NOTE DAL CORSO DI

SISTEMI A EVENTI DISCRETI

ANNO ACCADEMICO 2019/20

16 novembre 2019

CAPITOLO 3° :CATENE DI MARKOV

Docente: Riccardo Minciardi

SOMMARIO

1 *I processi di Poisson e la distribuzione esponenziale*

In questo paragrafo vengono richiamate due distribuzioni di variabile aleatoria (una discreta e l'altra continua) di particolare interesse per quanto verrà sviluppato nel seguito.

1.1 I processi di Poisson

Si consideri una sequenza di eventi discreti. Sia $\{N(t)\}$ il processo stocastico che conta gli eventi che accadono in $(0,t]$. Si ha allora $N(t_1) \leq N(t_2)$ $\forall\, t_1, t_2$: $\forall\, t_1 \leq t_2$. Si definisce $N(t_1,t_2) = N(t_2) - N(t_1)$. Si introducono ora le seguenti assunzioni:

1. in ogni istante può verificarsi al massimo un singolo evento (cioè non ci sono eventi contemporanei);
2. presi due intervalli arbitrari $(t_1,t_2]$ e $(t_3,t_4]$, tali che $t_1 < t_2 \leq t_3 < t_4$, le variabili aleatorie $N(t_1,t_2)$ e $N(t_3,t_4)$ sono indipendenti;
3. la probabilità $\Pr\{N(t_1,t_2) = n\}$ dipende solamente dalla differenza $(t_2 - t_1)$, oltre che, ovviamente, da n.

Un processo stocastico come quello descritto e che soddisfa le assunzioni viste sopra si dice processo di conteggio (counting process) stazionario con incrementi indipendenti, o processo di Poisson.

Risultato 1 (senza dimostrazione) *In riferimento ad un processo di Poisson, la variabile aleatoria $N(t)$ ha una distribuzione data da*

$$P_n(t) = Pr\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \qquad t \geq 0 \qquad n=0,\ 1,\ 2,\ 3,\ \dots \tag{1}$$

essendo λ un parametro che caratterizza la distribuzione considerata.

Si prova che $E[N(t)] = \lambda t$ e che $Var[N(t)] = \lambda t$. Il parametro λ può essere quindi interpretato come la frequenza degli arrivi ($1/\lambda$ è il tempo medio tra due eventi successivi).

1.2 La distribuzione esponenziale

Risultato 2 *(senza dimostrazione)* I tempi di inter-evento in un processo di Poisson costituiscono una sequenza di variabili aleatorie indipendenti e identicamente distribuite (i.i.d.) in modo esponenziale; più precisamente , la distribuzione cumulativa di tali variabili è data da

$$G(t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \qquad (2)$$

essendo λ il parametro caratteristico del processo di Poisson considerato. La densità delle variabili in questione è quindi

$$g(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \qquad (3)$$

In maniera perfettamente duale, una sequenza di realizzazioni di una variabile aleatoria distribuita in modo esponen-ziale può essere vista come una sequenza di realizzazioni di variabili aleatorie i.i.d. corrispondenti ai tempi di inter-evento di un processo di Poisson.

Risultato 3 (senza dimostrazione) *Detto V il generico tempo di inter-evento in un processo di Poisson*

$$Pr\{V \leq z + t | V > z\} = 1 - e^{-\lambda t} = Pr\{V \leq t\} =$$

in altre parole la variabile aleatoria "tempo residuo di inter-evento" ha la stessa distribuzione della variabile aleatoria "tempo di inter-evento". Il condizionamento V > z non ha pertanto contenuto informativo ("memoryless property"). Si può provare inoltre che la distribuzione esponenziale è l'unica distribuzione per la quale valga tale proprietà.

1.3 Sovrapposizione di processi di Poisson

Si consideri un processo puntuale (cioè una sequenza di eventi) costituito dalla sovrapposizione di *m* processi di Poisson mutuamente indipendenti, caratterizzati, rispettivamente, dai valori $\lambda_1, \ldots \lambda_m$ del parametro caratterizzante. Si può mostrare che in questo caso il processo puntuale così ottenuto è ancora un processo di Poisson, con valore del parametro pari a

$$\Lambda = \sum_{i=1}^{m} \lambda_i$$

2 *Catene di Markov a tempo discreto*

2.1 Definizioni e prime proprietà

Una catena di Markov a tempo discreto (DTMC) è un sistema a stato e tempo discreti in cui l'evoluzione della variabile di stato X_k è caratterizzata dal fatto che vale la seguente proprietà (Markov property):

$$Pr\{X_{k+1} = x_{k+1} | X_k = x_k, \dots, X_0 = x_0\} = Pr\{X_{k+1} = x_{k+1} | X_k = x_k\} \tag{4}$$

I possibili valori dello stato sono elementi di un insieme discreto S (finito o numerabile). Essi possono quindi essere individuati dagli interi 0, 1, 2, …

L'evoluzione di una DTMC è quindi completamente governata dall'insieme delle probabilità di transizione $p_{i,j}(k) = \Pr\{X_{k+1} = j \mid X_k = i\}$, $\forall\ i,j \in S$, $k = 0, 1, 2, \dots$ Ovviamente deve risultare

$$\sum_{j \in S} p_{ij}(k) = 1$$

Possono essere definite anche la le probabilità di transizione ad n passi

$$p_{ij}(k, k+n) = Pr\{X_{k+n} = j | X_k = i\} \qquad n \geq 1$$

Per il teorema della probabilità totale, si ha, se $k < u < k + n$,

$$p_{ij}(k, k+n) = \sum_{r \in S} Pr\{X_{k+n} = j | X_u = r, \dots, X_k = i\} \cdot Pr\{X_u = r | X_k = i\}$$

e, sulla base della definizione di DTMC,

$$Pr\{X_{k+n} = j | X_u = r, \dots, X_k = i\} = Pr\{X_{k+n} = j | X_u = r\} = p_{rj}(u, k+n)$$

Si ottiene allora

$$p_{ij}(k, k+n) = \sum_{r \in S} p_{ir}\,(k,u) p_{rj}(u, k+n) \qquad\qquad k < u < k+n \qquad (5)$$

che è detta <u>equazione di Chapman-Kolmogorov</u>.

Definendo la matrice

$$H(k, k+n) = \left[p_{ij}(k, k+n)\right]$$

l'equazione di Chapman-Kolmogorov si può scrivere come

$$H(k, k+n) = H(k, u)H(u, k+n)$$

Inoltre, se si sceglie $u = k + n - 1$, si ha

$$H(k, k+n) = H(k, k+n-1)H(k+n-1, k+n)$$

(che è nota come *equazione di Chapman-Kolmogorov forward*) e se si pone $u = k + 1$, si ha

$$H(k, k+n) = H(k, k+1)H(k+1, k+n)$$

(che è nota come *equazione di Chapman-Kolmogorov backward*).

Quando le <u>probabilità di transizione $p_{i,j}(k)$ sono indipendenti da k</u>, $\forall\ i,j \in S$, la DTMC si dice <u>omogenea</u> (o <u>tempo invariante, o stazionaria</u>). Quindi scriveremo semplicemente $p_{i,j}$ invece che $p_{i,j}(k)$. Se la DTMC è omogenea, anche $p_{i,j}(k,k+n)$ è indipendente da k, e si può indicare semplicemente con $p_{i,j,n}$. Inoltre, si indicherà con $H(n)$ la matrice $[p_{i,j,n}]$. Dall'equazione di Chapman-Kolmogorov forward si ha allora

$$H(n) = H(n-1)H(1)$$

Ovviamente $p_{i,j,n} = p_{i,j}$. $H(1) = [p_{i,j}]$ verrà indicata semplicemente come P . Si ha quindi

$H(2) = H(1)^2 = P^2$

$H(3) = H(1)^3 = P^3$ …

$H(n) = H(1)^n = P^n$

D'ora in poi saranno considerate solo DTMC omogenee. L'assunzione di omogeneità sarà di regola sottintesa.

Si consideri adesso la variabile aleatoria $V(i)$ corrispondente al numero di istanti di campionamento (o stadi) di permanenza nello stato i, quando tale stato viene visitato. E' immediato, sulla base della definizione di DTMC, dimostrare il seguente risultato.

Risultato 4 *La distribuzione della variabile aleatoria "tempo di permanenza nello stato* i*" è la distribuzione geometrica, con parametro* $p_{i,i}$

$$Pr\{V(i) = n\} = (1 - p_{ii})(p_{ii})^{n-1} \qquad n=1, 2, 3, \ldots.. \qquad (6)$$

Le probabilità di stato sono definite come

$$\pi_j\,(k) = Pr\,\{X_k = j\}$$

Il vettore di probabilità di stato è definito come

$$\pi\,(k) = [\pi_0\,(k), \pi_1\,(k), \ldots.]$$

e può essere un vettore a dimensione infinita. Si noti che il vettore di probabilità di stato è definito come un vettore riga. Dal teorema della probabilità totale si ha

$$\pi_j(k+1) = Pr\,\{X_{k+1} = j\} = \sum_{i \in S} Pr\{X_{k+1} = j | X_k = i\} \cdot Pr\{X_k = i\} = \sum_{i \epsilon S} p_{ij}\, \pi_i\,(k)$$

ovvero, in forma matriciale

$$\pi\,(k+1) = \pi\,(k)\,P \qquad (7)$$

da cui si ottiene immediatamente

$$\pi\,(k) = \ \pi\,(0)\,P^n \qquad\qquad k = 1,2, \dots . \qquad (8)$$

2.2 Classificazione degli stati in una DTMC

Definizione 1 *Uno stato j è detto* <u>*raggiungibile*</u> *dallo stato i (i≠j) se* $p_{i,j,n} > 0$ *per qualche n* = 1,2,...*Due stati si dicono* <u>*comunicanti*</u> *se sono mutualmente raggiungibili.*

Un sottoinsieme $\bar{S}$dello spazio degli stati S ;è detto <u>chiuso</u> se $p_{i,j} = 0\ \forall(i,j) : i \in \bar{S},\ j \notin \bar{S}$

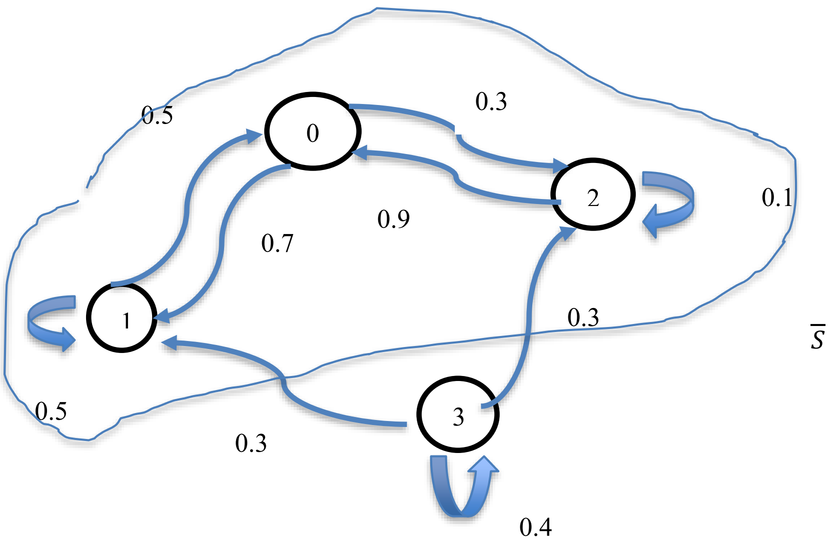


Fig. 1-Esempio di sottoinsieme chiuso.

Ad esempio, in Fig. 1, il sottoinsieme $\bar{S}\,\{0,1,2\}$ è chiuso. Se $\bar{S}$ è un sottoinsieme chiuso costituito da un solo stato i, si dice che i è uno stato assorbente (o stato trappola). In questo caso risulta $p_{i,i} = 1$.

Se $\bar{S}$ è un sottoinsieme chiuso in cui ogni coppia di stati è una coppia di stati comunicanti, allora si dice irriducibile. Se l'intero spazio degli stati è irriducibile, allora la DTMC stessa si dice irriducibile. Ovviamente, se la DTMC non è irriducibile, allora si dice riducibile.

Definizione 2 *Supponendo che la DTMC si trovi nello stato i allo stadio 0, il tempo di arrivo allo stato j è definito come*

$$T_{ij} = \min\{k > 0: X_k = j | X_0 = i\}$$

Se $j = i$, si ha $T_{i,i}$, che si dice tempo di ritorno allo stato i.

La variabile $T_{i,i}$ (come ovviamente anche qualunque variabile $T_{i,j}$) è evidentemente una variabile aleatoria (a valori discreti). Sia $\rho_i^k = Pr\{T_{i,i} = k\}$, $k = 1,2,\ldots\infty$. La probabilità che la DTMC prima o poi ritorni allo stato i è data da $\rho_i = \sum_{k=1}^{\infty} \rho_i^k$

Definizione 3 *Uno stato i è detto ricorrente se $\rho_i = 1$. Se $\rho_i < 1$, lo stato i è detto transiente.*

Nella DTMC in Figura 2, lo stato 2 è ricorrente mentre gli stati 0 e 1 sono transienti.

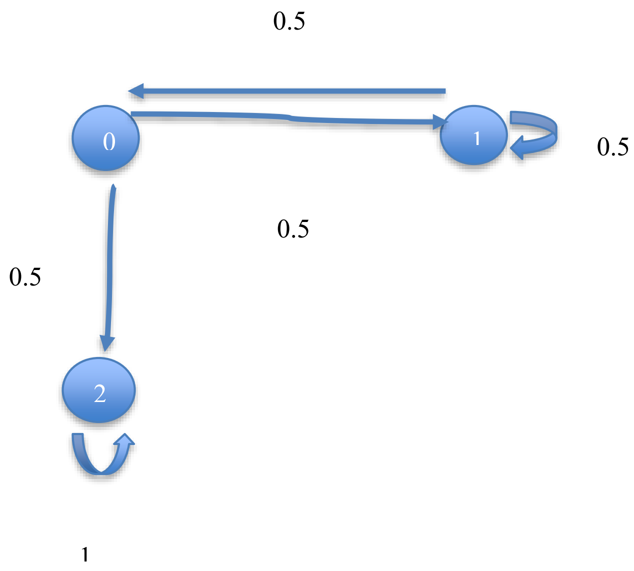


Figura 2: Esempio di DTMC con stati transienti e stati ricorrenti.

Valgono i seguenti risultati, le cui dimostrazioni, che qui non vengono fornite, sono però abbastanza semplici.

Risultato 5 *Se una DTMC ha uno spazio degli stati finito, almeno uno stato è ricorrente.*

Risultato 6 *Se i è uno stato ricorrente e j è raggiungibile da i, allora anche lo stato j è ricorrente.*

Risultato 7 *Se* $\bar{S}$ *è un sottoinsieme finito e irriducibile dello spazio degli stati, allora ogni stato in* $\bar{S}$ *è ricorrente.*

Definizione 4 *Sia i uno stato ricorrente. Il* <u>*tempo medio di ritorno*</u> *nello stato i è* $M_i = E[T_{i,i}] = \sum_{k=1}^{\infty} k\rho_i^k$.

Definizione 5 *Uno stato ricorrente i è detto* <u>*ricorrente positivo*</u> *se* $M_i < \infty$, *è detto* <u>*ricorrente nullo*</u> *se* $M_i = \infty$.

Risultato 8 *Se i è uno stato ricorrente positivo e j e raggiungibile da i, allora anche j è ricorrente positivo.*

Teorema 1 *Sia $\bar{S}$ un sottoinsieme irriducibile dello spazio degli stati; allora una e una sola delle seguenti affermazioni è vera:*

- *ogni stato in $\bar{S}$ è ricorrente positivo;*
- *ogni stato in $\bar{S}$ è ricorrente nullo;*
- *ogni stato in $\bar{S}$ è transiente.*

La dimostrazione di questo teorema è immediata, sulla base dei Risultati 8 e 6.

Teorema 2 *Sia $\bar{S}$ un sottoinsieme finito e irriducibile dello spazio degli stati. Allora ogni stato in $\bar{S}$ è ricorrente positivo.*

Dimostrazione. In forza del Risultato 7, ogni stato in $\bar{S}$ è ricorrente. In forza al Teorema 1, o ogni stato in $\bar{S}$ è ricorrente nullo oppure ogni stato in $\bar{S}$ è ricorrente positivo. Se fosse vera la prima ipotesi, allora partendo da un qualsiasi stato *i*, il tempo medio di ritorno sarebbe ∞. Quindi, mediamente, occorre attendere un numero infinito di stadi prima di ritornare allo stato *i*. Ma ciò è incompatibile con il fatto che gli stati per cui si dovrebbe passare sono in numero finito e sono tutti ricorrenti nulli. E' quindi vera la seconda possibilità, cioè gli stati sono tutti ricorrenti positivi.

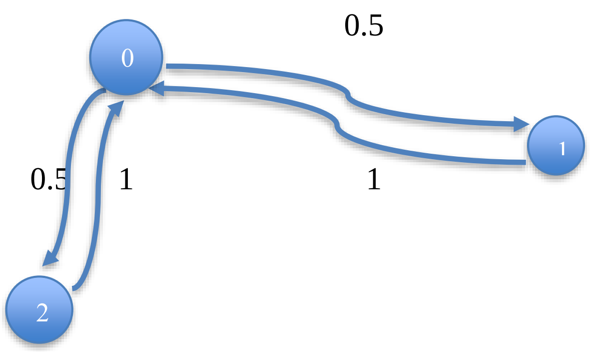
Si noti che nei teoremi e nei risultati precedenti $\bar{S}$ può anche coincidere con *S*.

Consideriamo ora l'insieme di interi $\{n > 0 : p_{i,i,n} > 0\}$. Sia d_i il massimo comune divisore in questo insieme.

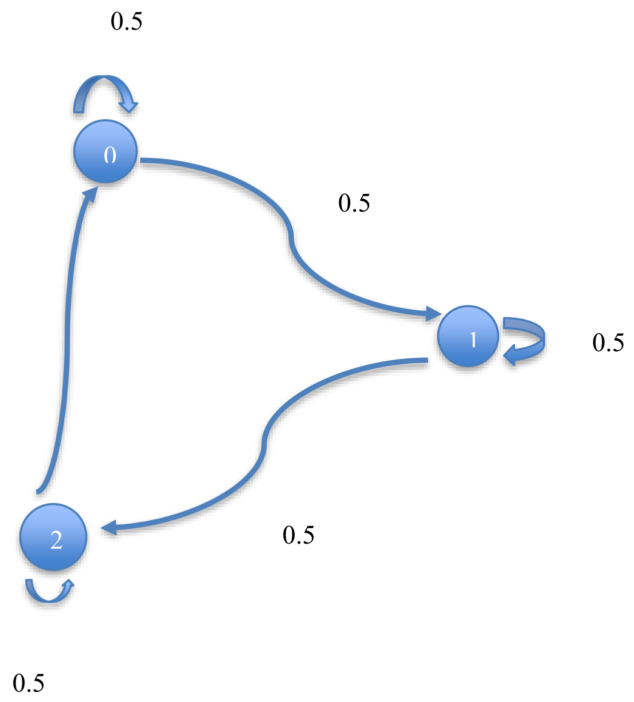
Definizione 6 *Uno stato ricorrente i è detto periodico (di periodo d_i) se $d_i \geq 2$. Se $d_i = 1$, lo stato i è detto aperiodico.*

Si noti che la precedente definizione si riferisce a stati transienti come a stati ricorrenti. Nella DTMC in Figura 3(a) tutti gli stati sono periodici con periodo 2, mentre nella DTMC in Figura 3(b) tutti gli stati sono aperiodici.

3a)



3b)



Teorema 3 (senza dimostrazione) *In una DTMC irriducibile, o tutti gli stati sono aperiodici, o tutti gli stati sono periodici con lo stesso periodo.*

2.3 Analisi a regime di una DTMC omogenea

Ci poniamo adesso il problema dell'<u>esistenza dei limiti</u> $\lim_{k\to\infty} \pi_j^k$ e del loro eventuale significato. Più precisamente ci poniamo i seguenti quesiti:

1. <u>sotto quali condizioni</u> i limiti $\pi_j = \lim_{k\to\infty} \pi_j^k$ <u>esistono</u> e sono <u>indipendenti da $\pi(0)$</u>?
2. se tali limiti esistono, essi <u>costituiscono una distribuzione di probabilità a regime ammissibile</u>, cioè tale che $\sum_{j\in S} \pi_j = 1$?
3. <u>come si determinano</u> tali limiti (ammesso che esistano)?

E' evidente che l'esistenza dei limiti in questione è connessa con l'<u>esistenza dei punti di equilibrio</u> dell'equazione (7), cioè di vettori π tali che $\pi = \pi\ P$. In effetti, a proposito delle precedenti domande, sussistono i seguenti risultati.

Teorema 4 (senza dimostrazione) *In una DTMC irriducibile aperiodica (cioè in cui tutti gli stati sono aperiodici), esistono i limiti* $\pi_j = \lim_{k\to\infty} \pi_j^k$, $\forall j \in S$, *e tali limiti sono indipendenti dal vettore delle probabilità di stato iniziali* $\pi(0)$.

Teorema 5 (senza dimostrazione) *In una DTMC irriducibile aperiodica in cui tutti gli stati sono transienti o ricorrenti nulli, risulta* $\pi_j = \lim_{k\to\infty} \pi_j^k = 0 \quad \forall j \in S$,

Teorema 6 (senza dimostrazione) *In una DTMC irriducibile aperiodica in cui tutti gli stati sono ricorrenti positivi, i limiti* π_j, $\forall j \in S$, *sono tutti* > 0 *e rappresentano la <u>distribuzione stazionaria di probabilità degli stati a regime</u>. Inoltre risulta* $\pi_j = 1/M_j$, $\forall j \in S$, *essendo* M_j *il tempo medio di ritorno allo stato j. <u>Il vettore</u>*

$$\pi = [\pi_0, \pi_1, \ldots]$$

<u>si determina in maniera univoca risolvendo il sistema di equazioni lineari</u>

$$\pi = \pi\ P \tag{9a}$$

$$\sum_{j\in S} \pi_j = 1 \tag{9b}$$

Una <u>DTMC irriducibile aperiodica in cui tutti gli stati sono ricorrenti positivi</u> si dice anche semplicemente <u>DTMC ergodica</u>. Si può notare che, sulla base del Teorema 2, <u>ogni DTMC irriducibile con spazio degli stati finito e con stati (tutti) aperiodici, è certamente una DTMC ergodica</u>, a cui può quindi essere applicato il Teorema 6.

2.4 Il processo birth-death a tempo discreto (DTMC-BD)

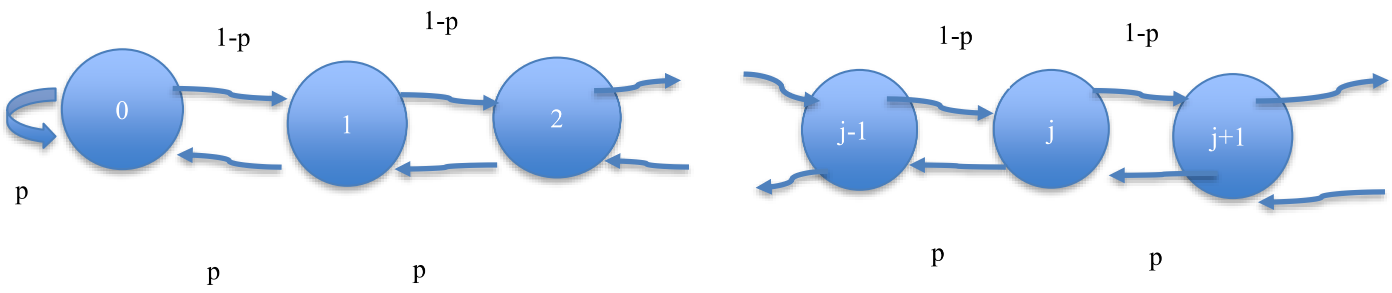


Figura 4: Processo birth-death a tempo discreto.

La matrice delle probabilità di transizione è data da

$$P = \begin{bmatrix} p & 1-p & 0 & 0 \ldots.. \\ p & 0 & 1-p & 0 \ldots.. \\ 0 & p & 0 & 1-p.. \\ 0 & 0 & p & 0.. \\ & \ldots & \ldots \ldots \ldots \ldots & \end{bmatrix}$$

E' facile mostrare (naturalmente supponendo $0 < p < 1$) che questa DTMC è irriducibile e aperiodica.

Consideriamo il sistema di equazioni (9) che in questo caso risulta essere

$$\pi_0 = \pi_0\, p + \pi_1\, p$$

$$\pi_j = \pi_{j-1}\,(1-p) + \pi_{j+1}\, p \qquad j=1,2,\ldots.$$

$$\sum_{j\in S} \pi_j = 1$$

Dalla prima delle precedenti equazioni si ha

$$\pi_1 = \frac{1-p}{p}\;\pi_0$$

Dalla seconda equazione, per $j = 1$, si ha, sostituendo π_1

$$\pi_2 = \left(\frac{1-p}{p}\right)^2 \pi_0$$

e in generale si ottiene

$$\pi_j = \left(\frac{1-p}{p}\right)^j \pi_0 \qquad\qquad j = 1,2,\ldots$$

L'ultima delle equazioni che compongono il sistema diventa quindi

$$\pi_0 + \sum_{j=1}^{\infty} \pi_j = \pi_0 + \pi_0 \sum_{j=1}^{\infty} \left(\frac{1-p}{p}\right)^j = 1$$

da cui si ricava $\pi_0 = \frac{1}{\sum_{i=0}^{\infty}\left(\frac{1-p}{p}\right)^i}$ e $\pi_{0j} = \frac{\left(\frac{1-p}{p}\right)^j}{\sum_{i=0}^{\infty}\left(\frac{1-p}{p}\right)^i}$ $\quad j = 1,2,\ldots$

Si hanno allora tre casi possibili:

a) $p > \frac{1}{2}$; in questo caso $\sum_{i=0}^{\infty}\left(\frac{1-p}{p}\right)^i$ è convergente e vale $\frac{p}{2p-1}$; si può dimostrare che in questo caso la DTMC è ergodica; le probabilità $\pi_0 = \frac{2p-1}{p}$ e $\pi_j = \frac{2p-1}{p}\left(\left(\frac{1-p}{p}\right)\right)^j$, $j = 1,2,\ldots$, rappresentano quindi la distribuzione stazionaria di probabilità degli stati a regime;

b) $p < \frac{1}{2}$; in questo caso $\sum_{i=0}^{\infty}\left(\frac{1-p}{p}\right)^i$ non è convergente; si può dimostrare che in questo caso tutti gli stati della DTMC sono transienti;

c) $p = \frac{1}{2}$; anche in questo caso la somma non converge; si può dimostrare che in questo caso tutti gli stati della DTMC sono ricorrenti nulli.

3 *Catene di Markov a tempo continuo*

3.1 Definizioni e prime proprietà

Definizione 7 *Una catena di Markov a tempo continuo (CTMC) è un sistema a stato discreto e tempo continuo in cui l'evoluzione della variabile di stato X(t) soddisfa la seguente proprietà (Markov property):*

$$Pr\{X(t_{k+1}) = x_{k+1} | X(t_k) = x_k, \dots, X(t_{k-1}) = x_{k-1}, \dots X(t_0) = x_0\} = Pr\{X_{k+1} = x_{k+1} | X_k = x_k\} \quad (10)$$

per ogni scelta di $t_0, t_1, \dots, t_{k-1}, t_k, t_{k+1}$ *tale che* $t_0 \leq t_1 \leq \dots \leq t_{k-1} \leq t_k \leq t_{k+1}$.

La precedente definizione stabilisce che $X(t)$ condensa tutta l'informazione passata (per $\tau \leq t$) ai fini della valutazione (probabilistica) dell'evoluzione della CTMC per $\tau > t$. E' immediato dedurre dalla definizione ora data che l'informazione sul tempo già trascorso (all'istante t) nello stato $i = x(t)$ (detto anche "età" dello stato i all'istante t) è superflua ai fini della suddetta valutazione.

L'evoluzione di una CTMC è governata dalle funzioni di transizione, definite come le probabilità di transizione da uno stato in un certo istante ad un altro stato in un altro istante

$$p_{ij}(s,t) = Pr\{X(t) = j | X(s) = i\} \qquad s \leq t \qquad \forall i, j \epsilon S \qquad (11)$$

Risulta, evidentemente, sulla base del teorema della probabilità totale e della precedente definizione (per $s \leq u \leq r$)

$$p_{ij}(s,t) = \sum_{r \in S} Pr\{X(t) = j | X(u) = r, X(s) = i\} \cdot Pr\{X(u) | \ X(s) = i\}$$
$$= \sum_{r \in S} Pr\{X(t) = j | X(u) = r\} \cdot Pr\{X(u) | \ X(s) = i\}$$

che può essere riscritta come

$$p_{ij}(s,t) = \sum_{r \in S} p_{ir}(s,u) p_{rj}(u,t) \qquad s \le u \le t \tag{12}$$

che è l'<u>equazione di Chapman-Kolmogorov nella sua versione a tempo continuo</u>.

Definiamo $\boldsymbol{P}(s,t) = [p_{i,j}(s,t)]$. Ovviamente risulterà $\boldsymbol{P}(s,s) = I$ e $\sum_{j \in S} p_{ij}(s,t) = 1$, $\forall i,j \in S$, $s \le t$. L'equazione (12) si può quindi scrivere

$$\boldsymbol{P}(s,t) = \boldsymbol{P}(s,u)\, \boldsymbol{P}(u,t) \qquad s \le u \le t \tag{13}$$

La (13) può essere scritta per $s \le t \le t + \Delta t$, *con* $\Delta t > 0$,

$$\boldsymbol{P}(s,t+\Delta t) = \boldsymbol{P}(s,t)\, \boldsymbol{P}(t,t+\Delta t)$$

ovvero, sottraendo $\boldsymbol{P}(s,t)$ ad ambo i membri,

$$\boldsymbol{P}(s,t+\Delta t) - \boldsymbol{P}(s,t) = \boldsymbol{P}(s,t)\, [\mathbf{P}(t, t+\Delta t) - I]$$

dividendo per Δt e passando al limite per $\Delta t \to 0$ si ha

$$\lim_{\Delta t \to 0} \frac{\boldsymbol{P}(s,t+\Delta t) - \boldsymbol{P}(s,t)}{\Delta t} = \boldsymbol{P}(s,t) \lim_{\Delta t \to 0} \frac{\boldsymbol{P}(t,t+\Delta t) - I}{\Delta t}$$

La matrice

$$Q = \lim_{\Delta t \to 0} \frac{\boldsymbol{P}(t,t+\Delta t) - \boldsymbol{I}}{\Delta t}$$

è chiamata <u>matrice dei rate di transizione</u>, o <u>generatore infinitesimale</u>. L'ultima equazione ottenuta può quindi riscriversi

$$\frac{\partial \mathbf{P}(s,t)}{\partial t} = \boldsymbol{P}(s,t)\, Q(t) \tag{14}$$

(con $s \leq t$) che viene chiamata equazione differenziale forward di Chapman-Kolmogorov. In maniera analoga, scrivendo la (13) per $s \leq s+\Delta s \leq t$, può essere ottenuta l'equazione differenziale backward di Chapman-Kolmogorov

$$\frac{\partial \mathbf{P}\ (s,t)}{\partial s} = -\,Q(s)\boldsymbol{P}\ (s,t) \tag{15}$$

(con $s \leq t$).

Il significato della matrice $Q(t)$ e delle equazioni differenziali di Chapman-Kolmogorov risulta chiaro nel caso delle CTMC omogenee (cioè tempo invarianti, o stazionarie). Una CTMC omogenea è una CTMC in cui le funzioni di transizione (12) sono dipendenti da s e t esclusivamente attraverso la differenza $\tau = t - s$. In altre parole, nel caso di CTMC omogenee, le funzioni (probabilità) di transizione

$$p_{ij}(s, t+\tau) = p_{ij}(\tau) Pr\{X(s+\tau) = j | X(s) = i\} \qquad \forall i, j \epsilon S$$

sono indipendenti da s. $\boldsymbol{P}\ (s,s+\tau)$ verrà allora indicata con $P(\tau)$. Nel seguito ci si riferirà esclusivamente a CTMC omogenee, lasciando tale assunzione di regola sottointesa.

Nel caso di una CTMC omogenea, ovviamente $Q(t)$ risulta indipendente da t. Verrà quindi indicata semplicemente con Q. La (14) diventa

$$\frac{dP(\tau)}{d\tau} = P(\tau)Q \tag{16}$$

che deve essere integrata con condizioni iniziali $p_{i,j}(0) = 0$, $i \neq j$, $p_{i,i}(0) = 1$, $\forall i \epsilon S$;. La soluzione della (16) è

$$P(\tau) = e^{Q\tau} \tag{17}$$

Essendo $\quad e^{Q\tau} = I + Q\tau + \frac{Q^2}{2!}\tau^2 + \cdots$..

Si consideri ora la variabile aleatoria $V(i)$ corrispondente al tempo di permanenza nello stato i, quando tale stato viene raggiunto. E' assai semplice dimostrare il seguente risultato.

Risultato 9 *La variabile aleatoria V (i) è distribuita in modo esponenziale, ovvero*

$$Pr\{V(i) \leq t\} = \begin{cases} 1 - e^{-\Lambda(i)t} & t \geq 0 \\ 0 & t < 0 \end{cases} \qquad (18)$$

Dimostrazione. Si osservi che, sulla base di definizione della CTMC, la lunghezza dell'intervallo di tempo per cui la CTMC è già stata nello stato presente non dà alcun contributo informativo. Poiché la distribuzione esponenziale è l'unica distribuzione che gode della memoryless property (si veda il Risultato 3), ne deriva che $V(i)$ è distribuito in modo esponenziale.

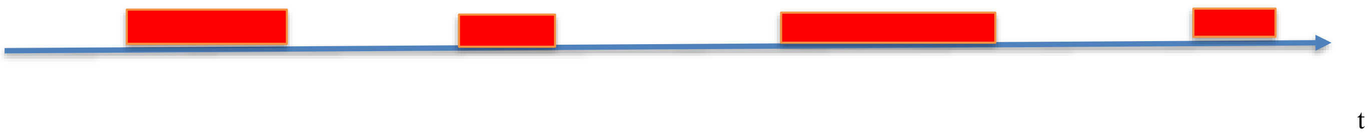


Figura 5: Intervalli di tempo nei quali il sistema si trova nello stato *i*.

Immaginiamo adesso di identificare, nell'evoluzione temporale della CTMC gli intervalli di tempo nei quali il sistema si trova nello stato *i* (si veda la Figura 5). Gli estremi a destra di tali intervalli corrispondono al verificarsi di eventi del tipo "la CTMC esce dallo stato *i*". Tali intervalli rappresentano realizzazioni successive della variabile aleatoria $V(i)$, distribuita in modo esponenziale. Essi possono essere quindi pensati come la sequenza dei tempi di interevento di un processo di Poisson con parametro $\Lambda(i)$. Evidentemente risulta $E[V(i)]=1/\Lambda(i)$. $\Lambda(i)$ ha quindi il significato di "rate di uscita dallo stato *i*".

Si può provare che il processo di Poisson ora considerato può essere pensato come la sovrapposizione di processi di Poisson indipendenti, ciascuno corrispondente alla sequenza di eventi "transizione dallo stato *i* allo stato *j*" ($j \neq i$). Sia $\lambda_{i,j}$ il parametro caratterizzante il generico processo di Poisson di tale tipo. Risulta allora, in base a quanto visto nel Paragrafo 1.3,

$$\Lambda(i) = \sum_{j \epsilon S, j \neq i} \lambda_{ij} \tag{19}$$

Naturalmente, se uno stato *j* non può essere raggiunto dallo stato *i* (cioè se l'evento "transizione dallo stato *i* allo stato *j*" non è un evento ammissibile), risulta $\lambda_{i,j} = 0$.

A questo punto è possibile tentare di fornire una interpretazione "fisica" dei coefficienti della matrice *Q*. A questo scopo, scriviamo la generica equazione scalare fornita dalla (16)

$$\frac{dp_{ij}(\tau)}{d\tau} = p_{ij}(\tau)q_{jj} \sum_{r \neq j} p_{ir}(\tau)q_{rj} \tag{20}$$

Consideriamo prima il caso $i = j$. La (20) diventa allora

$$\frac{dp_{ii}(\tau)}{d\tau} = p_{ii}(\tau)q_{ii} \sum_{r \neq i} p_{ir}(\tau)q_{ri}$$

Ponendo $\tau = 0$ ed usando le condizioni $p_{i,i}(0) = 1$ e $p_{i,r}(0) = 0$, $r \neq i$, si ottiene

$$\left.\frac{dp_{ii}(\tau)}{d\tau}\right|_{\tau=0} = q_{ii} \tag{21}$$

che può essere riscritta come

$$-q_{ii} = \frac{d}{d\tau}[1 - p_{ii}(\tau)]\Big|_{\tau=0} = \lim_{\tau \to 0} \frac{[1 - p_{ii}(\tau)] - [1 - p_{ii}(0)]}{\tau}$$

$1 - p_{i,i}(\tau)$ rappresenta la probabilità che la CTMC lasci lo stato *i* (per andare in un altro stato qualsiasi) in un intervallo di tempo di lunghezza τ. La precedente equazione ci consente quindi di affermare che tale probabilità, quando $\tau \to 0$, è data da $(-q_{i,i})\,\tau + O(\tau)$, essendo $O(\tau)$ un infinitesimo di ordine superiore a τ. Ne consegue che $-q_{i,i}$ può essere interpretato come la frequenza (rate) con cui si verifica la transizione "uscita dallo stato *i*".

Notiamo adesso che, sulla base della (18)

$$p_{ii}(\tau) = Pr\{V(i) > \tau\} = e^{-\Lambda(i)\tau}$$

e quindi

$$\frac{dp_{ii}(\tau)}{d\tau} = -\Lambda(i)e^{-\Lambda(i)\tau}$$

Per $\tau = 0$, tenendo presente la (21), si ha

$$-q_{ii} = \Lambda(i) \tag{22}$$

come è ovvio che sia, vista l'interpretazione del parametro $\Lambda(i)$. Si noti che è sempre $q_{ii} < 0, \forall i \in S$.

Consideriamo adesso il caso $i \neq j$ nella (20). In maniera analoga a quanto visto per il caso $i = j$, si ottiene ora facilmente

$$q_{ij} = \left.\frac{dp_{ij}(\tau)}{d\tau}\right|_{\tau=0} \tag{23}$$

ovvero

$$q_{ij} = \lim_{\tau \to 0} \frac{p_{ij}(\tau) - p_{ij}(0)}{\tau}$$

In maniera analoga a quanto visto sopra per il caso $i = j$, si comprende quindi che il parametro <u>$q_{i,j}$ può essere interpretato come la frequenza (rate) con cui si verifica la transizione "passaggio dallo stato i allo stato j"</u>. Risulterà allora, in analogia con quanto espresso dalla (22),

$$q_{ij} = \lambda_{ij} \tag{24}$$

Inoltre, tenendo presente che $\sum_{j \epsilon S} p_{ij}(\tau) = 1$, $\forall i \epsilon S$, $\forall \tau \geq 0$, derivando rispetto a τ, ponendo $\tau = 0$, e tenendo presente la (21) e la (23), si ha

$$\sum_{j \epsilon S} q_{ij} = 0$$

ovvero

$$-q_{ii} = \sum_{j \neq i} q_{ij} \tag{25}$$

come è ovvio che sia, date la (22) e la (24).

Infine, si definiscano per una CTMC le "probabilità di transizione a tempo discreto"come segue

$$\tilde{p}_{ij} = Pr\{X_{k+1}|X_k = i\} = probabilità\ che\ il\ prossimo\ stato\ sia\ j, dato\ lo\ stato\ corrente\ i$$

$\tilde{p}_{ij}$è la probabilità che si verifichi per primo l'evento "passaggio da *i* a *j*", quando il sistema si trova nello stato i. Risulterà quindi

$$\tilde{p}_{ij} = \frac{\lambda_{ij}}{\Lambda(i)} = -\frac{q_{ij}}{q_{ii}}$$

(ovviamente risulterà $\sum_{j\neq i} \tilde{p}_{ij} = 1$).

La DTMC definita dalle probabilità di transizione di una CTMC viene denominata "DTMC embedded nella CTMC". Essa rappresenta solo la probabilità di transizione da uno stato all'altro, e trascura l'informazione relativa alla frequenza delle transizioni, ai tempi medi di permanenza negli stati, ecc. Si noti che la DTMC così definita ha sempre una matrice delle probabilità di transizione in cui la diagonale principale è tutta fatta di zeri.

3.2 Analisi di transitorio e a regime di una CTMC omogenea

In maniera analoga a quanto visto per le DTMC, si possono definire le probabilità di stato come

$$\pi_j(t) = Pr\{X(t) = j\}$$

Il vettore delle probabilità di stato è $\pi(t) = [\pi_0(t), \pi_1(t),\ldots]$. Sempre sulla base del teorema della probabilità totale si ha

$$\pi_j(t) = Pr\{X(t) = j\} = \sum_{i\in S} Pr\{X(t) = j|X(0) = i\}\, Pr\{X(0) = i\} = \sum_{i\in S} p_{ij}(t)\pi_i(0)$$

$$t\geq 0$$

ovvero, in forma matriciale

$$\pi(t) = \pi(0)P(t) = \pi(0)e^{Q(t)} \tag{26}$$

Differenziando rispetto a *t* si ha

$$\frac{d\pi(t)}{dt} = \pi(t)Q \tag{27}$$

ovvero, in forma scalare,

$$\frac{d\pi_j(t)}{dt} = q_{jj}\pi_j(t) + \sum_{i\neq j} q_{ij}\pi_i\,(t) \tag{28}$$

j=0,1,2…

che riscritta come

$$d\pi_j(t) = \left(q_{jj}dt\right)\pi_j(t) + \sum_{i\neq j}\left(q_{ij}dt\right)\pi_i(t)$$

risulta più facilmente interpretabile. Si noti infatti che -($q_{j,j}dt$) è la probabilità che la CTMC lasci lo stato *j* (ammesso che vi si trovi) in un intervallo di tempo di lunghezza *d*t (ammesso che *dt* sia molto piccolo). Simile interpretazione ha la quantità ($q_{i,j}dt$).

In maniera analoga a quanto visto per il caso a tempo discreto, ci chiediamo ora quando esistono i limiti $\lim_{t\to\infty} \pi_j\,(t)$, indicati semplicemente con π_j, $j \in S$, e se ad essi è possibile attribuire il significato di distribuzione di probabilità dello stato a regime. A tale proposito, verrà qui fornito solo il risultato più importante. Per la comprensione di tale risultato, occorre premettere che, per quanto riguarda un CTMC, gli stati possono essere classificati in modo perfettamente analogo a quanto visto nel paragrafo 2.2, facendo riferimento alla classificazione che vale per la DTMC embedded.

Teorema 7 (senza dimostrazione) *In una CTMC irriducibile in cui tutti gli stati sono ricorrenti positivi (e quindi in ogni CTMC irriducibile con uno spazio degli stati finiti), esiste un'unica distribuzione di probabilità a regime dello stato tale che* $\pi_j > 0$ *e* $\pi_j = \lim_{t\to\infty} \pi_j\,(t)$. *Tale distribuzione è indipendente da* $\pi(0)$. *Inoltre* $\pi = [\pi_0, \pi_1, \dots]$ *può essere determinato in modo univoco risolvendo il sistema di equazioni*

$$\pi Q = 0 \tag{29a}$$

$$\sum_{j \in S} \pi_j = 1 \qquad (29b)$$

Le (29a) possono essere scritte in forma scalare come

$$q_{jj}\pi_j + \sum_{i \neq j} q_{ij}\pi_i = 0 \qquad j \in S$$

3.3 Il processo birth-death a tempo continuo (CTMC-BD)

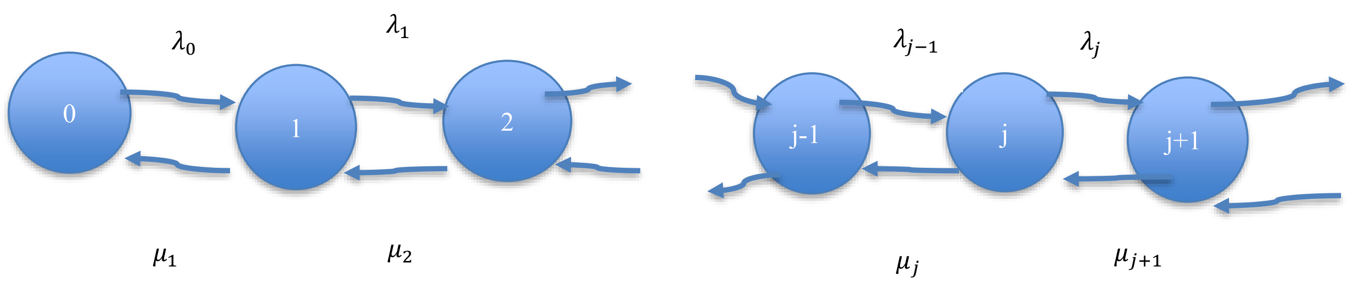


Figura 6. Processo birth-death a tempo continuo.

Si tratta di una CTMC (si veda la Figura 6) tale che

$$q_{ij} = 0 \quad per\ j > i+1 \ \ e \ \ per\ j < i-1$$

Il valore $\lambda_j = q_{j,j+1} > 0$ si chiama "<u>birth rate</u>" nello stato j, $j = 0,1,2,\ldots$, mentre il valore $\mu_j = q_{j,j-1} > 0$ si chiama "<u>death rate</u>" nello stato j, $j = 1,2,3, \ldots$

Evidentemente risulterà

$$q_{jj} = -(\lambda_j + \mu_j) \qquad j = 1,2,\ldots$$

$$q_{00} = -\lambda_0$$

e quindi la <u>matrice dei rate di transizione</u> risulta

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & \dots\dots\dots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 \\ \dots\dots & 0 & \mu_3 & \dots\dots \end{bmatrix}$$

Le equazioni (28) diventano in questo caso

$$\frac{d\pi_j(t)}{dt} = -(\lambda_j+\mu_j)\pi_j(t) + \lambda_{j-1}\pi_{j-1}(t) + \mu_{j+1}\pi_{j+1}(t) \qquad j=1,2,\dots \tag{30a}$$

$$\frac{d\pi_0(t)}{dt} = -\lambda_0\pi_0(t) + \mu_1\pi_1(t) \tag{30b}$$

Vediamo adesso di studiare la CTMC-BD all'<u>equilibrio</u> e cominciamo col considerare le condizioni (29), particolarizzate alla CTMC in questione. Si ha

$$-(\lambda_j+\mu_j)\pi_j + \lambda_{j-1}\pi_{j-1} + \mu_{j+1}\pi_{j+1} = 0 \qquad j = 1,2,\dots$$

$$-\lambda_0\pi_0(t) + \mu_1\pi_1(t) = 0$$

$$\sum_{j\in S} \pi_j = 1$$

Si ottiene facilmente

$$\pi_1 = \frac{\lambda_0}{\mu_1}\pi_0$$

$$\pi_2 = \frac{\lambda_0\lambda_1}{\mu_1\mu_2}\pi_0$$

e in generale

$$\pi_j = \frac{\lambda_0 \dots . \lambda_{j-1}}{\mu_1 \dots . \mu_j}\pi_0 \tag{31}$$

Per cui risulta

$$\pi_0 + \pi_0 \sum_{j=1}^{\infty} \frac{\lambda_0 \dots . \lambda_{j-1}}{\mu_1 \dots . \mu_j} = 1$$

ovvero

$$\pi_0 = \frac{1}{1 + \sum_{j=1}^{\infty} \frac{\lambda_0 \ldots . \lambda_{j-1}}{\mu_1 \ldots . \mu_j}}$$

(32)

Si può dimostrare che in una CTMC-BD, se $\exists \bar{j}$ tale che $\frac{\lambda_j}{\mu_j} < 1$ per $j \geq \bar{j}$, tutti gli stati sono ricorrenti positivi.

Si applica in tal caso il Teorema 7, che consente di affermare che esiste una distribuzione di probabilità a regime dello stato i cui valori sono forniti dalle (31) e (32).

RIFERIMENTI BIBLIOGRAFICI

C. G. Cassandras, S. Lafortune, *Introduction to Discrete Event Systems*, 2nd ed., Springer, 2008

G. Bolche, S. Greiner, H. de Meer, K.S. Trivedi, *Queueing Networks and Markov Chains*, 2nd ed, J. Wiley, 2006.