_{ij} = q_{ij} = \begin{cases} \lambda_i & \text{se } j = i+1 \\ \mu_i & \text{se } j = i-1 \\ -(\lambda_i + \mu_i) & \text{se } i = j \\ 0 & \text{altrimenti} \end{cases} \tag{2.40}$$

Ad esempio,

$$q_{kk} = \lim_{h \to 0^+} \frac{p_{kk}(h) - 1}{h} = \lim_{h \to 0^+} \frac{1 - (\lambda_k + \mu_k)h - 1}{h} = -(\lambda_k + \mu_k)$$

Riassumendo:

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \ldots & \ldots & \ldots & 0 \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \end{pmatrix} \tag{2.41}$$

Osservazione 2.11
Da quanto visto evince che:

• il generatore Q ($Q-$ matrice) è una matrice di *tassi*, non di probabilità .
• gli elementi sulla diagonale principale di Q sono gli opposti dei reciproci dei tempi medi di attesa residui negli stati, ovvero:

$$-q_{ii} = q_i = \eta_i$$

dove il tempo di permanenza nello stato i è una variabile aleatoria con distribuzione esponenziale di parametro η_i.
• fuori della diagonale, vi sono i tassi di passaggio nei vari stati.

Infatti, riferendoci per esempio al processo di nascita e morte, vediamo che:
• da uno stato $i \geq 1$ il sistema può passare nello stato $i-1$ (morte) oppure nello stato $i+1$ (nascita); se T_1 rappresenta il tempo di attesa prima di un evento "morte", allora T_1 ha distribuzione esponenziale di parametro μ_i; analogamente, se T_2 rappresenta il tempo di attesa prima di un evento "nascita", allora T_2 ha distribuzione esponenziale di parametro λ_i. Quindi, il tempo di attesa nello stato $i \geq 1$ è $T_i = \min(T_1, T_2)$, che ha distribuzione esponenziale di parametro $q_i = -q_{ii} = \lambda_i + \mu_i$.
Se invece $i = 0$, essendo impossibile un evento "morte", si trova che T_i ha distribuzione esponenziale di parametro $q_0 = -q_{00} = \lambda_0$.

Trascorso il tempo T_i di attesa nello stato i :
- se $i = 0$, il sistema passa nello stato 1 con probabilità $\hat{p}_{01} = q_{01}/(-q_{00}) = \lambda_0/\lambda_0 = 1$;
- se $i \geq 1$, il sistema passa nello stato $i-1$ con probabilità $\hat{p}_{i\ i-1} = q_{i\ i-1}/(-q_{ii}) = \mu_i/(\lambda_i + \mu_i)$, e nello stato $i+1$ con probabilità $\hat{p}_{i\ i+1} = q_{i\ i+1}/(-q_{ii}) = \lambda_i/(\lambda_i + \mu_i)$.

Si ottiene quindi, per $i \neq j$, $\hat{p}_{ij} = q_{ij}/(-q_{ii})$ e la matrice $\hat{P}$ assume la forma:

$$\hat{P} = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ \frac{\mu_1}{\lambda_1+\mu_1} & 0 & \frac{\lambda_1}{\lambda_1+\mu_1} & 0 & \dots & \dots & 0 \\ 0 & \frac{\mu_2}{\lambda_2+\mu_2} & 0 & \frac{\lambda_2}{\lambda_2+\mu_2} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

La matrice $\hat{P}$ descrive una CM "accelerata", nel senso che si trascura il tempo di permanenza nei vari stati, come fosse un random walk con la barriera di sinistra rifettente.

Esercizio 2.1 È data la $Q-$ matrice:

$$Q = \begin{pmatrix} -1/2 & 1/2 \\ 0 & 0 \end{pmatrix}$$

Si consideri la CM a tempo continuo con spazio degli stati $\{1, 2\}$ associata a Q e si determinino la matrice di transizione $P(t) = e^{tQ}$ e, se esiste, la distribuzione stazionaria per la CM.

Soluzione. Lo stato 2 è assorbente, poiché $q_{22} = 0$. Gli autovalori di Q sono 0 e $-1/2$ e quindi Q si diagonalizza nel seguente modo:

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -1/2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Dunque:

$$e^{tQ} = P(t) = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{-t/2} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} =$$

$$\begin{pmatrix} e^{-t/2} & 1 - e^{-t/2} \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \text{ per } t \to \infty$$

Si ha:

$$P(1) = \begin{pmatrix} e^{-1/2} & 1 - e^{-1/2} \\ 0 & 1 \end{pmatrix}$$

con autovalori 1 e $e^{-1/2}$; il fatto che $(P(1))_{22} = 1$ conferma che 2 sia stato assorbente.

Calcolando l'autovettore sinistro di $P(1)$ corrispondente all'autovalore 1, si trova $(\pi_1, \pi_2) = (0, 1)$, che conferma la distribuzione stazionaria trovata come $\lim_{t\to\infty} P(t)$. Quanto sopra discende anche da quanto trovato in generale per una $Q-$ matrice del tipo

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

con, in particolare, $\alpha = 1/2$ e $\beta = 0$ (v. Esempio 2.3).

Esercizio 2.2 Trovare a in modo che

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

sia il generatore di una CM a tempo continuo con spazio degli stati $\{1, 2, 3\}$. Dopo aver trovato il valore di a, trovare la distribuzione stazionaria della CM. Inoltre, calcolare approssimativamente $p_{ij}(t)$ $(i, j = 1, 2, 3)$ al tempo $t = 0.1$.

Soluzione. Affinché Q sia una $Q-$ matrice deve aversi $a+1/4+1/4 = 0$, ovvero $a = -1/2$. Si trova che gli autovalori di Q sono: 0, -0.5 e -0.9166 .
La distribuzione stazionaria $\pi = (\pi_1, \pi_2, \pi_3)$ è soluzione di:

$$\begin{cases} \pi Q = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

ovvero:

$$\begin{cases} -\frac{1}{2}\pi_1 + \frac{1}{3}\pi_3 = 0 \\ \frac{1}{4}\pi_1 - \frac{1}{4}\pi_2 + \frac{1}{3}\pi_3 = 0 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases}$$

Risolvendo, si trova $\pi = (2/11, 6/11, 3/11) \cong (0.18, 0.54, 0.27)$. Si ha poi:

$$P(h) = Id + hQ + o(h) =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + h \begin{pmatrix} -1/2 & 1/4 & 1/4 \\ 0 & -1/4 & 1/4 \\ 1/3 & 1/3 & -2/3 \end{pmatrix} + o(h) =$$

$$\approx \begin{pmatrix} 1 - h/2 & h/4 & h/4 \\ 0 & 1 - h/4 & h/4 \\ h/3 & h/3 & 1 - 2h/3 \end{pmatrix}, \text{ per } h \text{ piccolo}$$

Dunque:

$$P(0.1) \approx \begin{pmatrix} 0.95 & 0.025 & 0.025 \\ 0 & 0.975 & 0.025 \\ 0.033 & 0.033 & 0.933 \end{pmatrix}$$

Per ottenere approssimativamente $P(t) = e^{tQ}$ per t fissato si può usare la funzione $expm(\ )$ di Scilab, che calcola l'esponenziale di una matrice. Inserendo in ambiente Scilab la matrice Q e digitando $P = expm(0.1 * Q)$, si ottiene:

$$P(0.1) = \begin{pmatrix} 0.9516269 & 0.0244847 & 0.0238884 \\ 0.0003975 & 0.9757141 & 0.0238884 \\ 0.0314537 & 0.0322487 & 0.9362976 \end{pmatrix}$$

che differisce poco dall'approssimazione sopra trovata. Anche la distribuzione stazionaria può essere trovata in Scilab, osservando che, al crescere di t, $P(t)$ "converge" ad una matrice con le righe tutte uguali; in pratica, prendendo ad es. $t = 100$, ovvero calcolando $P = expm(100 * Q)$, si ottiene:

$$P(\infty) \cong P(100) = \begin{pmatrix} 0.1818182 & 0.5454545 & 0.2727273 \\ 0.1818182 & 0.5454545 & 0.2727273 \\ 0.1818182 & 0.5454545 & 0.2727273 \end{pmatrix}$$

Si osservi che

$$P(1) = e^Q = \begin{pmatrix} 0.6329765 & 0.2033461 & 0.1636774 \\ 0.0264458 & 0.8098768 & 0.1636774 \\ 0.1917906 & 0.2446823 & 0.5635270 \end{pmatrix}$$

è non degenere e regolare; i suoi autovalori si ottengono calcolando le potenze di e aventi per esponenti gli autovalori di Q, ovvero $e^0 = 1,\ e^{-0.5} = 0.6065307,\ e^{-0.9166} = 0.3998763$. A parte il primo autovalore, gli altri hanno modulo strettamente minore di 1.

Consideriamo ora la CM (a tempo discreto) "accelerata", ottenuta trascurando i tempi di permanenza nei vari stati; essa è descritta dalla matrice di transizione $\hat{P}$, dove

$$\hat{p}_{ij} = \begin{cases} \frac{q_{ij}}{-q_{ii}} & \text{se } i \neq j \\ 0 & \text{se } i = j \end{cases}$$

Si ottiene:

$$\hat{P} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

La matrice $\hat{P}$ è regolare ($P^4 > 0$), pertanto esiste la distribuzione stazionaria (ed invariante) $\hat{\pi}$ per la CM "accelerata", soluzione di $\hat{\pi} = \hat{\pi}\hat{P}$. Effettuando i calcoli, si trova:

$$\hat{\pi} = \left(\frac{4}{18}, \frac{1}{3}, \frac{4}{9}\right) \neq \pi = \left(\frac{2}{11}, \frac{6}{11}, \frac{3}{11}\right)$$

dove π è la distribuzione stazionaria della CM a tempo continuo, con generatore Q.

Esercizio 2.2 (continuato) Modifichiamo ora il generatore Q dell'Esercizio 2.2 nel modo seguente:

$$Q = \begin{pmatrix} -1/2 & 1/4 & 1/4 \\ 0 & -1/2 & 1/2 \\ 1/4 & 1/4 & -1/2 \end{pmatrix}$$

Si ottiene ancora per la matrice di transizione della CM "accelerata":

$$\hat{P} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

con distribuzione stazionaria $(\frac{4}{18}, \frac{1}{3}, \frac{4}{9})$ che è anche la distribuzione stazionaria della CM a tempo continuo con generatore Q, come si trova facilmente risolvendo l'equazione $\pi Q = 0$. Il fatto che ora la CM "accelerata" abbia lo stesso comportamento all'equilibrio della CM originale, si spiega considerando che i tempi di permanenza nei vari stati hanno tutti la stessa distribuzione, che è esponenziale di parametro $1/2$.

Esercizio 2.3 È data la $Q-$matrice:

$$Q = \begin{pmatrix} -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Si consideri la CM a tempo continuo con spazio degli stati $\{1, 2, 3\}$ e avente Q come generatore. Esiste la distribuzione stazionaria?

Soluzione. Gli stati 2 e 3 sono assorbenti, essendo $q_{22} = q_{33} = 0$; osserviamo che Q è doppiamente degenere, nel senso che i suoi autovalori sono: 0 contato due volte, $-1/2$. Dall'equazione $P(h) = hQ + Id + o(h)$, troviamo che, per h piccolo:

$$p_{11}(h) \cong 1 - h/2;\ p_{12}(h) \cong h/2;\ p_{12}(h) \cong 0;\ p_{21}(h) \cong 0;\ p_{22}(h) \cong 1;\ p_{23}(h) \cong 0;$$

$$p_{31}(h) \cong 0;\ p_{32}(h) \cong 0;\ p_{33}(h) \cong 1$$

Quindi, in un intervallo temporale di durata h, dallo stato 1 si passa nello stato 2 con probabilità $h/2$, ovvero con un tasso $\lambda = 1/2$ e si rimane in 1 con probabilità $1 - h/2$; dagli stati 2 e 3 non ci si può muovere, essendo essi stati assorbenti. Risolvendo l'equazione $(\pi_1, \pi_2, \pi_3)Q = (0, 0, 0)$ si trova che esistono infinite distribuzioni invarianti del tipo $(\pi_1, \pi_2, \pi_3) = (0, \theta, 1 - \theta)$, per $\theta \in [0, 1]$. Pertanto, non esiste la distribuzione stazionaria. D'altra parte, calcolando e^Q con Scilab, si trova:

$$P(1) = e^Q = \begin{pmatrix} 0.6065307 & 0.3934693 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La matrice stocastica $P(1)$ non è regolare ed i suoi autovalori sono: 1 contato due volte, $e^{-1/2}$; inoltre, calcolando le potenze ennesime di $P(1)$ si trova che la terza riga rimane invariata, mentre al crescere di n le prime due righe tendono a (0 1 0). Ad esempio, per $n = 1000$, si trova:

$$P(1)^{1000} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

La distribuzione stazionaria non esiste.

2.5 Il problema di Erlang

Consideriamo un sistema costituito da n sportelli (apparecchi di servizio) al quale arriva un flusso elementare di richieste (i clienti), secondo un processo di Poisson di intensità $\lambda > 0$. Ogni singolo sportello può soddisfare un solo cliente per volta, e un cliente che trova tutti gli sportelli occupati viene respinto (più tardi contempleremo la possibilità che egli possa mettersi in fila, ed aspettare che venga servito). Quando un cliente viene servito da uno sportello, esce dal sistema; supponiamo che i tempi di servizio siano v.a. indipendenti ed esponenziali di parametro $\mu > 0$. Ci proponiamo di trovare, ad ogni tempo t, il numero di clienti presenti nel sistema e di calcolare la probabilità di congestione.

Diremo che al tempo t il sistema si trova nello stato 0 se tutti gli sportelli sono liberi, nello stato 1 se uno sportello è occupato e tutti gli altri sono liberi, ..., lo stato n rappresenta il caso in cui tutti gli sportelli sono occupati.

Indichiamo con $p_k(t)$ la probabilità che all'istante t il sistema si trovi nello stato $k \in \{0, 1, \dots, n\}$. Troveremo delle semplici equazioni differenziali alle quali $p_k(t)$ deve soddi–sfare. Per prima cosa, calcoliamo la probabilità che al tempo $t + h$ il sistema si trovi nello stato 0 (tutti gli sportelli sono liberi). Tale evento è l'unione disgiunta degli eventi seguenti:

(i) all' istante t tutti gli sportelli sono liberi e non arriva alcun cliente tra t e $t + h$;

(ii) all'istante t un solo sportello è occupato, questo finisce di svolgere il servizio tra t e $t + h$ e, in tale intervallo non arriva alcun nuovo cliente;

(iii) all'istante t due o più sportelli sono occupati e finiscono il servizio nel successivo intervallo $(t, t + h)$, etc.

Facendo ipotesi analoghe a quelle fatte per il processo di Poisson, si trova che:

- la probabilità del primo evento è $p_0(t)e^{-\lambda h} = p_0(t)(1 - \lambda h + o(h))$;
- la probabilità del secondo evento è $p_1(t)(1 - e^{-\mu h})e^{-\lambda h} = p_1(t)\mu h + o(h)$
- la probabilità del terzo evento è $o(h)$.

Tenendo conto di quanto sopra, si trova che:

$$p_0(t + h) = p_0(t)(1 - \lambda h) + \mu h p_1(t) + o(h) \tag{2.42}$$

e, passando al limite per $h \to 0$ otteniamo l'equazione:

$$p_0'(t) = -\lambda p_0(t) + \mu p_1(t) \tag{2.43}$$

Facendo un ragionamento analogo per $1 \le k \le n$, si trova

$$p_k'(t) = \lambda p_{k-1}(t) - (\lambda + k\mu)p_k(t) + (k + 1)\mu p_{k+1}(t) \tag{2.44}$$

e, per $k = n$:

$$p_n'(t) = \lambda p_{n-1}(t) - n\mu p_n(t) \tag{2.45}$$

Come nel caso del processo di Poisson, il sistema di equazioni differenziali ottenute permette di calcolare le probabilità $p_k(t)$. Le costanti arbitrarie che compaiono quando si scrivono le soluzioni esplicite, si determinano tenendo conto delle condizioni iniziali

$$p_0(0) = 1, \;\; p_k(0) = 0, \;\; \text{per } k \ge 1 \tag{2.46}$$

(ovvero al tempo 0 tutti gli sportelli sono liberi).
Occorre aggiungere inoltre la condizione di normalizzazione $\sum_{k=0}^{n} p_k(t) = 1$.
Invece di trovare la soluzione esplicita delle equazioni differenziali di sopra, piuttosto complessa, vogliamo ora studiare il processo stazionario, ossia considerare la soluzione delle equazioni per $t \to \infty$. Se poniamo $\pi_k \doteq \lim_{t\to\infty} p_k(t)$, tali probabilità limite devono soddi–sfare il sistema seguente di equazioni che si ottengono da (2.43), (2.44), (2.45) sostituendo alle funzioni $p_k(t)$ le costanti π_k e alle derivate $p_k'(t)$ il valore zero:

$$\begin{cases} -\lambda\pi_0 + \mu\pi_1 = 0 \\ \lambda\pi_{k-1} - (\lambda + k\mu)\pi_k + (k+1)\mu\pi_{k+1} = 0 \ (1 \le k < n) \\ \lambda\pi_{n-1} - n\mu\pi_n = 0 \\ \sum_{k=0}^{n} \pi_k = 1 \end{cases} \tag{2.47}$$

Posto $\rho = \lambda/\mu$, si può dimostrare che le probabilità stazionare π_k soddisfano alle *formule di Erlang*:

$$\pi_k = \frac{\frac{\rho^k}{k!}}{\sum_{j=0}^{n} \frac{\rho^j}{j!}} \quad k = 0, 1, \ldots \tag{2.48}$$

Per $k = n$, si ottiene la probabilità che tutti gli sportelli siano occupati e quindi la probabilità che ogni nuovo cliente venga respinto; dunque, la probabilità di congestione è

$$\pi_n = \frac{\frac{\rho^n}{n!}}{\sum_{j=0}^{n} \frac{\rho^j}{j!}} \tag{2.49}$$

Il problema di Erlang rientra anch'esso nello schema dei processi di nascita e morte: si tratta di un processo di nascita e morte con $\lambda_k = \lambda$ per $0 \le k < n$ e $\lambda_k = 0$ per $k \ge n$; $\mu_k = 0$ per $k > n$ e $\mu_k = k\mu$ per $1 \le k \le n$ (cfr. con (2.12)).

2.6 Processi a coda

Consideriamo ora alcuni esempi di CM in tempo continuo, in particolare processi di nascita e morte, che sono utili come modelli per i processi a coda. Nel seguito supporremo sempre che il flusso di arrivi segua un processo di Poisson di intensità λ. I clienti che si trovano in uno sportello libero iniziano un tempo di servizio allo sportello, gli altri si mettono in coda. Quando uno sportello si libera, uno dei clienti in coda inizia il suo servizio. Non è importante l'ordine col quale i clienti accedono al servizio; facciamo l'ipotesi che i tempi di servizio siano indipendenti e identicamente distribuiti e indipendenti dal processo di Poissson che regola il flusso degli arrivi. Supponiamo inoltre che la distribuzione del tempo di servizio sia esponenziale di parametro μ.

Un processo di questo tipo viene denotato col simbolo $M/M/n$. La prima M indica che il flusso degli arrivi è Poissoniano, la seconda indica che il tempo di servizio è esponenziale ed n indica il numero degli sportelli, che può variare da 1 a ∞. Precisamente, utilizziamo la cosiddetta notazione di Kendall che indica le caratteristiche di una specifica coda con diversi campi separati da sbarre:

campo 1 / *campo* 2 / *campo* 3 / *campo* 4 / *campo* 5

dove:
campo 1 indica la distribuzione del tempo di interarrivo ($M =$ esponenziale, $G =$ di–stribuzione generica, $D =$ costante); *campo 2* indica la distribuzione del tempo di servizio; *campo 3* indica il numero di serventi; *campo 4* indica la lunghezza del buffer; *campo 5* indica il numero di clienti ammesso. Nel caso che il buffer e il numero dei clienti ammesso siano infiniti, si omettono i *campi 4* e *5*.

2.6.1 Coda M/M/n

Si tratta di un processo di nascita e morte con $\lambda_k = \lambda$ per ogni k, e $\mu_k = k\mu$ se $k \leq n$, $\mu_k = n\mu$ se $k \geq n$.

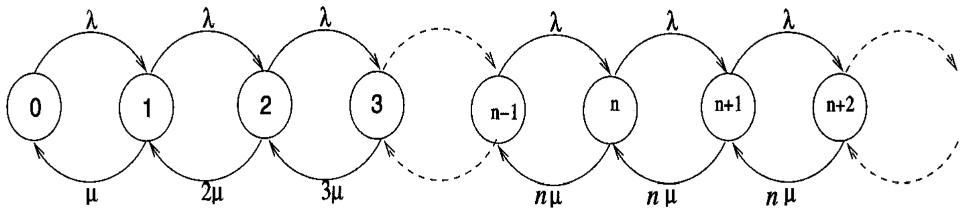

Fig. 10: Rappresentazione delle transizioni di stato di una coda M/M/n

Le probabilità stazionarie π_k che nel sistema siano presenti k clienti sono soluzioni del sistema di equazioni (2.11). Sostituendo i valori di λ_k e μ_k e posto $\rho = \lambda/\mu$, si trova che le soluzioni del sistema ottenuto sono:

$$\pi_k = \begin{cases} \frac{\rho^k}{k!}\pi_0 & \text{se } k \leq n \\ \frac{\rho^k}{n!n^{k-n}}\pi_0 & \text{se } k \geq n+1 \end{cases} \tag{2.50}$$

Per mostrare le (2.50), utilizziamo le relazioni (2.13), ovvero:

$$\lambda_0\pi_0 = \mu_1\pi_1 \ , \ \lambda_k\pi_k = \mu_{k+1}\pi_{k+1}$$

con $\lambda_k = \lambda \ \forall k$ e

$$\mu_k = \begin{cases} k\mu & \text{se } k = 1, \ldots, n \\ n\mu & \text{se } k = n+1, \ldots \end{cases}$$

Si ha allora:

$$\begin{cases} \pi_1 = \frac{\lambda}{\mu}\pi_0 \\ \pi_2 = \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2 \pi_0 \\ \pi_3 = \frac{1}{2\cdot 3}\left(\frac{\lambda}{\mu}\right)^3 \pi_0 \\ . \\ . \\ \pi_n = \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \pi_0 \end{cases}$$

mentre:

$$\pi_{n+1} = \frac{\lambda}{\mu_{n+1}}\pi_n = \frac{\lambda}{n\mu}\frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n \pi_0 =$$

$$= \left(\frac{\lambda}{\mu}\right)^{n+1} \frac{1}{n}\ \frac{1}{n!}\pi_0$$

$$\pi_{n+2} = \frac{\lambda}{\mu_{n+2}}\pi_{n+1} = \frac{\lambda}{n\mu}\left(\frac{\lambda}{\mu}\right)^{n+1} \frac{1}{n}\ \frac{1}{n!}\pi_0 =$$

$$= \left(\frac{\lambda}{\mu}\right)^{n+2} \frac{1}{n^2}\ \frac{1}{n!}\pi_0$$

e quindi:

$$\pi_k = \left(\frac{\lambda}{\mu}\right)^k \frac{1}{n^{k-n}}\ \frac{1}{n!}\pi_0\ , \ \text{per } k \geq n+1$$

Abbiamo in tal modo ottenuto le (2.50). Il valore di π_0 lo troveremo poi, imponendo che $\pi_0 + \sum_{k\geq 1} \pi_k = 1$.
Affinché questa soluzione sia compatibile, occorre naturalmente che sia soddisfatta la condizione:

$$\sum_{k=0}^{\infty} \pi_k < +\infty \tag{2.51}$$

ovvero:

$$\sum_{k=0}^{n} \frac{\rho^k}{k!} + \sum_{k=n+1}^{\infty} \frac{\rho^k}{n! n^{k-n}} < +\infty \tag{2.52}$$

Visto che la prima somma è finita, basta richiedere la convergenza della serie che, ponendo $j = k - n - 1$, si può riscrivere

$$\frac{1}{n\ n!}\rho^{n+1} \sum_{j=0}^{\infty} (\rho/n)^j \tag{2.53}$$

Dunque, la condizione di convergenza è $\rho/n < 1$, ovvero $\rho < n$. Tale condizione fornisce il numero di sportelli n da inserire in un sistema a coda con flusso di arrivi assegnato, affinché la coda si stabilizzi, cioé esista la distribuzione stazionaria. Per $\rho < n$, la somma della serie in (2.53) è

$$\frac{\rho^{n+1}}{n!(n-\rho)} \tag{2.54}$$

e si trova quindi infine, imponendo la condizione di normalizzazione $\sum_{k=0}^{\infty} \pi_k = 1$,:

$$\pi_0 = \left[\sum_{k=0}^{n} \frac{\rho^k}{k!} + \frac{\rho^{n+1}}{n!(n-\rho)}\right]^{-1} \tag{2.55}$$

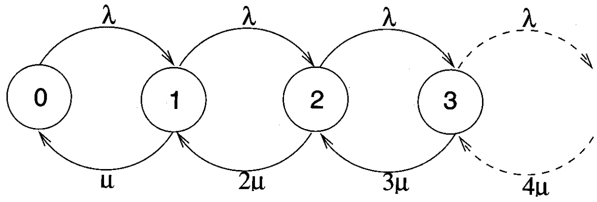

Fig. 11: Rappresentazione delle transizioni di stato di una coda $M/M/\infty$

2.6.2 Coda M/M/∞

Basta considerare il caso precedente con $n \to \infty$, ovvero $\lambda_k = \lambda,\ \mu_k = k\mu$, per ogni k.

Procedendo come nel caso della coda M/M/n, si ottiene:

$$\pi_k = \frac{\rho^k}{k!}\pi_0,\ k = 0, 1, \ldots \tag{2.56}$$

Imponendo la condizione di normalizzazione, si trova $\pi_0 = \left[\sum_{k=0}^{\infty} \frac{\rho^k}{k!}\right]^{-1}$, cioé $\pi_0 = e^{-\rho}$. Pertanto, si ha infine:

$$\pi_k = \frac{\rho^k}{k!}e^{-\rho},\ k = 0, 1, \ldots \tag{2.57}$$

e la distribuzione stazionaria esiste sempre, senza imporre alcuna condizione su ρ. Notare che (2.57) non è altro che la densità discreta di una v.a. di Poisson di parametro ρ.

2.6.3 Coda M/M/1

Basta considerare un processo di nascita e morte con $\lambda_k = \lambda$ e $\mu_k = \mu$, per ogni k.

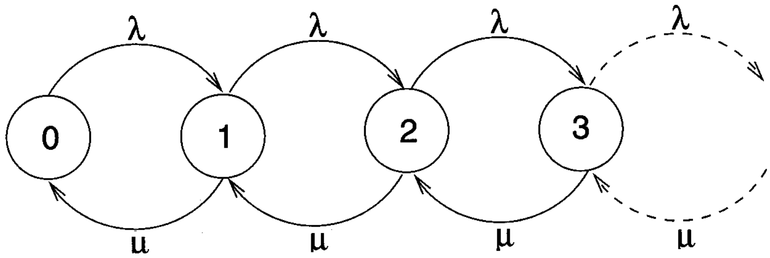

Fig. 12: Rappresentazione delle transizioni di stato di una coda M/M/1

Per trovare le probabilità stazionarie, occorre risolvere il sistema di equazioni (v. (2.11)), oppure usare il risultato per la coda M/M/n con $n = 1$:

$$\begin{cases} -\lambda\pi_0 + \mu\pi_1 = 0 \\ -(\lambda + \mu)\pi_k + \lambda\pi_{k-1} + \mu\pi_{k+1} = 0, \quad k \geq 1 \end{cases} \tag{2.58}$$

Alternativamente, usando le (2.13), si ottiene:

$$\begin{cases} \pi_1 = \frac{\lambda}{\mu}\pi_0 \\ \pi_2 = \left(\frac{\lambda}{\mu}\right)^2 \pi_0 \\ . \\ . \end{cases}$$

e quindi, in generale, si ha:

$$\pi_k = \left(\frac{\lambda}{\mu}\right)^k \pi_0$$

Imponendo la condizione di normalizzazione:

$$\pi_0 \left(\sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^k + 1 \right) = 1$$

si trova:

$$\pi_0 = \left(\frac{1}{1 - \lambda/\mu} \right)^{-1} = 1 - \frac{\lambda}{\mu} \tag{2.59}$$

e quindi, posto $\rho = \lambda/\mu$, si ottiene la distribuzione stazionaria:

$$\pi_k = \rho^k (1 - \rho), \ k = 0, 1, \dots \tag{2.60}$$

Naturalmente, la serie di sopra converge soltanto se $\rho < 1$; dunque la distribuzione stazionaria esiste solo se $\rho < 1$, cioé $\lambda < \mu$, il che significa che l'intensità di arrivo deve essere strettamente minore del parametro della distribuzione esponenziale del tempo di servizio. Si noti che (2.60) è la densità discreta di una v.a. Geometrica di parametro $1 - \rho$.

Riassumendo:

per la coda $M/M/1$ si prende $\mu_k = \mu = cost$
per la coda $M/M/n$ si prende $\mu_k = \mu \cdot \min(k, n)$
per la coda $M/M/\infty$ si prende $\mu_k = k\mu$

I relativi generatori (matrici Q) corrispondenti sono:

per $M/M/1$

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & \mu & -(\lambda+\mu) & \lambda & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{2.61}$$

per $M/M/n$ con, ad es. $n = 2$

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & \dots \\ 0 & 0 & 2\mu & -(\lambda+2\mu) & \lambda \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{2.62}$$

per $M/M/\infty$

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & \dots \\ 0 & 0 & 3\mu & -(\lambda+3\mu) & \lambda \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{2.63}$$

2.6.4 Distribuzione del tempo totale che un cliente trascorre nel sistema per una coda M/M/1

Se un cliente entra nel sistema quando il numero dei clienti è $N = n$, dovrà attendere che tutte le n persone presenti completino il servizio all'unico sportello, più il tempo necessario ad espletare il proprio servizio allo sportello (per semplicità, supponiamo che il cliente entra nel sistema esattamente nell'istante in cui inizia il servizio del primo dei clienti in fila allo sportello). Dunque, condizionatamente a $N = n$, il tempo T che il cliente attende nel sistema è la somma di $n+1$ v.a. esponenziali indipendenti di parametro μ, quindi ha distribuzione $\Gamma(n+1,\mu)$. Allora:

$$P(T \leq t|N = n) = \int_0^t \frac{\mu}{n!} e^{-\mu s} (\mu s)^n ds$$

da cui:

$$\begin{aligned} P(T \leq t) &= \sum_{n=0}^{\infty} P(T \leq t|N = n) P(N = n) = \\ &= \sum_{n=0}^{\infty} \left(\int_0^t \frac{\mu}{n!} e^{-\mu s} (\mu s)^n \cdot \rho^n (1-\rho) ds \right) = \\ &= \mu(1-\rho) \int_0^t e^{-\mu s} \sum_{n=0}^{\infty} \frac{(\mu s \rho)^n}{n!} \; ds = \\ &= \mu(1-\rho) \int_0^t e^{-\mu s} e^{\mu \rho s} \; ds = \mu(1-\rho) \int_0^t e^{-\mu(1-\rho)s} \; ds = \\ &= 1 - e^{-\mu(1-\rho)t} \end{aligned}$$

e quindi T ha distribuzione esponenziale di parametro $\mu(1-\rho) = \mu - \lambda$.
Ricordando che la media di una v.a esponenziale è il reciproco del parametro, si ottiene $W \doteq E(T) = 1/(\mu - \lambda)$.

2.6.5 Distribuzione del tempo T_c che un cliente trascorre in fila per una coda M/M/1

Osserviamo che $P(T_c = 0) = P(N = 0) = \pi_0 = (1 - \rho)$. Se il cliente entra nel sistema quando il numero dei clienti è $N = n$, con $n \geq 1$, il tempo che trascorre nella fila, prima che venga iniziato il proprio servizio, uguaglia il tempo di n tempi di servizio. Dunque, condizionatamente a $N = n \geq 1$, il tempo T_c che il cliente attende in fila è la somma di n v.a. esponenziali indipendenti di parametro μ, quindi ha distribuzione $\Gamma(n, \mu)$. Allora, per $n \geq 1$, se $t \geq 0$:

$$P(T_c \leq t | N = n) = \int_0^t \frac{\mu^n}{(n-1)!} s^{n-1} e^{-\mu s} ds.$$

Quindi, per $t \geq 0$:

$$P(T_c \leq t) = P(T_c = 0) + \sum_{n=1}^{\infty} P(T_c \leq t | N = n) P(N = n)$$

(ricordando che la distribuzione stazionaria è geometrica di parametro $1 - \rho$, ovvero $P(N = n) = \rho^n (1 - \rho)$)

$$= 1 - \rho + \sum_{n=1}^{\infty} \left(\int_0^t \frac{\mu^n}{(n-1)!} s^{n-1} e^{-\mu s} ds \right) \rho^n (1 - \rho)$$

$$= 1 - \rho + \rho(1 - \rho) \int_0^t \mu e^{-\mu s} \sum_{n=1}^{\infty} \frac{(\mu s \rho)^{n-1}}{(n-1)!} ds$$

$$= 1 - \rho + \rho(1 - \rho) \int_0^t \mu e^{-\mu s} e^{\mu s \rho} ds$$

$$= 1 - \rho + \rho \int_0^t \mu(1 - \rho) e^{-\mu s (1-\rho)} ds;$$

osserviamo che l'integrale non è altro che la probabilità che una v.a. esponenziale con parametro $\mu(1 - \rho)$ sia $\leq t$, pertanto vale $1 - e^{-\mu(1-\rho)t}$. Dunque, riprendendo il calcolo, otteniamo per $t \geq 0$:

$$P(T_c \leq t) = 1 - \rho + \rho(1 - e^{-\mu(1-\rho)t}) = 1 - \rho e^{-\mu(1-\rho)t}.$$

La v.a. T_c ha massa in $t = 0$, avendosi $P(T_c = 0) = 1 - \rho > 0$ (poiché $\rho < 1$) e, derivando $P(T_c \leq t)$ rispetto a t, si ottiene che la densità della parte continua di T_c è:

$$f_{T_c}(t) = \mu\rho(1 - \rho) e^{-\mu(1-\rho)t}, \ t > 0.$$

Si noti che l'integrale da 0 a $+\infty$ di $f_{T_c}(t)$ non vale 1, ma $\rho < 1$; si ha, però:

$$P(T_c < +\infty) = P(T_c = 0) + \int_0^{+\infty} f_{T_c}(t) dt = 1 - \rho + \rho = 1,$$

come deve essere.
La media di T_c è:

$$E(T_c) = 0 \cdot P(T_c = 0) + \int_0^{+\infty} t \cdot f_{T_c}(t)dt$$

$$= \rho \int_0^{+\infty} \mu(1-\rho)e^{-\mu(1-\rho)t}tdt = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu(\mu-\lambda)}.$$

Quindi:

$$W_c \doteq E(T_c) = \frac{\lambda}{\mu(\mu-\lambda)}.$$

2.6.6 Sistemi a coda in regime stazionario e relazioni di Little

Per i sistemi a coda Markoviani visti, l'esistenza della distribuzione stazionaria permette di considerare un regime stazionario per il processso $N(t)$ del numero di clienti. Basta considerare il processo con distribuzione iniziale uguale a quella stazionaria. Si può dimostrare che vale *l'ergodicità*, ovvero in tale situazione un sistema a code si evolve verso il regime stazionario e le medie temporali di quantità di interesse calcolate in $[0, t]$, tendono, per $t \to \infty$ alla media calcolata nel regime stazionario.

Per valutare l'efficienza di un sistema a code si introduce il *fattore di utilizzazione* $\eta = \rho/n = \lambda/(\mu n)$. Questa quantità è data dal tasso medio di arrivo dei clienti λ per la media del tempo di servizio, diviso il numero n degli sportelli. Si può dimostrare che η è uguale alla percentuale media di utilizzazione degli sportelli. Per un sistema non deterministico in regime stazionario si trova che $\eta < 1$, ovvero non è possibile avere una utilizzazione completa degli sportelli (nel regime stazionario vi sarà una probabilità positiva per uno sportello di essere libero). Quantità di notevole interesse sono:
1) il numero medio L di clienti nel sistema;
2) il numero medio L_c di clienti in attesa nelle code;
3) il tempo medio W che un cliente trascorre nel sistema;
4) il tempo medio W_c che un cliente trascorre in attesa nelle code.

Le due ultime quantità sono collegate dalla relazione

$$W = W_c + \frac{1}{\mu} \tag{2.64}$$

Ci accingiamo ora a dimostrare la prima legge di Little, ovvero che $L = \lambda W$. Per fare ciò, abbiamo bisogno del seguente Lemma.

Lemma Sia X_t un processo di Poisson omogeneo, con intensità λ.
Allora, per $t \to +\infty$ si ha, quasi certamente:

$$\frac{X_t}{t} \to \lambda.$$

Dim. Osserviamo che, se $t = n$ si ha $X_t = X_n = Z_1 + Z_2 + \ldots + Z_n$, dove $Z_i, i = 1, \ldots, n$ sono v.a. indipendenti con distribuzione di Poisson di parametro λ (infatti, ritorna $X_n \sim Poisson(n\lambda)$). Per la legge dei grandi numeri, per $n \to \infty$ si ha q.c.:

$$\frac{X_n}{n} = \frac{Z_1 + Z_2 + \ldots + Z_n}{n} \to E(Z_1) = \lambda.$$

Per $t > 0$, abbiamo $n = [t] \leq t < [t] + 1 = n + 1$, dove [ ] denota la parte intera; siccome il processo di Poisson X_t è non decrescente, si ha $X_n \leq X_t < X_{n+1}$. Abbiamo, allora:

$$\frac{X_n}{n+1} \leq \frac{X_t}{t} < \frac{X_{n+1}}{n} \qquad (*)$$

ma:

$$\frac{X_{n+1}}{n} = \frac{X_{n+1}}{n+1} \cdot \frac{n+1}{n} \to \lambda \cdot 1 = \lambda.$$

e

$$\frac{X_n}{n+1} = \frac{X_n}{n} \cdot \frac{n}{n+1} \to \lambda \cdot 1 = \lambda.$$

Pertanto, da $(*)$ otteniamo, per $t \to +\infty$:

$$\frac{X_t}{t} \to \lambda, \text{ q.c.}$$

e la dimostrazione del Lemma è conclusa.

Dimostrazione della prima legge di Little, $L = \lambda W$.
Indichiamo con $N(t)$ il numero di clienti presenti nel sistema all'istante t e poniamo

$$N_{[0,T]} = \int_0^T N(t)dt \qquad (2.65)$$

Allora, la media temporale di $N(t)$ nell'intervallo $[0, T]$ è :

$$\bar{N}_T = \frac{N_{[0,T]}}{T} \qquad (2.66)$$

Per $T \to +\infty$ grazie all'ergodicità, $\lim_{T\to+\infty} \bar{N}_T$ è uguale alla media di stato del numero di clienti, cioé L. Dunque, si ha:

$$L = \lim_{T\to+\infty} \bar{N}_T = \lim_{T\to+\infty} \frac{N_{[0,T]}}{T} = \lim_{T\to+\infty} \frac{N_{[0,T]}}{X_T} \cdot \frac{X_T}{T} \qquad (**)$$

dove X_t è il processo di Poisson degli arrivi. $N_{[0,T]}$ rappresenta il tempo cumulativo speso da tutti i clienti del sistema nell'intervallo $[0, T]$ e X_T è il numero di clienti che sono arrivati nell' intervallo $[0, T]$; quindi $\lim_{T\to+\infty} N_{[0,T]}/X_T$ rappresenta il tempo che un cliente trascorre nel sistema. Inoltre, per il Lemma X_T/T, tende a λ, per $T \to +\infty$.

Quindi, passando alla media in $(**)$, utilizzando anche il Teorema di convergenza dominata, otteniamo:

$$L = E(L) = E(\text{tempo che un cliente trascorre nel sistema}) \cdot \lambda = W \cdot \lambda.$$

La dimostrazione della *prima relazione di Little*

$$L = \lambda W \tag{2.69}$$

è conclusa.

Facendo un analogo ragionamento per il numero $N_c(t)$ di clienti in fila al tempo t, nell'ipotesi di ergodicità si ottiene la *seconda relazione di Little*:

$$L_c = \lambda W_c \tag{2.73}$$

Le due relazioni di Little sono valide per un'ampia classe di sistemi a coda in regime stazionario. Ad esempio, per una coda $M/M/1$ la distribuzione stazionaria esiste se e solo se $\rho = \lambda/\mu < 1$, cioé $\lambda < \mu$. Da (2.60), la distribuzione stazionaria del numero di clienti nel sistema è :

$$\pi_k = \left(\frac{\lambda}{\mu}\right)^k \left(1 - \frac{\lambda}{\mu}\right) \tag{2.74}$$

Si ha quindi:

$$L = \sum_{k=1}^{\infty} k\pi_k = \sum_{k=1}^{\infty} k \left(\frac{\lambda}{\mu}\right)^k \left(1 - \frac{\lambda}{\mu}\right) = \frac{\lambda}{\mu - \lambda} \tag{2.75}$$

e

$$\begin{aligned} L_c &= \sum_{k=1}^{\infty}(k-1)\pi_k = \left(1 - \frac{\lambda}{\mu}\right) \sum_{k=1}^{\infty}(k-1) \left(\frac{\lambda}{\mu}\right)^k = \\ &= \left(1 - \frac{\lambda}{\mu}\right) \sum_{k=1}^{\infty} k \left(\frac{\lambda}{\mu}\right)^k - \left(1 - \frac{\lambda}{\mu}\right) \sum_{k=1}^{\infty} \left(\frac{\lambda}{\mu}\right)^k = \\ &= \frac{\lambda^2}{\mu(\mu - \lambda)} \end{aligned} \tag{2.76}$$

Quindi, usando le relazioni di Little, otteniamo:

$$W = \frac{1}{\mu - \lambda}, \quad W_c = \frac{\lambda}{\mu(\mu - \lambda)} \tag{2.77}$$

che soddisfano la relazione $W = W_c + \frac{1}{\mu}$. Si noti che l'espressione per W è stata già trovata in 2.6.4, dove si è mostrato che il tempo che un cliente trascorre nel sistema ha distribuzione esponenziale di parametro $\mu - \lambda$, mentre l'espressione per W_c è stata già trovata in 2.6.5 Osserviamo che quando $\eta = \rho$ tende a 1, il numero medio dei clienti nel sistema ed in coda ed il tempo medio di un cliente nel sistema ed in coda tendono all'infinito. Questa è una caratteristica generale dei sistemi a coda stocastici. Se si cerca di aumentare il fattore di utilizzazione, si deve pagare il prezzo di far crescere il numero dei clienti in coda ed i loro tempi di attesa. Il valore $\eta = 1$ non è raggiungibile da un sistema a coda stocastico in regime stazionario; può però essere ottenuto da un sistema deterministico, ad esempio con uno sportello in cui i clienti arrivano ad intervalli di tempo regolari, fissi ed uguali al loro tempo di servizio.

Bibliografia

– P. Baldi, "Calcolo delle Probabilità", McGraw-Hill, 2007
– W. Feller, "An Introduction to Probability Theory and its Applications", vol. 1, Wiley, New York, 1968.
– B. V. Gnedenko, "Teoria della Probabilità", Editori Riuniti
– F. Biagini, M. Campanino, "Elementi di Probabilità e Statistica", Springer, 2006.