Probabilità e Processi Stocastici (455AA)

Lezione 14

Dario Trevisan – *https://web.dm.unipi.it/trevisan*

05/11/2025

Legge dei processi di Markov a salti

Richiami sulle catene di Markov

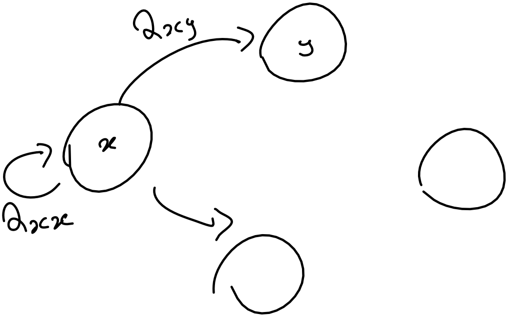
- $(X_n)_{n=0}^{N}$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.

Richiami sulle catene di Markov

- $(X_n)_{n=0}^N$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.
- Matrice di transizione $Q_{x \to y} = P(X_1 = y | X_0 = x)$ (stocastica)

Richiami sulle catene di Markov

- $(X_n)_{n=0}^N$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.
- Matrice di transizione $Q_{x\to y} = P(X_1 = y | X_0 = x)$ (stocastica)
- Rappresentazione grafica



Richiami sulle catene di Markov

- $(X_n)_{n=0}^N$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.
- Matrice di transizione $Q_{x \to y} = P(X_1 = y | X_0 = x)$ (stocastica)
- Rappresentazione grafica
- Peso di un cammino $\gamma = x_0 \to x_1 \to \ldots \to x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \ldots Q_{x_{n-1} x_n}$$

Richiami sulle catene di Markov

- $(X_n)_{n=0}^N$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.
- Matrice di transizione $Q_{x \to y} = P(X_1 = y | X_0 = x)$ (stocastica)
- Rappresentazione grafica
- Peso di un cammino $\gamma = x_0 \to x_1 \to \ldots \to x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \ldots Q_{x_{n-1} x_n}$$

- Legge del processo

$$P(X = \gamma) = P(X_0 = x_0) Q_\gamma$$

Richiami sulle catene di Markov

- $(X_n)_{n=0}^N$, un processo stocastico a **tempi discreti, stati discreti, di Markov, omogeneo**.
- Matrice di transizione $Q_{x \to y} = P(X_1 = y | X_0 = x)$ (stocastica)
- Rappresentazione grafica
- Peso di un cammino $\gamma = x_0 \to x_1 \to \ldots \to x_n$,

$$Q_\gamma = Q_{x_0 x_1} Q_{x_1 x_2} \ldots Q_{x_{n-1} x_n}$$

- Legge del processo

$$P(X = \gamma) = P(X_0 = x_0) Q_\gamma$$

- Legge marginale $\mu_n = \mu_{n-1} Q$, da cui $\mu_n = \mu_0 Q^n$.

Richiami sui processi a salti

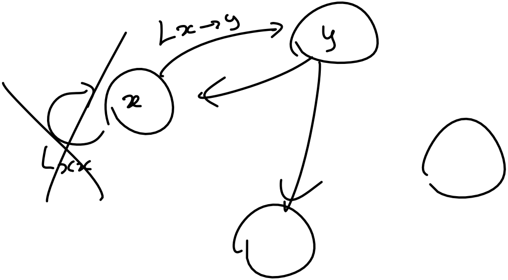
- $(X_t)_{t \in [0,T]}$, un processo stocastico a **tempi continui, stati discreti, di Markov, omogeneo**.

Richiami sui processi a salti

- $(X_t)_{t\in[0,T]}$, un processo stocastico a **tempi continui, stati discreti, di Markov, omogeneo**.
- Matrice di intensità di salto $L_{x\to y} = \frac{d}{dt}\Big|_{t=0} P(X_t = y|X_0 = x)$

Richiami sui processi a salti

- $(X_t)_{t\in[0,T]}$, un processo stocastico a **tempi continui, stati discreti, di Markov, omogeneo**.
- Matrice di intensità di salto $L_{x\to y} = \frac{d}{dt}\Big|_{t=0} P(X_t = y | X_0 = x)$
- Rappresentazione grafica



Richiami sui processi a salti

- $(X_t)_{t\in[0,T]}$, un processo stocastico a **tempi continui, stati discreti, di Markov, omogeneo**.
- Matrice di intensità di salto $L_{x\to y} = \frac{d}{dt}\big|_{t=0} P(X_t = y | X_0 = x)$
- Rappresentazione grafica
- Legge marginale $\mu_t = \mu_0 \exp(tL)$, da cui la *master equation*

$$\frac{d}{dt}\mu_t = \mu_t L.$$

$(\mu_t)_t \rightsquigarrow P(X=\gamma) \overset{?}{=}$

Legge del processo di Markov a salti

Consideriamo la probabilità di osservare un cammino che visiti nell'ordine gli stati $x_0 \to x_1 \to \ldots \to x_n$.

- Trattandosi di tempi continui, dobbiamo specificare i *tempi di permanenza* in ciascuno stato $t_1, \ldots, t_n$, in modo che il processo "salti" al tempo t_1 dallo stato x_0 verso x_1, al tempo $t_1 + t_2$ da x_1 verso x_2, e così via.

Legge del processo di Markov a salti

Consideriamo la probabilità di osservare un cammino che visiti nell'ordine gli stati $x_0 \to x_1 \to \ldots \to x_n$.

- Trattandosi di tempi continui, dobbiamo specificare i *tempi di permanenza* in ciascuno stato $t_1, \ldots, t_n$, in modo che il processo "salti" al tempo t_1 dallo stato x_0 verso x_1, al tempo $t_1 + t_2$ da x_1 verso x_2, e così via.
- Fissato δ tale che $t_1 = \delta h_1$, $t_2 = \delta h_2$ ecc., si trova l'approssimazione

$$
P(X^\delta = (x_0 \to \ldots \to x_n)) \approx P(X_0 = x_0) \overbrace{\left(Id_{x_0x_0} + \frac{t_1 L_{x_0x_0}}{h_1} \right)^{h_1}}^{\text{permanenza in } x_0} \underbrace{\delta L_{x_0x_1}}_{\text{salto}}
$$

$$
\cdot \underbrace{\left(Id_{x_1x_1} + \frac{t_2 L_{x_1x_1}}{h_2} \right)^{h_2}}_{\text{permanenza in } x_1} \underbrace{(\delta L_{x_1x_2})}_{\text{salto}} \cdot \ldots
$$

$$
\cdot \left(Id_{x_{n-1}x_{n-1}} + \frac{t_n L_{x_{n-1}x_{n-1}}}{h_n} \right)^{h_n} (\delta L_{x_{n-1}x_n}).
$$

$= \delta$

$\approx \delta^n$

- Al tendere di δ a zero si trova che

$$P(X^\delta = (x_0 \to \ldots \to x_n)) \approx P(X_0 = x_0) \left(1 + \frac{t_1 L_{x_0 x_0}}{h_1}\right)^{h_1} (\delta L_{x_0 x_1}) \cdot$$
$$\cdot \left(1 + \frac{t_2 L_{x_1 x_1}}{h_2}\right)^{h_2} (\delta L_{x_1 x_2}) \cdot \ldots$$
$$\cdot \left(1 + \frac{t_n L_{x_{n-1} x_{n-1}}}{h_n}\right)^{h_n} (\delta L_{x_{n-1} x_n}) \to 0$$

- Al tendere di δ a zero si trova che

$$P(X^\delta = (x_0 \to \ldots \to x_n)) \approx P(X_0 = x_0) \left(1 + \frac{t_1 L_{x_0 x_0}}{h_1}\right)^{h_1} (\delta L_{x_0 x_1}) \cdot$$
$$\cdot \left(1 + \frac{t_2 L_{x_1 x_1}}{h_2}\right)^{h_2} (\delta L_{x_1 x_2}) \cdot \ldots$$
$$\cdot \left(1 + \frac{t_n L_{x_{n-1} x_{n-1}}}{h_n}\right)^{h_n} (\delta L_{x_{n-1} x_n}) \to 0$$

- Dividendo per δ^n si ottiene una “densità di probabilità” non nulla, data dall’espressione

salti

$$p(X = \gamma) = P(X_0 = x_0) \prod_{k=1}^{n} \overbrace{\exp\left(t_k L_{x_{k-1} x_{k-1}}\right)}^{\text{densità esponenziale}} \prod_{k=1}^{n} L_{x_{k-1} x_k} \frac{(-L_{x_{k-1}x_{k-1}})}{(-L_{x_{k-1}x_{k-1}})}$$

tempi di permanenza in x_{k-1}

La derivazione rigorosa si ottiene introducendo

- le variabili aleatorie $T_1, T_2, \ldots, T_n$ come tempi di permanenza del processo in ciascuno degli stati visitati $x_0, x_1, \ldots, x_{n-1}$

La derivazione rigorosa si ottiene introducendo

- le variabili aleatorie $T_1, T_2, \ldots, T_n$ come tempi di permanenza del processo in ciascuno degli stati visitati $x_0, x_1, \ldots, x_{n-1}$
- tempi di salto sono dati dalle somme $S_1 = T_1$, $S_2 = T_1 + T_2$, $\ldots$, $S_n = T_1 + T_2 + \ldots + T_n$.

La derivazione rigorosa si ottiene introducendo

- le variabili aleatorie T_1, T_2, ..., T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, ..., x_{n-1}
- tempi di salto sono dati dalle somme $S_1 = T_1$, $S_2 = T_1 + T_2$, ..., $S_n = T_1 + T_2 + \ldots + T_n$.
- La densità è quindi

$$
\begin{aligned}
p(T_1 = t_1, X_{S_1} = x_1, T_2 = t_2, X_{S_2} = x_2 \ldots, T_n = t_n, X_{S_n} = x_n) \\
= P(X_0 = x_0) \prod_{k=1}^{n} \exp\left(t_k L_{x_{k-1}x_{k-1}}\right) L_{x_{k-1}x_k},
\end{aligned}
$$

La derivazione rigorosa si ottiene introducendo

- le variabili aleatorie T_1, T_2, ..., T_n come tempi di permanenza del processo in ciascuno degli stati visitati x_0, x_1, ..., x_{n-1}
- tempi di salto sono dati dalle somme $S_1 = T_1$, $S_2 = T_1 + T_2$, ..., $S_n = T_1 + T_2 + \ldots + T_n$.
- La densità è quindi

$$
\begin{aligned}
p(T_1 = t_1, X_{S_1} = x_1, T_2 = t_2, X_{S_2} = x_2 \ldots, T_n = t_n, X_{S_n} = x_n) \\
= P(X_0 = x_0) \prod_{k=1}^{n} \exp\left(t_k L_{x_{k-1}x_{k-1}}\right) L_{x_{k-1}x_k},
\end{aligned}
$$

- le variabili T_1, T_2, ..., T_n sono continue, mentre le variabili X_{S_1}, X_{S_2}, ..., X_{S_n} sono discrete.

a valori negli stati

a valori in $[0, +\infty)$

Una intepretazione alternativa dei processi a salti

La formula si può riscrivere separando le variabili continue da quelle discrete:

peso discreto

$$P(X_{S_1} = x_1, X_{S_2} = x_2 \dots, X_{S_n} = x_n) = P(X_0 = x_0) \prod_{k=1}^{n} \frac{L_{x_{k-1}x_k}}{-L_{x_{k-1}x_{k-1}}},$$

e

$$p(T_1 = t_1, T_2 = t_2, \dots, T_n = t_n | X_{S_1} = x_1, \dots X_{S_n} = x_n)$$
$$= P(X_0 = x_0) \prod_{k=1}^{n} (-L_{x_{k-1}x_{k-1}}) \exp\left(t_k L_{x_{k-1}x_{k-1}}\right).$$

Variabili esponenziali indipendenti

$$P(X_{S_1} = x_1, X_{S_2} = x_2 \ldots, X_{S_n} = x_n) = P(X_0 = x_0) \prod_{k=1}^{n} \frac{L_{x_{k-1}x_k}}{-L_{x_{k-1}x_{k-1}}},$$

- La prima equazione mostra che le variabili X_0, X_{S_1}, ..., X_{S_n} che indicano gli stati visitati dal processo sono una catena di Markov (a tempi discreti) con probabilità di transizione (per $x \neq y$)

$$Q_{xy} = \frac{L_{xy}}{-L_{xx}},$$

$$p(T_1 = t_1, T_2 = t_2, \ldots, T_n = t_n | X_{S_1} = x_1, \ldots X_{S_n} = x_n)$$

$$= P(X_0 = x_0) \prod_{k=1}^{n} (-L_{x_{k-1}x_{k-1}}) \exp\left(t_k L_{x_{k-1}x_{k-1}}\right).$$

- la seconda mostra che, noti gli stati visitati, i *tempi di permanenza* T_1, T_2, ..., T_n sono variabili aleatorie indipendenti tra loro, e ciascuna T_k ha densità continua esponenziale di parametro $-L_{x_{k-1}x_{k-1}}$.

Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:

Esempio

$$L = \begin{array}{cc} & \begin{array}{ccc} ON & SB & OFF \end{array} \\ \begin{array}{c} ON \\ SB \\ OFF \end{array} & \begin{pmatrix} -4 & 3 & 1 \\ 5 & -7 & 2 \\ 10 & 0 & -10 \end{pmatrix} \end{array}$$

$$Q = \begin{pmatrix} 0 & 3/4 & 1/4 \\ 5/7 & 0 & 2/7 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\Rightarrow (4, 7, 10)$$

Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \to x_i}$

Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \to x_i}$
  - stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \to y}$, $y \neq x_i$

Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \to x_i}$
  - stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \to y}$, $y \neq x_i$
- *Processi a salti*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:

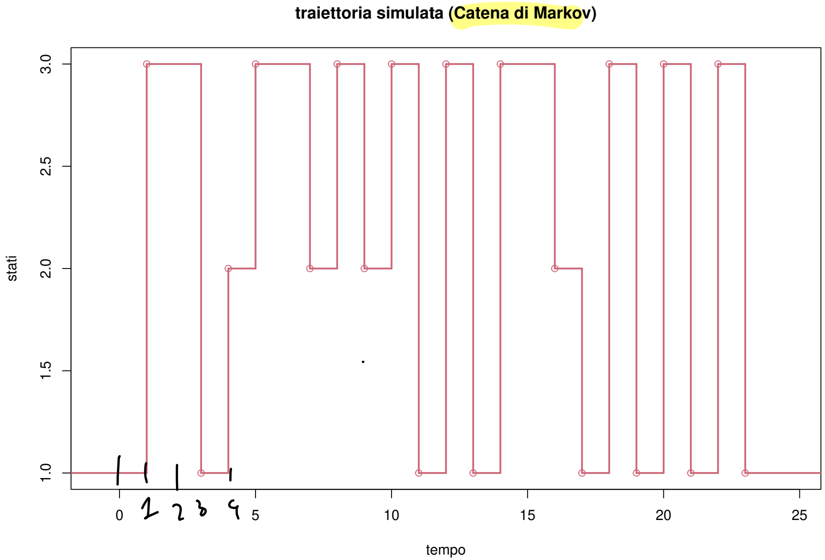
Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \to x_i}$
  - stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \to y}$, $y \neq x_i$
- *Processi a salti*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. esponenziale di parametro $-L_{x_i \to x_i}$

Simulazione di processi di Markov

- *Catene*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. geometrica (modificata) di parametro $1 - Q_{x_i \to x_i}$
  - stato al tempo X_{T_i} è scelta con probabilità $\propto Q_{x_i \to y}$, $y \neq x_i$
- *Processi a salti*: campionare X_0 con probabilità μ_0, poi per $i \geq 0$:
  - tempo di permanenza T_i in X_i è v.a. esponenziale di parametro $-L_{x_i \to x_i}$
  - stato al tempo X_{S_i} è scelta con probabilità $\propto L_{x_i \to y}$, $y \neq x_i$

Esempi


traiettoria simulata (Catena di Markov)
stati
tempo
1.0
1.5
2.0
2.5
3.0
0
5
10
15
20
25
1
2
3
4

Distribuzioni invarianti e stazionarietà

Distribuzione invariante

Negli esempi che abbiamo considerato si vede che una catena di Markov (o un processo di Markov a salti) tende verso un **equilibrio** in cui le densità **marginali sono costanti nel tempo**.

- Lo studio delle possibili densità *limite* è particolarmente rilevante.

Distribuzione invariante

Negli esempi che abbiamo considerato si vede che una catena di Markov (o un processo di Markov a salti) tende verso un **equilibrio** in cui le densità **marginali sono costanti nel tempo**.

- Lo studio delle possibili densità *limite* è particolarmente rilevante.
- Per definire tali densità, basta considerare rispettivamente l'equazione di evoluzione

$$\mu_n = \mu_{n-1} Q \quad \text{oppure} \quad \frac{d}{dt}\mu_t = \mu_t L,$$

e imporre che la densità marginale *non cambi* nel tempo.

$\boxed{\mu = \mu Q}$ $\boxed{0 = \mu L}$

Definizione: caso catene di Markov

Sia Q una matrice di transizione. Si dice che un vettore riga $\mu \in \mathbb{R}^E$ corrispondente ad una densità discreta sull'insieme degli stati,

$$\mu_x \in [0,1] \quad \text{per ogni } x \in E$$

e

$$\sum_{x \in E} \mu_x = 1$$

è una **distribuzione invariante** per Q se vale

$$\mu = \mu Q.$$

Definizione: caso processi di Markov a salti

Sia L una matrice di intensità di salto. Si dice che un vettore riga $\mu \in \mathbb{R}^E$ corrispondente ad una densità discreta sull'insieme degli stati,

$$\mu_x \in [0,1] \quad \text{per ogni } x \in E$$

e

$$\sum_{x \in E} \mu_x = 1$$

è una **distribuzione invariante** per L se vale

$$0 = \mu L.$$

- μ è invariante se e solo se, qualora si consideri un processo X (catena o a salti) tale che la legge marginale al tempo iniziale sia μ, allora tutte le leggi marginali coincidono con μ.

- μ è invariante se e solo se, qualora si consideri un processo X (catena o a salti) tale che la legge marginale al tempo iniziale sia μ, allora tutte le leggi marginali coincidono con μ.

- La condizione di invarianza si può riscrivere anche come (nel caso delle catene) segue: per ogni $x \in E$,

$$\sum_{y \neq x} \mu_x Q_{x \to y} = \sum_{y \neq x} \mu_y Q_{y \to x}. \qquad \forall x$$

Bilancio di flusso

$\sum_{y \neq x} \mu_x Q_{x \to y}$ = flusso uscente

$\sum_{y \neq x} \mu_y Q_{y \to x}$ = entrante

- μ è invariante se e solo se, qualora si consideri un processo X (catena o a salti) tale che la legge marginale al tempo iniziale sia μ, allora tutte le leggi marginali coincidono con μ.
- La condizione di invarianza si può riscrivere anche come (nel caso delle catene) segue: per ogni $x \in E$,

$$\sum_{y \neq x} \mu_x Q_{x \to y} = \sum_{y \neq x} \mu_y Q_{y \to x}.$$

- In questa formulazione il membro a sinistra si interpreta come flusso (di probabilità) uscente dallo stato x, mentre il membro a destra è un flusso entrante. L'equazione esprime quindi un *bilancio di flusso*.

- μ è invariante se e solo se, qualora si consideri un processo X (catena o a salti) tale che la legge marginale al tempo iniziale sia μ, allora tutte le leggi marginali coincidono con μ.
- La condizione di invarianza si può riscrivere anche come (nel caso delle catene) segue: per ogni $x \in E$,

$$\sum_{y \neq x} \mu_x Q_{x \to y} = \sum_{y \neq x} \mu_y Q_{y \to x}.$$

- In questa formulazione il membro a sinistra si interpreta come flusso (di probabilità) uscente dallo stato x, mentre il membro a destra è un flusso entrante. L'equazione esprime quindi un *bilancio di flusso*.
- Nel caso di processi a salti, l'equazione diventa

$$\sum_{y \neq x} \mu_x L_{xy} = \sum_{y \neq x} \mu_y L_{yx}.$$

Stazionarietà e distribuzione invariante

μ invariante $\Rightarrow$ se $\mu = P(X_0) \Rightarrow P(X_n) = \mu$ $\forall n$

È facile vedere che se è stazionario allora le marginali devono essere invarianti. Vale un risultato più preciso.

- Sia X una catena di Markov o un processo di Markov a salti. Allora X è stazionario se e solo se la marginale al tempo iniziale X_0 ha come densità una distribuzione invariante.

Domande fondamentali

- data Q (oppure L) le distribuzioni invarianti esistono sempre?

Domande fondamentali

- data Q (oppure L) le distribuzioni invarianti esistono sempre?
- se sì, quante sono?

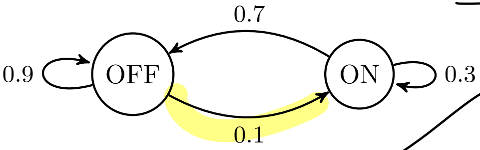
Domande fondamentali

- data Q (oppure L) le distribuzioni invarianti esistono sempre?
- se sì, quante sono?
- Dal lato pratico è invece importante disporre di algoritmi efficienti per poter calcolare, almeno in modo approssimato, le distribuzioni invarianti.

Teorema di esistenza

Se l'insieme degli stati di una catena di Markov (o processo di Markov a salti) E è finito allora esiste sempre almeno una distribuzione invariante μ.

Un esempio



$$\mu = (\mu_{OFF}, \mu_{ON})$$

Bilancio di flusso

$$\begin{cases} \mu_{OFF}\, 0{,}1 = \mu_{ON} \cdot 0{,}7 \\ \cancel{\mu_{ON} \cdot 0{,}7 = \mu_{OFF} \cdot 0{,}1} \end{cases}$$

$$\boxed{\mu_{OFF} + \mu_{ON} = 1}$$

$$\mu_{OFF} = 7\mu_{ON}$$

$$\hookrightarrow \mu = (7\mu_{ON}, \mu_{ON}) / 8\mu_{ON} = \left(\frac{7}{8}, \frac{1}{8}\right)$$

Lo stesso esempio in R

$\mu = \mu Q \xrightarrow{\text{traspongo}} \boxed{v = Q^T v}$ $\lambda = 1$ autovalore

Il comando `eigen()` determina autovalori e autovettori di una matrice: in questo caso ci interessano infatti gli autovettori di Q^T con autovalore 1, o equivalentemente gli autovettori di $Id - Q^t$ con autovalore 0 (il nucleo di $Id - Q^T$).

```
## [1] 1.0 0.2

##             [,1]       [,2]
## [1,] 0.9899495 -0.7071068
## [2,] 0.1414214  0.7071068

## [1] 0.875 0.125
```

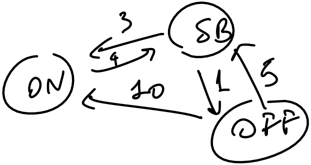
$\boxed{w = Qw}$ $w = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ $Qw = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = w$

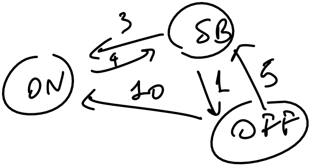
Determiniamo le distribuzioni invarianti nel caso della matrice di intensità di salto dell'esempio della lezione precedente.

$E = \{\text{Off}, \text{Standby}, \text{On}\}$, con matrice delle intensità di salto

$$L = \begin{array}{c c} & \begin{array}{ccc} \text{OFF} & \text{SB} & \text{ON} \end{array} \\ \begin{array}{c} \text{OFF} \\ \\ \\ \end{array} & \begin{pmatrix} * & 5 & 10 \\ 1 & * & 3 \\ 0 & 4 & * \end{pmatrix} \end{array} = \begin{pmatrix} -15 & 5 & 10 \\ 1 & -4 & 3 \\ 0 & 4 & -4 \end{pmatrix}$$

<u>Bilancio di flusso</u>





$$\begin{cases} \mu_{ON} = \mu_{SB} \cdot \frac{3}{4} + \mu_{OFF} \frac{10}{4} \\ \mu_{OFF}\, 15 = \mu_{SB} \end{cases}$$

$$\left(\mu_{OFF},\ 15\mu_{OFF},\ \frac{3}{4} \cdot 15\mu_{OFF} + \frac{10}{4}\mu_{OFF} \right)\alpha$$

$$\alpha \left(1,\ 15,\ \frac{55}{4} \right)$$

$$\frac{55+4+60}{4} = \frac{119}{4}$$

$$\mu = \left(\frac{4}{119},\ \frac{60}{119},\ \frac{55}{119} \right) \approx \left(0,\ \frac{1}{2},\ \frac{1}{2} \right)$$

```
## [1] -1.514005e+01 -7.859945e+00  8.881784e-16

##              [,1]        [,2]        [,3]
## [1,]  0.7539667 -0.09196364 -0.04908437
## [2,] -0.1055968 -0.65662548 -0.73626560
## [3,] -0.6483699  0.74858912 -0.67491013

## [1] 0.03361345 0.50420168 0.46218487
```

Attenzione!

Non confondete l'equazione $\mu(Id - Q) = 0$, oppure $\mu L = 0$ con le equazioni $(Id - Q)v = 0$ o $Lv = 0$.

- Infatti queste ultime hanno sempre come soluzione la densità uniforme $v_x = 1/d$, dove d è il numero degli elementi di E.

Attenzione!

Non confondete l'equazione $\mu(Id - Q) = 0$, oppure $\mu L = 0$ con le equazioni $(Id - Q)v = 0$ o $Lv = 0$.

- Infatti queste ultime hanno sempre come soluzione la densità uniforme $v_x = 1/d$, dove d è il numero degli elementi di E.
- Se non si specifica l'operazione di trasposizione si trova sempre tale soluzione, che però non è quella cercata.

Attenzione!

Non confondete l'equazione $\mu(Id - Q) = 0$, oppure $\mu L = 0$ con le equazioni $(Id - Q)v = 0$ o $Lv = 0$.

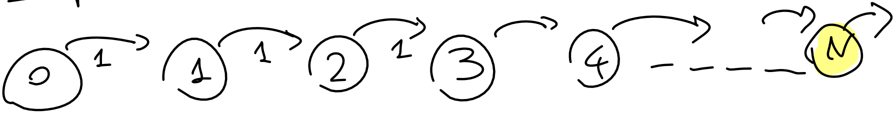
- Infatti queste ultime hanno sempre come soluzione la densità uniforme $v_x = 1/d$, dove d è il numero degli elementi di E.
- Se non si specifica l'operazione di trasposizione si trova sempre tale soluzione, che però non è quella cercata.
- Se la matrice Q è *bistocastica* allora μ uniforme è distribuzione invariante. Un caso particolare è quando Q sia simmetrica.

→ In particolare se $\underline{Q = Q^T}$

Cosa accade con infiniti stati?

Non necessariamente esiste μ invariante

Esempio $E = \mathbb{N}$



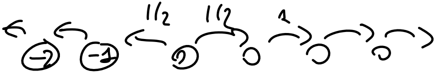
$\mu_0 \cdot 1 = 0 \rightarrow \mu_0 = 0$

$\mu_1 \cdot 1 = \cancel{\mu_0 \cdot 1} = 0 \rightarrow \mu_1 = 0$

$\vdots$

$\boxed{\mu_n = 0}$

$\mathbb{Z}$



$\mu_0 1 = 0 \quad \mu_0 = 0$

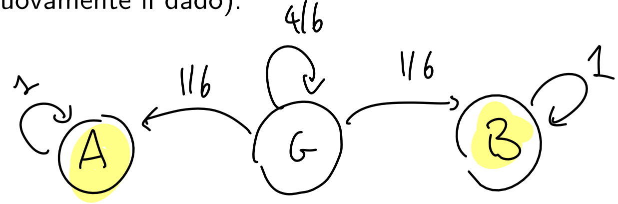
$\mu_1 1 = \mu_0 \cdot \frac{1}{2} = 0 \quad \mu_1 = 0$

Distribuzioni invarianti: classificazione

Un gioco

Alice e Bruno consiste lanciano un dado ripetutamente fintanto che non esca il numero 1 (e in tal caso vince Alice) oppure il numero 6 (e e in tal caso vince Bob).

- Possiamo rappresentare una partita tramite una catena di Markov sugli stati $E = \{A, G, B\}$, dove A indica che Alice ha vinto, B Bruno ha vinto, e G il gioco continua (si deve lanciare nuovamente il dado).



Vi sono almeno due distribuzioni stazionarie, corrispondenti rispettivamente al caso in cui Alice vinca, $\mu_A = (1, 0, 0)$ oppure Bruno vinca, $\mu_B = (0, 0, 1)$.

- Ma in realtà sono infinite, perché ogni combinazione

$$\alpha\mu_A + (1-\alpha)\mu_B = (\alpha, 0, 1-\alpha), \quad \text{con } \alpha \in [0, 1]$$

  è pure una distribuzione invariante (corrispondente al fatto che Alice vinca con probabilità α e Bruno con probabilità $1-\alpha$).

Se la catena di Markov al tempo iniziale si trova in G, la densità limite sarà corrispondente ad $\alpha = 1/2$, perché le regole del gioco non favoriscono né Alice né Bruno.

Classificazione degli stati

Data una matrice di transizione Q, si dice che lo stato $y \in E$ è **accessibile** (o raggiungibile) da $x \in E$ se esiste un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con $x_0 = x$ e $x_n = y$ e peso strettamente positivo:

$$Q_\gamma = \prod_{k=1}^{n} Q_{x_{k-1}x_k} > 0.$$

- Nel caso di una matrice di intensità di salto L, il peso Q_γ va sostituito con

$$L_\gamma = \prod_{k=1}^{n} L_{x_{k-1}x_k} > 0$$

Classificazione degli stati

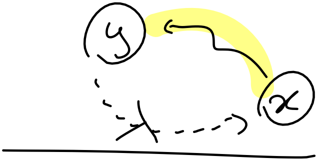
Data una matrice di transizione Q, si dice che lo stato $y \in E$ è **accessibile** (o raggiungibile) da $x \in E$ se esiste un cammino $\gamma = (x_0 \to x_1 \to \ldots \to x_n)$ con $x_0 = x$ e $x_n = y$ e peso strettamente positivo:

$$Q_\gamma = \prod_{k=1}^{n} Q_{x_{k-1}x_k} > 0.$$

- Nel caso di una matrice di intensità di salto L, il peso Q_γ va sostituito con
$$L_\gamma = \prod_{k=1}^{n} L_{x_{k-1}x_k}.$$
- Q^n_{xy} è la somma dei pesi di tutti i cammini lunghi n che collegano x a y: y è raggiungibile da x se esiste $n \geq 1$ tale che $Q^n_{xy} > 0$.

Se $y \in E$ è raggiungibile da $x \in E$, e $z \in E$ è raggiungibile da y, allora z è raggiungibile da x.

- Non è detto che anche x sia raggiungibile da y: se questo appunto non accade, lo stato x è detto **transitorio**: uno stato transitorio è tale che esista y raggiungibile da x ma x non è raggiungibile da y.



Se x è tale che
esiste y raggiungibile da x
ma da y non si raggiunge x
$\Rightarrow$ x è transitorio

Se $y \in E$ è raggiungibile da $x \in E$, e $z \in E$ è raggiungibile da y, allora z è raggiungibile da x.

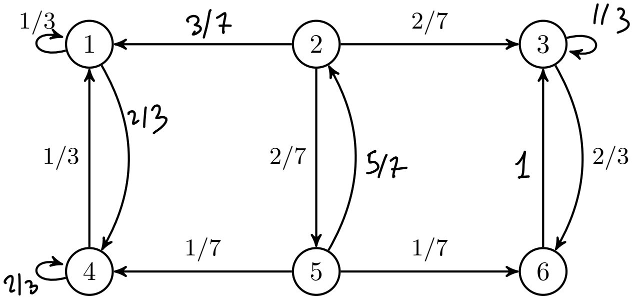
- Non è detto che anche x sia raggiungibile da y: se questo appunto non accade, lo stato x è detto **transitorio**: uno stato transitorio è tale che esista y raggiungibile da x ma x non è raggiungibile da y.
- Se x non è transitorio, è detto **ricorrente**.

Se $y \in E$ è raggiungibile da $x \in E$, e $z \in E$ è raggiungibile da y, allora z è raggiungibile da x.

- Non è detto che anche x sia raggiungibile da y: se questo appunto non accade, lo stato x è detto **transitorio**: uno stato transitorio è tale che esista y raggiungibile da x ma x non è raggiungibile da y.
- Se x non è transitorio, è detto **ricorrente**.
- Se l'insieme degli stati E è finito, non possono essere tutti transitori, deve esserci almeno uno stato ricorrente.

Un esempio

2 transitorio 1 ricorrente



Classi chiuse

Un sottoinsieme $C \subseteq E$ di stati è detto **classe chiusa** se, per ogni $x \in C$ e $y \in E$ raggiungibile da x, anche $y \in C$.

- Una classe chiusa C è detta **irriducibile** se non contiene altre classi chiuse $C' \subseteq C$ (diverse dai casi banali $C' = \emptyset$ oppure $C' = C$ stessa).

Classi chiuse

Un sottoinsieme $C \subseteq E$ di stati è detto **classe chiusa** se, per ogni $x \in C$ e $y \in E$ raggiungibile da x, anche $y \in C$.

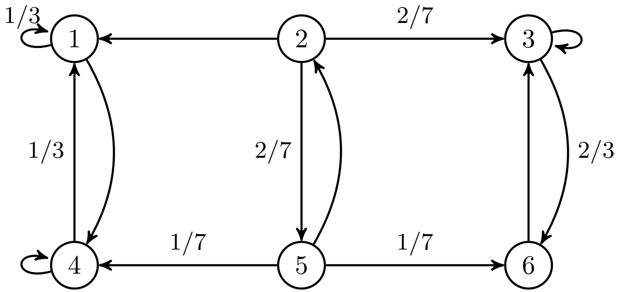
- Una classe chiusa C è detta **irriducibile** se non contiene altre classi chiuse $C' \subseteq C$ (diverse dai casi banali $C' = \emptyset$ oppure $C' = C$ stessa).
- La matrice Q (oppure L) è detta **irriducibile** se tutto l'insieme degli stati E è una classe chiusa irriducibile.

Da uno stato in una classe chiusa non è possibile raggiungere stati al di fuori di essa (mentre è possibile entrarvi)

- una classe chiusa è irriducibile quando da ogni stato in essa si può raggiungere qualsiasi altro stato in essa.

Da uno stato in una classe chiusa non è possibile raggiungere stati al di fuori di essa (mentre è possibile entrarvi)

- una classe chiusa è irriducibile quando da ogni stato in essa si può raggiungere qualsiasi altro stato in essa.
- Data una classe chiusa è ben definita la *restrizione* della matrice Q su $C \times C$, perché $Q_{x \to y} = 0$ per $x \in C$ e $y \notin C$, e quindi $(Q_{x \to y})_{y \in C}$ sono densità discrete di probabilità (la somma delle righe vale ancora 1). Un ragionamento analogo vale nel caso di matrici di intensità di salto L.


1/3
1
2
2/7
3
1/3
2/7
2/3
1/7
1/7
4
5
6

Il teorema di unicità

Sia Q una matrice di transizione (oppure L di intensità di salto) irriducibile su un insieme di stati E finito. Allora esiste una e una sola distribuzione invariante.

Classificazione delle distribuzioni invarianti

Se E è finito, si può sempre partizionare in insiemi a due a due disgiunti

$$E = \{\text{stati transitori}\} \cup C^1 \cup C^2 \cup \ldots \cup C^k,$$

dove le C^i sono classi chiuse irriducibili.

- Consideriamo la restrizione di Q su ciascuna classe chiusa irriducibile C^i: esiste una e una sola distribuzione invariante associata μ^i,

Classificazione delle distribuzioni invarianti

Se E è finito, si può sempre partizionare in insiemi a due a due disgiunti

$$E = \{\text{stati transitori}\} \cup C^1 \cup C^2 \cup \ldots \cup C^k,$$

dove le C^i sono classi chiuse irriducibili.

- Consideriamo la restrizione di Q su ciascuna classe chiusa irriducibile C^i: esiste una e una sola distribuzione invariante associata μ^i,
- Si può dimostrare (non lo faremo) che **tutte** le distribuzioni invarianti sono ottenibili come combinazioni

  $$\mu = \alpha_1 \mu^1 + \alpha_2 \mu^2 + \ldots, + \alpha_k \mu^k,$$

  dove α_1, α_2, ..., $\alpha_k \in [0, 1]$ sono tali che

  $$\alpha_1 + \alpha_2 + \ldots + \alpha_k = 1.$$

Classificazione delle distribuzioni invarianti

Se E è finito, si può sempre partizionare in insiemi a due a due disgiunti

$$E = \{\text{stati transitori}\} \cup C^1 \cup C^2 \cup \ldots \cup C^k,$$

dove le C^i sono classi chiuse irriducibili.

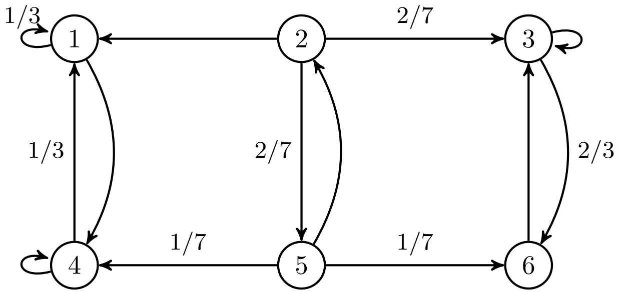
- Consideriamo la restrizione di Q su ciascuna classe chiusa irriducibile C^i: esiste una e una sola distribuzione invariante associata μ^i,
- Si può dimostrare (non lo faremo) che **tutte** le distribuzioni invarianti sono ottenibili come combinazioni

  $$\mu = \alpha_1 \mu^1 + \alpha_2 \mu^2 + \ldots, + \alpha_k \mu^k,$$

  dove α_1, α_2, $\ldots$, $\alpha_k \in [0, 1]$ sono tali che

  $$\alpha_1 + \alpha_2 + \ldots + \alpha_k = 1.$$

- In particolare ogni distribuzione invariante è nulla sugli stati transitori.


1/3
1
2
2/7
3
1/3
2/7
2/3
1/7
1/7
4
5
6