Intorduzione alla teoria delle Catene di Markov

Michele Gianfelice

a.a. 2014/2015

Definizione 1 *Sia* $\left(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq 0},\mathbb{P}\right)$ *uno spazio di probabilità filtrato. Una successione di v.a.* $\{\xi_n\}_{n\geq 0}$ *definita su* $\left(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq 0},\mathbb{P}\right)$ *ed adattata a* $\{\mathcal{F}_n\}_{n\geq 0}$ *è detta* catena di Markov *rispetto a* $\mathbb{P}$, *se* $\forall n\geq m\geq 0$, $B\in\mathcal{B}(\mathbb{R})$

$$\mathbb{P}\left(\{\omega\in\Omega:\xi_n(\omega)\in B\}\,|\mathcal{F}_m\right)=\mathbb{P}\left(\{\omega\in\Omega:\xi_n(\omega)\in B\}\,|\xi_m\right)\quad \mathbb{P}\text{-}q.c., \tag{1}$$

ovvero, se per ogni funzione f $\mathcal{B}(\mathbb{R})$*-mis., limitata*

$$\mathbb{E}\left[f(\xi_n)\,|\mathcal{F}_m\right]=\mathbb{E}\left[f(\xi_n)\,|\xi_m\right]\quad \mathbb{P}\text{-}q.c.. \tag{2}$$

La (1) è detta proprietà di Markov.

Ciò implica che, se $n\geq m\geq 0$, dati due eventi F e P tali che:

- F appartiene alla σalgebra rispetto alla quale sono misurabili le v.a. ξ_i, $i\geq m$;
- $P\in\mathcal{F}_k$, $0\leq k\leq m$;

si ha

$$\mathbb{P}(F\cap P|\mathcal{F}_m)=\mathbf{1}_P\mathbb{P}(F|\mathcal{F}_m)=\mathbb{P}(F|\mathcal{F}_m)\,\mathbb{P}(P|\mathcal{F}_m)\quad \mathbb{P}\text{-q.c.} \tag{3}$$

Ovvero, nel caso in cui s'interpreti $\{\xi_n\}_{n\geq 0}$ come l'evoluzione di un sistema reale, l'evento F rappresenta un evento futuribile rispetto allo stato attuale del sistema ξ_m, mentre P rappresenta un evento passibile di essere accaduto nel passato relativamente allo stato presente del sistema. Pertanto, la relazione precedente implica che per una catena di Markov, subordinatamente allo stato attuale del sistema l'evoluzione futura di questo non è afflitta da quanto è già accaduto.

Se si interpreta l'indice che definisce la sequenza n come indice temporale, una sequenza stocastica può essere considerata come una modellizzazione dell'evoluzione di un sistema dinamico fisico, biologico, ecc..

Osservazione 2 *Se* $\{\xi_n\}_{n\geq0}$ *definita su* $\left(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geq0},\mathbb{P}\right)$ *è una catena di Markov rispetto a* $\mathbb{P}$, *nel caso in cui la filtrazione* $\{\mathcal{F}_n\}_{n\geq0}$ *coincide con la filtrazione generata dalla successione di v.a.* $\{\xi_n\}_{n\geq0}$, *ovvero tale che* $\forall n\geq0$, $\mathcal{F}_n=\mathcal{F}_n(\xi):=\mathcal{F}_n^{\xi}$ *cioè la* σ*algebra generata dalla collezione di v.a.* $\{\xi_i\}_{i=0}^{n}$ [1], *si dirà che* $\{\xi_n\}_n\geq0$ *su* $(\Omega,\mathcal{F},\mathbb{P})$ *è una catena di Markov rispetto a* $\mathbb{P}$.

0.1 Probabilità di transizione

Definizione 3 *Sia* $\{\xi_n\}_{n\geq0}$ *su* $(\Omega,\mathcal{F},\mathbb{P})$ *una catena di Markov rispetto a* $\mathbb{P}.\forall n\geq0$, *La distribuzione probabilità condizionata della v.a.* ξ_{n+1} *rispetto alla v.a.* $\xi_n,\mathbb{P}(\xi_{n+1}(\cdot)|\xi_n)$, *ovvero tale che,* $\forall B\in\mathcal{B}(\mathbb{R})$,

$$\mathbb{P}\left(\xi_{n+1}^{-1}(B)|\xi_n\right)=\mathbb{P}(\{\omega\in\Omega:\xi_{n+1}(\omega)\in B\}|\xi_n) \tag{4}$$

è detta probabilità di transizione *al passo n-esimo.*

Inoltre, un nucleo di probabilità[2]

$$\mathcal{B}(\mathbb{R})\times\mathbb{R}\ni(B,x)\longmapsto P_{n+1}(x;B):=\mathbb{P}_{\xi_{n+1}}(B|\xi_n=x)\in[0,1] \tag{5}$$

da $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *in sé, che mappa la distribuzione di probabilità* $\mathbb{P}_{\xi_n}$ [3] *in* $\mathbb{P}_{\xi_{n+1}}$, *cioè tale che,* $\forall B\in\mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}}d\mathbb{P}_{\xi_n}(x)P_{n+1}(x;B)&=\int_{\mathbb{R}}d\mathbb{P}_{\xi_n}(x)\mathbb{P}_{\xi_{n+1}}(B|\xi_n=x)\\&=\mathbb{E}[\mathbb{P}(\xi_{n+1}(B)|\xi_n)]=\mathbb{P}_{\xi_{n+1}}(B)\quad,\end{aligned} \tag{6}$$

che sia una versione regolare di $\mathbb{P}(\xi_{n+1}(\cdot)|\xi_n)$, *ovvero tale che,* $\forall B\in\mathcal{B}(\mathbb{R})$,

$$\mathbb{P}\left(\xi_{n+1}^{-1}(B)|\xi_n\right)=P_{n+1}(\xi_n;B)=\int_B P_{n+1}(\xi_n;dx)\quad\mathbb{P}\text{-}q.c., \tag{7}$$

è detto funzione di transizione *al passo n-esimo.*

Quindi, $\forall n\geq m\geq0$ e per ogni f $\mathcal{B}(\mathbb{R})$-mis. limitata, la (2) si scrive

$$\begin{aligned}\mathbb{E}[f(\xi_n)|\xi_m]&=\mathbb{E}[f(\xi_n)|\mathcal{F}_m]=\mathbb{E}[\cdots\mathbb{E}[\mathbb{E}[f(\xi_n)|\mathcal{F}_{n-1}]|\mathcal{F}_{n-2}]\cdots|\mathcal{F}_m]\\&=\int_{\mathbb{R}}P_{m+1}(\xi_m;dx_{m+1})\cdots\int_{\mathbb{R}}P_n(x_{n-1};dx_n)f(x_n)\quad\mathbb{P}\text{-q.c.}.\end{aligned} \tag{8}$$

Definizione 4 *Sia* $\{\xi_n\}_{n\geq0}$ *su* $(\Omega,\mathcal{F},\mathbb{P})$ *una catena di Markov rispetto a* $\mathbb{P}$. *La misura di probabilità* $\pi:=\mathbb{P}_{\xi_0}$ *su* $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ *è detta* distribuzione iniziale *della catena.*

[1] cf. dispense par. 1.1 $\mathcal{F}_n^{\xi}=\sigma\{\xi_0,..,\xi_n\}:=\bigvee_{k=0}^{n}\xi_k^{-1}(\mathcal{B}(\mathbb{R}))$.
[2] cf. *Appunti del corso di Probabilità e Processi Stocastici* Definizione 80 e Osservazione 83.
[3] cf. *Appunti del corso di Probabilità e Processi Stocastici* par. 5.1.

Consideriamo una Catena di Markov $\xi := \{\xi_n\}$ come applicazione misurabile da $(\Omega, \mathcal{F})$ a $(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$ [4].

Proposizione 5 *Sia $\xi = \{\xi_n\}_{n\geq 0}$ su $(\Omega, \mathcal{F}, \mathbb{P})$ una catena di Markov rispetto a $\mathbb{P}$. Se $\{P_{n+1}(\cdot; dx)\}_{n\geq 0}$ e $\pi(dx)$ sono rispettivamente la famiglia delle funzioni di transizione associate alla catena e la sua distribuzione iniziale, la coppia $\left(\pi(ds), \{P_{n+1}(\cdot; ds)\}_{n\geq 0}\right)$ determina completamente la distribuzione di probabilità delle traiettorie della catena, cioè $\mathbb{P}_\xi$ su $(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$.*

Dimostrazione: E' sufficiente considerare il generico insieme cilindrico $C\left(\times_{i=0}^{n-1} B_i\right)$ dove, $\forall n \geq 0, B_i \in \mathcal{B}(\mathbb{R}), i = 0, .., n-1$, poiché $\mathcal{C}$ è generata dalla collezione di questi eventi. Allora, siccome

$$\begin{aligned} \{\omega \in \Omega : \xi_0(\omega) \in B_0, .., \xi_{n-1}(\omega) \in B_{n-1}\} = \\ \{\omega \in \Omega : \xi_0(\omega) \in B_0, .., \xi_{n-2}(\omega) \in B_{n-2}\} \cap \{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\} \end{aligned} \tag{9}$$

e $\{\omega \in \Omega : \xi_0(\omega) \in B_0, .., \xi_{n-2}(\omega) \in B_{n-2}\} \in \mathcal{F}_{n-2}$,

$$\begin{aligned} \mathbb{P}_\xi\left(C\left(\times_{i=0}^{n-1} B_i\right)\right) &= \mathbb{P}\{\omega \in \Omega : \xi_0(\omega) \in B_0, .., \xi_{n-1}(\omega) \in B_{n-1}\} \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbf{1}_{\{\omega\in\Omega:\xi_{n-1}(\omega)\in B_{n-1}\}}\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbf{1}_{\{\omega\in\Omega:\xi_{n-1}(\omega)\in B_{n-1}\}}|\mathcal{F}_{n-2}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_{n-1}(\omega)\in B_{n-1}\}}|\mathcal{F}_{n-2}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)\right] \end{aligned} \tag{10}$$

Ripetendo la procedura scrivedo l'ultima espressione dell'equazione precedente come valore atteso dell'attesa condizionata rispetto a $\mathcal{F}_{n-3}$, dato che $\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)$ è indipendente da ξ_{n-3}, per la definizione di catena di Markov, si ha

$$\begin{aligned} &\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)|\mathcal{F}_{n-3}\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_{n-2}(\omega)\in B_{n-2}\}}\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)|\xi_{n-3}\right] \\ &= \mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_{n-2}(\omega)\in B_{n-2}\}}|\xi_{n-3}\right] \\ &= \mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-2}(\omega) \in B_{n-2}\}|\xi_{n-3}\right) \ , \end{aligned} \tag{11}$$

ovvero

$$\begin{aligned} &\mathbb{P}\{\omega \in \Omega : \xi_0(\omega) \in B_0, .., \xi_{n-1}(\omega) \in B_{n-1}\} \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-3}(\omega)\in B_{n-3}\}}\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_{n-2}(\omega)\in B_{n-2}\}}\times\right.\right. \\ &\left.\left.\times\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)|\mathcal{F}_{n-3}\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-3}(\omega)\in B_{n-3}\}}\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-2}(\omega) \in B_{n-2}\}|\xi_{n-3}\right)\times\right. \\ &\left.\times\mathbb{P}\left(\{\omega \in \Omega : \xi_{n-1}(\omega) \in B_{n-1}\}|\xi_{n-2}\right)\right] \ . \end{aligned} \tag{12}$$

[4] Per la definizione dello spazio di misura $(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$ cf. *Leggi 0-1, successioni di v.a. stazionarie in senso stretto ed introduzione alla teoria ergodica.*

Iterando, si ha

$$\begin{aligned}
&\mathbb{P}\{\omega\in\Omega:\xi_0(\omega)\in B_0,..,\xi_{n-1}(\omega)\in B_{n-1}\} \qquad (13)\\
&=\mathbb{E}\left[\mathbf{1}_{\{\omega\in\Omega:\xi_0(\omega)\in B_0\}}\prod_{k=2}^{n}\mathbb{P}\left(\{\omega\in\Omega:\xi_{n-k+1}(\omega)\in B_{n-k+1}\}\,|\xi_{n-k}\right)\right]\\
&=\int_{\{\omega'\in\Omega\,:\,\xi_0(\omega')\in B_0\}}d\mathbb{P}(\omega)\prod_{k=2}^{n}\mathbb{P}\left(\{\omega'\in\Omega:\xi_{n-k+1}(\omega')\in B_{n-k+1}\}\,|\xi_{n-k}\right)(\omega)\ ,
\end{aligned}$$

cioè, per la (6) e per la definizione di distribuzione iniziale della catena,

$$\mathbb{P}_\xi\left(C\left(\times_{i=0}^{n-1}B_i\right)\right)=\int_{B_0}\pi(dx_0)\int_{B_1}P_1(x_0;dx_1)\cdots\int_{B_{n-1}}P_n(x_{n-1};B_n) \qquad (14)$$

e la tesi segue dal teorema di Kolmogorov sull'estensione di una misura su $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C}\right)$. ■

Viceversa, una distribuzione di probabilità $\mathbb{P}$ su $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C}\right)$ della forma 14 definisce una catena di Markov $\xi=\{\xi_n\}_{n\geq0}$ sullo spazio di probabilità filtrato $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C},\{\mathcal{F}_n\}_{n\geq0},\mathbb{P}\right)$ tale che, $\forall n\geq0,\omega\in\mathbb{R}^{\mathbb{N}},\xi_n(\omega)=\omega_n,\mathcal{F}_n:=\bigvee_{k=0}^{n}\xi_k^{-1}(\mathcal{B}(\mathbb{R}))$ e, $\forall B\in\mathcal{B}(\mathbb{R})$,

$$\mathbb{P}\left(\left\{\omega\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(\omega)\in B\right\}|\xi_n\right)=\mathbb{P}\left(\left\{\omega\in\mathbb{R}^{\mathbb{N}}:\omega_{n+1}\in B\right\}|\xi_n\right)=P_{n+1}(\cdot;B)\quad\mathbb{P}\text{-q.c. .} \qquad (15)$$

Infatti, vale il seguente risultato.

Proposizione 6 *Sia* $(\Omega,\mathcal{F})=\left(\mathbb{R}^{\mathbb{N}},\mathcal{C}\right)$ *e* $\mathbb{P}$ *la misura di probabilità su* $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C}\right)$ *tale che,* $\forall n\geq0,\ B\in\mathcal{B}\left(\mathbb{R}^{n+1}\right)$,

$$\begin{aligned}
\mathbb{P}(C(B))&=\mathbb{P}\left\{x\in\mathbb{R}^{\mathbb{N}}:(x_0,..,x_n)\in B\right\} \qquad (16)\\
&=\int_{\mathbb{R}}\pi(dx_0)\int_{\mathbb{R}}P_1(x_0;dx_1)\cdots\int_{\mathbb{R}}P_n(x_{n-1};dx_n)\,\mathbf{1}_B(x_0,..,x_n)\ .
\end{aligned}$$

Allora, esiste una catena di Markov $\{\xi_n\}_{n\geq0}$ *su* $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C},\mathbb{P}\right)$ *di distribuzione iniziale* $\pi(dx)$ *e famiglia di funzioni di transizione* $\{P_{n+1}(\cdot;dx)\}_{n\geq0}$.

Dimostrazione: Sia $\xi=\{\xi_n\}_{n\geq0}$ la successione di v.a. su $\left(\mathbb{R}^{\mathbb{N}},\mathcal{C}\right)$ tale che, $\forall n\geq0,x\in\mathbb{R}^{\mathbb{N}},\xi_n(x):=x_n$ e sia $\mathcal{F}_n:=\mathcal{F}_n^{\xi}$. Allora, per ipotesi, $\forall n\geq0,B\in\mathcal{B}\left(\mathbb{R}^{n+1}\right)$

$$\begin{aligned}
\mathbb{P}\left\{x\in\mathbb{R}^{\mathbb{N}}:(\xi_0(x),..,\xi_n(x))\in B\right\}&=\mathbb{P}\left\{x\in\mathbb{R}^{\mathbb{N}}:(x_0,..,x_n)\in B\right\} \qquad (17)\\
&=\int_{\mathbb{R}}\pi(dx_0)\int_{\mathbb{R}}P_1(x_0;dx_1)\cdots\int_{\mathbb{R}}P_n(x_{n-1};dx_n)\,\mathbf{1}_B(x_0,..,x_n)\ .
\end{aligned}$$

Inoltre, dato $B'\in\mathcal{B}(\mathbb{R})$, da (6) segue che

$$\begin{aligned}
&\mathbb{P}\left\{x\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(x)\in B',(\xi_0(x),..,\xi_n(x))\in B\right\}=\\
&\int_{\mathbb{R}}\pi(dx_0)\int_{\mathbb{R}}P_1(x_0;dx_1)\cdots\int_{\mathbb{R}}P_n(x_{n-1};dx_n)\,\mathbf{1}_B(x_0,..,x_n)\int_{B'}P_{n+1}(x_n;dx_{n+1})\\
&\qquad=\mathbb{E}\left[\mathbf{1}_{\{x\in\mathbb{R}^{\mathbb{N}}\,:\,(x_0,..,x_n)\in B\}}((\xi_0,..,\xi_n))\,P_{n+1}(\xi_n;B')\right]\ ,
\end{aligned}$$

ovvero, per la definizione di probabilità condizionata,

$$\begin{aligned}\mathbb{P}\left(\left\{x\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(x)\in B'\right\}|\xi_n,...,\xi_0\right) &= \mathbb{P}\left(\left\{x\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(x)\in B'\right\}|\mathcal{F}_n\right) \\ &= P_{n+1}(\xi_n;B') \quad \mathbb{P}\text{-q.c. .}\end{aligned} \tag{18}$$

D'altra parte però, $\forall B''\in\mathcal{B}(\mathbb{R})$,

$$\begin{aligned}&\mathbb{P}\left\{x\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(x)\in B',\xi_n(x)\in B''\right\} = \\ &\quad\int_{\mathbb{R}}\pi(dx_0)\int_{\mathbb{R}}P_1(x_0;dx_1)\cdots\int_{\mathbb{R}}P_n(x_{n-1};dx_n)\,\mathbf{1}_{B''}(x_n)\int_{B'}P_{n+1}(x_n;dx_{n+1}) \\ &\quad= \mathbb{E}\left[\mathbf{1}_{\{x\in\mathbb{R}^{\mathbb{N}}\,:\,x_n\in B''\}}(\xi_n)\,P_{n+1}(\xi_n;B')\right] \ ,\end{aligned}$$

cioè

$$\mathbb{P}\left(\left\{x\in\mathbb{R}^{\mathbb{N}}:\xi_{n+1}(x)\in B'\right\}|\xi_n\right) = P_{n+1}(\xi_n;B') \quad \mathbb{P}\text{-q.c. .} \tag{19}$$

da cui segue la (1). ■

1 Catene di Markov omogenee

Definizione 7 *Una catena di Markov che ammette una descrizione regolare della probabilità di transizione ad ogni passo si dice* omogenea *se esiste una funzione di transizione $P(\cdot;dx)$ tale che $\forall n\geq 0, P_{n+1}(\cdot;dx)=P(\cdot;dx)$, ovvero la probabilità di transizione ad ogni passo è identica a quella del primo.*

Se s'interpreta una catena di Markov come l'evoluzione di un sistema reale, la condizione di omogeneità della catena è equivalente ad affermare che l'evoluzione del sistema da un dato istante di tempo a quello successivo non dipende esplicitamente dal tempo in distribuzione. Infatti, se $\xi=\{\xi_n\}_{n\geq 0}$, è una catena di Markov omogenea con funzione di transizione $P, \forall B\in\mathcal{B}(\mathbb{R})$, e $\forall n\geq k\geq m\geq 0$, dalla (14) si ha

$$\begin{aligned}\mathbb{P}(\{\omega\in\Omega:\xi_n(\omega)\in B\}|\xi_m) &= \mathbb{P}(\{\omega\in\Omega:\xi_{n-m}(\omega)\in B\}|\xi_0) \\ &= P^{(n-m)}(\xi_0;B) = \int_{\mathbb{R}}P^{(k-m)}(\xi_0;dx)\,P^{(n-k)}(x;B)\end{aligned} \tag{20}$$

dove $P^{(k)}$ è il nucleo della potenza k-sima dell'operatore integrale associato a P, ovvero

$$P^{(k)}(x;dy) := \int_{\mathbb{R}}P(x;dx_1)\cdots\int_{\mathbb{R}}P(x_{k-1};dy) \ . \tag{21}$$

La (20) è detta *equazione di Chapman-Kolmogorov.*

Questo è quindi l'analogo stocastico della condizione di autonomia, cioè di non dipendenza dal tempo, dell'endomorfismo che determina l'evoluzione di un sistema deterministico.

1.0.1 Convergenza

Proposizione 8 *Se esiste una distribuzione di probabilità su ν su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tale che*

$$\int_{\mathbb{R}} \nu(dx) P(x; dy) = \nu(dy) \ , \tag{22}$$

allora la catena di Markov omogenea $\xi = \{\xi_n\}_{n\geq 0}$ determinata dalla distribuzione iniziale ν e dalla funzione di transizione P per cui vale la (22) è detta stazionaria, *ovvero è una successione di v.a. stazionaria in senso stretto*[5].

Dimostrazione: Infatti, analogamente alla (20), $\forall n \geq 0, B_i \in \mathcal{B}(\mathbb{R}), \ i = 1, .., n,$ si ha

$$\begin{aligned}
&\mathbb{P}(\{\omega \in \Omega : \xi(\omega) \in C((\times_{i=1}^n B_i))\} |\xi_0) \qquad (23)\\
&= \mathbb{P}(\{\omega \in \Omega : \xi_n(\omega) \in B_n, .., \xi_1(\omega) \in B_1\} |\xi_0)\\
&= \int_{B_1} P(\xi_0; dx_1) \cdots \int_{B_{n-1}} P(x_{n-1}; B_n)\\
&= \int_{B_1} P(\xi_1; dx_2) \cdots \int_{B_{n-1}} P(x_n; B_n)\\
&= \mathbb{P}(\{\omega \in \Omega : \xi_{n+1}(\omega) \in B_n, .., \xi_2(\omega) \in B_1\} |\xi_1)\\
&= \mathbb{P}(\{\omega \in \Omega : \mathbf{S}\xi(\omega) \in C((\times_{i=1}^n B_i))\} |\xi_1) \ .
\end{aligned}$$

Integrando la prima e l'ultima espressione rispetto a ν su $B_0 \in \mathcal{B}(\mathbb{R})$ si ha

$$\mathbb{P}\{\omega \in \Omega : \xi(\omega) \in C((\times_{i=0}^n B_i))\} = \mathbb{P}\{\omega \in \Omega : \mathbf{S}\xi(\omega) \in C((\times_{i=0}^n B_i))\} \tag{24}$$

la quale implica che ξ è stazionaria. ■

Consideriamo la catena di Markov omogenea $\xi = \{\xi_n\}_{n\geq 0}$, descritta dalla funzione di transizione P il cui valore iniziale ξ_0 è fissato, ovvero consideriamo su $(\mathbb{R}^{\mathbb{N}}, \mathcal{C})$ la distribuzione di probabilità $\mathbb{P}^{\xi_0}$ tale che $\forall n \geq 1, B \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}^{\xi_0}\{\omega \in \Omega : \xi_0(\omega) \in B\} = \mathbf{1}_B(\xi_0) \ , \tag{25}$$

$$\mathbb{P}^{\xi_0}\{\omega \in \Omega : \xi_n(\omega) \in B\} = P^{(n)}(\xi_0; B) \ , \tag{26}$$

cioè se ξ_0 è fissato uguale a $x \in \mathbb{R}$,

$$\mathbb{P}^x := \mathbb{P}^{\xi_0} = \mathbb{E}[\cdot|\xi_0 = x] \tag{27}$$

Come dimostra il seguente risultato, se $\{\xi_n\}_{n\geq 0}$ converge in distribuzione ad una v.a. ξ_∞ percui $\mathbb{P}_{\xi_\infty}$ è indipendente dal valore assunto dalla condizione iniziale ξ_0, e soddisfa la (22) e la successione $\{\mathbf{S}^n \xi\}_{n\geq 0} \subseteq \mathbb{R}^{\mathbb{N}}$ tale che $\mathbf{S}^0\xi := \xi$ converge in distribuzione alla catena di Markov stazionaria $\eta = \{\eta_n\}$ descritta dalla distribuzione iniziale $\nu = \mathbb{P}_{\xi_\infty}$ e dalla funzione di transizione P.

[5] cf. note Leggi 0-1 par. 3 def. 11.

Proposizione 9 *Se per ogni valore assunto dalla condizione iniziale ξ_0 la catena di Markov omogenea $\xi = \{\xi_n\}_{n\geq 0}$ descritta dalla distribuzione di probabilità data dalle (25) e (26) associata alla funzione di transizione P converge in distribuzione ad una v.a. ξ_∞ tale che $\nu := \mathbb{P}_{\xi_\infty}$ verifica la (22), allora $\forall B \in \mathcal{C}, \{\mathbf{S}^n\xi\}_{n\geq 0}$ converge in distribuzione alla catena di Markov stazionaria avente ν come distribuzione iniziale e P come funzione di transizione.*

Dimostrazione: Sia ν una distribuzione di probabilità su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ che soddisfa la (22) e, $\forall x \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})$ sia

$$\lim_{n\to\infty} \mathbb{P}^x \{\omega \in \Omega : \xi_n(\omega) \in B\} = \lim_{n\to\infty} P^{(n)}(x; B) = \nu(B) \ . \tag{28}$$

Poniamo inoltre $\forall A \in \mathcal{C}$

$$\phi(x) := \mathbb{P}^x \{\omega \in \Omega : \mathbf{S}\xi(\omega) \in A\} \ . \tag{29}$$

Allora, se $\eta = \{\eta_n\}$ è la catena di Markov omogenea descritta dalla distribuzione iniziale ν e dalla funzione di transizione P, dalla Proposizione 5 si ha

$$\begin{aligned}\mathbb{E}\left[\mathbb{P}^x\left\{\omega \in \Omega : \mathbf{S}^{n+1}\xi(\omega) \in A\right\}|\xi_n\right] &= \mathbb{E}\left[\mathbb{P}\left\{\omega \in \Omega : \mathbf{S}^{n+1}\eta(\omega) \in A\right\}|\eta_n\right] \\ &= \mathbb{E}\left[\mathbb{P}\left\{\omega \in \Omega : \mathbf{S}\eta(\omega) \in A\right\}|\eta_0\right]\end{aligned} \tag{30}$$

quindi

$$\mathbb{E}\left[\mathbb{P}\left\{\omega \in \Omega : \mathbf{S}\eta(\omega) \in A\right\}|\eta_0 = x\right] = \mathbb{P}^x\left\{\omega \in \Omega : \mathbf{S}\xi(\omega) \in A\right\} = \phi(x) \quad \nu\text{-q.c.} \ . \tag{31}$$

Dunque,

$$\begin{aligned}\mathbb{P}^x\left\{\omega \in \Omega : \mathbf{S}^{n+1}\xi(\omega) \in A\right\} &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{P}^x\left\{\omega \in \Omega : \mathbf{S}^{n+1}\xi(\omega) \in A\right\}|\xi_n\right]\right] \qquad (32)\\ &= \int_{\mathbb{R}} P^{(n)}(x; dy)\, \mathbb{E}\left[\mathbb{P}^{\xi_0}\left\{\omega \in \Omega : \mathbf{S}^{n+1}\xi(\omega) \in A\right\}|\xi_n = y\right] \\ &= \int_{\mathbb{R}} P^{(n)}(x; dy)\, \phi(y) \ . \qquad (33)\end{aligned}$$

Poiché $\phi \in [0,1]$ ed è $\mathcal{B}(\mathbb{R})$-mis. esiste una successione $\{\phi_n\}_{n\geq 1}$ di funzioni semplici che converge uniformemente a ϕ. Allora, $\forall m \geq 1, \lim_{n\to\infty} \int_{\mathbb{R}} P^{(n)}(x; dy)\, \phi_m(y) = \int_{\mathbb{R}} \nu(dy)\, \phi_m(y)$ quindi, per la (30),

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}^x\left\{\omega \in \Omega : \mathbf{S}^{n+1}\xi(\omega) \in A\right\} &= \int_{\mathbb{R}} \nu(dy)\, \phi(y) \\ &= \int_{\mathbb{R}} \nu(dy)\, \mathbb{E}\left[\mathbb{P}\left\{\omega \in \Omega : \mathbf{S}\eta(\omega) \in A\right\}|\eta_0\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\mathbb{P}\left\{\omega \in \Omega : \mathbf{S}\eta(\omega) \in A\right\}|\eta_0\right]\right] \\ &= \mathbb{P}\left\{\omega \in \Omega : \eta(\omega) \in A\right\} \ ,\end{aligned} \tag{34}$$

cioè la tesi. ■

Pertanto il problema di determinare il comportamento asintotico di una catena di Markov omogenea ξ descritta dalla funzione di transizione P è ridotto a determinare:

- Se $\forall B \in \mathcal{B}(\mathbb{R}), x \in \mathbb{R}, \lim_{n\to\infty} P^{(n)}(x;B)$ esiste ed è indipendente da x;
- Quante sono le *distribuzioni iniziali stazionarie* per ξ, cioè le distribuzioni di probabilità ν su $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ che soddisfano la (22) che non sono a.c. l'una rispetto all'altra; infatti, qualora ce ne fosse più di una, una qualsiasi combinazione convessa sarebbe ancora soluzione della (22).

Definizione 10 *Sia $\xi = \{\xi_n\}_{n\geq 0}$ una catena di Markov omogenea descritta da una funzione di transizione P. Un evento $A \in \mathcal{B}(\mathbb{R})$ è detto* chiuso per ξ *se $\forall x \in A, P(x;A) = 1$.*

Definizione 11 *Una catena di Markov omogenea $\xi = \{\xi_n\}_{n\geq 0}$ descritta da una funzione di transizione P, è detta* indecomponibile *se non esistono due eventi disgiunti chiusi per ξ.*

E' chiaro che qualora esistessero A_1, A_2 chiusi per ξ, allora, $\forall n \geq 1$,

$$P^{(n)}(x;A_i) = \mathbf{1}_{A_i}(x), \; i = 1,2 \;, \tag{35}$$

cioè $\left\{P^{(n)}(x;A_i)\right\}_{n\geq 0}$ non convergerebbe. Tuttavia potrebbe esistere una distribuzione di probabilità ν che soddisfa (22). Qualora la catena di Markov stazionaria $\eta = \{\eta_n\}_{n\geq 0}$ descritta dalla distribuzione iniziale ν e dalla funzione di transizione P sia ergodica, allora, $\forall B \in \mathcal{B}(\mathbb{R})$, per il teorema ergodico per successioni di v.a. stazionarie, la successione delle medie empiriche $\left\{\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_B(\eta_i)\right\}_{n\geq 1}$ converge a $\nu(B)$ ν-q.c., da cui segue, per il teorema di Lebesgue della convergenza monotona, che, $\forall x \in \mathbb{R}$,

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} P^{(n)}(x;B) &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mathbf{1}_B(\eta_i)\,|\eta_0 = x\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_B(\eta_i)\right|\left.\eta_0 = x\right] \\
&= \mathbb{E}\left[\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_B(\eta_i)\right|\left.\eta_0 = x\right] = \nu(B) \quad \nu\text{-q.c. .}
\end{aligned} \tag{36}$$

Teorema 12 *Se la catena di Markov omogenea $\xi = \{\xi_n\}_{n\geq 0}$ descritta da una funzione di transizione P è indecomponibile e ammette una distribuzione iniziale stazionaria ν, allora questa è unica e la catena di Markov stazionaria $\eta = \{\eta_n\}_{n\geq 0}$ descritta dalla distribuzione iniziale ν e dalla funzione di transizione P è ergodica.*

Dimostrazione: Sia ν una distribuzione iniziale stazionaria per $\xi, \eta = \{\eta_n\}_{n\geq 0}$ la catena di Markov stazionaria descritta dalla distribuzione iniziale ν e dalla funzione di transizione P e $\mathbb{P}$ la distribuzione di probabilità su $(\Omega, \mathcal{F})$ associata a ν e P. Sia inoltre A un evento invariante relativamente a η, ovvero tale che $\exists B \in \mathcal{C}$ tale che $\forall n \geq 0, A = A_{\mathbf{S}^{-n}B}$[6]. Posto $\phi(\eta_0) := \mathbb{E}(\mathbf{1}_A|\eta_0)$, poiché

$$\mathbb{P}(A|\eta_0) = \mathbb{E}[\mathbb{P}(A|\mathcal{F}_1(\eta))|\eta_0] \ , \tag{37}$$

per la proprietà di Markov, siccome $\forall n \geq 0, A \in \mathcal{F}^n(\eta)$ (la σalgebra generata dalla successione $\mathbf{S}^n\eta$)[7],

$$\mathbb{P}(A|\mathcal{F}_1(\eta)) = \mathbb{P}(A|\eta_1) \quad \mathbb{P}\text{-q.c.} \ , \tag{38}$$

ma, per la Proposizione 8,

$$\mathbb{P}(A|\eta_0 = x) = \mathbb{P}(A_{\mathbf{S}^{-1}B}|\eta_1 = x) \quad \nu(dx)\text{-q.c.} \tag{39}$$

e per l'invarianza di A,

$$\mathbb{P}(A_{\mathbf{S}^{-1}B}|\eta_1) = \mathbb{P}(A|\eta_1) = \phi \ . \tag{40}$$

Perciò,

$$\phi(x) = \mathbb{E}[\mathbb{P}(A|\eta_1)|\eta_0 = x] = \mathbb{E}_\eta[\phi|\eta_0 = x] = \int_{\mathbb{R}} P(x;dy)\phi(y) \ , \quad \nu(dx)\text{-q.c.} \ . \tag{41}$$

Allo stesso modo si ha

$$\mathbb{P}(A|\mathcal{F}_n(\eta)) = \mathbb{P}(A|\eta_n) = \mathbb{P}(A_{\mathbf{S}^{-n}B}|\eta_n) = \mathbb{P}(A|\eta_0) = \phi \ , \quad \mathbb{P}\text{-q.c.} \tag{42}$$

e

$$\begin{aligned}\mathbb{E}[\phi(\eta_n)|\mathcal{F}_{n-1}(\eta)] &= \mathbb{E}[\phi(\eta_n)|\eta_{n-1}] = \mathbb{E}[\mathbb{P}(A|\eta_n)|\eta_{n-1}] \\ &= \mathbb{P}(A|\eta_{n-1}) = \phi(\eta_{n-1}) = \mathbb{E}[\mathbf{1}_A|\mathcal{F}_{n-1}(\eta)] \ , \quad \mathbb{P}\text{-q.c.} \ ,\end{aligned} \tag{43}$$

cioè, essendo $\phi \in [0,1]$, la successione $\{\phi(\eta_n)\}_{n\geq 0}$ è una martingala regolare che quindi, converge a $\mathbf{1}_A$ $\mathbb{P}$-q.c. Quindi, poiché, $\forall n \geq 0$, la distribuzione di η_n è $\nu, \forall \varepsilon \in (0,1)$

$$\nu\{x \in \mathbb{R} : \phi(x) \in [\varepsilon, 1-\varepsilon]\} = 0 \tag{44}$$

quindi ν-q.c. $\phi \in \{0,1\}$. Poniamo

$$C_0 := \{x \in \mathbb{R} : \phi(x) = 0\} \ ; \ C_1 := \{x \in \mathbb{R} : \phi(x) = 1\} \ . \tag{45}$$

Allora, eccetto eventualmente sull'evento $D \in \mathcal{B}(\mathbb{R})$ tale che $\nu(D) = 0$,

$$0 = \int_{\mathbb{R}} P(x;dy)\phi(y) \ , \quad x \in C_0 \tag{46}$$

$$1 = \int_{\mathbb{R}} P(x;dy)\phi(y) \ , \quad x \in C_1 \tag{47}$$

[6]cf. Note Leggi 0-1 sezione 3 Definizione 9.
Poniamo $\mathbf{S}^0$ uguale all'applicazione identità su $\mathbb{R}^{\mathbb{N}}$.
[7]cf. Note Leggi 0-1 sezione 3 Proposizione 13

Pertanto,

$$P\left(x;C_i\right)=1\ ,\quad x\in C_i\backslash D\ ,\ i=0,1\ . \tag{48}$$

Definiamo $\forall i=0,1$

$$C_i^{(0)}:=C_i \tag{49}$$

$$C_i^{(1)}:=C_i^{(0)}\backslash D \tag{50}$$

$$C_i^{(n+1)}:=C_i^{(n)}\cap\left\{x\in\mathbb{R}:P\left(x;C_i^{(n)}\right)=1\right\}\ . \tag{51}$$

Se $P\left(x;C_i^{(n)}\right)=1$ allora, $x\in C_i^{(n)}$ ν-q.c., quindi $\nu\left(C_i^{(n)}\backslash C_i^{(n+1)}\right)=0$. Pertanto, ponendo $A_i:=\bigcap\limits_{n\geq 0}C_i^{(n)}$, si ha $\nu\left(A_i\right)=\nu\left(C_i\right)$, ma gli eventi C_0,C_1 sono chiusi e disgiunti, dunque, per l'ipotesi di indecomponibilità: uno dei due è vuoto, ϕ vale uno o zero ν-q.c., $\mathbb{P}\left(A\right)$ è uguale a zero o uno e η è ergodica. L'unicità segue dal fatto che se esistessero ν_1,ν_2 tali che valga per ciascuna quanto appena affermato, si avrebbe necessariamente che se $\nu_1\left(A\right)=1$ allora $\nu\left(A\right)=0$. Allora, considerando la misura di probabilità $\mathbb{Q}$ su $(\Omega,\mathcal{F})$ associata catena di Markov descritta dalla funzione di transizione P e dalla distribuzione $\frac{1}{2}\nu_1+\frac{1}{2}\nu_2$, che è stazionaria in quanto verifica la (22), così come ogni combinazione convessa di due distribuzioni di probabilità stazionarie e quindi, per quanto appena mostrato, ergodica, si avrebbe $\mathbb{Q}\left(A\right)=\frac{1}{2}$ invece che 0 o 1. ■