PROCESSI STOCASTICI: ALCUNE INTEGRAZIONI TEORICHE

1. A SIMPLE IDENTITY CONCERNING CONDITIONING

Lemma 1.1. *Let E, F, G be event and let $F \subset G$. Let $\mathbb{Q}$ be the probability measure defined as $\mathbb{Q} = \mathbb{P}(\cdot | G)$. Then*

$$\mathbb{P}(E|F) = \mathbb{Q}(E|F). \quad (1.1)$$

Proof. We have (in the 2nd identity use that $EF \subset G$ and $F \subset G$).

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(EF)}{\mathbb{P}(F)} = \frac{\mathbb{P}(EFG)}{\mathbb{P}(FG)} = \frac{\mathbb{P}(EFG)/\mathbb{P}(G)}{\mathbb{P}(FG)/\mathbb{P}(G)} = \frac{\mathbb{P}(EF|G)}{\mathbb{P}(F|G)} = \mathbb{Q}(E|F).$$

□

2. PROOF OF THM. 1.4.2

Sia B un evento determinato da $X_0, \dots, X_T$. Poniamo in generale

$$A_k := \{X_k = j_0, X_{k+1} = j_1, \dots, X_{k+n} = j_n\}.$$

Allora

$$\mathbb{P}(A_T \cap B \cap \{T = m\} \cap \{X_T = i\}) = \mathbb{P}(A_m \cap B \cap \{T = m\} \cap \{X_m = i\}). \quad (2.1)$$

Osservo che $B \cap \{T = m\}$ è determinato da $X_0, X_1, \dots, X_m$. Per la proprietà di Markov al tempo m (usata nella seconda e terza identità) abbiamo

$$\begin{aligned} & \mathbb{P}(A_m \cap B \cap \{T = m\} \cap \{X_m = i\}) \\ &= \mathbb{P}(A_m \cap B \cap \{T = m\} | X_m = i) \mathbb{P}(X_m = i) \\ &= \mathbb{P}(A_m | X_m = i) \mathbb{P}(B \cap \{T = m\} | X_m = i) \mathbb{P}(X_m = i) \\ &= \mathbb{P}_i(A_0) \mathbb{P}(B \cap \{T = m\} | X_m = i) \mathbb{P}(X_m = i) \\ &= \mathbb{P}_i(A_0) \mathbb{P}(B \cap \{T = m\} \cap \{X_m = i\}) \\ &= \mathbb{P}_i(A_0) \mathbb{P}(B \cap \{T = m\} \cap \{X_T = i\}) \end{aligned} \quad (2.2)$$

By combining (2.1) and (2.2) we get

$$\mathbb{P}(A_T \cap B \cap \{T = m\} \cap \{X_T = i\}) = \mathbb{P}_i(A_0) \mathbb{P}(B \cap \{T = m\} \cap \{X_T = i\}).$$

Sommando com $m = 0$ ad ∞ abbiamo

$$\mathbb{P}(A_T \cap B \cap \{T < \infty\} \cap \{X_T = i\}) = \mathbb{P}_i(A_0) \mathbb{P}(B \cap \{T < \infty\} \cap \{X_T = i\}).$$

Dividiamo per $\mathbb{P}(T < \infty, X_T = i)$, otteniamo

$$\mathbb{P}(A_T \cap B | T < \infty, X_T = i) = \mathbb{P}_i(A_0) \mathbb{P}(B | T < \infty, X_T = i). \quad (2.3)$$

Prendendo B come l'evento certo, abbiamo

$$\mathbb{P}(A_T | T < \infty, X_T = i) = \mathbb{P}_i(A_0), \quad (2.4)$$

e quindi $(X_{T+n})_{n \geq 0}$ è $MC(\delta_i, P)$. Combinando (2.3) e (2.4) otteniamo l'indipendenza tra $(X_{T+n})_{n \geq 0}$ e B rispetto a $\mathbb{P}(\cdot | T < \infty, X_T = i)$.

3. STAZIONARIETÀ

Sia P matrice stocastica.

Definition 3.1. Una misura $\lambda : I \rightarrow [0, \infty)$ si dice invariante (o stazionaria) rispetto a P se $\lambda P = \lambda$ e $\lambda \not\equiv 0$.

Definition 3.2. Una distribuzione $\lambda : I \rightarrow [0, 1]$ si dice invariante (o stazionaria) rispetto a P se $\lambda P = \lambda$.

Definition 3.3. Una catena di Markov $(X_n)_{n \geq 0}$ di parametri (λ, P) si dice invariante (o stazionaria) se per ogni $m \geq 0$ intero il processo stocastico $(X_{n+m})_{n \geq 0}$ è anch'esso una catena di Markov di parametri (λ, P) .

Proposition 3.4. Una catena di Markov $(X_n)_{n \geq 0}$ di parametri (λ, P) è stazionaria se e solo se λ è distribuzione stazionaria rispetto a P .

Proof. Supponiamo che $(X_n)_{n \geq 0}$ sia stazionaria. Allora per ogni $i \in I$ deve valere $P(X_0 = i) = P(X_1 = i)$. Per le note regole di calcolo questo equivale al fatto che $\lambda(i) = (\lambda P)(i)$ per ogni $i \in I$. Quindi λ è distribuzione stazionaria rispetto a P .

Supponiamo ora che λ sia distribuzione stazionaria rispetto a P . Fissiamo $m \geq 0$ intero e dimostriamo che $(Y_n)_{n \geq 0}$ è catena di Markov di parametri (λ, P) , dove $Y_n := X_{m+n}$. Sappiamo che ci basta provare che

$$P(Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n) = \lambda(i_0)p_{i_0, i_1} \cdots p_{i_{n-1}, i_n} \quad (3.1)$$

per ogni $n \geq 0$ e ogni $i_0, i_1, \dots, i_n \in I$. Per definizione di Y_k e poi per le note regole di calcolo otteniamo

$$P(Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n) = P(X_m = i_0, X_{m+1} = i_1, \dots, X_{m+n} = i_n) = (\lambda P^m)_{i_0} p_{i_0, i_1} \cdots p_{i_{n-1}, i_n}.$$

Al fine di ottenere (3.1) ci basta osservare che per la stazionarietà di λ vale $(\lambda P^m)_{i_0} = \lambda(i_0)$. $\square$

4. REVERSIBILITÀ

Sia P matrice stocastica.

Definition 4.1. Una misura $\lambda : I \rightarrow [0, \infty)$ si dice reversibile (o che soddisfa l'equazione del bilancio dettagliato) rispetto a P se $\lambda \not\equiv 0$ e

$$\lambda(i)p_{i,j} = \lambda(j)p_{j,i} \quad \forall i, j \in I. \quad (4.1)$$

Definition 4.2. Una distribuzione $\lambda : I \rightarrow [0, 1]$ si dice reversibile (o che soddisfa l'equazione del bilancio dettagliato) rispetto a P se

$$\lambda(i)p_{i,j} = \lambda(j)p_{j,i} \quad \forall i, j \in I. \quad (4.2)$$

Definition 4.3. Una catena di Markov $(X_n)_{n \geq 0}$ di parametri (λ, P) si dice reversibile se per ogni $N \geq 0$ intero il processo stocastico $(Y_n)_{0 \leq n \leq N}$ dove $Y_n = X_{N-n}$ è una catena di Markov di parametri (λ, P) a tempi in $\{0, 1, \dots, N\}$.

Proposition 4.4. Una distribuzione λ reversibile rispetto a P è anche invariante rispetto a P .

Proof. Sia λ reversible. Per l'equazione del bilancio dettagliato e usando che P è stocastica abbiamo

$$\sum_{j \in I} \lambda_j P_{j,i} = \sum_{j \in I} \lambda_i P_{i,j} = \lambda_i \sum_{j \in I} P_{i,j} = \lambda_i.$$

□

Proposition 4.5. *Una catena di Markov $(X_n)_{n \geq 0}$ di parametri (λ, P) è reversibile se e solo se λ è distribuzione reversibile rispetto a P .*

Proof. Supponiamo che $(X_n)_{n \geq 0}$ sia reversibile. Prendendo $N = 1$ otteniamo per ogni $i, j \in I$ che

$$P(Y_0 = i, Y_1 = j) = P(X_{1-0} = i, X_{1-1} = j) = P(X_0 = j, X_1 = i).$$

Per le note regole di calcolo possiamo riscrivere il primo e l'ultimo membro come

$$\lambda(i)p_{i,j} = \lambda(j)p_{j,i}.$$

Abbiamo quindi verificato che λ soddisfa l'equazione del bilancio dettagliato.

Supponiamo ora che valga l'equazione del bilancio dettagliato. Per il Teo. 1.1.1 del Norris, dobbiamo dimostrare che per ogni $N \geq 0$, $n \geq 0$, $i_0, i_1, \dots, i_n \in I$ vale

$$P(Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n) = \lambda(i_0)p_{i_0, i_1} \cdots p_{i_{n-1}, i_n} \quad (4.3)$$

dove $Y_k := X_{N-k}$. Per tale definizione di Y_k e per le note regole di calcolo vale

$$P(Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n) = P(X_0 = i_n, X_{n-1} = i_{n-1}, \dots, X_{n-1} = i_1, X_n = i_0) = \lambda(i_n)p_{i_n, i_{n-1}}p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0} \quad (4.4)$$

Per l'equazione del bilancio dettagliato abbiamo

$$\lambda(i_k)p_{i_k, i_{k-1}} = p_{i_{k-1}, i_k} \lambda(i_{k-1}). \quad (4.5)$$

Applicando (4.5) con $k = n$ abbiamo

$$\lambda(i_n)p_{i_n, i_{n-1}}p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0} = p_{i_{n-1}, i_n} \lambda(i_{n-1})p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0}.$$

Applicando iterativamente (4.5) con $k = n - 1$, poi $k = n - 2, \dots$, fino a $k = 1$ otteniamo

$$\lambda(i_n)p_{i_n, i_{n-1}}p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0} = p_{i_{n-1}, i_n} p_{i_{n-2}, i_{n-1}} \cdots p_{i_0, i_1} \lambda(i_0). \quad (4.6)$$

A tal punto la tesi (4.3) segue da (4.4) e (4.6). □

5. TEOREMI DI STRUTTURA PER LE DISTRIBUZIONI INVARIANTI

Da scrivere...

Da includere il seguente risultato:

Theorem 5.1. *Consideriamo una catena di Markov $(X_n)_{n \geq 0}$ con spazio degli stati finito I e matrice di transizione P . Siano $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_N$ le classi comunicanti chiuse (sappiamo che $1 \leq N < \infty$). Allora le distribuzioni invarianti di P sono tutte e sole della forma $\lambda = \sum_{k=1}^N \alpha_k \lambda_k$ dove $\alpha_k \geq 0$, $\sum_{k=1}^N \alpha_k = 1$ e λ_k è l'unica distribuzione invariante con supporto in $\mathcal{C}_k$. λ_k è nulla fuori di $\mathcal{C}_k$ e ristretta a $\mathcal{C}_k$ è l'unica distribuzione invariante della matrice ridotta $(P_{ij})_{i,j \in \mathcal{C}_k}$, che è matrice stocastica irriducibile. In particolare, esiste sempre almeno una distribuzione invariante. Essa è unica se e solo se vi è un'unica classe comunicante chiusa.*

6. PERIODO

Sia P matrice stocastica. Di seguito useremo molto la seguente proprietà:

$$P_{i,j}^{m+r} \geq P_{i,z}^m P_{z,j}^r \quad \forall m, r \geq 0, \forall i, j, z \in I. \quad (6.1)$$

Dimostrazione di (6.1): $P_{i,j}^{m+r} = \sum_{k \in I} P_{i,k}^m P_{k,j}^r \geq P_{i,z}^m P_{z,j}^r$ siccome gli addendi sono non-negativi.

Definition 6.1. Dato uno stato $i \in I$ definiamo $\mathcal{T}(i) := \{n \geq 0 : p_{i,i}^{(n)} > 0\}$ e definiamo il periodo di i come $\text{per}(i) := \text{massimo comune denominatore (MCD) di } \mathcal{T}(i)$. Se $\text{per}(i) = 1$ diciamo che lo stato i è aperiodico, altrimenti diciamo che è periodico.

Proposition 6.2. Se P è irriducibile, allora $\text{per}(i) = \text{per}(j) \forall i, j \in I$.

Proof. Siccome P è irriducibile, abbiamo $i \leftrightarrow j$ e quindi esistono $r, \ell \geq 0$ con $p_{i,j}^{(r)} > 0$ e $p_{j,i}^{(\ell)} > 0$. Poniamo $m := r + \ell$.

Claim 1. $m \in \mathcal{T}(i) \cap \mathcal{T}(j)$.

Proof. Per (6.1) abbiamo $p_{i,i}^{(m)} \geq p_{i,j}^{(r)} p_{j,i}^{(\ell)} > 0$ e perciò $m \in \mathcal{T}(i)$. Analogamente $p_{j,j}^{(m)} \geq p_{j,i}^{(\ell)} p_{i,j}^{(r)} > 0$ e perciò $m \in \mathcal{T}(j)$.

Claim 2. $\mathcal{T}(i) \subset \mathcal{T}(j) - m$.

Proof. Sia $a \in \mathcal{T}(i)$, allora $p_{i,i}^{(a)} > 0$. Allora, usando anche (6.1),

$$p_{j,j}^{(a+m)} \geq p_{j,i}^{(\ell)} p_{i,i}^{(a)} p_{i,j}^{(r)} > 0.$$

Hence, $a + m \in \mathcal{T}(j)$, i.e. $a \in \mathcal{T}(j) - m$. Questo prova che $\mathcal{T}(i) \subset \mathcal{T}(j) - m$.

Claim 3. Ogni divisore comune di $\mathcal{T}(j)$ è divisore comune di $\mathcal{T}(i)$.

Proof. Sia d un divisore comune di $\mathcal{T}(j)$. Allora $d|m$ siccome $m \in \mathcal{T}(j)$ per il Claim 1. Quindi d divide ogni elemento di $\mathcal{T}(j) - m$ e quindi d divide ogni elemento di $\mathcal{T}(i)$ dato che $\mathcal{T}(i) \subset \mathcal{T}(j) - m$.

Conclusione. Per il Claim 3 il MCD di $\mathcal{T}(j)$ è divisore comune di $\mathcal{T}(i)$. Questo implica che

$$\text{MCD di } \mathcal{T}(j) \leq \text{MCD di } \mathcal{T}(i),$$

che equivale a $\text{per}(j) \leq \text{per}(i)$. Per simmetria $\text{per}(i) \leq \text{per}(j)$ e quindi $\text{per}(i) = \text{per}(j)$. $\square$

Per la suddetta proposizione, quando P è irriducibile, chiamiamo periodo di P il periodo dei suoi elementi (che è lo stesso per tutti gli elementi). Quando P è irriducibile diciamo che P è aperiodica se P ha periodo 1, altrimenti diciamo che P è periodica.

Ricordiamo il seguente risultato di teoria dei numeri (senza dimostrarlo):

Lemma 6.3. Ogni sottinsieme di $\mathbb{N} = \{0, 1, 2, \dots\}$ chiuso per addizione ed avente MCD pari a 1 contiene tutti i numeri di $\mathbb{N}$ tranne al più un sottinsieme finito.

Osserviamo che per ogni stato i abbiamo che $\mathcal{T}(i)$ è chiuso per addizione. Infatti, se $a, b \in \mathcal{T}(i)$ allora $p_{i,i}^{(a)} > 0$ e $p_{i,i}^{(b)} > 0$ e per (6.1) concludiamo che $p_{i,i}^{(a+b)} \geq p_{i,i}^{(a)} p_{i,i}^{(b)} > 0$, che equivale a $a + b \in \mathcal{T}(i)$.

Spesso viene data una definizione diversa di stato aperiodico, come in [N], che però per il seguente lemma è equivalente a quella data in questa sezione:

Lemma 6.4. *Uno stato i è aperiodico se e solo se esiste $m \geq 0$ tale che $p_{i,i}^{(k)} > 0$ per ogni $k \geq m$.*

Proof. Supponiamo che i sia aperiodico. Allora, per quanto osservato prima, $\mathcal{T}(i) \subset \mathbb{N}$ è chiuso per addizione e, per la definizione di stato aperiodico, MCD di $\mathcal{T}(i)$ è pari a 1. Per il Lemma 6.3 concludiamo che esiste $m \geq 0$ tale che $k \in \mathcal{T}(i)$ per ogni $k \geq m$, che equivale al fatto che $p_{i,i}^{(k)} > 0$ per ogni $k \geq m$.

Supponiamo ora che esista $m \geq 0$ tale che $p_{i,i}^{(k)} > 0$ per ogni $k \geq m$. Allora $\mathcal{T}(i) \supset \{m, m+1, m+2, \dots\}$. Siccome il MCD di m e $m+1$ è pari a 1, abbiamo che il MCD di $\mathcal{T}(i)$ è pari a 1, e questo equivale al fatto che i è aperiodico. $\square$

L'interesse per le matrici stocastiche aperiodico deriva dal problema della convergenza all'equilibrio (vedasi Thm. 1.8.3 in [N]).

7. SIMULAZIONE DI CATENE DI MARKOV

Fissiamo P matrice stocastica $I \times I$ e λ distribuzione su I . Sia $U_1, U_2, U_3, \dots$ una successione di variabili aleatorie indipendenti e uniformemente distribuite su $[0, 1]$. Facciamo una partizione di $[0, 1]$ in sottintervalli disgiunti $[0, 1] = \cup_{i \in I} A_i$ dove A_i ha lunghezza λ_i (se $\lambda_i = 0$ si può prendere come A_i l'intervallo degenere $\emptyset$). Per ogni i facciamo una partizione di $[0, 1]$ in sottintervalli disgiunti $[0, 1] = \cup_{j \in I} A_{i,j}$ dove $A_{i,j}$ ha lunghezza $P_{i,j}$ (se $P_{i,j} = 0$ si può prendere come $A_{i,j}$ l'intervallo degenere $\emptyset$). Introduciamo le funzioni

$$\begin{cases} G_0 : [0, 1] \rightarrow I, & G_0(u) = i \text{ se } u \in A_i, \\ G : I \times [0, 1] \rightarrow I & G(i, u) = j \text{ se } u \in A_{i,j}. \end{cases}$$

Proposition 7.1. *Definendo $X_0 = G_0(U_0)$ ed iterativamente $X_n = G(X_{n-1}, U_n)$ per $n \geq 1$ abbiamo che $(X_n)_{n \geq 0}$ è $CM(\lambda, P)$.*

Nota: siccome i computer sanno simulare variabili aleatorie indipendenti uniformemente distribuite su $[0, 1]$ questo è il modo con cui si simula al computer la catena $CM(\lambda, P)$ se $|I| < \infty$.

Proof. Dato $n \geq 0$ e $i_0, i_1, \dots, i_n \in I$ abbiamo

$$\begin{aligned} \mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) &= \mathbb{P}(G_0(U_0) = i_0, G(i_0, U_1) = i_1, \dots, G(i_{n-1}, U_n) = i_n) \\ &= \mathbb{P}(U_0 \in A_{i_0}, U_1 \in A_{i_0, i_1}, \dots, U_n \in A_{i_{n-1}, i_n}) \end{aligned}$$

Per indipendenza abbiamo

$$\begin{aligned} &\mathbb{P}(U_0 \in A_{i_0}, U_1 \in A_{i_0, i_1}, \dots, U_n \in A_{i_{n-1}, i_n}) \\ &= \mathbb{P}(U_0 \in A_{i_0}) \prod_{k=1}^n \mathbb{P}(U_k \in A_{i_{k-1}, i_k}) \\ &= m(A_{i_0}) \prod_{k=1}^n m(A_{i_{k-1}, i_k}) = \lambda_{i_0} \prod_{k=1}^n P_{i_{k-1}, i_k}. \end{aligned}$$

Siccome

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \lambda_{i_0} \prod_{k=1}^n P_{i_{k-1}, i_k}$$

per Teorema 1.1.1 in [N] (con $N = \infty$) otteniamo che $(X_n)_{n \geq 0}$ è $CM(\lambda, P)$. $\square$

8. MISURA DI GIBBS CANONICA

Consideriamo un sistema fisico a contatto con un bagno termico avente temperatura assoluta T . Sia I lo spazio delle configurazioni microscopiche del sistema. Ci limitiamo a sistemi discretizzati per cui I è numerabile. Sia $H : I \rightarrow \mathbb{R}$ l'Hamiltoniana del sistema ($H(\sigma)$ è l'energia associata alla configurazione σ). La fisica statistica ci dice che il sistema fisico, raggiunto l'equilibrio termodinamico, ha temperatura assoluta T e la misura di probabilità sullo spazio delle configurazioni I che descrive l'equilibrio termodinamico è la cosiddetta misura canonica di Gibbs μ_β :

$$\mu_\beta(\sigma) := \frac{e^{-\beta H(\sigma)}}{\sum_{\sigma' \in I} e^{-\beta H(\sigma')}}, \quad \forall \sigma \in I$$

dove $\beta := \frac{1}{kT}$ e k è la costante di Boltzmann.

9. RICHIAMI SULLE V.A. ESPONENZIALI

Definition 9.1. Una variabile aleatoria $T : \Omega \rightarrow [0, +\infty)$ è detta esponenziale di parametro $\lambda > 0$ (scriviamo $T \sim \text{Exp}(\lambda)$) se

$$\mathbb{P}(T > t) = e^{-\lambda t} \quad \text{per ogni } t \geq 0. \quad (9.1)$$

Equivalentemente, T è una v.a. continua con funzione di densità $f_\lambda(x) = \mathbf{1}(x \geq 0)\lambda e^{-\lambda x}$. Come caso degenere una variabile aleatoria T con $\mathbb{P}(T = +\infty) > 0$ verrà considerata variabile aleatoria esponenziale di parametro $\lambda = 0$. Si noti che così (9.1) è valida anche nel caso $\lambda = 0$.

Dalla definizione è facile provare che $\mathbb{E}[T] = 1/\lambda$ se $T \sim \text{Exp}(\lambda)$ e $\lambda \geq 0$.

Proposition 9.2. (Proprietà di perdita di memoria) Sia $T \sim \text{Exp}(\lambda)$ con $\lambda \geq 0$. Allora vale

$$\mathbb{P}(T > s + t | T > s) = \mathbb{P}(T > t) \quad \forall s, t \geq 0. \quad (9.2)$$

Proof. Sappiamo che (9.1) vale per ogni $\lambda \geq 0$. Abbiamo allora

$$\mathbb{P}(T > s + t | T > s) = \frac{\mathbb{P}(T > s + t)}{\mathbb{P}(T > s)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(T > t).$$

□

► Fare [N, Theorem 2.3.2] con proof. Attenzione, modificare la seconda riga dell'enunciato richiedendo solo $\lambda_n \geq 0$. Il teorema resta vero per $\lambda_n \geq 0$, con la convenzione che $1/\lambda = +\infty$ per $\lambda = 0$. Infatti, se per qualche k abbiamo $\lambda_k = 0$ allora necessariamente abbiamo $1/\lambda_k = +\infty$ e $S_k = +\infty$ q.c. e quindi $\sum_{n=1}^{\infty} \lambda_n^{-1} = +\infty$ e $\mathbb{P}(\sum_{n=1}^{\infty} S_n = \infty) = 1$. Questo prova il teorema nel caso qualche λ_n sia zero. Nel caso in cui tutti i λ_n siano positivi, vale la dimostrazione del libro.

Lemma 9.3. Se $T \sim \text{Exp}(\lambda)$ con $\lambda > 0$ e $\gamma \geq 0$, allora $T/\gamma \sim \text{Exp}(\lambda\gamma)$ (con la convenzione che $T/0 := +\infty$).

Proof. Se $\gamma = 0$ abbiamo $T/\gamma := +\infty$ e quindi $T/\gamma \sim \text{Exp}(0) = \text{Exp}(\lambda\gamma)$.

Se $\gamma > 0$, allora abbiamo per (9.1)

$$\mathbb{P}(T/\gamma > x) = \mathbb{P}(T > \gamma x) = e^{-\lambda\gamma x}$$

e quindi $T/\gamma \sim \text{Exp}(\lambda\gamma)$ (sempre per (9.1) ma ora con $\lambda\gamma$ al posto di λ). □

10. Q -MATRICE E MATRICE DI SALTO

Definition 10.1. Una matrice $Q = (q_{i,j})_{i,j \in I}$ è una Q -matrice se soddisfa le seguenti condizioni:

- (i) $0 \leq -q_{i,i} < +\infty \forall i \in I$;
- (ii) $q_{i,j} \geq 0$ per ogni $i, j \in I$ con $i \neq j$;
- (iii) $-q_{i,i} = \sum_{j \in I: j \neq i} q_{i,j}$

Si noti che quando I è infinito la somma $\sum_{j \in I: j \neq i} q_{i,j}$ è una serie a termini non negativi (per (ii)) e deve essere sommabile dato che (cf. (iii)) il risultato deve dare $-q_{i,i}$ che per (i) è finito.

Una Q -matrice ammette una rappresentazione grafica tramite grafo orientato pesato con vertici dati dagli elementi di I e lati orientati (freccie) pesati costruiti come segue: per ogni coppia (i, j) con $i \neq j$ e $q_{i,j} > 0$ viene disegnata una freccia da i a j con peso $q_{i,j}$ (vedi Ex. 2.6.1 in [N])

Nel seguito scriveremo

$$q_i = q(i) := -q_{i,i}, \quad i \in I.$$

Ad una Q matrice Q associamo la cosiddetta matrice di salto $\Pi = (\pi_{i,j})_{i,j \in I}$, data dalla matrice stocastica su I definita come segue:

$$\pi_{i,j} = \begin{cases} q_{i,j}/q_i & \text{se } j \neq i, q_i \neq 0, \\ 0 & \text{se } j = i, q_i \neq 0, \\ 0 & \text{se } j \neq i, q_i = 0, \\ 1 & \text{se } j = i, q_i = 0. \end{cases}$$

Esempi di Q -matrice e matrice di salto associata:

$$Q = \begin{pmatrix} -3 & 1 & 2 \\ 4/5 & -4/5 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \Pi = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -3 \end{pmatrix}, \quad \Pi = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1 & 0 & 0 \\ 2/3 & 1/3 & 0 \end{pmatrix}.$$

11. CATENE DI MARKOV A TEMPO CONTINUO

Sia V insieme numerabile, finito o infinito, con la topologia discreta. Una funzione $f : [0, +\infty) \rightarrow V$ è continua a destra se $\lim_{s \downarrow t} f(s) = f(t)$ for ogni $t \geq 0$. Siccome V ha la topologia discreta abbiamo che f è continua a destra se e solo se per ogni $t \geq 0$ esiste $\varepsilon > 0$ tale che $f(s) = f(t)$ per ogni $s \in [t, t + \varepsilon]$.

Data una funzione f continua a destra definiamo i tempi $(J_n)_{n \geq 0}$ e $(S_n)_{n \geq 0}$ come segue: posto $J_0 = 0$, definiamo iterativamente

$$J_{n+1} = \inf\{t \geq J_n : f_t \neq f_{J_n}\} \text{ per } n = 0, 1, 2, \dots,$$

e definiamo

$$S_n := \begin{cases} J_n - J_{n-1} & \text{se } J_{n-1} < \infty, \\ \infty & \text{altrimenti.} \end{cases}$$

Chiamiamo $(J_n)_{n \geq 0}$ i tempi di salto e $(S_n)_{n \geq 0}$ i tempi di permanenza, anche se - strettamente parlando - i salti avvengono solo ai tempi J_n tali che $n \geq 1$ e $J_n < +\infty$. Si noti che $S_n \in (0, +\infty]$.

Definiamo il tempo di esplosione

$$\zeta := \sup\{J_n : n \geq 0\} = \sum_{n=1}^{\infty} S_n.$$

e il numero totale di salti

$$N := \sup\{n : J_n < \infty\}$$

Abbiamo tre casi

(C1) $\zeta = \infty$, $N = \infty$. In questo caso $J_n < \infty$ per ogni n e

$$[0, \infty) = [0, \zeta) = \sqcup_{n=0}^{\infty} [J_n, J_{n+1}), \quad f = \text{costante su } [J_n, J_{n+1}).$$

(C2) $\zeta = \infty$, $N < \infty$. In questo caso f fa N salti e

$$[0, \infty) = [J_0, J_1) \sqcup [J_1, J_2) \sqcup \cdots \sqcup [J_{N-1}, J_N) \cup [J_N, \infty)$$

e su ciascun intervallo o semiretta nel r.h.s. f è costante.

(C3) $\zeta < \infty$, $N = \infty$. In questo caso $J_n < \infty$ per ogni n e

$$[0, \zeta) = \sqcup_{n=0}^{\infty} [J_n, J_{n+1}), \quad f = \text{costante su } [J_n, J_{n+1}).$$

I tre suddetti casi sono illustrati dalle tre figure a pagina 68,69 in [N].

Allo spazio degli stati I aggiungiamo uno stato astratto ∂ (non appartenente ad I), detto stato cimitero.

Definition 11.1. Un processo $(X_t)_{t \geq 0}$ definito sullo spazio di probabilità $(\Omega, \mathcal{F}, P)$ e con spazio degli stati l'insieme numerabile I è detto essere continuo a destra e minimale se

- $X_t(\omega) \in I \cup \{\partial\}$ per ogni $t \geq 0$, $\omega \in \Omega$;
- per ogni $\omega \in \Omega$ la funzione

$$[0, +\infty) \ni t \mapsto X_t(\omega) \in I \cup \{\partial\}$$

è continua a destra;

- denotato con $\zeta(\omega)$ il tempo di esplosione della funzione $(X_t(\omega))_{t \geq 0}$, abbiamo che

$$X_t(\omega) \in I \text{ per ogni } t \in [0, \zeta(\omega)), \quad X_t(\omega) = \partial \text{ per ogni } t \geq \zeta(\omega).$$

Nel seguito, dato un processo continuo a destra e minimale, indicheremo con J_n, S_n, ζ, N le quantità $J_n(\omega), S_n(\omega), \zeta(\omega), N(\omega)$ riferite alla traiettoria continua a destra $(X_t(\omega))_{t \geq 0}$

Notiamo che $X_t \in I$ per ogni t se non c'è esplosione nella traiettoria in oggetto.

Definition 11.2. Dato un processo $(X_t)_{t \geq 0}$ con spazio degli stati I , continuo a destra e minimale, definiamo il processo di salto $(Y_n)_{n \geq 0}$ come segue:

$$Y_n := \begin{cases} X_{J_n} & \text{se } n < N, \\ X_{J_N} & \text{se } n \geq N \text{ e } N < \infty. \end{cases}$$

Definition 11.3. Dati un insieme numerabile I , una Q -matrice Q e una distribuzione λ su I , il processo a tempi continui $(X_t)_{t \geq 0}$ è detto catena di Markov con distribuzione iniziale λ e generatore Q se

- (i) $(X_t)_{t \geq 0}$ è un processo continuo a destra e minimale,
- (ii) la catena di salto $(Y_n)_{n \geq 0}$ è $CM(\lambda, \Pi)$ dove Π è la matrice di salto associata a Q ,
- (iii) per ogni $n \geq 1$, condizionatamente a $Y_0, Y_1, \dots, Y_{n-1}$, i tempi di permanenza $S_1, S_2, \dots, S_n$ sono variabili aleatorie indipendenti esponenziali di parametri $q(Y_0), q(Y_1), \dots, q(Y_{n-1})$ rispettivamente, dove $q(i) := -q_{i,i}$.

Il punto (iv) significa che, fissato $n \geq 1$ e fissati $i_0, i_1, \dots, i_{n-1}$, rispetto alla probabilità condizionata

$$\mathbb{P}(\cdot | Y_0 = i_0, Y_1 = i_1, \dots, Y_{n-1} = i_{n-1}),$$

i tempi di permanenza $S_1, S_2, \dots, S_n$ sono variabili aleatorie indipendenti esponenziali di parametri $q(i_0), q(i_1), \dots, q(i_{n-1})$ rispettivamente. Per variabile esponenziale di parametro 0 intendiamo una variabile aleatoria T con $\mathbb{P}(T = +\infty) = 1$.

La dinamica stocastica della catena di Markov $(X_t)_{t \geq 0}$ fino al tempo di esplosione puo' essere descritta come segue. Si sceglie lo stato iniziale i_0 con probabilità λ_{i_0} . In i_0 la catena di Markov aspetta un tempo esponenziale di parametro $q(i_0)$ e, se tale tempo di attesa è finito, salta in uno stato i_1 scelto con probabilità $i_1 \mapsto \pi_{i_0, i_1}$. Poi la catena resta in i_1 un tempo esponenziale di parametro $q(i_1)$ (indipendente dai precedenti oggetti aleatori) e, se tale tempo di attesa è finito, salta in uno stato i_2 scelto con probabilità $i_2 \mapsto \pi_{i_1, i_2}$, e così via.

In modo più preciso possiamo costruire la catena di Markov $(X_t)_{t \geq 0}$ come segue. Sia $(Y_n)_{n \geq 0}$ una catena di Markov a tempo discreto di parametri (λ, Π) e siano $T_1, T_2, \dots, T_n$ della v.a. i.i.d. esponenziali di parametro 1, indipendenti da $(Y_n)_{n \geq 0}$. Poniamo $S_n := T_n/q(Y_{n-1})$. Per Lemma 9.3 S_n è v.a. esponenziale di parametro $q(Y_{n-1})$. Definiamo

$$J_n := S_1 + S_2 + \dots + S_n$$

e

$$X_t := \begin{cases} Y_n & \text{se } J_n \leq t < J_{n+1} \text{ per qualche } n, \\ \partial & \text{altrimenti.} \end{cases} \quad (11.1)$$

Proposition 11.4. *Il processo $(X_t)_{t \geq 0}$ definito in (11.1) è $CM(\lambda, Q)$.*

Proof. Per la definizione (11.1) le v.a. J_n e le v.a. S_n sono proprio i tempi di salto e i tempi di permanenza associati a $(X_t)_{t \geq 0}$. Le proprietà (i) and (ii) della Definizione 11.3 sono banalmente verificate. Proviamo la proprietà (iii). Fissiamo $n \geq 1$ e $i_0, i_1, \dots, i_{n-1} \in I$. Rispetto alla probabilità condizionata $\mathbb{P}(\cdot | Y_0 = i_0, Y_1 = i_1, \dots, Y_{n-1} = i_{n-1})$ i tempi di permanenza $S_1, S_2, \dots, S_n$ sono dati rispettivamente da $\frac{T_1}{q(i_0)}, \frac{T_2}{q(i_1)}, \dots, \frac{T_n}{q(i_{n-1})}$. Siccome le v.a. $T_1, T_2, \dots$ sono indipendenti, concludo che $S_1, S_2, \dots, S_n$ sono indipendenti (rispetto alla suddetta probabilità condizionata). Per il Lemma 9.3 ho che $\frac{T_1}{q(i_0)}, \frac{T_2}{q(i_1)}, \dots, \frac{T_n}{q(i_{n-1})}$ sono v.a. esponenziali di parametri $q(i_0), q(i_1), \dots, q(i_{n-1})$, rispettivamente. Quindi lo stesso vale per $S_1, S_2, \dots, S_n$ (rispetto alla suddetta probabilità condizionata). Ciò conclude la verifica della proprietà (iii). □

Definition 11.5. *Diciamo che la catena di Markov $(X_t)_{t \geq 0}$ non esplosione se il tempo di esplosione è infinito q.c.*

► Fare [N, Theorem 2.7.1, p.90] con proof che i casi (i) e (ii) implicano che la catena non esplosione (tralasciare (iii)).

11.1. Esempio. Consideriamo per esempio la catena di Markov $(X_t)_{t \geq 0}$ con spazio degli stati $I = \{1, 2, 3\}$ e Q matrice

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

Arrivati nello stato 1, indipendentemente dal passato, si aspetta lì un tempo esponenziale di parametro 3 e poi, indipendentemente da quanto fatto fino a quel momento, si salta in 2 o 3 con probabilità $1/3$, $2/3$ rispettivamente.

Arrivati nello stato 2, indipendentemente dal passato, si aspetta lì un tempo esponenziale di parametro $4/5$ e poi, indipendentemente da quanto fatto fino a quel momento, si salta in 1 o 3 con probabilità $1/4$, $3/4$ rispettivamente.

Arrivati nello stato 3, si resta lì per sempre.

Esempio: sia $\lambda = \delta_1$. Abbiamo

$$\begin{aligned}\mathbb{P}(J_1 > t) &= e^{-3t}, \\ \mathbb{P}(J_1 < t, J_2 > t, Y_1 = 2) &= \int_0^t ds 3e^{-3s} \frac{1}{3} e^{-\frac{4}{5}(t-s)} = e^{-\frac{4}{5}t} \int_0^t ds e^{-3s + \frac{4}{5}s} \\ &= e^{-\frac{4}{5}t} \int_0^t ds e^{-\frac{11}{5}s} = e^{-\frac{4}{5}t} \frac{5}{11} \int_0^t ds \frac{11}{5} e^{-\frac{11}{5}s} = e^{-\frac{4}{5}t} \frac{5}{11} (1 - e^{-\frac{11}{5}t}) \\ &= \frac{5}{11} e^{-\frac{4}{5}t} - \frac{5}{11} e^{-3t}, \\ \mathbb{P}(Y_1 = 2, Y_2 = 3) &= \frac{1}{3} \frac{3}{4} = \frac{1}{4}.\end{aligned}$$

Esercizio: calcolare $\mathbb{P}(J_2 > t)$

11.2. Esempio: Processo di Poisson.

Definition 11.6. Il processo di Poisson di intensità $\lambda > 0$ è la catena di Markov a tempo continuo con spazio degli stati $\mathbb{N} := \{0, 1, \dots\}$ e generatore

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \dots \\ 0 & -\lambda & \lambda & \dots \\ 0 & 0 & -\lambda & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Più precisamente, $Q = (q_{ij})_{i,j \geq 0}$ e

$$q_{i,j} = \begin{cases} -\lambda & \text{se } i = j \\ \lambda & \text{se } j = i + 1 \\ 0 & \text{altrimenti.} \end{cases}$$

Oss: Per [N, Theorem 2.7.1, p.90] il processo di Poisson non esplose.

Di solito il processo di Poisson ha distribuzione iniziale δ_0 . In questo caso la dinamica è così ottenuta: si parte in 0, si aspetta un tempo esponenziale di parametro λ e si salta in 1, si aspetta un tempo esponenziale di parametro λ (indipendente dal precedente) e si salta in 2, si aspetta un tempo esponenziale di parametro λ (indipendente dal precedente) e si salta in 3, ... Un modo più formale è il seguente: siano $S_1, S_2, \dots$ variabili i.i.d. con $S_k \sim \text{Exp}(\lambda)$, allora (posto $S_0 := 0$)

$$X_t = n \quad \text{se } S_0 + \dots + S_n \leq t < S_0 + \dots + S_n + S_{n+1}.$$

L'insieme aleatorio dei punti della semiretta $[0, +\infty)$ dato da $\{0 = J_0, J_1, J_2, \dots\}$ è detto Processo Puntuale di Poisson di intensità λ contenente l'origine. L'insieme aleatorio dei punti della semiretta $[0, +\infty)$ dato da $\{J_1, J_2, \dots\}$ è detto Processo Puntuale di Poisson di intensità λ (a volte si dice anche "stationario", oppure "omogeneo").

11.3. Esempio: processo di nascita. È una generalizzazione del processo di Poisson. Le traiettorie aumentano sempre di un'unità quando c'è un salto ma il tasso di nascita non è uniforme:

Definition 11.7. Un processo di nascita con tassi di nascita $q_j \geq 0$ per $j \geq 0$ è una catena di Markov a tempo continuo su $\mathbb{N}$ avente generatore

$$Q = \begin{pmatrix} -q_0 & q_0 & 0 & \cdots \\ 0 & -q_1 & q_1 & \cdots \\ 0 & 0 & -q_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

Più precisamente, $Q = (q_{ij})_{i,j \geq 0}$ e

$$q_{i,j} = \begin{cases} -q_i & \text{se } i = j \\ q_i & \text{se } j = i + 1 \\ 0 & \text{altrimenti.} \end{cases}$$

► Fare [N, Theorem 2.5.2, p.83] con proof.

11.4. Esempio: processo di nascita e morte.

Definition 11.8. Un processo di nascita e morte con tassi di nascita $b_j \geq 0$ per $j \geq 0$ e tassi di morte $d_j \geq 0$ con $j \geq 1$ è una catena di Markov a tempo continuo su $\mathbb{N}$ avente generatore

$$Q = \begin{pmatrix} -b_0 & b_0 & 0 & 0 & \cdots \\ d_1 & -(b_1 + d_1) & b_1 & 0 & \cdots \\ 0 & d_2 & -(b_2 + d_2) & b_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

Più precisamente, $Q = (q_{ij})_{i,j \geq 0}$ e

$$q_{i,j} = \begin{cases} -(b_i + d_i) & \text{se } i = j \\ d_i & \text{se } j = i - 1 \\ b_i & \text{se } j = i + 1 \\ 0 & \text{altrimenti} \end{cases}$$

con la convenzione che $d_0 := 0$.

La dinamica è così descritta: da quando la popolazione ha raggiunto k individui si attende un tempo esponenziale di parametro $b_k + d_k$ (indipendentemente dai precedenti tempi di attesa). A tale tempo con probabilità $\frac{d_k}{b_k + d_k}$ un individuo muore, altrimenti (quindi con probabilità $\frac{b_k}{b_k + d_k}$) un individuo nasce.

11.5. Osservazione sulla continuità a destra. La scelta di lavorare con processi continui a destra ha una motivazione tecnica profonda. Per risultati di teoria della misura (discussi nella sec.6.6 di [N], per chi fosse interessato) il vantaggio dei processi continui a destra è che la probabilità di ogni evento può essere determinata dalle distribuzioni finito-dimensionali, cioè dalle probabilità della forma

$$\mathbb{P}(X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n)$$

per $n \geq 0$, $0 \leq t_0 \leq t_1 \leq \dots \leq t_n$ e $i_0, i_1, \dots, i_n \in I$.

Osserviamo che ogni traiettoria continua a destra senza esplosione fa un numero finito di salti in ogni intervallo temporale finito. In particolare, ha limite sinistro in ogni tempo t . Le traiettorie continue a destra e aventi limite finito si denotano spesso come “càdlàg” dal francese “continue à droite, limite à gauche”. Se una catena di Markov non ha esplosione (ad esempio quando I è finito) allora le traiettorie sono càdlàg quasi certamente. In alcuni testi le catene di Markov vengono definite direttamente come processi càdlàg, cioè con traiettorie càdlàg (escludendo così l’esplosione).

12. PERDITA DI MEMORIA IN CTMC

CTMC= continuous time Markov chain

Definition 12.1. Data una CTMC $(X_t)_{t \geq 0}$ una variabile aleatoria $T : \Omega \rightarrow [0, +\infty]$ è detta tempo di arresto per la suddetta CTMC se, per ogni $t \in [0, +\infty)$, l’evento $\{T \leq t\}$ è determinato dal tratto di traiettoria $(X_s)_{0 \leq s \leq t}$ fino al tempo t .

Esempi:

- 1) Ogni tempo deterministico è un tempo di arresto (es. $T = 25$).
- 2) Dato $A \subset I$, $H_A := \inf\{t \geq 0 : X_t \in A\}$ è un tempo di arresto. Infatti

$$\begin{aligned} \{H_A \leq t\} &= \{H_A < t\} \cup \{H_A = t\}, \\ \{H_A < t\} &= \{\exists s \in [0, t) : X_s \in A\}, \\ \{H_A = t\} &= \{X_s \notin A \forall s < t \text{ e } \exists \text{ successione } (s_n) \text{ con } X_{s_n} \in A, s_n \geq t, s_n \rightarrow t\}. \end{aligned}$$

Sia (s_n) una successione con $s_n \geq t$, $s_n \rightarrow t$ e $X_{s_n} \in A$. Per continuità a destra in t della traiettoria e siccome I ha la topologia discreta, abbiamo $X_{s_n} = X_t$ per n grande. Quindi $X_t \in A$. Concludiamo che

$$\{H_A = t\} = \{X_s \notin A \forall s < t \text{ e } X_t \in A\}.$$

Quindi, grazie alla continuità a destra!!!, vale

$$\{H_A \leq t\} = \{\exists s \in [0, t] : X_s \in A\},$$

e perciò H_A è un tempo d’arresto.

► Fare [N, Theorem 2.8.1, p.93] senza proof (proprietà di Markov forte per CTMC)

Commenti:

- (i) “conditional on $T < \infty$ and $X_T = i$ ” significa “rispetto alla probabilità condizionata $\mathbb{P}(\cdot | T < \infty, X_T)$ ”.
- (ii) I tempi deterministici sono tempi di arresto, quindi la proprietà di Markov forte implica la proprietà di Markov (che sarebbe la proprietà di Markov forte ristretta ai tempi deterministici).

Non daremo la dimostrazione generale della proprietà di Markov forte per le CTMC. Proviamo la proprietà di Markov per il processo di Poisson (vedi sezione successiva).

13. PROCESSO DI POISSON

Ricordare la definizione di Processo di Poisson data nella Sezione 11.2. Facciamo notare che se $(Y_t)_{t \geq 0}$ è processo di Poisson che inizia nell’origine di intensità λ , allora un generico processo di Poisson $(X_t)_{t \geq 0}$ con distribuzione iniziale ν si realizza ponendo

$$X_t := Y_t + Z, \tag{13.1}$$

dove Z è variabile aleatoria a valori in I di legge ν e Z è indipendente da $(X_t)_{t \geq 0}$.

Theorem 13.1. *Sia $(X_t)_{t \geq 0}$ un processo di Poisson di intensità $\lambda > 0$, distribuzione iniziale qualsiasi, e sia $s \geq 0$. Valgono i seguenti fatti:*

- (i) *Per ogni $i \in \mathbb{N}$, rispetto a $\mathbb{P}(\cdot | X_s = i)$ il processo $(X_{s+t})_{t \geq 0}$ è di Poisson con intensità λ e stato iniziale i , ed è indipendente da $(X_t)_{0 \leq t \leq s}$.*
- (ii) *Il processo $(X_{s+t} - X_s)_{t \geq 0}$ è un processo di Poisson con intensità λ e stato iniziale 0, ed è indipendente da $(X_t)_{0 \leq t \leq s}$.*

Proof. Per semplicità diamo la dimostrazione del punto (i) nel caso di processo di Poisson che inizia nell'origine. La dimostrazione di [N, Theorem 2.4.1, p.75] mostra il seguente Fatto A: *rispetto a $\mathbb{P}(\cdot | X_s = i)$, il processo $(X_{s+t} - X_s)_{t \geq 0}$ è un processo di Poisson con intensità λ e stato iniziale 0, ed è indipendente da $(X_t)_{0 \leq t \leq s}$.* ► Rinviamo a [N] per tale dimostrazione.

Siccome rispetto a $\mathbb{P}(\cdot | X_s = i)$ abbiamo $X_s = i$, il suddetto risultato è equivalente al fatto che, rispetto a $\mathbb{P}(\cdot | X_s = i)$, il processo $(X_{s+t} - i)_{t \geq 0}$ è un processo di Poisson con intensità λ e stato iniziale 0, ed è indipendente da $(X_t)_{0 \leq t \leq s}$. Questo risultato equivale al punto (i).

Proviamo (ii). Dati insiemi A, B nello spazio delle traiettorie abbiamo

$$\begin{aligned}
& \mathbb{P}((X_{s+t} - X_s)_{t \geq 0} \in A, (X_t)_{0 \leq t \leq s} \in B) \\
&= \sum_{i=0}^{\infty} \mathbb{P}((X_{s+t} - X_s)_{t \geq 0} \in A, (X_t)_{0 \leq t \leq s} \in B | X_s = i) \mathbb{P}(X_s = i) \\
&= \sum_{i=0}^{\infty} \mathbb{P}((X_{s+t} - X_s)_{t \geq 0} \in A | X_s = i) \mathbb{P}((X_t)_{0 \leq t \leq s} \in B | X_s = i) \mathbb{P}(X_s = i) \quad (13.2) \\
&= \sum_{i=0}^{\infty} \mathbb{P}_0((X_t)_{t \geq 0} \in A) \mathbb{P}((X_t)_{0 \leq t \leq s} \in B | X_s = i) \mathbb{P}(X_s = i) \\
&= \mathbb{P}_0((X_t)_{t \geq 0} \in A) \mathbb{P}((X_t)_{0 \leq t \leq s} \in B).
\end{aligned}$$

Infatti, la prima identità è un'applicazione della legge della probabilità totale. La seconda e la terza identità seguono dal suddetto Fatto A. La quarta segue dalla legge della probabilità totale. $\square$

Definition 13.2. *Un processo a tempi continui è detto essere ad incrementi indipendenti se, dati $n \geq 2$ e tempo $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, gli incrementi*

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}} \quad (13.3)$$

sono variabili aleatorie indipendenti.

Definition 13.3. *Un processo a tempi continui è detto essere ad incrementi stazionari se la legge del generico incremento $X_t - X_s$, con $t \geq s \geq 0$ è determinata solo da $t - s$. Equivalentemente, $X_t - X_0$ ha la stessa legge di $X_{t+a} - X_a$ per ogni $a \geq 0$.*

Proposition 13.4. *Sia $(X_t)_{t \geq 0}$ un processo di Poisson di intensità $\lambda > 0$. Allora $(X_t)_{t \geq 0}$ è un processo ad incrementi stazionari ed indipendenti e il singolo incremento $X_t - X_s$, dove $0 \leq s \leq t$, è variabile aleatoria di Poisson di parametro $\lambda(t - s)$.*

Proof. Se ν è la distribuzione iniziale di $(X_t)_{t \geq 0}$ allora possiamo realizzare $(X_t)_{t \geq 0}$ tramite (13.1), dove $(Y_t)_{t \geq 0}$ è processo di Poisson di intensità $\lambda > 0$ che inizia nell'origine. Siccome $X_t - X_s = Y_t - Y_s$ e tutte le affermazioni riguardano gli incrementi, possiamo ridurci al caso di processi di Poisson che iniziano nell'origine. Quindi, di seguito, $(X_t)_{t \geq 0}$ è un processo di Poisson di intensità $\lambda > 0$ che inizia nell'origine.

Siano $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$. Per la proprietà di Markov al tempo t_{n-1} (v. Theorem 13.1–(ii) con $s = t_{n-1}$) l'incremento $X_{t_n} - X_{t_{n-1}}$ è indipendente da $(X_u)_{0 \leq u \leq t_{n-1}}$ e perciò è indipendente da $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_{n-1}} - X_{t_{n-2}}$. In modo analogo, $X_{t_{n-1}} - X_{t_{n-2}}$ è indipendente da $X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_{n-2}} - X_{t_{n-3}}$, e così via. In generale, otteniamo che per ogni k l'incremento $X_{t_{k+1}} - X_{t_k}$ è indipendente dai precedenti, per $k = 1, 2, \dots, n-1$. Per tale osservazione e la regola del prodotto abbiamo quindi

$$\begin{aligned}
& \mathbb{P}(X_{t_2} - X_{t_1} \in A_1, X_{t_3} - X_{t_2} \in A_2, \dots, X_{t_n} - X_{t_{n-1}} \in A_{n-1}) \\
&= \mathbb{P}(X_{t_2} - X_{t_1} \in A_1) \prod_{k=2}^{n-1} \mathbb{P}(X_{t_{k+1}} - X_{t_k} \in A_k | X_{t_2} - X_{t_1} \in A_1, \dots, X_{t_k} - X_{t_{k-1}} \in A_{k-1}) \\
&= \mathbb{P}(X_{t_2} - X_{t_1} \in A_1) \prod_{k=2}^{n-1} \mathbb{P}(X_{t_{k+1}} - X_{t_k} \in A_k) \\
&= \prod_{k=1}^{n-1} \mathbb{P}(X_{t_{k+1}} - X_{t_k} \in A_k).
\end{aligned} \tag{13.4}$$

Così si prova l'indipendenza degli incrementi in (13.3).

Per Theorem 13.1–(ii) $X_t - X_s$ ha la stessa legge di X_{t-s} (ricordare che il process $(X_t)_{t \geq 0}$ inizia nell'origine). Quindi gli incrementi sono stazionari.

Per provare che $X_t - X_s$ è variabile aleatoria di Poisson di parametro $\lambda(t-s)$, per la suddetta stazionarietà degli incrementi ci basta provare che X_t è v.a. di Poisson di parametro λt . Abbiamo

$$\begin{aligned}
& \mathbb{P}_0(X_t = n) = \mathbb{P}_0(J_1, J_2, \dots, J_n \leq t, J_{n+1} > t) \\
&= \int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-2}}^t dt_{n-1} \int_{t_{n-1}}^t dt_n \lambda e^{-\lambda t_1} \lambda e^{-\lambda(t_2-t_1)} \cdots \lambda e^{-\lambda(t_n-t_{n-1})} e^{-\lambda(t-t_n)} \\
&= \lambda^n e^{-\lambda t} \int_0^t dt_1 \int_{t_1}^t dt_2 \cdots \int_{t_{n-2}}^t dt_{n-1} \int_{t_{n-1}}^t dt_n 1 \\
&= \lambda^n e^{-\lambda t} t^n P(Z_1 < Z_2 < Z_3 < \cdots < Z_n)
\end{aligned} \tag{13.5}$$

dove $Z_1, Z_2, \dots, Z_n$ sono variabili i.i.d. uniformemente distribuite in $[0, t]$. Per simmetria, per ogni σ permutazione di $\{1, 2, \dots, n\}$ abbiamo

$$P(Z_1 < Z_2 < Z_3 < \cdots < Z_n) = P(Z_{\sigma(1)} < Z_{\sigma(2)} < Z_{\sigma(3)} < \cdots < Z_{\sigma(n)}). \tag{13.6}$$

On the other hand,

$$1 = P(Z_1, Z_2, \dots, Z_n \text{ sono tutte distinte}) = \sum_{\sigma} P(Z_{\sigma(1)} < Z_{\sigma(2)} < Z_{\sigma(3)} < \cdots < Z_{\sigma(n)}) \tag{13.7}$$

dove la somma su σ è sulle permutazioni di $\{1, 2, \dots, n\}$, che sono $n!$.

Combinando (13.6) e (13.7) concludiamo che

$$P(Z_1 < Z_2 < Z_3 < \cdots < Z_n) = 1/n!$$

e quindi, per (13.5),

$$\mathbb{P}_0(X_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

□

14. CARATTERIZZAZIONI EQUIVALENTI ALLA DEFINIZIONE DI CTMC PER SPAZIO DEGLI STATI FINITO

Ricordiamo che data una catena di Markov con spazio degli stati I finito, non c'è esplosione, Quindi quasi certamente non appare lo stato cimitero, e a meno di cambiare il processo su un evento di misura nulla possiamo pensare che le traiettorie siano a valori in I (e non in $I \cup \{\partial\}$).

Theorem 14.1. *Sia $Q = (q_{ij})_{i,j \in I}$ una Q -matrice su I . Sia $(X_t)_{t \geq 0}$ un processo continuo a destra con $X_t \in I$ per ogni $t \geq 0$, e supponiamo di avere tale processo per ogni distribuzione iniziale di λ , con la proprietà che*

$$\mathbb{P}_\lambda(\cdot | X_0 = i) = \mathbb{P}_i(\cdot), \quad (14.1)$$

dove $\mathbb{P}_\lambda$ si riferisce al processo con distribuzione iniziale λ e $\mathbb{P}_i$ è l'abbreviazione di $\mathbb{P}_{\delta_i}$. Allora i seguenti fatti sono equivalenti:

- (a) Per ogni distribuzione iniziale λ $(X_t)_{t \geq 0}$ è CM(λ, Q).
- (b) Per ogni distribuzione iniziale λ , per ogni $t, h \geq 0$, condizionatamente all'evento $X_t = i$ (cioè rispetto a $\mathbb{P}_\lambda(\cdot | X_t = i)$), X_{t+h} è indipendente da $(X_s)_{0 \leq s \leq t}$ e per ogni $j \in I$ si ha

$$\mathbb{P}_\lambda(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij}h + o(h) \quad (14.2)$$

per $h \downarrow 0$, uniformemente in t .

- (c) Per ogni distribuzione iniziale λ , per ogni $n = 0, 1, 2, \dots$, per ogni famiglia di tempi $0 \leq t_0 \leq t_1 \leq \dots \leq t_{n+1}$ e per ogni famiglia di stati $i_0, i_1, \dots, i_{n+1}$ vale

$$\mathbb{P}_\lambda(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = p_{i_n, i_{n+1}}(t_{n+1} - t_n) \quad (14.3)$$

dove la mappa $t \mapsto P(t) = (p_{ij}(t))_{i,j \in I}$ è la soluzione del problema di Cauchy

$$\begin{cases} P'(t) = P(t)Q, \\ P(0) = \mathbb{I}. \end{cases} \quad (14.4)$$

Alcuni commenti:

- (1) la frase “ $\mathbb{P}_\lambda(X_{t+h} = j | X_t = i) = \delta_{ij} + q_{ij}h + o(h)$ per $h \downarrow 0$, uniformemente in t ” significa

$$\lim_{h \downarrow 0} \sup_{t \geq 0} \left| \frac{\mathbb{P}_\lambda(X_{t+h} = j | X_t = i) - \delta_{ij}}{h} - q_{ij} \right| = 0. \quad (14.5)$$

- (2) La soluzione del problema di Cauchy (14.4) è unica ed è data da

$$P(t) = e^{tQ} := \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n.$$

La suddetta funzione matriciale $(P(t))_{t \geq 0}$ è anche l'unica soluzione del problema di Cauchy

$$\begin{cases} P'(t) = QP(t), \\ P(0) = \mathbb{I}. \end{cases}$$

- (3) Per I finito, grazie al suddetto teorema, (b) e (c) sono definizioni equivalenti di catena di Markov a tempo continuo. (b) è la cosiddetta *definizione infinitesimale*, mentre (c) è la cosiddetta *definizione tramite le probabilità di transizione*.

Il suddetto teorema corrisponde a [N, Theorem 2.8.2, p.94] con alcune precisazioni aggiuntive. La dimostrazione che diamo segue molto quella di [N, Theorem 2.8.2, p.94] con alcune integrazioni.

Proof.

• Proviamo che (a) $\Rightarrow$ (b).

Come nella parte a pagina 94 della proof di [N, Theorem 2.8.2, p.94] si ottiene, per ogni coppia $i, j \in I$, che

$$\mathbb{P}_i(X_h = j) \geq \delta_{ij} + q_{ij}h + o(h). \quad (14.6)$$

Di seguito mostriamo che (14.6) implica che

$$\mathbb{P}_i(X_h = j) = \delta_{ij} + q_{ij}h + o(h). \quad (14.7)$$

cioè

$$\lim_{h \downarrow 0} \frac{\mathbb{P}_i(X_h = j) - \delta_{ij}}{h} = q_{ij}. \quad (14.8)$$

A tal fine osserviamo che (14.6) implica

$$\liminf_{h \downarrow 0} \frac{\mathbb{P}_i(X_h = j) - \delta_{ij}}{h} \geq q_{ij}. \quad (14.9)$$

Supponiamo per assurdo che

$$\exists j_0 \text{ tale che } \limsup_{h \downarrow 0} \frac{\mathbb{P}_i(X_h = j_0) - \delta_{ij_0}}{h} > q_{ij_0}. \quad (14.10)$$

Allora per qualche $a > 0$ e per una sotto-successione $h_n \rightarrow 0$ abbiamo

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}_i(X_{h_n} = j_0) - \delta_{ij_0}}{h_n} \geq q_{ij_0} + a. \quad (14.11)$$

Notare che, date due successioni a_n e b_n , vale

$$\liminf_{n \rightarrow \infty} (a_n + b_n) \geq \liminf_{n \rightarrow \infty} a_n + \liminf_{n \rightarrow \infty} b_n. \quad (14.12)$$

Analogo risultato vale se uno considera una generica famiglia di successioni. Affermiamo ora che

$$\begin{aligned} 0 = \liminf_{n \rightarrow \infty} 0 &= \liminf_{n \rightarrow \infty} \sum_j \frac{\mathbb{P}_i(X_{h_n} = j) - \delta_{ij}}{h_n} \\ &\geq \sum_j \liminf_{n \rightarrow \infty} \frac{\mathbb{P}_i(X_{h_n} = j) - \delta_{ij}}{h_n} \geq q_{ij_0} + a + \sum_{j \neq j_0} q_{ij} = a > 0. \end{aligned} \quad (14.13)$$

Infatti la prima disuguaglianza segue da (14.12), mentre per ottenere la seconda disuguaglianza abbiamo applicato (14.11) per $j = j_0$ e (14.9) per $j \neq j_0$. Siccome in (14.13) abbiamo ottenuto che $0 \geq a > 0$, cosa assurda, (14.10) deve essere falso e quindi per ogni j deve valere

$$\limsup_{h \downarrow 0} \frac{\mathbb{P}_i(X_h = j) - \delta_{ij}}{h} \leq q_{ij}. \quad (14.14)$$

As a byproduct of (14.9) and (14.14) we finally get (14.8).

Avendo (14.8) mostriamo come concludere la derivazione di (b). Per la proprietà di Markov abbiamo che, rispetto a $\mathbb{P}_\lambda(\cdot | X_t = i)$, X_{t+h} è indipendente da $(X_s)_{0 \leq s \leq t}$ e per ogni $j \in I$ si ha

$$\mathbb{P}_\lambda(X_{t+h} = j | X_t = i) = \mathbb{P}_i(X_h = j). \quad (14.15)$$

Quindi

$$\sup_{t \geq 0} \left| \frac{\mathbb{P}_\lambda(X_{t+h} = j | X_t = i) - \delta_{ij}}{h} - q_{ij} \right| = \left| \frac{\mathbb{P}_i(X_h = j) - \delta_{ij}}{h} - q_{ij} \right|. \quad (14.16)$$

Dalla suddetta identità e da (14.8) segue immediatamente (14.5). Questo conclude la dimostrazione di (b).

• Proviamo che (b) $\Rightarrow$ (c).

Definiamo $p_{ij}(t) := \mathbb{P}_i(X_t = j)$ e proviamo che $P(t) := (p_{ij}(t))_{i,j \in I}$ soddisfa il sistema di Cauchy (14.4). Affermiamo che, dati $t, h \geq 0$, vale

$$\begin{aligned} p_{ij}(t+h) &= \sum_{k \in I} \mathbb{P}_i(X_t = k) \mathbb{P}_i(X_{t+h} = j | X_t = k) \\ &= \sum_{k \in I} p_{ik}(t) (\delta_{kj} + q_{kj}h + o(h)) = \sum_{k \in I} p_{ik}(t) (\delta_{kj} + q_{kj}h) + o(h). \end{aligned} \quad (14.17)$$

Infatti la prima uguaglianza segue dalla legge della probabilità totale, la seconda uguaglianza segue dalla definizione di $p_{ik}(t)$ e da (14.2), la terza uguaglianza segue dal fatto che I sia finito.

Grazie a (14.17) abbiamo, per $t \geq 0$ e $h > 0$,

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in I} p_{ik}(t) q_{kj} + O(h). \quad (14.18)$$

Prendendo il limite $h \downarrow 0$ otteniamo che

$$\frac{d}{dt_+} p_{ij}(t) = \sum_{k \in I} p_{ik}(t) q_{kj} = (P(t)Q)_{ij}, \quad (14.19)$$

dove $\frac{d}{dt_+}$ denota la derivata destra. Vogliamo provare ora che lo stesso vale per la derivata sinistra:

$$\frac{d}{dt_-} p_{ij}(t) = \sum_{k \in I} p_{ik}(t) q_{kj} = (P(t)Q)_{ij}, \quad (14.20)$$

A tal fine fissiamo $t > 0$ e prendiamo $h > 0$ e come per (14.17) otteniamo

$$\begin{aligned} p_{ij}(t) &= \sum_{k \in I} \mathbb{P}_i(X_{t-h} = k) \mathbb{P}_i(X_t = j | X_{t-h} = k) \\ &= \sum_{k \in I} p_{ik}(t-h) (\delta_{kj} + q_{kj}h + o(h)) = \sum_{k \in I} p_{ik}(t-h) (\delta_{kj} + q_{kj}h) + o(h). \end{aligned} \quad (14.21)$$

Notiamo che abbiamo applicando (14.2) dopo il tempo che appare nel condizionamento è $t-h$. Siccome lo sviluppo infinitesimale (14.2) vale uniformemente nel tempo del condizionamento possiamo scrivere $o(h)$ senza preoccuparci del fatto che il tempo $t-h$ varia quando prendiamo il limite $h \downarrow 0$.

Dividendo per h abbiamo quindi

$$\frac{p_{ij}(t) - p_{ij}(t-h)}{h} = \sum_{k \in I} p_{ik}(t-h) q_{kj} + O(h). \quad (14.22)$$

Siccome $p_{ik}(s) \in [0, 1]$ abbiamo che il membro destro di (14.22) è limitato uniformemente con $h \downarrow 0$, quindi concludiamo che

$$\limsup_{h \downarrow 0} \left| \frac{p_{ij}(t) - p_{ij}(t-h)}{h} \right| < +\infty. \quad (14.23)$$

Questo implica che $\lim_{h \downarrow 0} (p_{ij}(t) - p_{ij}(t-h)) = 0$. Siccome tale limite vale per ogni $t > 0$ abbiamo dimostrato che $p_{ij}(t)$ è una funzione continua a sinistra. Come conseguenza abbiamo

$$\lim_{h \downarrow 0} \sum_{k \in I} p_{ik}(t-h) q_{kj} = \sum_{k \in I} p_{ik}(t) q_{kj}. \quad (14.24)$$

Usando il suddetto limite e prendendo il limite $h \downarrow 0$ nell'identità (14.22) otteniamo (14.20).

Combinando (14.19) con (14.20) otteniamo che $p_{ij}(t)$ è derivabile in t e vale $P'(t) = P(t)Q$. Siccome banalmente $P(0) = \mathbb{I}$, abbiamo provato che $P(t)$ risolve il sistema di Cauchy (14.4).

Fissato $a \geq 0$ definitamo ora, per $t \geq 0$, $\bar{p}_{ij}(t) := \mathbb{P}_\lambda(X_{a+t} = j | X_a = i)$ e $\bar{P}(t) := (\bar{p}_{ij}(t))_{i,j \in I}$. Con la stessa algebra usata da (14.17) in poi, e con l'accortezza di rimpiazzare il generico termine X_s con X_{a+s} , otteniamo che $\bar{P}'(t) = \bar{P}(t)Q$ e $\bar{P}(0) = \mathbb{I}$. Ne deriva (per unicità della soluzione di (14.4)) che $\bar{P}(t) = P(t)$. Questo prova che

$$\mathbb{P}_\lambda(X_{a+t} = j | X_a = i) = p_{ij}(t), \quad \forall t \geq 0, a \geq 0. \quad (14.25)$$

Possiamo finalmente provare (14.3). Per il Lemma 1.1 con $\mathbb{Q}(\cdot) := \mathbb{P}(\cdot | G)$, $E := \{X_{t_{n+1}} = i_{n+1}\}$, $F := \{X_{t_0} = i_0, \dots, X_{t_n} = i_n\}$, $G = \{X_{t_n} = i_n\}$, abbiamo

$$\mathbb{P}_\lambda(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = \mathbb{Q}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n). \quad (14.26)$$

Per (b), rispetto a $\mathbb{Q}$, l'evento $\{X_{t_{n+1}} = i_{n+1}\}$ è indipendente da $(X_s)_{0 \leq s \leq t_n}$ e quindi da $\{X_{t_0} = i_0, \dots, X_{t_n} = i_n\}$. Quindi il membro destro di (14.26) è semplicemente $\mathbb{Q}(X_{t_{n+1}} = i_{n+1}) = \mathbb{P}_\lambda(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n)$. Abbiamo così provato che

$$\mathbb{P}_\lambda(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \dots, X_{t_n} = i_n) = \mathbb{P}_\lambda(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n). \quad (14.27)$$

Combinando (14.27) e (14.25) (con $a = t_n$ e $a+t = t_{n+1}$) arriviamo a (14.3) e abbiamo vinto.

• Proviamo che (c) $\Rightarrow$ (a).

Fissiamo $n \geq 1$, tempi $0 = t_0 \leq t_1 \leq \dots \leq t_n$ e stati $i_0, i_1, \dots, i_n$. Per la regola del prodotto abbiamo

$$\mathbb{P}_\lambda(X_{t_0} = i_0, \dots, X_{t_n} = i_n) = \mathbb{P}_\lambda(X_{t_0} = i_0) \prod_{k=0}^{n-1} \mathbb{P}_\lambda(X_{t_{k+1}} = i_{k+1} | X_{t_0} = i_0, \dots, X_{t_k} = i_k). \quad (14.28)$$

La suddetta identità e (14.3) implicano che

$$\begin{aligned} \mathbb{P}_\lambda(X_0 = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) &= \lambda(i_0) \prod_{k=0}^{n-1} p_{i_k, i_{k+1}}(t_{k+1} - t_k) \\ &= \lambda(i_0) p_{i_0, i_1}(t_1) \prod_{k=1}^{n-1} p_{i_k, i_{k+1}}(t_{k+1} - t_k). \end{aligned} \quad (14.29)$$

Sommando su i_0 abbiamo che per ogni $n \geq 1$, per ogni $0 \leq t_1 \leq \dots \leq t_n$ e per ogni $i_1, i_2, \dots, i_n \in I$ vale

$$\mathbb{P}_\lambda (X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_n} = i_n) = [\lambda P(t_1)]_{i_1} \prod_{k=1}^{n-1} P(t_{k+1} - t_k)_{i_k, i_{k+1}}. \quad (14.30)$$

Ne deriva che le distribuzioni a tempi finiti del generico processo $(X_t)_{t \geq 0}$ che soddisfa (c) sono univocalmente determinate. Tali distribuzioni sono le stesse della catena di Markov $CM(\lambda, Q)$ dato che $(a) \Rightarrow (b) \Rightarrow (c)$ e quindi una $CM(\lambda, Q)$ soddisfa (c). Abbiamo già discusso nella sottosezione (11.5) che per processi continui a destra le distribuzioni a tempi finiti determinano univocalmente il processo. Quindi deve essere $(X_t)_{t \geq 0} = CM(\lambda, Q)$. $\square$

Remark 14.2. *Come spiegato nella dimostrazione (c) $\Rightarrow$ (a) (v. (14.30)) abbiamo che data $(X_t)_{t \geq 0}$ una $CM(\lambda, Q)$ su spazio degli stati finito I , vale*

$$\mathbb{P}_\lambda (X_{t_1} = i_1, \dots, X_{t_n} = i_n) = [\lambda P(t_1)]_{i_1} \prod_{k=1}^{n-1} P(t_{k+1} - t_k)_{i_k, i_{k+1}}$$

dove $P(t) = (p_{ij}(t))_{i,j \in I}$ soddisfa il problema di Cauchy (14.4), $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ e $i_1, \dots, i_n \in I$.

Remark 14.3. *Per I infinito il Teorema 14.1 non è vero. Si prova che (a) equivale a (c) aggiungendo che $P(t)$ è la soluzione minimale non negativa del problema di Cauchy (14.4). Le persone interessate possono trovare i dettagli nei teoremi [N, Theorems 2.8.3, 2.8.4, 2.8.6].*

Remark 14.4. *Il punto (b) del Teorema 14.1 con $t = 0$ implica che, per $h \geq 0$,*

$$\mathbb{P}_i(X_h = j) = \delta_{i,j} + q_{i,j}h + o(h). \quad (14.31)$$

Remark 14.5. (14.3) nel punto (c) del Teorema 14.1 implica che $p_{ij}(t) = \mathbb{P}_i(X_t = j)$. Ne deriva che $P(t)$ è una matrice stocastica. Notiamo inoltre che, per $t, s \geq 0$,

$$\begin{aligned} P(t+s)_{ij} &= p_{ij}(t+s) = \mathbb{P}_i(X_{t+s} = j) \\ &= \sum_k \mathbb{P}_i(X_t = k, X_{t+s} = j) = \sum_k p_{ik}(t)p_{kj}(s) = (P(t)P(s))_{ij} \end{aligned} \quad (14.32)$$

Abbiamo quindi che $(P(t))_{t \geq 0}$ è un semigruppò di matrici stocastiche, dove la proprietà di semigruppò è data da

$$\begin{cases} P(0) = \mathbb{I}, \\ P(t+s) = P(t)P(s) \quad \forall t, s \geq 0. \end{cases} \quad (14.33)$$

15. CLASSI COMUNICANTI IN CATENE DI MARKOV A TEMPO CONTINUO

► Fare tutta sezione 3.2, pag.111 di [N]. In particolare, fare [N, Thm.3.2.1 p.111] con proof.

16. TEMPI DI RAGGIUNGIMENTO E PROBABILITÀ DI ASSORBIMENTO

► Fare l'inizio della Sez. 3.3 in [N] fino a [N, Thm.3.1.1] incluso.

Diamo la dimostrazione del suddetto teorema:

Proof of Thm. 3.1.1 in [N]. Siccome $h_i^A = \mathbb{P}_i(D_A < \infty) = \mathbb{P}_i(H_A < \infty)$, per [N, Thm.1.3.2] sappiamo che $(h_i^A)_{i \in I}$ è la minima soluzione non-negativa del sistema

$$\begin{cases} h_i^A = 1 & \text{per } i \in A, \\ h_i^A = \sum_{j \in J} \pi_{ij} h_j^A & \text{per } i \notin A. \end{cases} \quad (16.1)$$

Quindi, per dimostrare la tesi, ci basta provare che per ogni $i \notin A$ l'equazione

$$h_i^A = \sum_{j \in J} \pi_{ij} h_j^A \quad (16.2)$$

equivale all'equazione

$$\sum_{j \in I} q_{ij} h_j^A = 0. \quad (16.3)$$

Consideriamo prima il caso $q_i > 0$. Allora $\pi_{ij} = q_{ij}/q_i$ per $i \neq j$ e $\pi_{ii} = 0$. Abbiamo quindi che

$$\sum_{j \in J} \pi_{ij} h_j^A = \sum_{j \in J: j \neq i} \pi_{ij} h_j^A = \sum_{j \in J: j \neq i} \frac{q_{ij}}{-q_{ii}} h_j^A.$$

Otteniamo quindi che (16.2) equivale a

$$-q_{ii} h_i^A = \sum_{j \in J: j \neq i} q_{ij} h_j^A. \quad (16.4)$$

Banalmente (16.4) equivale a (16.3).

Supponiamo ora che $q_i = 0$. Allora $q_{ij} = 0$ per ogni $j \in I$ e perciò (16.3) corrisponde alla tautologia $0 = 0$. D'altro canto abbiamo $\pi_{ij} = \delta_{ij}$ e quindi (16.2) corrisponde alla tautologia $h_i^A = h_i^A$, perciò (16.2) e (16.3) sono equivalenti anche quando $q_i = 0$. $\square$

17. RICORRENZA E TRANSIENZA IN CATENE DI MARKOV A TEMPO CONTINUO

► Fare la sezione 3.4 fino a [N, Thm.3.4.1] con dimostrazione inclusa.

Facciamo notare il seguente corollario:

Corollary 17.1. *Data $(X_t)_{t \geq 0}$ una catena di Markov a tempo continuo, uno stato i è ricorrente/transiente se e solo se lo è per la catena di salto associata. In particolare, se lo spazio degli stati è finito, uno stato i è ricorrente se e solo se appartiene ad una classe comunicante chiusa, altrimenti è transiente.*

18. DISTRIBUZIONI INVARIANTI

Definition 18.1. *Una distribuzione λ su I si dice invariante per la Q se $\lambda Q = 0$, dove λ è pensata come vettore riga.*

Theorem 18.2. *Sia Q una Q -matrice con matrice di salto associata Π . Sia λ una distribuzione e sia μ la misura non-negativa su I definita come $\mu_i = \lambda_i q_i$. Allora i seguenti fatti sono equivalenti:*

- (i) λ è una distribuzione invariante per Q ;
- (ii) μ è una misura invariante per Π (includiamo qui la misura nulla tra le misure invarianti)

Per la dimostrazione vedere la dimostrazione del [N, Thm. 3.5.1, p.117].

Proposition 18.3. *Sia I finito e Q irriducibile. Allora esiste un'unica distribuzione invariante per Q . Inoltre, vale $\lambda_i = \frac{\pi_i/q_i}{\sum_{j \in I} \pi_j/q_j}$ dove π è l'unica distribuzione invariante per la matrice di salto Π . In particolare, $\lambda_i > 0$ per ogni i .*

Proof. Se per qualche i fosse $q_i = 0$, allora lo stato i sarebbe uno stato assorbente. Siccome Q è irriducibile, c'è un'unica classe comunicante, e perciò deve essere $I = \{i\}$. Allora $\lambda = \delta_0$ è l'unica distribuzione invariante.

Possiamo quindi supporre che $q_i > 0$ per ogni i . Poniamo $\mu_i = \lambda_i q_i$. Per il Teorema 18.2, la distribuzione λ è distribuzione invariante per Q se e solo se μ è misura (eventualmente nulla) invariante per Π . Notiamo che μ non può essere la misura nulla perchè deve essere $1 = \sum_i \lambda_i = \sum \mu_i/q_i$. Quindi abbiamo che la distribuzione λ è invariante per Q se e solo se $\hat{\mu}$ è distribuzione invariante per Π , dove $\hat{\mu}_i := \mu_i/\mu(I)$. Dato che Q è irriducibile, abbiamo che Π è irriducibile. Siccome I è finito, sappiamo che Π ha un'unica distribuzione invariante π . Ne deriva che la distribuzione λ è invariante per Q se e solo se $\hat{\mu} = \pi$, che equivale a: la distribuzione λ è invariante per Q se e solo se $\frac{\lambda_i q_i}{\sum_{j \in I} \lambda_j q_j} = \pi_i$ per ogni $i \in I$. Abbiamo che $\frac{\lambda_i q_i}{\sum_{j \in I} \lambda_j q_j} = \pi_i$ per ogni $i \in I$ se e solo se $\lambda_i q_i = C \pi_i$ per qualche $C > 0$. In tal caso deve essere $1 = \sum_i \lambda_i = C \sum_i \pi_i/q_i$. Quindi l'unica distribuzione invariante λ è data da $\lambda_i = \frac{\pi_i/q_i}{\sum_{j \in I} \pi_j/q_j}$. Si noti che vale $\lambda_i > 0$ dato che $\pi_i = \frac{1}{\mathbb{E}_i[T_i]} > 0$. $\square$

Lemma 18.4. *Se Q è una Q -matrice e $\mathcal{C}$ è classe comunicante chiusa, allora $\hat{Q} := (Q_{ij})_{i,j \in \mathcal{C}}$ è una Q -matrice su $\mathcal{C}$ irriducibile.*

Proof. Verifichiamo prima che $\hat{Q}$ è una Q -matrice. I segni dei coefficienti di $\hat{Q}$ sono gli stessi di Q per le entrate parametrizzate da $i, j \in \mathcal{C}$. Quindi $\hat{Q}_{i,i} \leq 0$ per ogni $i \in \mathcal{C}$ e $\hat{Q}_{i,j} \geq 0$ for ogni $i \neq j$ in $\mathcal{C}$. Dobbiamo provare che per ogni $i \in \mathcal{C}$ vale $\sum_{j \in \mathcal{C}} Q_{ij} = 0$. Siccome $\mathcal{C}$ è classe comunicante chiusa per Q abbiamo che $Q_{ij} = 0$ se $i \in \mathcal{C}$ e $j \notin \mathcal{C}$ (v. [N, Thm. 3.2.1 p.111]). Quindi, dato $i \in \mathcal{C}$, abbiamo

$$\sum_{j \in \mathcal{C}} Q_{ij} = \sum_{j \in I} Q_{ij} = 0$$

dove l'ultima uguaglianza segue dal fatto che Q è una Q matrice.

Proviamo l'irriducibilità di $\hat{Q}$. Siccome $\mathcal{C}$ è classe comunicante per Q , dati $i, j \in \mathcal{C}$ abbiamo che esistono stati $i_0, i_1, \dots, i_n$ con $i_0 = i, i_n = j$ e $Q_{i_0, i_1} > 0, Q_{i_1, i_2} > 0, \dots, Q_{i_{n-1}, i_n} > 0$ per l'equivalenza tra (i) e (iii) in [N, Thm. 3.2.1 p.111]. Sempre per l'equivalenza tra (i) e (iii) in [N, Thm. 3.2.1 p.111] otteniamo quindi che $i \rightarrow i_k$ per $\mathcal{C}$ per ogni $k = 0, 1, \dots, n$. Siccome $\mathcal{C}$ è chiusa per Q abbiamo che $i_k \in \mathcal{C}$ per ogni k e quindi $\hat{Q}_{i_0, i_1} = Q_{i_0, i_1} > 0, \hat{Q}_{i_1, i_2} = Q_{i_1, i_2} > 0, \dots, \hat{Q}_{i_{n-1}, i_n} = Q_{i_{n-1}, i_n} > 0$. Applicando ancora l'equivalenza tra (i) e (iii) in [N, Thm. 3.2.1 p.111] otteniamo che $i \rightarrow j$ rispetto a $\hat{Q}$. $\square$

Ricordare il Teorema 5.1 per $|I|$ finito.

Theorem 18.5. *Consideriamo una catena di Markov $(X_t)_{t \geq 0}$ con spazio degli stati finito I e generatore Q . Siano $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_N$ le classi comunicanti chiuse (deve essere $N \geq 1$). Allora le distribuzioni invarianti di Q sono tutte e sole della forma $\lambda = \sum_{k=1}^N \alpha_k \lambda^{(k)}$ dove $\alpha_k \geq 0, \sum_{k=1}^N \alpha_k = 1$ e $\lambda^{(k)}$ è l'unica distribuzione invariante con supporto in $\mathcal{C}_k$. $\lambda^{(k)}$ è nulla fuori di $\mathcal{C}_k$ e ristretta a $\mathcal{C}_k$ è l'unica distribuzione invariante della Q -matrice*

$(Q_{ij})_{i,j \in C_k}$. In particolare, esiste sempre almeno una distribuzione invariante. Essa è unica se e solo se vi è un'unica classe comunicante chiusa.

Anche se semplice, non abbiamo dimostrato il suddetto teorema (non è quindi richiesta la dimostrazione).

19. CATENE STAZIONARIE

Definition 19.1. La catena di Markov $(X_t)_{t \geq 0}$ è detta stazionaria se per ogni $T \geq 0$ il processo $(X_{T+t})_{t \geq 0}$ ha la stessa legge di $(X_t)_{t \geq 0}$, cioè è una catena di Markov con la stessa distribuzione iniziale e lo stesso generatore di $(X_t)_{t \geq 0}$.

Di seguito ci limiteremo al caso I finito.

Lemma 19.2. Sia I finito e sia λ una distribuzione su I . Sia Q una Q -matrice e sia $(P(t))_{t \geq 0}$ il semigruppato con generatore Q , cioè $P(t) = e^{tQ}$. Allora vale: $\lambda Q = 0$ (cioè λ è invariante per Q) se e solo se $\lambda P(t) = \lambda$ (cioè λ è invariante per la matrice stocastica $P(t)$ per ogni $t \geq 0$).

Proof. Supponiamo che $\lambda Q = 0$. Allora $\lambda Q^n = 0$ e quindi $\lambda P(t) = \lambda \sum_{n \geq 0} \frac{(tQ)^n}{n!} = \sum_{n \geq 0} \frac{t^n}{n!} (\lambda Q^n) = 0$.

Supponiamo che $\lambda P(t) = \lambda$ per ogni $t \geq 0$. Derivando entrambi i membri in t abbiamo $\lambda P'(0) = 0$. Siccome $P'(0) = Q$ concludiamo che $\lambda Q = 0$. $\square$

Theorem 19.3. Sia I finito e sia $(X_t)_{t \geq 0} = CM(\lambda, Q)$. Allora $(X_t)_{t \geq 0}$ è catena di Markov stazionaria se e solo se λ è invariante per Q .

Proof. Supponiamo che $(X_t)_{t \geq 0}$ sia catena di Markov stazionaria. Allora, per ogni $t \geq 0$ e per ogni $i \in I$, abbiamo $\mathbb{P}(X_t = i) = \mathbb{P}(X_0 = i)$, cioè $(\lambda P(t))_i = \lambda_i$. Quindi abbiamo per ogni $t \geq 0$ che $\lambda P(t) = \lambda$ e per il Lemma 19.2 concludiamo che λ è invariante per Q .

Supponiamo che sia λ invariante per Q . Allora, dati $T \geq 0$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ e $i_1, i_2, \dots, i_n \in I$, per il Teorema 14.1 abbiamo

$$\begin{aligned} & \mathbb{P}(X_{T+t_1} = i_1, X_{T+t_2} = i_2, \dots, X_{T+t_n} = i_n) \\ &= (\lambda P(T+t_1))_{i_1} P(t_2 - t_1)_{i_1, i_2} \cdots P(t_n - t_{n-1})_{i_{n-1}, i_n} \\ &= P(t_1)_{i_1} P(t_2 - t_1)_{i_1, i_2} \cdots P(t_n - t_{n-1})_{i_{n-1}, i_n}. \end{aligned} \quad (19.1)$$

Si noti che l'ultima identità segue dal Lemma 19.2 e dal fatto che λ è invariante per Q . Dalla suddetta identità (19.1), per il Teorema 14.1, otteniamo che il processo $(X_{T+t})_{t \geq 0}$ è $CM(\lambda, Q)$. Per l'arbitrarietà di $T \geq 0$ abbiamo che $(X_t)_{t \geq 0}$ è catena di Markov stazionaria. $\square$

20. INVERSIONE TEMPORALE DI CATENE STAZIONARIE

Lemma 20.1. Sia I finito. Sia Q una Q -matrice e sia λ una distribuzione invariante per Q con entrate positive. Allora la matrice $\hat{Q}$ definita come

$$\hat{Q}_{ij} = \frac{\lambda_j}{\lambda_i} Q_{ji}, \quad \forall i, j \in I, \quad (20.1)$$

è una Q -matrice avente λ come distribuzione invariante. Inoltre posto $P(t) := e^{tQ}$ e $\hat{P}(t) := e^{t\hat{Q}}$, abbiamo

$$\hat{P}(t)_{ij} = \frac{\lambda_j}{\lambda_i} P(t)_{ji}, \quad \forall i, j \in I. \quad (20.2)$$

Proof. Proviamo che $\hat{Q}$ è una Q -matrice. $\hat{Q}$ ha gli stessi segni di Q quindi gli elementi diagonali sono ≤ 0 e gli elementi fuori diagonale sono ≥ 0 . Infine abbiamo per ogni i fissato

$$\sum_j \hat{Q}_{ij} = \sum_j \frac{\lambda_j}{\lambda_i} Q_{ji} = \frac{1}{\lambda_i} \sum_j \lambda_j Q_{ji} = \frac{1}{\lambda_i} (\lambda Q)_i = 0,$$

dove nella prima identità abbiamo usato (20.1) e nell'ultima il fatto che λ è invariante per Q .

Proviamo che λ è invariante per $\hat{Q}$. Abbiamo

$$(\lambda \hat{Q})_j = \sum_i \lambda_i \hat{Q}_{ij} = \sum_i \lambda_i \frac{\lambda_j}{\lambda_i} Q_{ji} = \lambda_j \sum_i Q_{ji} = 0,$$

dove nella prima identità abbiamo usato (20.1) e nell'ultima il fatto che Q è una Q -matrice.

Proviamo (20.2). Sia U la matrice diagonale con $U_{ii} = \lambda_i$. Allora (20.1) si può scrivere come $\hat{Q} = UQU^{-1}$. Siccome l'esponenziale di matrici commuta con la coniugazione abbiamo $e^{t\hat{Q}} = Ue^{tQ}U^{-1}$, ovvero $\hat{P}(t) = UP(t)U^{-1}$, che corrisponde a (20.2). $\square$

Theorem 20.2. *Sia I finito. Sia Q una Q -matrice e sia λ una distribuzione invariante per Q con entrate positive. Sia $(X_t)_{t \geq 0} = CM(\lambda, Q)$. Allora per ogni $T \geq 0$ il processo invertito nel tempo $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, \hat{Q})$ sulla finestra temporale $[0, T]$.*

Proof. Dati $T \geq 0$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T$ e $i_1, i_2, \dots, i_n \in I$, per il Teorema 14.1 abbiamo

$$\begin{aligned} & \mathbb{P}(X_{T-t_1} = i_1, X_{T-t_2} = i_2, \dots, X_{T-t_{n-2}} = i_{n-2}, X_{T-t_{n-1}} = i_{n-1}, X_{T-t_n} = i_n) = \\ & \mathbb{P}(X_{T-t_n} = i_n, X_{T-t_{n-1}} = i_{n-1}, X_{T-t_{n-2}} = i_{n-2}, \dots, X_{T-t_2} = i_2, X_{T-t_1} = i_1) = \quad (20.3) \\ & (\lambda P(T - t_n))_{i_n} P(t_n - t_{n-1})_{i_n, i_{n-1}} P(t_{n-1} - t_{n-2})_{i_{n-1}, i_{n-2}} \cdots P(t_2 - t_1)_{i_2, i_1} \end{aligned}$$

Siccome λ è invariante per Q per il Lemma 19.2 abbiamo $(\lambda P(T - t_n))_{i_n} = \lambda_{i_n}$, e quindi (20.3) ci dà

$$\begin{aligned} & \mathbb{P}(X_{T-t_1} = i_1, X_{T-t_2} = i_2, \dots, X_{T-t_{n-1}} = i_{n-1}, X_{T-t_n} = i_n) = \\ & \lambda_{i_n} P(t_n - t_{n-1})_{i_n, i_{n-1}} P(t_{n-1} - t_{n-2})_{i_{n-1}, i_{n-2}} \cdots P(t_2 - t_1)_{i_2, i_1}. \quad (20.4) \end{aligned}$$

Per (20.2) abbiamo $\lambda_{i_n} P(t_n - t_{n-1})_{i_n, i_{n-1}} = P(t_n - t_{n-1})_{i_{n-1}, i_n} \lambda_{i_{n-1}}$. Abbiamo quindi

$$\begin{aligned} & \mathbb{P}(X_{T-t_1} = i_1, X_{T-t_2} = i_2, \dots, X_{T-t_{n-1}} = i_{n-1}, X_{T-t_n} = i_n) = \\ & P(t_n - t_{n-1})_{i_{n-1}, i_n} \lambda_{i_{n-1}} P(t_{n-1} - t_{n-2})_{i_{n-1}, i_{n-2}} \cdots P(t_2 - t_1)_{i_2, i_1}. \quad (20.5) \end{aligned}$$

Continuando ad applicare a cascata (20.2) otteniamo alla fine che

$$\begin{aligned} & \mathbb{P}(X_{T-t_1} = i_1, X_{T-t_2} = i_2, \dots, X_{T-t_{n-1}} = i_{n-1}, X_{T-t_n} = i_n) = \\ & P(t_n - t_{n-1})_{i_{n-1}, i_n} P(t_{n-1} - t_{n-2})_{i_{n-2}, i_{n-1}} \cdots P(t_2 - t_1)_{i_1, i_2} \lambda_{i_1}. \quad (20.6) \end{aligned}$$

Siccome $\lambda_{i_1} = (\lambda P(t_1))_{i_1}$ dato che λ è invariante per Q (applicare Lemma 19.2), abbiamo che

$$\begin{aligned} & P(t_n - t_{n-1})_{i_{n-1}, i_n} P(t_{n-1} - t_{n-2})_{i_{n-2}, i_{n-1}} \cdots P(t_2 - t_1)_{i_1, i_2} \lambda_{i_1} \\ & = (\lambda P(t_1))_{i_1} P(t_2 - t_1)_{i_1, i_2} \cdots P(t_{n-1} - t_{n-2})_{i_{n-2}, i_{n-1}} P(t_n - t_{n-1})_{i_{n-1}, i_n}. \quad (20.7) \end{aligned}$$

Per il Teorema 14.1 concludiamo che $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, \hat{Q})$ sulla finestra temporale $[0, T]$. $\square$

21. CATENE REVERSIBILI

Definition 21.1. Data una Q -matrice Q e una distribuzione λ , diciamo che λ soddisfa l'equazione del bilancio dettagliato (o che λ è reversibile per Q) se

$$\lambda_i Q_{ij} = \lambda_j Q_{ji} \quad \forall i, j \in I. \quad (21.1)$$

Proposition 21.2. Data una Q -matrice Q e una distribuzione λ , se λ è reversibile per Q allora λ è invariante per Q .

Proof. Abbiamo $(\lambda Q)_i = \sum_j \lambda_j Q_{ji} = \sum_j \lambda_i Q_{ij} = \lambda_i \sum_j Q_{ij} = 0$, dove nella seconda identità abbiamo usato il bilancio dettagliato e nell'ultima il fatto che Q è una Q -matrice. $\square$

Definition 21.3. Una catena di Markov $(X_t)_{t \geq 0} = CM(\lambda, Q)$ è detta reversibile se per ogni $T \geq 0$ il processo invertito nel tempo $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, Q)$ sulla finestra temporale $[0, T]$.

Theorem 21.4. Sia I finito, sia Q una Q -matrice, sia λ una distribuzione con entrate tutte positive e sia $(X_t)_{t \geq 0} = CM(\lambda, Q)$. Allora i seguenti fatti sono equivalenti:

- (i) $(X_t)_{t \geq 0}$ è catena di Markov reversibile;
- (ii) λ è reversibile per Q .

Proof. Supponiamo valga (i). Allora per ogni $i \in I$ e $T \geq 0$ abbiamo $\mathbb{P}(X_0 = i) = \mathbb{P}(X_{T-0} = i)$, cioè $\lambda_i = (\lambda P(T))_i$. Abbiamo quindi provato che $\lambda = \lambda P(T)$ per ogni $T \geq 0$ e quindi, per il Lemma 19.2, λ è distribuzione invariante per Q . Possiamo quindi applicare il Teorema 20.2. Per Teorema 20.2, dato $T > 0$, sappiamo che $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, \hat{Q})$ ma siccome $(X_t)_{t \geq 0}$ è reversibile abbiamo che $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, Q)$. Concludiamo che $Q = \hat{Q}$, e quindi vale il bilancio dettagliato (= λ è reversibile per Q).

Supponiamo ora valga (ii). Per la Prop. 21.2 abbiamo che λ è distribuzione invariante per Q . Possiamo quindi applicare il Teorema 20.2 che ci dice che per ogni $T \geq 0$ il processo $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, \hat{Q})$. Siccome $Q = \hat{Q}$ per il bilancio dettagliato, abbiamo che per ogni $T \geq 0$ il processo $(X_{T-t})_{0 \leq t \leq T}$ è $CM(\lambda, Q)$, cioè $(X_t)_{t \geq 0}$ è catena di Markov reversibile. $\square$

22. GENERATORE DI UNA CATENA DI MARKOV

Di seguito supponiamo lo spazio degli stati I essere finito.

Abbiamo visto che, data una catena di Markov a tempo continuo su spazio degli stati finito, il suo generatore Q è una Q -matrice che genera il semigruppò $(P(t))_{t \geq 0}$, infatti abbiamo $P(t) = e^{tQ}$.

Se la catena di Markov ha distribuzione iniziale λ , allora la distribuzione $\lambda(t)$ di X_t è data da $\lambda P(t)$, cioè $\lambda(t)_i := \mathbb{P}(X_t = i) = (\lambda P(t))_i$. Siccome $P'(t) = QP(t)$, abbiamo quindi che la mappa $t \mapsto \lambda(t)$ soddisfa il problema di Cauchy

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

L'equazione $\frac{d}{dt} \lambda(t) = \lambda(t)Q$ è un esempio di equazione di Fokker-Plack.

Usiamo una notazione più diffusa e scriviamo $p_i(t) := \mathbb{P}(X_t = i)$. Allora l'equazione $\frac{d}{dt}\lambda(t) = \lambda(t)Q$ diventa

$$\frac{d}{dt}p_i(t) = \sum_{j \in I} p_j(t)Q_{ji}. \quad (22.1)$$

Usando che $Q_{ii} = -\sum_{j \in I} Q_{ij}$, possiamo riscrivere (22.1) nella forma più comune:

$$\boxed{\frac{d}{dt}p_i(t) = \sum_{j \neq i} p_j(t)Q_{ji} - \sum_{j \neq i} p_i(t)Q_{ij} \quad \forall i \in I.} \quad (22.2)$$

Il generatore Q ci permette di derivare anche delle equazioni differenziali per il valore atteso degli osservabili. Dato un'osservabile, e cioè una mappa $f : I \rightarrow \mathbb{R}$, consideriamo il valore atteso $\mathbb{E}_i[f(X_t)]$ dell'osservabile f applicato alla catena di Markov al tempo t quando lo stato iniziale è i .

Le funzioni da I in $\mathbb{R}$ sono date dallo spazio $\mathbb{R}^I$, e possono essere pensate come vettori colonna. Abbiamo

Proposition 22.1. *Vale*

$$\mathbb{E}_i[f(X_t)] = (P(t)f)_i, \quad i \in I, \quad (22.3)$$

dove a destra f è pensato come vettore colonna. In particolare, abbiamo che

$$\frac{d}{dt}[f(X_t)] = QP(t)f = P(t)Qf, \quad (22.4)$$

e quindi vale il seguente problema di Cauchy

$$\boxed{\begin{cases} \frac{d}{dt}\mathbb{E}_i[f(X_t)] = \sum_j Q_{ij}\mathbb{E}_j[f(X_t)], \\ \mathbb{E}_i[f(X_0)] = f(i). \end{cases}} \quad (22.5)$$

cioè

$$\boxed{\begin{cases} \frac{d}{dt}\mathbb{E}_i[f(X_t)] = \sum_j Q_{ij}(\mathbb{E}_j[f(X_t)] - \mathbb{E}_i[f(X_t)]), \\ \mathbb{E}_i[f(X_0)] = f(i). \end{cases}} \quad (22.6)$$

Proof. Per definizione di valore atteso abbiamo

$$\mathbb{E}_i[f(X_t)] = \sum_j \mathbb{P}_i[X_t = j]f(j) = P(t)_{ij}f(j) = (P(t)f)_i.$$

Questo prova (22.3).

Per le proprietà del semigruppato $(P(t))_{t \geq 0}$ abbiamo $\frac{d}{dt}P(t) = QP(t) = P(t)Q$. Questa identità e (22.3) implicano (22.4).

Per (22.3) e (22.4) abbiamo

$$\begin{aligned} \frac{d}{dt}\mathbb{E}_i[f(X_t)] &= \frac{d}{dt}(P(t)f)_i = \left(\frac{d}{dt}P(t)f\right)_i = (QP(t)f)_i \\ &= \sum_j Q_{ij}(P(t)f)_j = \sum_j Q_{ij}\mathbb{E}_j[f(X_t)]. \end{aligned}$$

Questo prova (22.5). Per avere (22.6) osserviamo che $Q_{ii} = -\sum_{j \neq i} Q_{ij}$. $\square$

Consideriamo la passeggiata semplice simmetrica su $\mathbb{Z}^d$ a tempi continui. In questo caso il generatore Q è dato da

$$Q_{i,j} = \begin{cases} \frac{1}{2d} & \text{if } |i - j| = 1, \\ -1 & \text{if } i = j. \end{cases}$$

Siccome lo spazio degli stati è infinito uno non può lavorare con osservabili generici f . Per assicurarci che $\mathbb{E}_i[f(X_t)]$ sia ben definita, prendiamo $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ limitata. Poniamo

$$g(t, x) := \mathbb{E}_x[f(X_t)], \quad x = (x_1, \dots, x_d) \in \mathbb{Z}^d.$$

Allora (22.6) diventa

$$\begin{cases} \partial_t g(t, x) = \frac{1}{2d} \sum_{y: y \neq x} [g(t, y) - g(t, x)] = \frac{1}{2d} \sum_{k=1}^d \tilde{\partial}_k g(t, x), \\ g(0, x) = f(x), \end{cases} \quad (22.7)$$

dove

$$\begin{aligned} \tilde{\partial}_k^2 \Delta_k g(t, x) &= (g(t, x + e_k) - g(t, x)) + (g(t, x - e_k) - g(t, x)) \\ &= (g(t, x + e_k) - g(t, x)) - (g(t, x - e_k) - g(t, x)) \end{aligned} \quad (22.8)$$

è la derivata seconda discreta in x bella direzione k . Quindi

$$\begin{cases} \partial_t g(t, x) = \frac{1}{2} \tilde{\Delta} g(t, x), \\ g(0, x) = f(x), \end{cases} \quad (22.9)$$

dove $\tilde{\Delta}$ è il Laplaciano discreto.

23. FUNZIONE GENERATRICE DEI MOMENTI

Definition 23.1. Data una v.a. reale X , la funzione generatrice dei momenti di X (denotata con $M_X(t)$) è definita come $M_X(t) := E(e^{tX}) \in [0, +\infty]$ per $t \in \mathbb{R}$.

La suddetta definizione si estende a più variabili aleatorie reali:

Definition 23.2. Date variabili aleatorie reali $X_1, X_2, \dots, X_m$, la loro funzione generatrice dei momenti congiunta è la funzione

$$M_{X_1, X_2, \dots, X_m} : \mathbb{R}^m \ni (t_1, t_2, \dots, t_m) \mapsto E(e^{t_1 X_1 + t_2 X_2 + \dots + t_m X_m}) \in [0, +\infty].$$

Ricordiamo che la funzione generatrice dei momenti determina univocamente la distribuzione:

Proposition 23.3. Se $M_{X_1, X_2, \dots, X_m}(t_1, t_2, \dots, t_m) = M_{Y_1, Y_2, \dots, Y_m}(t_1, t_2, \dots, t_m)$ per ogni $(t_1, t_2, \dots, t_m) \in \mathbb{R}^m$, allora i vettori aleatori $(X_1, X_2, \dots, X_m)$ e $(Y_1, Y_2, \dots, Y_m)$ hanno la stessa distribuzione. In particolare se $M_X(t) = M_Y(t)$ per ogni $t \in \mathbb{R}$, allora X e Y hanno la stessa distribuzione.

24. VARIABILI ALEATORIE GAUSSIANE

Definition 24.1. Una v.a. X si dice gaussiana (o normale) di media $\mu \in \mathbb{R}$ e varianza $\sigma^2 > 0$ se X è v.a. continua con funzione di densità

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

In tal caso scriveremo $X \sim \mathcal{N}(\mu, \sigma^2)$.

Ricordiamo che, se $X \sim \mathcal{N}(0, 1)$, X si dice v.a. gaussiana (normale) standard.

Convieni nel seguito includere il caso degenere di v.a. gaussiana di media μ e varianza 0 definita come la v.a. deterministica costantemente uguale a μ .

Ricordiamo alcuni fatti fondamentali sulle v.a. gaussiane (vedasi per esempio il libro [R]):

- Se $X \sim \mathcal{N}(\mu, \sigma^2)$ e $Y := \alpha X + \beta$, allora

$$Y \sim \mathcal{N}(\alpha\mu + \beta, \alpha^2\sigma^2). \quad (24.1)$$

In particolare, se X è gaussiana standard, allora $Y := \sigma X + \mu$ è $\mathcal{N}(\mu, \sigma^2)$. Vedasi Cap. 5 in [R]

- Se $X \sim \mathcal{N}(0, 1)$ allora la funzione $M_X(t)$ generatrice dei momenti di X , definita come $M_X(t) := E(e^{tX})$ per $t \in \mathbb{R}$, è data da $M_X(t) = e^{\frac{t^2}{2}}$ (vedasi cap. 7 in [R]).

Come conseguenza delle suddette proprietà abbiamo:

Proposition 24.2. Se $Y \sim \mathcal{N}(\mu, \sigma^2)$ allora

$$M_Y(t) = e^{t\mu + \frac{t^2\sigma^2}{2}}.$$

Infatti possiamo realizzare Y come $Y := \sigma X + \mu$ con X normale standard, e quindi

$$M_Y(t) = E(e^{t(\sigma X + \mu)}) = e^{t\mu} E(e^{t\sigma X}) = e^{t\mu + \frac{t^2\sigma^2}{2}}.$$

Altra conseguenza:

Proposition 24.3. *Se X_1, X_2 siano v.a. gaussiane indipendenti con $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$. Allora, dati $\alpha_1, \alpha_2 \in \mathbb{R}$, $Y := \alpha_1 X_1 + \alpha_2 X_2$ è ancora v.a. gaussiana e si ha*

$$Y \sim \mathcal{N}(\alpha_1 \mu_1 + \alpha_2 \mu_2, \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2).$$

Proof. $\alpha_1 X_1$ e $\alpha_2 X_2$ sono ancora indipendenti. La funzione generatrice dei momenti di una somma v.a. aleatorie indipendenti è il prodotto delle singole funzioni generatrici, quindi

$$E_Y(t) = E(e^{t\alpha_1 X_1 + t\alpha_2 X_2}) = E(e^{t\alpha_1 X_1})E(e^{t\alpha_2 X_2})$$

Siccome $\alpha_i X_i \sim \mathcal{N}(\alpha_i \mu_i, \alpha_i^2 \sigma_i^2)$, concludiamo che

$$E(e^{t\alpha_1 X_1})E(e^{t\alpha_2 X_2}) = e^{t\alpha_1 \mu_1} e^{t^2 \frac{\alpha_1^2 \sigma_1^2}{2}} e^{t\alpha_2 \mu_2} e^{t^2 \frac{\alpha_2^2 \sigma_2^2}{2}} = e^{t(\alpha_1 \mu_1 + \alpha_2 \mu_2) + t^2 \frac{\alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2}{2}}$$

Combinando le suddette formule otteniamo

$$E_Y(t) = e^{t(\alpha_1 \mu_1 + \alpha_2 \mu_2) + t^2 \frac{\alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2}{2}}. \quad (24.2)$$

Siccome v.a. con la stessa funzione generatrice dei momenti hanno la stessa distribuzione (legge), e siccome il membro destro di (24.2) è la funzione generatrice di una gaussiana $\mathcal{N}(\alpha_1 \mu_1 + \alpha_2 \mu_2, \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2)$, concludiamo che $Y \sim \mathcal{N}(\alpha_1 \mu_1 + \alpha_2 \mu_2, \alpha_1^2 \sigma_1^2 + \alpha_2^2 \sigma_2^2)$. Combinando la suddetta identità con le proposizioni 23.3 e 24.2 otteniamo la tesi. $\square$

Corollary 24.4. *Se $X_1, X_2, \dots, X_n$ sono v.a. gaussiane indipendenti con $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ e $a_1, a_2, \dots, a_n$ sono in $\mathbb{R}$, allora*

$$a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

è v.a. gaussiana di tipo $\mathcal{N}(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$.

Proof. Dimostriamo il fatto per induzione su n . Per $n = 1, 2$ segue dalla Prop. 24.3. Supponiamo la tesi sia vera fino a n_1 . Allora $a_1 X_1 + a_2 X_2 + \dots + a_{n_1} X_{n_1}$ è v.a. gaussiana per ipotesi induttiva, inoltre è indipendente da X_n siccome $X_1, X_2, \dots, X_n$ sono v.a. indipendenti. Applicando Prop. 24.3 alle due v.a. gaussiane indipendenti $(a_1 X_1 + a_2 X_2 + \dots + a_{n_1} X_{n_1})$ e X_n otteniamo che $a_1 X_1 + a_2 X_2 + \dots + a_n X_n$ è v.a. gaussiana. Il calcolo della relativa media e varianza segue dalla linearità del valore atteso e dalle proprietà della varianza nel caso di somma di v.a. indipendenti. $\square$

25. VETTORI GAUSSIANI E PROCESSI GAUSSIANI

Definition 25.1. *Un vettore aleatorio $(X_1, \dots, X_d)$ in $\mathbb{R}^d$ è detto vettore gaussiano standard (equivalentemente, è detto avere distribuzione gaussiana standard d -dimensionale) se le sue coordinate $X_1, X_2, \dots, X_d$ sono v.a. gaussiane standard indipendenti.*

Un vettore aleatorio $(X_1, \dots, X_d)$ in $\mathbb{R}^d$ è un vettore gaussiano standard se è un vettore aleatorio continuo con funzione di densità

$$f(x_1, x_2, \dots, x_d) = \prod_{i=1}^d \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} = \frac{1}{(\sqrt{2\pi})^d} e^{-(x_1^2 + x_2^2 + \dots + x_d^2)/2}.$$

Definition 25.2. *Un vettore aleatorio Y in $\mathbb{R}^n$ è detto vettore gaussiano se esiste una matrice A di tipo $n \times d$, un vettore gaussiano standard d -dimensionale e un vettore deterministico $b \in \mathbb{R}^d$ per cui $Y = AX + b$ (in tale scrittura, Y, X, b sono da pensare come vettori colonna).*

Proposition 25.3. *Sia Y un vettore vettore gaussiano in $\mathbb{R}^n$, con $Y = AX + b$ dove A è matrice $n \times d$, $b \in \mathbb{R}^d$ è deterministico, X è vettore gaussiano standard d -dimensionale. Allora vale*

$$\mathbb{E}[Y_i] = b_i \quad \forall i = 1, \dots, d, \quad (25.1)$$

$$\text{Cov}(Y_i, Y_j) = (AA^T)_{i,j} \quad \forall i, j = 1, \dots, d. \quad (25.2)$$

Proof. Abbiamo

$$\mathbb{E}[Y_i] = \mathbb{E}\left[\sum_j A_{ij}X_j + b_i\right] = \sum_j A_{ij}\mathbb{E}[X_j] + b_i = b_i$$

e questo prova (25.1).

Abbiamo

$$\begin{aligned} \text{Cov}(Y_i, Y_j) &= \text{Cov}\left(\sum_k A_{ik}X_k, \sum_r A_{jr}X_r\right) = \\ &= \sum_k \sum_r A_{ik}A_{jr}\text{Cov}(X_k, X_r) = \sum_k \sum_r A_{ik}A_{jr}\delta_{kr} = (AA^T)_{ij}, \end{aligned}$$

e questo prova (25.2). $\square$

La matrice $\text{Cov}(Y) := (\text{Cov}(Y_i, Y_j))_{i,j=1}^n$ è detta *matrice di covarianza* del vettore Y . $b = \mathbb{E}[Y]$ è detto *vettore medio* di Y . Si noti che $\text{Cov}(Y)$ è matrice simmetrica, perchè $\text{Cov}(Y_i, Y_j) = \text{Cov}(Y_j, Y_i)$. Inoltre è semidefinita positiva, cioè $t \cdot \text{Cov}(Y)t \geq 0$ per ogni $t \in \mathbb{R}^d$. Infatti,

$$t \cdot \text{Cov}(Y)t = \sum_i \sum_j t_i \text{Cov}(Y_i, Y_j) t_j = \text{Var}\left(\sum_i t_i Y_i\right) \geq 0.$$

Proposition 25.4. *Sia Y vettore gaussiano n -dimensionale di media b e matrice di covarianza Γ . Allora la sua funzione generatrice dei momenti è data da*

$$M_Y(t_1, t_2, \dots, t_n) = \exp\left\{\bar{t} \cdot b + \frac{1}{2}\bar{t} \cdot \Gamma \bar{t}\right\}, \quad (25.3)$$

dove $\cdot$ indica il prodotto scalare in $\mathbb{R}^d$ e $\bar{t}$ è il vettore $\bar{t} = (t_1, \dots, t_n)$ pensato come vettore colonna.

Proof. Abbiamo

$$\begin{aligned} M_Y(t_1, t_2, \dots, t_n) &= \mathbb{E}\left[\exp\left\{\sum_{i=1}^n t_i Y_i\right\}\right] = \mathbb{E}\left[\exp\left\{\sum_{i=1}^n \sum_{j=1}^d t_i A_{ij} X_j + \sum_i t_i b_i\right\}\right] = \\ &= \exp\left\{\sum_i t_i b_i\right\} \mathbb{E}\left[\exp\left\{\sum_j \left(\sum_i t_i A_{ij}\right) X_j\right\}\right] = \exp\left\{\sum_i t_i b_i\right\} \prod_j \mathbb{E}\left[\exp\left\{\left(\sum_i t_i A_{ij}\right) X_j\right\}\right]. \end{aligned} \quad (25.4)$$

Si noti che nell'ultima identità abbiamo usato l'indipendenza di $X_1, \dots, X_d$. Siccome $X_j \sim \mathcal{N}(0, 1)$, abbiamo

$$\mathbb{E}\left[\exp\left\{\left(\sum_i t_i A_{ij}\right) X_j\right\}\right] = \exp\left\{\frac{1}{2}\left(\sum_i t_i A_{ij}\right)^2\right\}.$$

Quindi

$$\prod_j \mathbb{E} \left[\exp \left\{ \left(\sum_i t_i A_{ij} \right) X_j \right\} \right] = \exp \left\{ \frac{1}{2} \sum_j \left(\sum_i t_i A_{ij} \right)^2 \right\}. \quad (25.5)$$

Possiamo scrivere

$$\begin{aligned} \sum_j \left(\sum_i t_i A_{ij} \right)^2 &= \sum_j \sum_i \sum_r t_i A_{ij} t_r A_{rj} \\ &= \sum_i \sum_r t_i \left(\sum_j A_{ij} A_{rj} \right) t_r = \sum_i \sum_r t_i (AA^T)_{ir} t_r = \bar{t} \cdot \Gamma \bar{t}. \end{aligned} \quad (25.6)$$

Combinando (25.4), (25.5) e (25.6) otteniamo (25.3). $\square$

La suddetta proposizione e la Prop. 23.3 implicano:

Corollary 25.5. *Due vettori gaussiani n -dimensionali aventi lo stesso vettore medio e la stessa matrice di covarianza hanno necessariamente la stessa distribuzione.*

Proposition 25.6. *Un vettore gaussiano n dimensionale avente vettore medio b e matrice di covarianza Γ invertibile ha funzione di densità*

$$f(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{d/2} (\det \Gamma)^{1/2}} \exp \left\{ -\frac{1}{2} (x - b) \cdot \Gamma^{-1} (x - b) \right\}, \quad (25.7)$$

dove x è il vettore di $\mathbb{R}^d$ dato da $x = (x_1, \dots, x_n)$ (pensato come vettore colonna).

Nella dimostrazione sotto si prova che $\det(\Gamma) > 0$ e quindi (25.7) ha senso.

Proof. Γ è matrice simmetrica e quindi esiste trasformazione ortogonale O per cui $\Gamma = ODO^{-1}$, cioè $\Gamma = ODO^T$, con D matrice diagonale.

Siccome Γ è semidefinita positiva, abbiamo che $D_{ii} \geq 0$ per ogni i . Infatti, chiamato e_i il versore i -esimo della base canonica, abbiamo

$$D_{ii} = e_i \cdot D e_i = e_i \cdot O^T \Gamma O e_i = (O e_i) \cdot \Gamma (O e_i) \geq 0.$$

Chiamo $\sqrt{D}$ la matrice diagonale con $(\sqrt{D})_{ii} = \sqrt{D_{ii}}$. Allora

$$\Gamma = ODO^T = O\sqrt{D}\sqrt{D}O^T = (O\sqrt{D}O^T)(O\sqrt{D}O^T)$$

Poniamo $R = O\sqrt{D}O^T$. R è matrice simmetrica e, per sopra, $R^2 = RR^T = \Gamma$. Ne deriva che $(\det R)^2 = \det(\Gamma)$. Sappiamo che $\det(\Gamma) \neq 0$ dato che Γ è invertibile. Quindi $\det(R) \neq 0$ e $\det(\Gamma) = (\det R)^2 > 0$.

Sia $X = (X_1, \dots, X_n)$ vettore gaussiano standard (pensato come vettore colonna) e sia $Y' = RX + b$. Per la proposizione 25.3 Y' è un vettore gaussiano con media b e matrice di covarianza $RR^T = \Gamma$. Applicando il Corollario 25.5 concludiamo che Y' ha la stessa legge di Y . Siccome l'affermazione che vogliamo provare riguarda solo la legge di Y , senza perdita di generalità possiamo pensare che $Y = RX + b$.

Notiamo che $X = R^{-1}(Y - b)$. Consideriamo quindi la mappa affine invertibile $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita come $f(y) = R^{-1}(y - b)$. Il suo jacobiano è R^{-1} e quindi il determinante dello jacobiano è $\det(R^{-1}) = 1/\det(R) = 1/\sqrt{\det(\Gamma)} > 0$. Quindi $|\det(\text{Jac}(f))| = 1/\sqrt{\det(\Gamma)}$.

Abbiamo

$$\begin{aligned} \mathbb{P}(Y \in B) &= \mathbb{P}(RX + b \in B) = \mathbb{P}(X \in f(B)) = \int_{f(B)} g(x) dx \\ &= \int_B g(f(y)) |\det(\text{Jac}(f))| dy = \int_B g(f(y)) \frac{1}{\sqrt{\det(\Gamma)}} dy, \end{aligned} \quad (25.8)$$

dove $g(x) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}x \cdot x}$. Quindi otteniamo

$$\mathbb{P}(Y \in B) = \int_B \frac{1}{(2\pi)^{d/2}} \exp \left\{ -\frac{1}{2} R^{-1}(y-b) \cdot R^{-1}(y-b) \right\} \frac{1}{\sqrt{\det(\Gamma)}} dy.$$

Per concludere basta osservare che per simmetria $R^{-1}(y-b) \cdot R^{-1}(y-b) = (y-b) \cdot R^{-1} R^{-1}(y-b) = (y-b) \cdot \Gamma^{-1}(y-b)$. $\square$

Remark 25.7. In alcuni testi il contenuto della Prop. 25.6 è preso come definizione di vettore gaussiano quando la matrice di covarianza è invertibile.

26. MOTO BROWNIANO 1-DIMENSIONALE

Definition 26.1. Un processo stocastico $(B_t)_{t \geq 0}$ a valori reali è detto moto Browniano con parametro di drift μ , coefficiente di diffusione σ^2 e punto iniziale x (scriviamo $(B_t)_{t \geq 0} = BM(x, \mu, \sigma^2)$) se valgono le seguenti proprietà:

- (P1) $B_0 = x$.
- (P2) Il processo ha incrementi indipendenti, cioè dati $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$ gli incrementi $B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$ sono variabili aleatorie indipendenti.
- (P3) per ogni $t \geq 0$ e per ogni $h \geq 0$ gli incrementi $B_{t+h} - B_t$ sono gaussiane $\mathcal{N}(\mu t, \sigma^2 t)$.
- (P4) quasi certamente le traiettorie sono continue, cioè $1 = \mathbb{P}$ (la mappa $t \mapsto B_t$ è continua).

Il moto Browniano è detto anche *processo di Wiener*. Quando $x = 0, \mu = 0$ e $\sigma^2 = 1$, parliamo di moto Brownian standard.

Theorem 26.2. Dati $x \in \mathbb{R}, \mu \in \mathbb{R}$ e $\sigma^2 \geq 0$, il moto Browniano $B(x, \mu, \sigma^2)$ esiste.

Non diamo la dimostrazione del suddetto teorema. Una dimostrazione basata sulla costruzione di Paul Lévy si può trovare ad esempio in [MP]. In alternativa, il moto Browniano si può costruire come processo limite della passeggiata aleatoria simmetrica opportunamente riscalata nel tempo e nello spazio (vedasi il *principio di invarianza di Donsker* ad esempio in [MP]).

Apparentemente il moto Browniano sembra essere qualcosa di molto specifico perché abbiamo specificato la legge degli incrementi. Notiamo che per la proprietà (P3) gli incrementi del moto Browniano sono stazionari, cioè la legge di $B_{t+h} - B_t$, con $t, h \geq 0$, dipende solo dalla lunghezza h dell'intervallo $[t, t+h]$. Il seguente teorema mostra invece il carattere universale del moto Browniano (v. ad esempio [MP]):

Theorem 26.3. Sia $(B_t)_{t \geq 0}$ un processo che inizia in x (cioè $B(0) = x$), con incrementi indipendenti e stazionari e con traiettorie continue quasi certamente. Allora $(B_t)_{t \geq 0}$ è un moto Browniano: esistono $\mu \in \mathbb{R}$ e $\sigma^2 \geq 0$ tali che $(B_t)_{t \geq 0} = BM(x, \mu, \sigma^2)$.

27. ALCUNE PROPRIETÀ DEL MOTO BROWNIANO

Proposition 27.1. Se $(B_t)_{t \geq 0}$ è un moto Browniano standard, allora il processo $(B'_t)_{t \geq 0}$ definito come

$$B'_t = x + \mu t + \sigma B_t$$

è $BM(x, \mu, \sigma^2)$.

Proof. Abbiamo che $(B'_t)_{t \geq 0}$ soddisfa le proprietà (P1), (P2), (P3), (P4) nella definizione di $BM(x, \mu, \sigma^2)$. Infatti $B'_0 = x + \sigma B_0 = x$ (vale (P1)!). Inoltre abbiamo

$$B'_{t+h} - B'_t = \mu h + \sigma(B_{t+h} - B_t).$$

Quindi, essendo immagine di incrementi indipendenti, anche gli incrementi di B' sono indipendenti. Inoltre, essendo immagine affine, anche gli incrementi di B' sono gaussiane e allora vale per (24.1) $B'_{t+h} - B'_t = \mathcal{N}(\mu h, \sigma^2 h)$. Essendo B' immagine affine di B , le traiettorie di B' sono continue q.c. $\square$

Proposition 27.2 (Proprietà di riscaldamento). *Sia $a > 0$. Se $(B_t)_{t \geq 0}$ è $BM(x, \mu, \sigma^2)$, allora il processo $(B_{at})_{t \geq 0}$ è $BM(x, \mu a, \sigma^2 a)$. In particolare, se $(B_t)_{t \geq 0}$ è BM standard, allora il processo riscaldato $B_{at}/\sqrt{a}$ è BM standard.*

Proof. È banale verificare che $(B_{at})_{t \geq 0}$ verifica (P1), (P2), (P4) dalle analoghe proprietà di $(B_t)_{t \geq 0}$. Per (P3) di B applicato con at e ah al posto di t e h , $B_{a(t+h)} - B_{at}$ è $\mathcal{N}(\mu a, \sigma^2 a)$. $\square$

Proposition 27.3. *Il moto Browniano è un processo gaussiano.*

Proof. Fissiamo tempi $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$. Sia B un $BM(x, \mu, \sigma^2)$. Se $\sigma^2 = 0$ allora $B_t = x + \mu t$ ed è banale verificare che il processo è un processo gaussiano dato che il vettore $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ è deterministico. Poniamo $\Delta_i := B_{t_i} - B_{t_{i-1}}$. Sappiamo che $\Delta_1, \dots, \Delta_n$ sono v.a. indipendenti e $\Delta_i \sim \mathcal{N}(\mu(t_i - t_{i-1}), \sigma^2(t_i - t_{i-1}))$ (per le proprietà (P2) e (P3)). Allora

$$\begin{cases} B_{t_1} = x + \Delta_1 \\ B_{t_2} = x + \Delta_1 + \Delta_2, \\ \dots \\ B_{t_n} = x + \Delta_1 + \Delta_2 + \dots + \Delta_n. \end{cases} \quad (27.1)$$

Sia Z_i v.a. tale che $\Delta_i = \mu(t_i - t_{i-1}) + \sigma Z_i$. Allora $(Z_1, Z_2, \dots, Z_n)$ è un vettore gaussiano standard. Abbiamo quindi

$$\begin{cases} B_{t_1} = x + \mu t_1 + \sigma Z_1 \\ B_{t_2} = x + \mu t_2 + \sigma(Z_1 + Z_2), \\ \dots \\ B_{t_n} = x + \mu t_n + \sigma(Z_1 + Z_2 + \dots + Z_n). \end{cases} \quad (27.2)$$

Questo prova che il vettore $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ è immagine affine di un vettore gaussiano standard. $\square$

Proposition 27.4. *Sia B un $BM(x, \mu, \sigma^2)$. Allora $Cov(B_t, B_s) = \sigma^2 \min\{t, s\}$.*

Proof. Siccome $B_r = B_r - B_0 + x$ abbiamo che $B_r \sim \mathcal{N}(x + \mu r, \sigma^2 r)$. Siccome la covarianza trascura le costanti additive possiamo pensare che $x = \mu = 0$ senza perdita di generalità e quindi $B_r \sim \mathcal{N}(0, \sigma^2 r)$. Ne deriva che $\mathbb{E}[B_t] = 0$ e $\mathbb{E}[B_s] = 0$. Quindi $Cov(B_t, B_s) = \mathbb{E}[B_t B_s]$. Supponiamo per esempio che $s \leq t$. Allora

$$Cov(B_t, B_s) = \mathbb{E}[B_t B_s] = \mathbb{E}[(B_t - B_s)B_s + B_s^2] = \mathbb{E}[(B_t - B_s)B_s] + \mathbb{E}[B_s^2]. \quad (27.3)$$

Usiamo ora che $B_t - B_s$ e $B_s - B_0 = B_s$ sono incrementi indipendenti e sono gaussiane di media zero. Quindi otteniamo che $\mathbb{E}[(B_t - B_s)B_s] = \mathbb{E}[B_t - B_s]\mathbb{E}[B_s] = 0$. Combinando questo con (27.3) concludiamo che

$$Cov(B_t, B_s) = \mathbb{E}[B_s^2] = \sigma^2 s. \quad \square$$

Proposition 27.5 (Inversione temporale). *Se $(B_t)_{t \geq 0}$ è BM standard, allora il processo $(X_t)_{t \geq 0}$ definito come*

$$X_t := \begin{cases} tB(1/t) & \text{se } t > 0, \\ 0 & \text{se } t = 0, \end{cases} \quad (27.4)$$

è pure un BM standard.

Proof. Dati $t, h \geq 0$ abbiamo che, dati tempi $0 \leq t_1 \leq t_2 \leq \dots \leq t_n$, $X = (X_{t_1}, \dots, X_{t_n})$ è un vettore gaussiano. Ne deriva che il process X è un processo gaussiano. Verifichiamo che ha stesse covarianze del moto Browniano standard: dati $t, h \geq 0$ per la prop. precedente abbiamo per la prop. 27.4

$$\begin{aligned} \text{Cov}(X_{t+h}, X_t) &= \text{Cov}((t+h)B(1/(t+h)), tB(1/t)) \\ &= (t+h)t \text{Cov}(B(1/(t+h)), B(1/t)) = (t+h)t \frac{1}{t+h} = t. \end{aligned}$$

Inoltre banalmente $\mathbb{E}[X_t] = 0$ per ogni $t \geq 0$. Quindi X è un processo gaussiano con le stesse covarianze e medie del moto Browniano standard. Questo implica che (P1), (P2) e (P3) valgono pure per X . Rest da provare che X ha traiettorie continue q.c. Per la continuità delle traiettorie di B , $(X_t)_{t > 0}$ è q.c. una traiettoria continua. Resta da provare che

$$\lim_{t \rightarrow 0} X_t = 0 = X_0 \text{ quasi certamente.} \quad (27.5)$$

Siccome $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ e $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ hanno la stessa legge (essendo vettori gaussiani della stessa media e con la stessa matrice di covarianze), otteniamo che $(X_t)_{t \geq 0, t \in \mathbb{Q}}$ ha la stessa legge di $(B_t)_{t \geq 0, t \in \mathbb{Q}}$. Siccome $\lim_{t \rightarrow 0} B_t = 0$ q.c., allora $\lim_{t \rightarrow 0, t \in \mathbb{Q}} B_t = 0$ q.c. Ma essendo $(X_t)_{t \geq 0, t \in \mathbb{Q}} \sim (B_t)_{t \geq 0, t \in \mathbb{Q}}$, concludiamo che

$$\lim_{t \rightarrow 0, t \in \mathbb{Q}} X_t = 0 \text{ q.c.} \quad (27.6)$$

(27.6) può essere ottenuto anche così. Numeriamo i numeri razionali come $q_1, q_2, \dots$. Allora, per la continuità della funzione di probabilità rispetto a successioni monotone di eventi, abbiamo

$$\begin{aligned} 1 &= \mathbb{P}\left(\lim_{t \rightarrow 0, t \in \mathbb{Q}} B_t = 0\right) = \mathbb{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{r=1}^{\infty} \{|B_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\}\right) \\ &= \lim_{m \rightarrow \infty} \lim_{k \rightarrow \infty} \lim_{r \rightarrow \infty} \mathbb{P}\left(\{|B_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\}\right) \end{aligned} \quad (27.7)$$

Siccome l'evento riguarda un numero finito di tempi abbiamo che

$$\{|B_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\} = \{|X_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\}$$

Quindi otteniamo

$$\lim_{m \rightarrow \infty} \lim_{k \rightarrow \infty} \lim_{r \rightarrow \infty} \mathbb{P}\left(\{|X_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\}\right) = 1$$

ma il membro sinistro è proprio

$$\mathbb{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{r=1}^{\infty} \{|X_{q_j}| \leq 1/m \forall j \leq r \text{ with } q_j \leq 1/k\}\right) = \mathbb{P}\left(\lim_{t \rightarrow 0, t \in \mathbb{Q}} X_t = 0\right) = 1.$$

Questo completa la dimostrazione di pinco.

Grazie a (27.6), essendo X q.c. continua per $t \in (0, +\infty)$ allora deve essere $\lim_{t \rightarrow 0} X_t = 0$ q.c. □

Proposition 27.6 (Legge dei grandi numeri). *Se $(B_t)_{t \geq 0}$ è $BM(x, \mu, \sigma^2)$, allora*

$$\lim_{t \rightarrow \infty} \frac{B_t}{t} = 0 \text{ quasi certamente.} \quad (27.8)$$

Proof. Dimostriamo prima il risultato nel caso di B dato da BM standard. Consideriamo il processo X ottenuto da B come in (27.4). Abbiamo $X_t = tB(1/t)$ per $t > 0$. Siccome X è BM standard per Prop. 27.5, abbiamo che $\lim_{t \rightarrow 0} X_t = 0$ q.c. e quindi vale (27.8).

Per la Prop. 27.1, scrivendo $B_t = x + \mu t + \sigma Z_t$, abbiamo che Z è BM standard e per quanto appena provato vale $\lim_{t \rightarrow \infty} \frac{Z_t}{t} = 0$ quasi certamente. Quindi

$$\frac{B_t}{t} = \frac{x}{t} + \frac{\mu t}{t} + \sigma \frac{Z_t}{t} \rightarrow \mu \quad \text{q.c.}$$

□

28. IL LAPLACIANO COME GENERATORE DEL MOTO BROWNIANO

Consideriamo il moto Browniano $BM(0, 0, \sigma^2)$. Allora, dati $s \leq t$, allora

$$\mathbb{P}(X_{t+h} \in A | X_t = x) = \mathbb{P}(X_{t+h} - X_t + x \in A | X_s = x) = \mathbb{P}(X_{t+h} - X_t + x \in A).$$

Abbiamo che $X_{t+h} - X_t + x \sim \mathcal{B}(x, \sigma^2 h)$. Quindi

$$\mathbb{P}(X_{t+h} \in A | X_t = x) = \int_A \frac{1}{\sqrt{2\pi\sigma^2 h}} e^{-\frac{(y-x)^2}{2\sigma^2 h}} dy = \int_A p_h(x, y) dy.$$

$p_h(x, y)$ si dice densità della probabilità di transizione del moto Browniano.

Notiamo anche che, se B è moto Browniano $BM(x, 0, \sigma^2)$, allora

$$\mathbb{P}(X_t \in A) = \int_A p_t(x, y) dy.$$

Quando $\sigma^2 = 2$ abbiamo

$$p_t(x, y) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{(y-x)^2}{4t}}$$

Abbiamo $\partial_t p_t(x, y) = \Delta_y p_t(x, y)$, equivalentemente $\partial_t p_t(x, y) = \Delta_x p_t(x, y)$, per $t > 0$. Allora $p_t(x, y)$ corrisponde per CM a tempo continuo a $P(t)_{ij}$, Δ corrisponde al generatore Q e le equazioni $\partial_t p_t(x, y) = \Delta_y p_t(x, y)$ e $\partial_t p_t(x, y) = \Delta_x p_t(x, y)$ corrispondono a $\frac{d}{dt} P(t) = P(t)Q = QP(t)$ (tranne il fatto che ora deriviamo solo per $t > 0$).

Data una funzione f chiamiamo

$$g(t, x) := \mathbb{E}_x [f(B_t)] \quad (28.1)$$

dove $\mathbb{E}_x$ è il valore atteso rispetto a $B = BM(x, 0, \sigma^2)$. Allora $g(t, x) = \int p_t(x, y) f(y) dy$. Per avere che $g(t, x)$ è ben definito prendiamo f limitata. Ricordiamo il seguente teorema (v. [E, Thm.1, Sect. 2.3]):

Theorem 28.1. *Sia $f \in C(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^n)$ e definiamo g come (28.1). Allora*

- (i) $g \in C^\infty(\mathbb{R}^d \times (0, +\infty))$,
- (ii) $\partial_t g = \Delta g$ per $x \in \mathbb{R}^d$ e $t > 0$,
- (iii) $\lim_{\substack{(x,t) \rightarrow (x^0, 0) \\ x \in \mathbb{R}^d, t > 0}} g(x, t) = g(x^0) \quad \forall x^0 \in \mathbb{R}^d.$

Se $\sigma^2 = 1$ le considerazioni sopra restano vera ma il generatore sarà $\frac{1}{2}\Delta$. Come già osservato il generatore della passeggiata aleatoria simmetrica semplice su $\mathbb{Z}^d$ è $\frac{2}{2}\tilde{\Delta}$, dove $\tilde{\Delta}$ è il Laplaciano discreto. Non è un caso. In effetti, la passeggiata aleatoria simmetrica semplice su $\mathbb{Z}^d$ opportunamente riscalata nel tempo e nello spazio converge al moto Browniano standard (principio di invarianza di Donsker). Notiamo che il suddetto teorema per $\frac{1}{2}\Delta$ diventa la controparte continua di (22.6).

29. MOTO BROWNIANO MULTIDIMENSIONALE

Definiamo il moto Browniano d -dimensionale standard $(B(t))_{t \geq 0}$ come

$$B(t) = (B_1(t), B_2(t), \dots, B_d(t))$$

dove $(B_1(t))_{t \geq 0}, (B_2(t))_{t \geq 0}, \dots, (B_d(t))_{t \geq 0}$ sono moti Browniani standard indipendenti. Un generico moto Browniano d -dimensional è ottenuto componendo $B(t)$ con una mappa affine $\mathcal{A} : \mathbb{R}^d \rightarrow \mathbb{R}^d$.

APPENDIX A. INVERSIONE TEMPORALE PER CATENE DI MARKOV A TEMPO DISCRETO

Theorem A.1. Sia $(X_n)_{n \geq 0}$ $CM(\lambda, P)$ con λ distribuzione invariante tale che $\lambda_i > 0$ per ogni i .

- Sia $\hat{P}$ è la matrice data da

$$\hat{P}_{ij} := \frac{\lambda_j}{\lambda_i} P_{ji} \quad \forall i, j \in I. \quad (\text{A.1})$$

Allora $\hat{P}$ è una matrice stocastica e λ è invariante per $\hat{P}$.

- Dato $N \geq 0$ definiamo il processo invertito nel tempo $(X_n^{(N)})_{0 \leq n \leq N}$ come $X_n^{(N)} := X_{N-n}$. Allora $(X_n^{(N)})_{0 \leq n \leq N}$ è $CM(\lambda, \hat{P})$ sulla finestra temporale $[0, N]$

Proof. Verifichiamo che $\hat{P}$ è matrice stocastica (si noti che è ben definita dato che λ ha entrate positive). Banalmente $\hat{P}$ ha entrate non negative. Dato $i \in I$ abbiamo

$$\sum_j \hat{P}_{ij} = \sum_j \frac{\lambda_j}{\lambda_i} P_{ji} = \frac{1}{\lambda_i} \sum_j \lambda_j P_{ji} = \frac{1}{\lambda_i} (\lambda P)_i = \frac{\lambda_i}{\lambda_i} = 1,$$

dove la penultima identità segue dal fatto che $\lambda P = \lambda$ essendo λ invariante per P .

Verifichiamo che λ è distribuzione invariante per $\hat{P}$. Per (A.1) abbiamo

$$(\lambda \hat{P})_j = \sum_i \lambda_i \hat{P}_{i,j} = \sum_i \lambda_i \frac{\lambda_j}{\lambda_i} P_{ji} = \lambda_j \sum_i P_{ji} = \lambda_j.$$

Si noti che nell'ultima identità abbiamo usato che P è matrice stocastica.

Verifichiamo che $(X_n^{(N)})_{0 \leq n \leq N}$ è $CM(\lambda, \hat{P})$. Fissati $i_0, i_1, \dots, i_N \in I$ abbiamo

$$\begin{aligned} & \mathbb{P} \left(X_0^{(N)} = i_0, X_1^{(N)} = i_1, \dots, X_N^{(N)} = i_N \right) \\ &= \mathbb{P} (X_0 = i_N, X_1 = i_{N-1}, \dots, X_N = i_0) \\ &= \lambda_{i_N} P_{i_N, i_{N-1}} P_{i_{N-1}, i_{N-2}} P_{i_{N-2}, i_{N-3}} \cdots P_{i_1, i_0} \\ &= P_{i_{N-1}, i_N} \lambda_{i_{N-1}} P_{i_{N-1}, i_{N-2}} P_{i_{N-2}, i_{N-3}} \cdots P_{i_1, i_0} \\ &= P_{i_{N-1}, i_N} P_{i_{N-2}, i_{N-1}} \lambda_{i_{N-2}} P_{i_{N-2}, i_{N-3}} \cdots P_{i_1, i_0} \\ &= \dots \\ &= P_{i_{N-1}, i_N} P_{i_{N-2}, i_{N-1}} P_{i_{N-3}, i_{N-2}} \cdots P_{i_0, i_1} \lambda_{i_0}. \end{aligned} \quad (\text{A.2})$$

Si noti che la terza identità segue da (A.1) che implica $\lambda_{i_N} P_{i_N, i_{N-1}} = P_{i_{N-1}, i_N} \lambda_{i_{N-1}}$. Analogamente abbiamo le identità successive. Per (A.2) abbiamo

$$\mathbb{P} \left(X_0^{(N)} = i_0, X_1^{(N)} = i_1, \dots, X_N^{(N)} = i_N \right) = \lambda_{i_0} P_{i_0, i_1} \cdots P_{i_{N-3}, i_{N-2}} P_{i_{N-2}, i_{N-1}} P_{i_{N-1}, i_N},$$

e questo implica che $(X_n^{(N)})_{0 \leq n \leq N}$ è $CM(\lambda, \hat{P})$. $\square$

REFERENCES

- [E] L.C. Evans; *Partial Differential Equations*. American Mathematical Society. Graduate Studies in Mathematics. Volume 19.
- [MP] P. Mörters, Y. Peres; *Brownian motion*. Cambridge University Press. Disponibile online gratuitamente.
- [N] J. Norris; *Markov chains*. Cambridge University Press.

[R] S.M. Ross; *Calcolo delle probabilità*. Casa Ed. Apogeo.