Probabilistic construction of continuous-time Markov chains

P.-F. Rodriguez

Recall our definition of a continuous-time Markov chain (CTMC - really: pure jump process with no explosion)

Def A CTMC on a countable state space S is a collection $\{P_x\}_{x\in S}$ of prob. measures on Ω, the space of right-continuous S-valued trajectories with finitely many jumps in finite intervals (endowed with canonical coordinates $X_t : \Omega \to S$, $\omega \mapsto \omega(t)$, canonical σ-algebra $\mathcal{F}$, filtration $(\mathcal{F}_t)_{t\geq 0}$, and shifts $(\theta_t)_{t\geq 0}$) such that

(1) $P_x[X_0 = x] = 1$

(2) $E_x[f(X_{t+h}) \mid \mathcal{F}_t] \overset{P_x\text{-a.s.}}{=} E_{X_t}[f(X_h)]$ $\quad \forall x\in S,\ t,h\geq 0$, $f: S\to \mathbb{R}$ bded.

The goal of this note is to prove the following existence result:

Theorem Suppose that $c(\cdot)$ and $(p_{x,y})_{x,y\in S}$ satisfy

(3) $c: S \to [0,\infty)$

(4) $p_{x,y}\geq 0 \ \forall x,y\in S$, $\sum_{y\in S} p_{x,y} = 1$, $x\in S$ and $p_{x,x} = \begin{cases} 1, & \text{if } c(x)=0 \\ 0, & \text{if } c(x)>0 \end{cases}$

and that

(5) $\sum_{n\geq 0} c(Z_n)^{-1} = \infty$, P_x^Z-a.s., for $x\in S$

where $\{Z_n ; n\geq 0\}$ is the unique discrete-time Markov chain with transition probabilities $(p_{x,y})_{x,y\in S}$, and P_x^Z stands for its canonical law started at x.

Then, there exists a CTMC (as defined above) with transition rates $c(x)$, $x\in S$ and transition probabilities $(p_{x,y})_{x,y\in S}$, i.e., such that

(6) $P_x(T>t) = e^{-c(x)t}$, for $x\in S$, $t\geq 0$
(understood as: $T=\infty$, P_x-a.s. if $c(x)=0$)

(7) $P_x[X_T = y] = p_{x,y}$, for $x,y\in S$

where $T = \inf\{s\geq 0,\ X_s \neq X_0\}$ (with the convention $\inf\emptyset = \infty$) denotes the time of the first jump.

Remark(s)

1) Assumption (5) prevents explosions (i.e. accumulations of jumps in finite time). It is in fact equivalent to requiring

(8) $$\lim_{n\to\infty} T_n = \infty, \quad \tilde{P}_x\text{-a.s.}$$ (with $\{T_n, n\geq 1\}$, $\tilde{P}_x$ as, as defined below in (11), (12)).

For a proof of this equivalence, see for instance [L] Thm 2.33, p.75. It is however immediately clear that (3), (4) alone are not sufficient in general (a notable exception being $c(x) = \lambda \in [0,\infty)$, all x, see 2) below) to guarantee existence. To see this, consider the following

Example: $S = \mathbb{N}$, $p_{k,k+1} = 1$ $\forall k \in \mathbb{N}$, $c(k) = (k+1)^2$

Then, defining $\tau_k \sim \mathrm{Exp}(c(k))$, indep., on some $(\Omega, \mathcal{A}, P)$ and $N_t = \sup\{n \geq 1 : \sum_{k=1}^{n} \tau_k \leq t\}$, $t \geq 0$ ($\sup\emptyset = 0$)

we see that $\{N_t : t \geq 0\}$ escapes to infinity (in finite time):

$$P(N_t = \infty \text{ for some } t \in [0,\infty)) = P\Big(\sum_{k=1}^{\infty} \tau_k < \infty\Big) = 1 \quad \text{since } \sum_{k\geq 1} E[\tau_k] = \sum_{k\geq 1} \frac{1}{(k+1)^2} < \infty,$$

$$\sum_{k\geq 1} \mathrm{var}(\tau_k) = \sum_{k\geq 1} \frac{1}{(k+1)^4} < \infty$$ (see e.g. [Du] Thm. 2.5.3.).

2) As discussed in class, a drastic simplification occurs in the homogenous case ($c(x) = \lambda \in [0,\infty)$, all x). Then, the spatial and temporal dynamics decouple, and one simply sets $P_x = (P_x^Z \otimes P) \circ \tilde{X}_\cdot^{-1}$, where

$\tilde{X}_t = Z_{N_t}$ with Z the discrete time chain with trans. prob. $(p_{x,y})_{x,y\in S}$, and canon. law P_x^Z (s.t. $P_x^Z(Z_0 = x) = 1$) and $(N_t)_{t\geq 0}$ a Poisson process of rate λ under P).

(for $\lambda > 0$)

3) As further discussed in class, the CTMC constructed in the above Thm is necessarily unique.

4) Condition (5) is easily seen to hold if for instance $\sup_{x\in S} c(x) < \infty$, or if Z is irreducible and recurrent. $\square$

Proof of the Theorem

We will assume that $c(x) > 0$ for all $x \in S$ (no absorbing states). The proof in the general case is similar, albeit notationally more cumbersome, and left as an exercise. We consider, on an auxiliary prob. space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$, a family $\{\tau_n(x), n \geq 0, x \in S\}$ of <u>independent</u> RV's satisfying

(9) $$\tilde{P}(\tau_n(x) > t) = e^{-c(x)t}, \quad t \geq 0,$$

as well as, on $(\Omega^Z, \mathcal{A}^Z, P^Z_x)$, $x \in S$,

(10) $\{Z_n;\, n \geq 0\}$ the canonical discrete-time MC with trans prob. $(p_{x,y})_{x,y \in S}$

and write

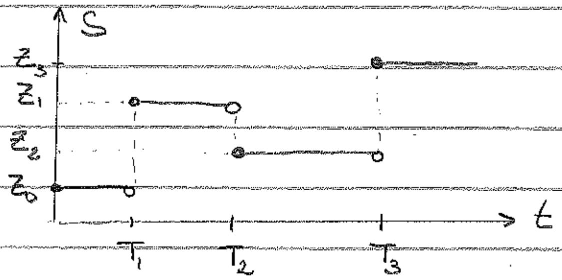
(11) $$\tilde{P}_x \overset{\text{def.}}{=} P^Z_x \otimes \tilde{P} \quad (\text{on } (\Omega^Z \times \tilde{\Omega}, \mathcal{A}^Z \times \tilde{\mathcal{A}})).$$

We then define the random variables (under $\tilde{P}_x$)

(12) $$T_0 = 0, \quad T_1 = \tau_0(Z_0), \ldots, T_n = \sum_{k=0}^{n-1} \tau_k(Z_k), \quad n \geq 1,$$

and set

(13) $$\tilde{X}_t = \sum_{n \geq 0} Z_n \, 1\{T_n \leq t < T_{n+1}\}, \quad \text{for } t \geq 0.$$



We will show that $\{P_x\}_{x \in S}$ with

(14) $$P_x = \tilde{P}_x \circ (\tilde{X}_\cdot)^{-1} \quad (\text{the law of } (\tilde{X}_t)_{t \geq 0} \text{ under } \tilde{P}_x)$$

has the desired properties. First, by (5) and Remark 1), $\tilde{X}_\cdot$ maps into Ω (cf above (1)). Moreover,

$$P_x[X_0 = x] \overset{(14)}{=} \tilde{P}_x[\tilde{X}_0 = x] \overset{(13)}{=} P^Z_x[Z_0 = x] = 1,$$

so (1) holds, and

$$P_x[T > t] \overset{(13)}{=} \tilde{P}_x[T_1 > t] \overset{(12)}{=} \tilde{P}_x[\tau_0(\underset{\substack{\| \\ x,\; P^Z_x\text{-a.s.}}}{Z_0}) > t] \overset{(9)}{=} e^{-c(x)t},$$

for $t \geq 0$, i.e. (6) holds, and (7) follows similarly. It thus remains to show $\{P_x\}_{x \in S}$ in (14) satisfy (2), which is the main task.

Using Dynkin's π-λ-theorem, we see that (2) follows if we show

$$E_x\left[f_0(X_{s_0})\cdots f_k(X_{s_k})\, f(X_{s_k+h})\right]$$

$$(15)\qquad = E_x\left[f_0(X_{s_0})\cdots f_k(X_{s_k})\, E_{X_{s_k}}\left[f(X_h)\right]\right]$$

$\forall x\in S$, $0\le s_0<\dots<s_k$, $h>0$, $f_i : S\to\mathbb{R}$ bded.

Let $N_t=\sup\{n\ge 0,\ T_n\le t\}$ (so that $\tilde{X}_t = Z_{N_t}$, see (13)), then with $C=\{N_{s_i}=n_i,\ i=0,\dots,k,\ N_{s_k+h}=n_k+n\}$ the left-hand side of (15) can be recast as

$$(16)\qquad LHS(15) \overset{(14)}{=} \tilde{E}_x\left[f_0(Z_{N_{s_0}})\cdots f_k(Z_{N_{s_k}})\, f(Z_{N_{s_k+h}})\right]$$

$$= \sum_{\substack{0\le n_0\le n_1\le\cdots\le n_k\\ n\ge 0}} \tilde{E}_x\Big[\underbrace{\prod_{i=0}^{k} f_i(Z_{n_i})\, f(Z_{n_k+n})}_{\overset{def}{=}F}\, 1_C\Big]$$

The key is now to note that, conditionally on the Z's, the dependence structure of the RV's N_{s_i} appearing in C is rather explicit, cf. (12): namely, cond on $Z_0, Z_1, \dots, Z_{n_k+n}$, the RV's $\tau_i(Z_i)$, $0\le i\le n_k+n$ are indep., see (12) and (9) and $\tau_i(Z_i)\sim \mathrm{Exp}(\lambda(Z_i))$, $0\le i\le n_k+n$. Hence, we can write

$$(17)\qquad \tilde{E}_x\left[1_C \mid \sigma(Z_0,\dots,Z_{n_k+n})\right] = h(Z_0,\dots,Z_{n_k+n})$$

where, for $x_0,\dots,x_{n_k+n}\in S$,

$$(18)\qquad h(x_0,\dots,x_{n_k+n}) = \int_A \overbrace{\Big(\prod_{i=0}^{n_k+n}\lambda(x_i)\,e^{-\lambda(x_i)t_i}\,dt_i\Big)}^{\overset{def}{=} d\mu_{x_0,\dots,x_{n_k+n}}(t)}$$

$$(19)\qquad \text{with } A\overset{def}{=}\Big\{(t_0,\dots,t_{n_k+n})\in\mathbb{R}_+^{n_k+n} : t_0+\dots+t_{n_i-1}\le s_i< t_0+\dots+t_{n_i},\ 0\le i\le k$$
$$\text{and } t_0+\dots+t_{n_k+n-1}\le s_k+h< t_0+\dots+t_{n_k+n}\Big\}$$

We now integrate over t_{n_k} in (18) using the memoryless property of the exponential, which can be stated as

$$(20)\qquad \int_u^\infty \lambda e^{-\lambda t}\varphi(t-u)\,dt = e^{-\lambda u}\int_0^\infty \lambda e^{-\lambda t}\varphi(t)\,dt,\quad \text{for } u\ge 0,\ \lambda>0,$$
φ bded, meas.

so that, with $t=t_{n_k}$, $\lambda=\lambda(x_{n_k})$, $u=s_k-(t_0+\dots+t_{n_k-1})$ (which satisfies $u\ge 0$ on A), and noting moreover that the second line in (19) can be recast as

$$h-(t_{n_{k+1}}+\dots+t_{n_{k+n}}) < t_{n_k}-u \le h-(t_{n_{k+1}}+\dots+t_{n_{k+n-1}})$$

we obtain:

(21) $h(x_0,\dots,x_{n_k+n}) = h_1(x_0,\dots,x_{n_k})\, h_2(x_{n_k},\dots,x_{n_k+n})$

where

$$h_1(x_0,\dots,x_{n_k}) = \mu_{x_0,\dots,x_{n_k}}\Big(\{(t_0,\dots,t_{n_k}) : t_0+\dots+t_{n_i-1} \le s_i < t_0+\dots+t_{n_i},\ 0\le i\le k\}\Big)$$

$$h_2(x_{n_k},\dots,x_{n_k+n}) = \mu_{x_{n_k},\dots,x_{n_k+n}}\Big(\Big\{(t_0,\dots,t_n) : \sum_{\ell=0}^{n-1} t_\ell \le h < \sum_{\ell=0}^{n} t_\ell\Big\}\Big).$$

Substituting (21) into (17) and returning to (16) yields

$$\tilde{E}_x[F 1_C] = E_x^Z[\underbrace{F}_{= f_0(Z_{n_0})\cdots f_k(Z_{n_k}) f(Z_{n_k+n})} h_1(Z_0,\dots,Z_{n_k})\, h_2(Z_{n_k},\dots,Z_{n_k+n})]$$

(only the randomness of $Z_.$ remains!)

(22) simple Markov for $Z_.$ at time n_k

$$= E_x^Z\Big[\prod_{i=0}^{k} f_i(Z_{n_i})\, \overbrace{h_1(Z_0,\dots,Z_{n_k})}\ \underbrace{E^Z_{Z_{n_k}}[f(Z_n)\, h_2(Z_0,\dots,Z_n)]}_{\overset{\text{def}}{=} G_n(Z_{n_k})}\Big]$$

By a similar calculation as in (17), (18), but simpler, we find that, for $\hat{C} = \{N_{s_0}=n_0,\dots,N_{s_k}=n_k\}$,

(23) $\tilde{E}_x[1_{\hat{C}} \mid \sigma(Z_0,\dots,Z_{n_k})] = h_1(Z_0,\dots,Z_{n_k})$

so that, going back to (22), we obtain that

$$\tilde{E}_x[F 1_C] = \tilde{E}_x\Big[\prod_{i=0}^{k} f_i(Z_{n_i})\, G_n(Z_{n_k})\, 1_{\hat{C}}\Big]$$

and, inserting this into (16), we finally get that

(24) $\text{LHS}(15) = \sum_{n\ge 0} \tilde{E}_x\Big[\prod_{i=1}^{k} f_i(Z_{N_{s_i}})\, G_n(Z_{N_{s_k}})\Big]$

Yet another calculation as in (23) (but even simpler!) identifies $h_2(Z_0,\dots,Z_n) = \tilde{E}_x[1_{\{N_h=n\}} \mid \sigma(Z_0,\dots,Z_n)]$, which implies (cf. the def of $G_n(\cdot)$ in (22)) that

(25) $G_n(y) = \tilde{E}_y[f(Z_{N_h})\, 1\{N_h=n\}]$, for $n\ge 0$, $y\in S$.

Substituting (25) into (24) and summing over n, we see that

$$\text{LHS}(15) = \tilde{E}_x\Big[\prod_{i=0}^{k} f_i(Z_{N_{s_i}})\, \tilde{E}_{Z_{N_{s_k}}}[f(Z_{N_h})]\Big]$$

$$\overset{\text{below}(15)}{=} \tilde{E}_x\Big[\prod_{i=0}^{k} f_i(\tilde{X}_{s_i})\, \tilde{E}_{\tilde{X}_{s_k}}[f(\tilde{X}_h)]\Big],$$

which is the RHS of (15), on account of (14)! $\square$

References:

[Li] T.M. Liggett, Continuous-time Markov Processes - An Introduction. GSM, Vol 113, AMS, 2010.

[Du] R. Durrett, Probability, Theory and Examples. Cambridge University Press, 2010.