Chaînes de Markov en temps continu et génétique des populations

Chapitre 1 : Chaînes de Markov en temps continu

Nicolas Champagnat

12 mars 2015

Table des matières

Version provisoire. Merci de me signaler les coquilles !

Dans toute la suite, on adopte la convention de noter X un processus $(X_t, t \in \mathbb{R}_+)$ ou $(X_n, n \in \mathbb{N})$. La notation $:=$ signifie que le membre de gauche de l'égalité est défini par le membre de droite.

Par souci de concision, on écrira dans toute la suite CMTC pour *chaîne de Markov en temps continu* et CMTD pour *chaîne de Markov en temps discret*.

Les preuves en plus petits caractères peuvent être omises dans une première lecture.

On recommande de se remémorer les propriétés et résultats principaux sur les CMTD (voir par exemple [2]) avant de poursuivre la lecture.

Pour des références plus détaillées sur les chaînes de Markov à temps continu, on renvoie à [2, 1]. La présentation proposée dans ce document est en grande partie adaptée de [2].

1 Les matrices de taux de transition et leurs exponentielles

Soit I un ensemble dénombrable ou fini, qui jouera le rôle de l'espace d'état de notre CMTC.

Définition 1.1 *On note $\mathcal{M}_I(\mathbb{R})$ l'ensemble des matrices $A = (a_{i,j})_{i,j\in I}$ à coefficients réels. Une matrice $Q = (q_{i,j})_{i,j\in I} \in \mathcal{M}_i(\mathbb{R})$ est un **matrice de taux de transitions sur** I (ou **Q-matrice** sur I) si*

(i) $0 \leq -q_{i,i} < \infty$ *pour tout* $i \in I$ *;*

(ii) $q_{i,j} \geq 0$ *pour tout* $i \neq j$ *dans* I *;*

(iii) $\sum_{j\in I} q_{i,j} = 0$ *pour tout* $i \in I$.

Dans la définition précédente, le terme général de la série $\sum_{j\in I} q_{i,j}$ n'est pas de signe constant, donc il pourrait y avoir une amiguïté sur le sens à lui donner. Nous adoptons la convention

$$\sum_{j\in I} q_{i,j} = q_{i,i} + \sum_{j\neq i} q_{i,j},$$

dans lequel la seconde série est bien définie car son terme général est positif ou nul.

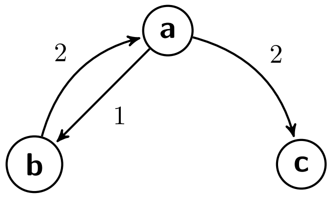
On utilisera dans toute la suite les notations $q(i)$ ou q_i pour

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

Pour tout $i \neq j$, on appelle $q_{i,j}$ le **taux de transition de** i **vers** j, et $q_i = \sum_{j\neq i} q_{i,j}$ le **taux total de saut depuis** i.

Exemple : Soit $I = \{a, b, c\}$ et Q la matrice de taux de transition

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



où la première ligne correspond à l'indice a, la seconde à b et la troisième à c. Il est souvent commode de représenter cette matrice par le diagramme ci-dessus, où nous verrons que les flèches correspondent aux sauts possibles de la CMTC, et les chiffres correspondent aux taux de saut correspondants.

Définition 1.2 *On suppose I fini. Pour toute matrice $Q \in \mathcal{M}_I(\mathbb{R})$, on note*

$$\exp(Q) = e^Q = \sum_{k \geq 0} \frac{Q^k}{k!}.$$

La convergence de cette série est classique [on peut par exemple montrer que la série est de Cauchy pour une norme d'algèbre sur $\mathcal{M}_I(\mathbb{R})$], ainsi que la proposition suivante, dont nous omettons la démonstration.

Proposition 1.3 *Si $Q_1, Q_2 \in \mathcal{M}_I(\mathbb{R})$ satisfont $Q_1Q_2 = Q_2Q_1$, alors*

$$e^{Q_1+Q_2} = e^{Q_1}e^{Q_2}.$$

Théorème 1.4 *Soit I fini et $Q \in \mathcal{M}_I(\mathbb{R})$ quelconque. Pour tout $t \geq 0$, on pose $P(t) = (p_{i,j}(t))_{i,j\in I} = e^{tQ}$. Alors*

(i) **(propriété de semigroupe)** *$P(t + s) = P(t)P(s)$ pour tout $s, t \geq 0$;*

(ii) *$(P(t), t \geq 0)$ est l'unique solution de l'**équation (de Kolmogorov) forward** définie par*

$$\begin{cases} \frac{d}{dt}P(t) = P(t)Q, & \forall t \geq 0, \\ P(0) = Id; \end{cases}$$

(iii) *$(P(t), t \geq 0)$ est l'unique solution de l'**équation backward** définie par*

$$\begin{cases} \frac{d}{dt}P(t) = QP(t), & \forall t \geq 0, \\ P(0) = Id; \end{cases}$$

Démonstration (i) est une conséquence immédiate de la proposition 1.3.

Pour (ii) et (iii), on observe que l'on peut appliquer le théorème de dérivation sous le signe somme pour obtenir

$$P'(t) = \sum_{k\geq 1} \frac{t^{k-1}Q^k}{(k-1)!} = P(t)Q = QP(t).$$

L'unicité de la solution des équations forward et backward s'obtient par exemple avec le théorème de Cauchy-Lipschitz, ou bien en remarquant que $\frac{d}{dt}(M(t)e^{-tQ}) = 0$ pour toute solution $M(t)$ de l'une de ces équations. $\square$

Le résultat suivant fait le lien entre les matrices de taux de transition et les matrices de probabilités de transitions, c'est-à-dire les matrices stochastiques.

Théorème 1.5 *Soit I fini et $Q \in \mathcal{M}_I(\mathbb{R})$.*
Q est une matrice de taux de transition $\iff P(t) = (p_{i,j}(t))_{i,j\in I} = e^{tQ}$ est stochastique pour tout $t \geq 0$.

Démonstration Prouvons d'abord le sens $\Longrightarrow$. On note $\mathbf{1}$ le vecteur colonne dont toutes les coordonnées valent 1. $P(t)$ est stochastique ssi ses coordonnées sont toutes positives et $P(t)\mathbf{1} = \mathbf{1}$. Or, d'après le théorème 1.4 (ii),

$$\frac{d}{dt}(P(t)\mathbf{1}) = P(t)Q\mathbf{1} = 0,$$

car $Q\mathbf{1} = 0$ pour une matrice de taux de transition. De plus, pour a suffisamment grand, la matrice $Q + a\mathrm{Id}$ a toutes ses coordonnées positives, donc par la définition 1.2, la matrice $\exp(t(Q + a\mathrm{Id})) = e^{at}P(t)$ a toutes ses coordonnées positives. Ceci conclut la preuve de la première implication.

Prouvons maintenant le sens $\Longleftarrow$. Lorsque $t \to 0$, on a $P(t) = e^{tQ} = \mathrm{Id} + tQ + O(t^2)$, donc $\mathbf{1} = P(t)\mathbf{1} = \mathbf{1} + tQ\mathbf{1} + O(t^2)$, dont on déduit que $Q\mathbf{1} = 0$. De plus, pour $i \neq j$, $q_{i,j} = \lim_{t\to 0} \frac{p_{i,j}(t)}{t} \geq 0$. $\square$

2 Première construction (abstraite) d'une CMTC

Soit I fini et Q une matrice de taux de transition sur I. Le théorème précédent montre que l'on peut construire des CMTD à partir des matrices $P(t)$ pour n'importe quelle valeur de t. Par exemple, pour tout $k \geq 0$, on note $(X_n^k, n \geq 0)$ une CMTD de matrice de transition $P(2^{-k})$ et de loi initiales $\lambda = (\lambda_i)_{i\in I}$ donnée. D'après les propriétés de base des CMTD, le processus $(Y_n, n \in \mathbb{N})$, où $Y_n := X_{2n}^{k+1}$, est également une CMTD de loi initiale λ et de matrice de probabilités de transition $P(2^{k+1})^2 = e^{2^{-k}Q} = P(2^k)$. Ainsi, les processus $(Y_n, n \geq 0)$ et $(X_n^k, n \geq 0)$ ont même loi.

Autrement dit, si on définit $Z^k_{n2^{-k}} = X^k_n$ pour tout $k \geq 0$ et $n \geq 0$, alors pour tout $k' \geq k$ et pour tout s de la forme $n2^{-k}$,

$$Z^{k'}_s = Z^k_s \qquad \text{en loi.}$$

Ceci suggère qu'il existe un processus $(Z_t, t \in \mathbb{R}_+)$ qui serait une certaine limite des processus Z^k, tel que $Z_{n2^{-k}} = Z^k_{n2^{-k}}$ en loi pour tout $n, k \geq 0$. De plus, puisque Z^k est une CMTD, on a pour tout $k \geq 0$, $0 \leq n_0 \leq n_1 \leq \ldots \leq n_{m+1}$ et $i_0, \ldots, i_{m+1} \in I$,

$$\begin{aligned}
\mathbb{P}(Z^k_{n_{m+1}2^{-k}} = i_{m+1} \mid Z^k_{n_0 2^{-k}} = i_0, \ldots, Z^k_{n_m 2^{-k}} = i_m)& \\
&= \mathbb{P}(Z^k_{n_{m+1}2^{-k}} = i_{m+1} \mid Z^k_{n_m 2^{-k}} = i_m) \\
&= \left[\left(P(2^{-k})\right)^{n_{m+1}-n_m}\right]_{i_m, i_{m+1}} \\
&= \left[P\left((n_{m+1}-n_m)2^{-k}\right)\right]_{i_m, i_{m+1}} \\
&= p_{i_m, i_{m+1}}((n_{m+1}-n_m)2^{-k}).
\end{aligned}$$

On s'attend alors à ce que le processus limite satisfasse pour tout $0 \leq t_0 \leq \ldots \leq t_{n+1}$ et $i_0, \ldots, i_{n+1} \in I$,

$$\begin{aligned}
\mathbb{P}(Z_{t_{n+1}} = i_{n+1} \mid Z_{t_0} = i_0, \ldots, Z_{t_n} = i_n) &= \mathbb{P}(Z_{t_{n+1}} = i_{n+1} \mid Z_{t_n} = i_n) \\
&= p_{i_n, i_{n+1}}(t_{n+1} - t_n). \qquad (1)
\end{aligned}$$

La première égalité de cette équation s'appelle **propriété de Markov (faible**[1]**)**. La seconde égalité permet une première construction — plutôt abstraite — d'une CMTC.

Soit $\Omega_0 = I^{\mathbb{R}_+}$ l'ensemble de toutes les fonctions de $\mathbb{R}_+$ dans I, muni de sa tribu cylindrique $\mathcal{F}^0$ engendrée par l'ensemble $\mathcal{R}$ de tous les cylindres fini-dimensionnels de la forme

$$\Gamma_{t_0,\ldots,t_n,A_0,\ldots,A_n} := \{x \in \Omega_0 : x(t_0) \in A_0, \ldots, x(t_n) \in A_n\},$$

où $n \in \mathbb{N}$, $0 \leq t_0 \leq \ldots \leq t_n$ et $A_0, \ldots, A_n \subset I$. Lorsque $A_0 = \{i_0\}, \ldots, A_n = \{i_n\}$, on notera par abus de notation

$$\Gamma_{t_0,\ldots,t_n,i_0,\ldots,i_n} := \Gamma_{t_0,\ldots,t_n,\{i_0\},\ldots,\{i_n\}}.$$

Soit $i \in I$ fixé. La relation (1) suggère de définir une mesure μ sur $\mathcal{R}$ par la formule

$$\mu(\Gamma_{t_0,\ldots,t_n,i_0,\ldots,i_n}) := p_{i_{n-1},i_n}(t_n - t_{n-1}) \ldots p_{i_0,i_1}(t_1 - t_0) p_{i,i_0}(t_0),$$

et

$$\mu(\Gamma_{t_0,\ldots,t_n,A_0,\ldots,A_n}) := \sum_{i_0 \in A_0,\ldots,i_n \in A_n} \mu(\Gamma_{t_0,\ldots,t_n,i_0,\ldots,i_n}). \qquad (2)$$

1. La propriété de Markov forte apparaîtra dans le théorème 7.5.

Proposition 2.1 *La fonction μ définie sur $\mathcal{R}$ admet un prolongement sur $\mathcal{F}^0$ qui est une mesure, toujours notée μ.*

Démonstration Ce résultat découle du résultat suivant, qui est la reformulation du théorème de Daniell-Kolmogorov dans notre cas, que l'on admettra.

Théorème 2.2 (de Daniell-Kolmogorov) *Supposons que pour tout $0 \leq t_0 \leq \ldots \leq t_n$, la fonction définie pour tout $A_0, \ldots, A_n \subset I$ par*

$$\mu_{t_0,\ldots,t_n}(A_0 \times \ldots \times A_n) := \mu(\Gamma_{t_0,\ldots,t_n,A_0,\ldots,A_n})$$

est une mesure de probabilité sur I^{n+1} (muni de la tribu formée par tous ses sous-ensembles). Si la famille de mesures $(\mu_{t_0,\ldots,t_n}, n \geq 0, t_0 \leq \ldots \leq t_n)$ satisfait la propriété de

compatibilité : *pour tout $n \geq 1$, pour tous réels $0 \leq t_0 \leq \ldots \leq t_n$, pour tout $j \in \{0, 1, \ldots, n\}$ et $A_0, \ldots, A_{j-1}, A_{j+1}, \ldots, A_n \subset I$,*

$$\begin{aligned}\mu_{t_0,\ldots,t_n}&(A_0 \times \ldots A_{j-1} \times I \times A_{j+1} \times \ldots \times A_n)\\ &= \mu_{t_0,\ldots,t_{j-1},t_{j+1},\ldots,t_n}(A_0 \times \ldots A_{j-1} \times A_{j+1} \times \ldots \times A_n),\end{aligned}$$

alors μ peut être prolongée à $\mathcal{F}^0 = \sigma(\mathcal{R})$ par une mesure de probabilité.

Il suffit donc de vérifier que $\mu_{t_0,\ldots,t_n}$ définie comme dans (2) est une mesure de probabilité, et que ces mesures satisfont la propriété de compatibilité. Le fait que $\mu_{t_0,\ldots,t_n}$ est une mesure est trivial par (2). Le fait que c'est une mesure de probabilité découle du fait que les matrices $P(t_0), P(t_1 - t_0), \ldots, P(t_n - t_{n-1})$ sont stochastiques. La propriété de compatibilité est facilement impliquée par le cas où les A_i sont des singletons, c'est-à-dire par la relation

$$\begin{aligned}\mu_{t_0,\ldots,t_n}&(\{i_0\} \times \ldots \{i_{j-1}\} \times I \times \{i_{j+1}\} \times \ldots \times \{i_n\})\\ &= \mu_{t_0,\ldots,t_{j-1},t_{j+1},\ldots,t_n}(\{i_0\} \times \ldots \{i_{j-1}\} \times \{i_{j+1}\} \times \ldots \times \{i_n\}).\end{aligned}$$

Or, le terme de gauche vaut par définition

$$\begin{aligned}\sum_{i_j \in I} \prod_{k=0}^{n} p_{i_{k-1},i_k}(t_k - t_{k-1}) = \prod_{k=0}^{j-1} p_{i_{k-1},i_k}(t_k - t_{k-1}) \times \\ \left(\sum_{i_j \in I} p_{i_{j-1},i_j}(t_j - t_{j-1}) p_{i_j,i_{j+1}}(t_{j+1} - t_j) \right) \prod_{k=j+2}^{n} p_{i_{k-1},i_k}(t_k - t_{k-1}),\end{aligned}$$

avec la convention $t_{-1} = 0$ et $i_{-1} = i$. La relation souhaitée découle du fait que $P(t_j - t_{j-1})P(t_{j+1} - t_j) = e^{(t_{j+1} - t_{j-1})Q} = P(t_{j+1} - t_{j-1})$. □

On a donc obtenu une loi μ sur $\Omega_0 = I^{\mathbb{R}_+}$ telle que le processus canonique Z sur cet espace (défini par $Z_t(\omega) = \omega(t)$ pour tout $\omega \in \Omega_0$) satisfait (1). Mais il s'agit d'une construction abstraite, peu manipulable.

Tous les résultats de cette sous-section seraient également vrais pour I dénombrable (notamment le théorème de Daniell-Kolmogorov et donc la proposition 2.1), mais la difficulté dans ce cas consiste à définir convenablement les matrices stochastiques $P(t)$, car les exponentielles des matrices tQ ne sont pas nécessairement bien définies.

Ainsi, la construction précédente ne permet pas de construire une CMTC en toute généralité. De plus, le processus a été seulement caractérisé par sa loi, sur un espace abstrait $I^{\mathbb{R}_+}$ muni d'une tribu assez grossière (la tribu cylindrique). En particulier, elle ne permet pas facilement d'avoir des propriétés de régularité des trajectoires de la CMTC (par exemple, sont-elles càdlàg, c'est-à-dire continues à droite et admettant une limite à gauche en tout point ?). L'objet de la suite est de construire la CMTC de façon plus explicite, en commençant par son espace de trajectoires dans la sous-section suivante.

3 Espace des trajectoires d'une CMTC

Nous allons construire des CMTC dont les trajectoires $(X_t, t \geq 0)$ sont constantes par morceaux, continues à droite et admettant une limite à gauche (noté **càdlàg**, voir le chapitre sur la convergence des processes stochastiques pour une définition précise) en (presque) tout point, à valeur dans un espace d'état $I \cup \{\partial\}$, constitué d'un ensemble fini ou dénombrable I et d'un point cimetière ∂ n'appartenant pas à I. Une fonction constante par morceaux peut être caratérisée par la suite $(Y_n, n \geq 0)$ de ses valeurs prises dans I, appelée **chaîne incluse**, et par la suite des **temps de séjour** $(S_i, i \geq 1)$ à valeurs dans $]0, +\infty]$ pour chacune de ces valeurs (voir figure 1). Autrement dit, le processus X vaut Y_0 pendant une durée S_1, puis vaut Y_1 pendant une durée S_2, etc. Remarquons que les deux suites $(Y_n, n \geq 0)$ et $(S_i, i \geq 1)$ ne sont plus définies après le premier indice tel que $S_i = +\infty$. Dans ce cas, on convient que l'on complètera les deux suites avec des valeurs arbitraires. Avec la convention que le processus est càdlàg, ceci conduit à la formule

$$X_t = Y_n, \quad \text{pour tout } t \in [J_n, J_{n+1}[,$$

où $J_0 = 0$ et $J_n = S_1 + \ldots + S_n$, $n \geq 1$ sont les instants de sauts du processus X (J comme *jump*). Remarquons que, dans le cas où $\sup_n J_n < +\infty$, c'est-à-dire $\sum_{i \geq 1} S_i < +\infty$, la formule précédente ne spécifie pas la valeur de X_t pour $t \geq \sum_{i \geq 1} S_i$, qui n'est pas caractérisée par les suites (Y_n) et (S_i). On appelle

$$\zeta := \sum_{i \geq 1} S_i$$

le **temps d'explosion** de X, et on pose par convention $X_t = \partial$ pour tout

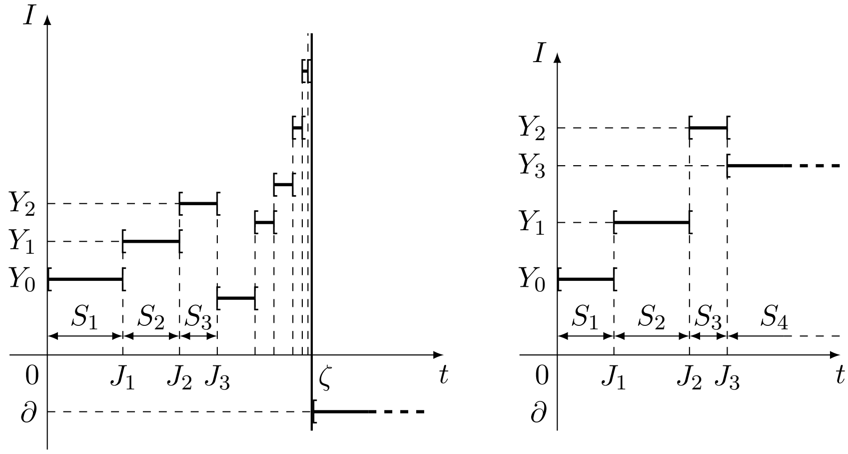


FIGURE 1 – Deux exemples de trajectoire d'une CMTC. À gauche, il y a explosion au temps $\zeta < \infty$ et donc une infinité de sauts sur tout intervalle $]\zeta - \varepsilon, \zeta[$, $\varepsilon > 0$. À droite, $S_4 = +\infty$, et dans ce cas les valeurs de $Y_4, Y_5, \ldots$ sont arbitraires et n'influent pas sur la trajectoire de $(X_t, t \geq 0)$.

$t \geq \zeta$. Finalement, on définit

$$X_t = \begin{cases} \sum_{n \geq 0} Y_n \mathbb{1}_{\{J_n \leq t < J_{n+1}\}}, & \text{pour } t < \zeta, \\ \partial & \text{pour } t \geq \zeta. \end{cases} \tag{3}$$

Remarquons que cette limite est continue à droite en tout $t \in \mathbb{R}_+$ et admet une limite à gauche en tout $t \in \mathbb{R}_+^* \setminus \{\zeta\}$. Remarquons également la relation $Y_n = X_{J_n}$ pour tout $n \geq 0$.

Définition 3.1 *On définit Ω l'ensemble des trajectoires X obtenues par la formule* (3) *à partir de n'importe-quelles suites (Y_n) dans I et (S_i) dans* $]0, +\infty]$. *Comme dans la sous-section précédente, on définit sur Ω la tribu cylindrique $\mathcal{F}$ engendrée par les cylindres fini-dimensionnels*

$$\Gamma_{t_0,\ldots,t_n,A_0,\ldots,A_n} := \{x \in \Omega : x(t_0) \in A_0, \ldots, x(t_n) \in A_n\},$$

et pour tout $t \geq 0$ la tribu $\mathcal{F}_t$ engendrée par les ensembles de la forme $\Gamma_{t_0,\ldots,t_n,A_0,\ldots,A_n}$ avec $t_n \leq t$. On obtient alors un espace mesurable filtré $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0})$.

Si on parvient à construire sur cet espace une mesure $\mathbb{P}$ telle que

$$\mathbb{P}(\Gamma_{t_0,\ldots,t_n,i_0,\ldots,i_n}) = p_{i,i_0}(t_0) p_{i_0,i_1}(t_1 - t_0) \ldots p_{i_{n-1},i_n}(t_n - t_{n-1}),$$

alors il s'agira d'un mesure prolongeant la mesure μ comme dans la proposition 2.1, avec la convention que $\mathbb{P}(\Omega_0 \setminus \Omega) = 0$. On aura alors montré qu'il existe une version de la CMTC de la sous-section précédente qui est càdlàg (sauf éventuellement en ζ).

Notons que la loi d'un processus $(X_t, t \geq 0)$ à valeurs dans Ω muni de la tribu $\mathcal{F}$ est caractérisée par ses marginales fini-dimensionnelles [2], c'est-à-dire les lois des vecteurs $(X_{t_0}, \dots, X_{t_n})$ pour tout $0 \leq t_0 \leq \dots \leq t_n$. Or, d'après la formule (3), la loi de ces vecteurs est caractérisée par la loi jointe des suites $(Y_n, n \geq 0)$ et $(S_i, i \geq 1)$. En d'autres termes, on a le résultat suivant.

Proposition 3.2 *La loi d'un processus à valeurs dans Ω muni de la tribu $\mathcal{F}$ est caractérisée par la loi jointe de sa chaîne incluse et de ses temps de séjour.*

4 Quelques propriétés des lois exponentielles

La proposition suivante est classique et nous laissons sa preuve en exercice.

Proposition 4.1 (absence de mémoire) *Les lois exponentielles et la loi $\delta_{+\infty}$ (qui sera par convention appelée dans la suite loi exponentielle de paramètre 0) sont les seules lois de v.a. T à valeurs dans $[0, +\infty]$ telles que*

$$\mathbb{P}(T > s + t \mid T > s) = \mathbb{P}(T > t) \qquad \textit{pour tout } s, t \geq 0.$$

Le résultat suivant précise les cas où une somme de v.a. exponentielles indépendantes est finie.

Théorème 4.2 *Soit $S_1, S_2, \dots$ des v.a. indépendantes telles que S_n suit la loi exponentielle de paramètre $0 < \lambda_n < \infty$ donné pour tout $n \geq 1$.*

(i) *Si $\sum_{n\geq 1} \frac{1}{\lambda^n} < +\infty$, alors $\mathbb{P}(\sum_{n\geq 1} S_n < +\infty) = 1$.*

(ii) *Si $\sum_{n\geq 1} \frac{1}{\lambda^n} = +\infty$, alors $\mathbb{P}(\sum_{n\geq 1} S_n = +\infty) = 1$.*

Démonstration Par convergence monotone, $\mathbb{E}(\sum S_n) = \sum \frac{1}{\lambda_n}$, ce qui prouve le point (i).

Par convergence dominée, $\mathbb{E} \exp(-\sum S_n) = \prod(1 + \frac{1}{\lambda_n})^{-1}$. En passant au logarithme, il est standard de vérifier que le produit infini du terme de droite vaut zéro lorsque la série $\sum_n \log(1 + \frac{1}{\lambda_n})$ diverge, ce qui est le cas lorsque $\sum_{n\geq 1} \frac{1}{\lambda_n} = +\infty$. Le point (ii) en découle immédiatement. $\square$

Le résultat suivant donne la loi jointe de l'infimum et de l'argmin d'une suite d'exponentielles indépendantes.

2. Puisque $\mathcal{F}$ est la tribu engendrée par les cylindres fini-dimensionnels et que l'ensemble de ces cylindres fini-dimensionnels est une algèbre de Boole (ou π-système), c'est une conséquence du théorème d'extension de Caratheodory.

Théorème 4.3 *Soit I dénombrable et $(T_k, k \in I)$ une famille de v.a. indépendantes telle que T_k suit la loi exponentielle de paramètre $q_k \in \mathbb{R}_+$. Soit $q := \sum_{k\in I} q_k$ supposé dans $\mathbb{R}_+^*$. Soit enfin $T = \inf_{k\in I} T_k$. Alors, p.s., l'infimum est atteint à un unique indice K tel que T et K sont indépendants, T suit la loi exponentielle de paramètre q et*

$$\mathbb{P}(K = k) = \frac{q_k}{q}, \qquad \textit{pour tout } k \in I.$$

Démonstration Posons $K = \partial$ sur l'événement où l'argmin n'est pas atteint ou est atteint simultanément pour deux indices $k \in I$ ou plus. Alors, pour tout $k \in I$ et $t \in \mathbb{R}_+$,

$$\begin{aligned}\mathbb{P}(K = k \text{ et } T \geq t) &= \mathbb{P}(T_k \geq t \text{ et } T_j > t_k \text{ pour tout } j \neq k) \\ &= \int_t^\infty q_k e^{-q_k s} \mathbb{P}(T_j > s \text{ pour tout } j \neq k)\, ds \\ &= \int_t^\infty q_k e^{-q_k t} \prod_{j\neq k} e^{-q_j s}\, ds \\ &= \int_t^\infty q_k e^{-qs}\, ds = \frac{q_k}{q}\, e^{-qs},\end{aligned}$$

où la seconde égalités est obtenue en conditionnant par rapport à T_k à l'intérieur de l'espérance et la troisième égalité découle de l'indépendance des v.a. T_j, $j \in I$. □

5 Construction d'une CMTC à partir de sa chaîne incluse et de ses temps de séjour

D'après la proposition 3.2, il suffit de caractériser la loi jointe de $(Y_n, n \geq 0)$ et $(S_i, i \geq 1)$ pour caractériser la loi du processus stochastique $(X_t, t \geq 0)$ à trajectoires dans Ω donné par la formule (3). La chaîne incluse Y sera en fait une CMTD, et les temps de séjours seront exponentiels conditionnellement au processus Y.

5.1 Définition d'une CMTC

Afin de construire la CMTD Y, définissons sa matrice de probabilités de transition.

Fixons I un ensemble fini ou dénombrable et Q une matrice de taux de transition sur I. On définit alors la matrice $\Pi = (\pi_{i,j})_{i,j\in I}$ par

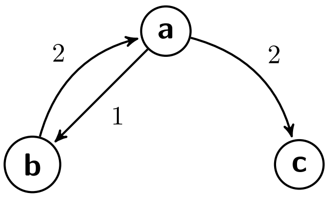
$$\pi_{i,j} := \begin{cases} q_{i,j}/q_i & \text{si } i \neq j \text{ et } q_i \neq 0, \\ 0 & \text{si } i \neq j \text{ et } q_i = 0; \end{cases} \quad \text{et} \quad \pi_{i,i} := \begin{cases} 0 & \text{si } q_i \neq 0, \\ 1 & \text{si } q_i = 0. \end{cases} \tag{4}$$

Proposition 5.1 *Si Q est une matrice de taux de transitions, alors Π est une matrice stochastique.*

Démonstration Triviale. $\square$

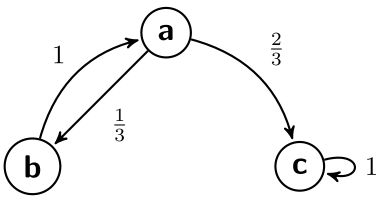
Exemple : On reprend l'exemple où $I = \{a, b, c\}$ et Q la matrice de taux de transition

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



Alors la matrice Π et le diagramme de transition de la CMTD correspondante sont donnés par

$$\Pi = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$



Définition 5.2 *Soit I fini ou dénombrable et Q une matrice de taux de transition sur I. Un processus $(X_t, t \geq 0)$ à trajectoires dans Ω (défini dans la section 3) est une CMTC de loi initiale $\lambda = (\lambda_i)_{i \in I}$ et de* ***générateur (infinitésimal)*** *Q (ou de matrice de taux de transition Q) ssi*

(i) *sa chaîne incluse $(Y_n)_{n \geq 0}$ est une CMTD de loi initiale λ et de matrice de probabilités de transition Π ;*

(ii) *pour tout $n \geq 1$, conditionnellement à $Y_0, \dots, Y_{n-1}$, ses temps de séjour $S_1, \dots, S_n$ sont indépendants de loi exponentielles de paramètres respectifs $q(Y_0), \dots, q(Y_{n-1})$;*

(iii) *si $\zeta = \sum_{i \geq 1} S_i < \infty$, $X_t = \partial$ pour tout $t \geq \zeta$.*

Pour tout $i \in I$, on note alors $\mathbb{P}_i$ la loi du processus $(X_t, t \geq 0)$ lorsque $\lambda = \delta_i$ (c.-à-d. lorsque $X_0 = i$ p.s.) et $\mathbb{E}_i$ l'espérance correspondante. Pour tout mesure de probabilité λ sur I, on note $\mathbb{P}_\lambda := \sum_{i \in I} \lambda_i \mathbb{P}_i$ et $\mathbb{E}_\lambda$ l'espérance correspondante.

Remarquons que le point (iii) de la définition est superflu car il est automatiquement vérifié pour un processus à trajectoires dans Ω. Puisque la chaîne incluse Y est à valeurs discrètes, le point (ii) se reformule comme : pour tout $n \geq 1$, $i_0, \dots, i_{n-1} \in I$ et $s_1, \dots, s_n \geq 0$,

$$\mathbb{P}(S_1 \geq s_1, \dots, S_n \geq s_n \mid Y_0 = i_0, \dots, Y_{n-1} = i_{n-1}) = \exp\left(-\sum_{j=1}^{n} q(i_{j-1}) s_j\right).$$

5.2 Constructions algorithmiques d'une CMTC

Cette définition suggère la méthode de simulation suivante pour la CMTC $(X_t, t \geq 0)$.

Construction algorithmique 1 Soit $(T_n)_{i \geq 1}$ une suite i.i.d. de v.a. exponentielles de paramètres 1.

1. On tire $X_0 = Y_0$ selon la loi λ ;
2. On simule $(Y_n, n \geq 0)$ CMTD de matrice de probabilités de transition Π ;
3. Pour tout $n \geq 1$, on pose $S_n = \frac{T_n}{q(Y_{n-1})}$.

On obtient alors le processus X par la formule (3), où $J_n = S_1 + \ldots + S_n$.

Cette première construction n'est pas pratique car elle sépare la simulation de la CMTD Y et des temps de séjour. Il est facile de vérifier que la suivante donne la même loi pour les suites $(Y_n)_{n \geq 0}$ et $(S_n)_{n \geq 1}$ (et donc la même loi pour le processus X).

Construction algorithmique 2 Soit $(T_n)_{n \geq 1}$ une suite i.i.d. de v.a. exponentielles de paramètres 1.

1. On tire $X_0 = Y_0$ selon la loi λ ;
2. On pose $k = 0$;
3. On pose $S_{k+1} = \frac{T_{k+1}}{q(Y_k)}$, et on tire Y_{k+1} selon la loi $(\pi_{Y_k, j})_{j \in I}$;
4. On pose $k = k + 1$ et on retourne à l'étape 3.

On obtient le processus X par la formule (3). Il s'agit d'une définition par récurrence, qui se réalise algorithmiquement par une boucle `for` ou `while`.

Le théorème 4.3 permet une construction alternative à la précédente, qui peut se comprendre de la façon suivante : on associe à chaque prochain état possible de la CMTC une horloge exponentielle indépendante des autres dont le paramètre est le taux de transition vers cet état. La première horloge qui sonne donne alors la date de saut et l'état vers lequel on saute.

Construction algorithmique 2 Soit $(T_n^i)_{n\geq 1,\, i\in I}$ une famille i.i.d. de v.a. exponentielles de paramètres 1.

1. On tire $X_0 = Y_0$ selon la loi λ ;
2. On pose $k = 0$;
3. Pour tout $j \in I$, $j \neq Y_k$, on pose $S_{k+1}^j = \frac{T_{k+1}^j}{q_{Y_k,j}}$;
4. On pose $S_{k+1} = \inf_{j\neq Y_k} S_{k+1}^j$;
5. On pose $Y_{k+1} = j$ si $S_{k+1}^j = S_{k+1}$;
6. On pose $k = k+1$ et on retourne à l'étape 3.

On obtient le processus X par la formule (3). Remarquons que, d'après le théorème 4.3, la valeur de Y_{k+1} à l'étape 5 est bien définie.

Ces constructions justifient les expressions « taux exponentiel de saut de i vers j » pour la quantité $q_{i,j}$, et taux exponentiel de saut depuis i » pour q_i.

6 Critères de non-explosion

Théorème 6.1 *Si $(X_t, t \geq 0)$ est une CMTC de générateur Q, alors $\zeta = +\infty$ p.s. si l'une des conditions suivantes est vérifiée :*

(i) *I est fini ;*

(ii) $\sup_{i\in I} q_i < \infty$ *;*

(iii) *$X_0 = i$ et i est un point récurrent pour la CMTD de matrice de probabilités de transition Π (c.-à-d. pour la chaîne incluse).*

Démonstration Pour tout $n \geq 1$, on pose $T_n = q(Y_{n-1})S_n$. Alors, conditionnellement à $Y_0, \ldots, Y_{n-1}$, les v.a. $T_1, \ldots, T_n$ sont i.i.d. de loi exponentielle de paramètre 1. Puisque cette loi ne dépend pas de $Y_0, \ldots, Y_{n-1}$, on déduit alors facilement de la formule des probabilités totales (exercice !) que les v.a. $T_1, \ldots, T_n$ et donc la suite $(T_n)_{n\geq 1}$ est i.i.d. de loi exponentielle de paramètre 1, et que cette suite est indépendante de la chaîne incluse $(Y_n, n \geq 0)$.

Pour (i) et (ii), puisque $\bar{q} = \sup_{i\in I} q_i < \infty$, on en déduit que

$$\bar{q}\zeta = \bar{q}\sum_{n\geq 1} S_n \geq \sum_{n\geq 1} T_n = \infty, \quad \text{p.s.},$$

d'après le théorème 4.2.

Pour (iii), l'hypothèse que i est récurrent pour Y signife que la suite $(Y_n, n \geq 0)$ visite p.s. infiniment souvent i. Soit $N_1, N_2, \ldots$ les instants de visite (aléatoires) successifs. On a alors

$$q_i\zeta \geq q_i \sum_{k\geq 1} S_{N_k} = \sum_{k\geq 1} T_{N_k}.$$

Or, puisque la suite $(N_k, k \geq 1)$ ne dépend que de $(Y_n, n \geq 0)$, elle est indépendante de $(T_n)_{n\geq 1}$, et dans la suite $(T_{N_k}, k \geq 0)$ est i.i.d. exponentielle de paramètre 1. On conclut donc de nouveau grâce au théorème 4.2. $\square$

7 Propriété de Markov forte

La propriété de Markov forte est une propriété fondamentale des CMTC, très utile en pratique. On rappelle que la filtration $(\mathcal{F}_t)_{t\geq 0}$ a été définie dans la section 3.

7.1 Définitions et propriétés préliminaires

Les définitions suivantes sont classiques.

Définition 7.1 (i) *Une v.a. T à valeurs dans $[0, +\infty]$ sur $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_i)_{i\in I})$ est un* ***temps d'arrêt*** *si, pour tout $t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$.*

(ii) *Soit T un temps d'arrêt. On note $\mathcal{F}_T$ et on appelle la tribu au temps T la tribu*

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t,\ \forall t \geq 0\}.$$

On commence par quelques propriétés simples et classiques.

Proposition 7.2 *Soit T un temps d'arrêt. Alors $\{T < t\} \in \mathcal{F}_t$ pour tout $t > 0$.*

Démonstration Ceci découle immédiatement de la relation $\{T < t\} = \bigcup_{n\geq 1}\{T \leq t - \frac{1}{n}\} \in \mathcal{F}_t$. $\square$

Proposition 7.3 *Pour tout $m \geq 0$, le m-ième temps de saut J_m est un temps d'arrêt.*

Démonstration On prouve ce résultat par récurrence sur m. $J_0 = 0$ est clairement un temps d'arrêt. Soit $m \geq 0$. Supposons que J_m est un temps d'arrêt. Alors

$$\{J_{m+1} \leq t\} = \bigcup_{s\in\mathbb{Q},\ s\leq t} \Big(\{J_m < s\} \cap \{X_s \neq X_{J_m}\}\Big) \in \mathcal{F}_t. \qquad \square$$

Proposition 7.4 *Soit S et T deux temps d'arrêt. Alors X_T et $\{S \leq T\}$ sont $\mathcal{F}_T$-mesurables.*

Démonstration Le processus $(X_t, t \geq 0)$ est p.s. càd constant par morceaux, donc p.s. sur $\{T < t\}$, $\exists n \geq 0$, $\forall m \geq n$, $\exists k \geq 0$ tel que $(k-1)2^{-m} \leq T < k2^{-m} \leq t$ et $X_{k2^{-m}} = X_T$. Autrement dit,

$$\{X_T = i\} \cap \{T \leq t\} = \Big(\{X_t = i\} \cap \{T = t\}\Big) \bigcup \left(\bigcup_{n\geq 0} \bigcap_{m \geq n} \bigcup_{k \geq 0} \Big(\{X_{k2^{-m}} = i\} \cap \{(k-1)2^{-m} \leq T < k2^{-m} \leq t\}\Big) \right)$$

Cet événement est $\mathcal{F}_t$-mesurable, donc X_T est $\mathcal{F}_T$-mesurable.

Pour montrer que $\{S \leq T\} \in \mathcal{F}_T$, on remarque que, pour tout $t \leq 0$,

$$\{S > T\} \cap \{T \leq t\} = \bigcup_{s \in \mathbb{Q},\ s \leq t} \Big(\{T \leq s\} \cap \{S > s\}\Big) \in \mathcal{F}_t. \qquad \square$$

7.2 Résultat principal

Théorème 7.5 (Propriété de Markov forte) *Soit I fini ou dénombrable et Q une matrice de taux de transition sur I. Soit $(X_t, t \geq 0)$ une CMTC de générateur Q et de loi intiiale λ. Soit T un temps d'arrêt. Alors, pour tout $i \in I$, conditionnellement à $\{X_T = i\}$, $X_{T+\cdot} = (X_{T+t}, t \geq 0)$ est une CMTC de générateur Q et de loi initiale δ_i, indépendante de $\mathcal{F}_T$, c.-à-d. pour tout $\Gamma \in \mathcal{F}$, $A \in \mathcal{F}_T$ et $i \in I$ tel que $\mathbb{P}_\lambda(X_T = i) > 0$,*

$$\mathbb{P}_\lambda(A \cap \{X_{T+\cdot} \in \Gamma\} \mid X_T = i) = \mathbb{P}_\lambda(A \mid X_T = i)\mathbb{P}_i(\Gamma),$$

ou de manière équivalente

$$\mathbb{P}_\lambda(A \cap \{X_{T+\cdot} \in \Gamma\} \cap \{X_T = i\}) = \mathbb{P}_\lambda(A \cap \{X_T = i\})\mathbb{P}_i(\Gamma).$$

Remarquons que si $X_T = i$, alors $T < \zeta$ (et $T < \infty$), et donc on peut reformuler la propriété de Markov comme suit : conditionnellement à X_T et $\{T < \zeta\}$, $X_{T+\cdot}$ est une CMTC de générateur Q issue de X_T, indépendante de $\mathcal{F}_T$.

Avant de donner la preuve de ce résultat, nous donnons plusieurs formulations équivalentes de la propriété de Markov, dont certaines seront plus faciles à manipuler en pratique.

Proposition 7.6 (Formulations équivalentes) *La propriété de Markov forte du théorème précédent est équivalente à chacune des propriétés suivantes :*

(a) *Pour tout temps d'arrêt T et tout $\Gamma \in \mathcal{F}$,*

$$\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid \mathcal{F}_T) = \mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid X_T), \quad \textit{p.s. sur } \{T < \zeta\}, \quad (5)$$

et pour tout $i \in I$ tel que $\mathbb{P}_\lambda(X_T = i) > 0$,

$$\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid X_T = i) = \mathbb{P}_i(\Gamma).$$

(b) *Pour tout temps d'arrêt T et tout $\Gamma \in \mathcal{F}$,*

$$\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid \mathcal{F}_T) = F(X_T), \quad \textit{p.s. sur } \{T < \zeta\},$$

où pour tout $i \in I$ tel que $\mathbb{P}_\lambda(X_T = i) > 0$,

$$F(i) = \mathbb{P}_i(\Gamma).$$

Rappelons que le point (a) est équivalent à

(a') Pour tout temps d'arrêt T et toute fonction $G : \Omega \to \mathbb{R}$ bornée mesurable,

$$\mathbb{E}_\lambda(G(X_{T+\cdot}) \in \Gamma \mid \mathcal{F}_T) = \mathbb{E}_\lambda(G(X_{T+\cdot}) \in \Gamma \mid X_T), \quad \text{p.s. sur } \{T < \zeta\},$$

et pour tout $i \in I$ tel que $\mathbb{P}_\lambda(X_T = i) > 0$,

$$\mathbb{E}_\lambda(G(X_{T+\cdot}) \in \Gamma \mid X_T = i) = \mathbb{E}_i(G(X)).$$

et similairement pour (b).

On peut également montrer (on l'admettra) que toutes ces formulations sont équivalentes à

(a") Pour tout temps d'arrêt T, pour tout $s \geq 0$ et toute fonction $g : I \to \mathbb{R}$ bornée,

$$\mathbb{E}_\lambda(g(X_{T+s}) \mid \mathcal{F}_T) = \mathbb{E}_\lambda(g(X_{T+s} \mid X_T), \quad \text{p.s. sur } \{T < \zeta\},$$

et pour tout $i \in I$ tel que $\mathbb{P}_\lambda(X_T = i) > 0$,

$$\mathbb{E}_\lambda(g(X_{T+s}) \in \Gamma \mid X_T = i) = \mathbb{E}_i(g(X_s)).$$

et similairement pour (b).

Démonstration de la proposition 7.6 L'équivalence entre (a) et (b) est claire (on rappelle qu'étant donnée une v.a. X, une autre v.a. $\sigma(X)$-mesurable est p.s. égale à une fonction mesurable de X).

On appelera (c) la propriété de Markov de l'énoncé du théorème 7.5.

Montrons que (c) implique (a). Puisque la v.a. X_T est à valeurs discrètes, la v.a. $\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid X_T)$ vaut par définition $\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid X_T = i)$ sur l'événement $\{X_T = i\}$, pour tout $i \in I$. Or, (d) implique que $\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid X_T = i) = \mathbb{P}_i(\Gamma) = F(i)$. Il nous reste donc à montrer (5). Pour tout $A \in \mathcal{F}_T$,

$$\mathbb{P}_\lambda(A \cap \{T < \zeta\} \cap \{X_{T+\cdot} \in \Gamma\}) = \mathbb{E}_\lambda \left[\mathbb{1}_A \mathbb{1}_{T<\zeta} \mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid \mathcal{F}_T) \right],$$

et cette relation caractérise la v.a. $\mathcal{F}_T$ mesurable $\mathbb{1}_{T<\zeta}\mathbb{P}_\lambda(X_{T+\cdot} \in \Gamma \mid \mathcal{F}_T)$. Or d'après (c), pour tout $i \in I$,

$$\mathbb{P}_\lambda(A \cap \{X_T = i\} \cap \{X_{T+\cdot} \in \Gamma\}) = \mathbb{P}_\lambda(A \cap \{X_T = i\})\mathbb{P}_i(\Gamma).$$

Donc

$$\mathbb{1}_{T<\zeta}\mathbb{P}_\lambda(X_{T+\cdot}\in\Gamma\mid\mathcal{F}_T)=\sum_{i\in I}\mathbb{1}_{X_T=i}\mathbb{P}_i(\gamma)$$

qui est bien une v.a. $\mathcal{F}_T$-mesurable. Or le membre de droite est en fait une fonction de X_T seulement, donc une v.a. $\sigma(X_T)$-mesurable. Donc

$$\begin{aligned}\mathbb{1}_{T<\zeta}\mathbb{P}_\lambda(X_{T+\cdot}\in\Gamma\mid X_T)&=\mathbb{E}_\lambda\left[\mathbb{1}_{T<\zeta}\mathbb{P}_\lambda(X_{T+\cdot}\in\Gamma\mid\mathcal{F}_T)\mid X_T\right]\\&=\mathbb{1}_{T<\zeta}\mathbb{P}_\lambda(X_{T+\cdot}\in\Gamma\mid\mathcal{F}_T),\end{aligned}$$

et on a prouvé (5).

Montrons maintenant que (a) implique (c). Il découle de (a) que pour tout $A\in\mathcal{F}_T$ et $B\subset I$,

$$\begin{aligned}\mathbb{E}_\lambda\left[\mathbb{1}_{X_T\in B}\mathbb{P}_\lambda(A\cap\{X_{T+\cdot}\in\Gamma\}\mid X_T)\right]&=\mathbb{P}_\lambda(A\cap\{X_{T+\cdot}\in\Gamma\}\cap\{X_T\in B\})\\&=\mathbb{E}_\lambda\left[\mathbb{1}_A\mathbb{1}_{X_T\in B}\mathbb{P}_\lambda(X_{T+\cdot}\in\Gamma\mid\mathcal{F}_T)\right]\\&=\mathbb{E}_\lambda\left[\mathbb{1}_A\mathbb{1}_{X_T\in B}\mathbb{P}_{X_T}(\Gamma)\right]\\&=\mathbb{E}_\lambda\left[\mathbb{1}_{X_T\in B}\mathbb{P}_{X_T}(\Gamma)\mathbb{P}_\lambda(A\mid X_T)\right].\end{aligned}$$

Puisque cette relation est vraie pour tout $B\subset I$, les v.a. $\sigma(X_T)$-mesurables $\mathbb{P}_\lambda(A\cap\{X_{T+\cdot}\in\Gamma\}\mid X_T)$ et $\mathbb{P}_{X_T}(\Gamma)\mathbb{P}_\lambda(A\mid X_T)$ coïncident p.s. sur $\{X_T\not\in I\}=\{T<\zeta\}$. On a donc prouvé (c). □

Démonstration du théorème 7.5 Pour tout $m\geq 0$, on pose

$$\mathcal{G}_m:=\sigma(Y_0,\dots,Y_m,S_1,\dots,S_m).$$

Cette tribu représente l'information disponible lorsque l'on obrserve que les m premiers sauts du processus. La preuve du lemme suivant sera donnée à la fin de cette preuve.

Lemme 7.7 *Soit T un temps d'arrêt et $A\in\mathcal{F}_T$. Alors pour tout $m\geq 0$, il existe une v.a. T_m et un événement A_m $\mathcal{G}_m$-mesurables tels que $T=T_m$ et $A=A_m$ p.s. sur l'événement $\{T<J_{m+1}\}$.*

Pour simplifier les écriture, on note $\widetilde{X}_t:=X_{T+t}$ pour tout $t\geq 0$. Le processus $\widetilde{X}$ a ses trajectoires dans Ω, et on note $(\widetilde{Y}_n,n\geq 0)$ sa chaîne incluse et $(\widetilde{S}_i,i\geq 1)$ ses temps de séjour.

On commence par réduire le problème : puisque $\mathcal{F}$ est engendré par les cylindres fini-dimensionnels (voir la note en bas de la page 9), eux-même étant engendrés par les événements de la forme

$$\Big\{Y_0=i_0,\dots,Y_n=i_n,S_1>s_1,\dots,S_n>s_n\Big\},$$

il suffit de démontrer que pour tout $A\in\mathcal{F}_T$, $i_0,\dots,i_n\in I$ et $s_1,\dots,s_n\in\mathbb{R}_+$,

$$\begin{aligned}&\mathbb{P}_\lambda(\{\widetilde{Y}_0=i_0,\dots,\widetilde{Y}_n=i_n,\widetilde{S}_1>s_1,\dots,\widetilde{S}_n>s_n\}\cap A\cap\{X_T=i\})\\&\qquad=\mathbb{P}_i(Y_0=i_0,\dots,Y_n=i_n,S_1>s_1\dots,S_n>s_n)\mathbb{P}_\lambda(A\cap\{X_T=i\}).\end{aligned}$$

Puisque $X_T = i$ implique que $T < \zeta$, il suffit en fait de le montrer en remplaçant $\{X_T = i\}$ par $\{X_T = i\} \cap \{J_m \leq T < J_{m+1}\}$ pour tout $m \geq 0$, puis sommer sur m.

D'après les propositions 7.3 et 7.4, on a $\{J_m \leq T\} \cap \{X_T = i\} \in \mathcal{F}_T$. Donc, si on pose $A' := A \cap \{J_m \leq T\} \cap \{X_T = i\}$, on doit montrer que

$$\begin{aligned} &\mathbb{P}_\lambda(\{\widetilde{Y}_0 = i_0, \dots, \widetilde{Y}_n = i_n, \widetilde{S}_1 > s_1, \dots, \widetilde{S}_n > s_n\} \cap A' \cap \{T < J_{m+1}\}) \\ &\qquad = \mathbb{P}_i(Y_0 = i_0, \dots, Y_n = i_n, S_1 > s_1 \dots, S_n > s_n)\mathbb{P}_\lambda(A' \cap \{T < J_{m+1}\}). \end{aligned}$$

Enfin, d'après le lemme 7.7, on peut écrire $T = T_m$ et $A' = A_m$ sur $\{T < J_{m+1}\}$, où T_m et A_m sont $\mathcal{G}_m$-mesurables. On arrive donc à la relation suivante, qui est celle que l'on démontre sans le reste de cette preuve :

$$\begin{aligned} &\mathbb{P}_\lambda(\{\widetilde{Y}_0 = i_0, \dots, \widetilde{Y}_n = i_n, \widetilde{S}_1 > s_1, \dots, \widetilde{S}_n > s_n\} \cap A_m \cap \{T_m < J_{m+1}\}) \\ &\quad = \mathbb{P}_i(Y_0 = i_0, \dots, Y_n = i_n, S_1 > s_1 \dots, S_n > s_n)\mathbb{P}_\lambda(A_m \cap \{T_m < J_{m+1}\}). \end{aligned} \tag{6}$$

Sur $\{J_m \leq T_m < J_{m+1}\}$, on a

$$\widetilde{Y}_n = Y_{m+n},\ \forall n \geq 0, \quad \widetilde{S}_1 = S_{m+1} - (T_m - J_m), \quad \widetilde{S}_n = S_{m+n},\ \forall n \geq 2.$$

De plus, d'après la définition 5.2, conditionnellement à $\{Y_m = i\}$, S_{m+1} est indépendant de la tribu $\mathcal{G}_m$, et donc de $T_m - J_m$ et A_m. Donc, d'après lampropriété d'absence de mémoire de la loi exponentielle,

$$\mathbb{P}_\lambda(S_{m+1} > s_1 + T_m - J_m \mid A_m \cap \{S_{m+1} > T_m - J_m\}) = e^{-q_i s_1} = \mathbb{P}_i(S_1 > s_1). \tag{7}$$

Finalement, par la propriété de Markov pour la chaîne incluse, puisque $S_{m+2}, \dots, S_{m+n}$ sont indépendants de $\mathcal{G}_m$ et de S_{m+1} conditionnellement à $Y_{m+1}, \dots, Y_{m+n}$ par la définition 5.2, et en utilisant (7), on obtient

$$\begin{aligned} &\mathbb{P}_\lambda(\{\widetilde{Y}_0 = i_0, \dots, \widetilde{Y}_n = i_n, \widetilde{S}_1 > s_1, \dots, \widetilde{S}_n > s_n\} \cap A_m \cap \{T_m < J_{m+1}\}) \\ &= \mathbb{P}_\lambda(\{Y_m = i_0, \dots, Y_{m+n} = i_n, S_{m+1} > s_1 + T_m - J_m, S_{m+2} > s_2, \dots, S_{m+n} > s_n\} \mid \\ &\qquad\qquad A_m \cap \{S_{m+1} > T_m - J_m\})\mathbb{P}_\lambda(A_m \cap \{S_{m+1} > T_m - J_m\}) \\ &= \mathbb{P}_i(Y_0 = i_0, \dots, Y_n = i_n, S_1 > s_1, \dots, S_n > s_n)\mathbb{P}_\lambda(A_m \cap \{T_m < J_{m+1}\}). \qquad \square \end{aligned}$$

Démonstration du lemme 7.7 Soit $t \geq 0$. On pose

$$\mathcal{A}_t = \{A \in \mathcal{F}_t : \exists A_m \in \mathcal{G}_m,\ A \cap \{t < J_{m+1}\} = A_m \cap \{t < J_{m+1}\}\}.$$

Il est clair que $\mathcal{A}_t$ est une tribu. De plus, pour tout $s \leq t$,

$$\begin{aligned} &\{X_s = i\} \cap \{t < J_{m+1}\} \\ &\qquad = \left\{ \left(\bigcup_{k=0}^{m-1} \{Y_k = i, J_k \leq s < J_{k+1}\} \right) \cup \{Y_m = i,\ J_m \leq s\} \right\} \cap \{t < J_{m+1}\}. \end{aligned}$$

On en déduit que $\{X_s = i\} \in \mathcal{A}_t$, et donc que

$$\mathcal{A}_t = \mathcal{F}_t. \tag{8}$$

Soit T un temps d'arrêt et $A \in \mathcal{F}_T$. On pose $B(t) = \{t < T\} \in \mathcal{F}_t$ et $A(t) = A \cap \{T \leq t\} \in \mathcal{F}_t$. Il suit de (8) qu'il existe $B_m(t) \in \mathcal{G}_m$ tel que $B(t) = B_m(t)$ sur $\{t < J_{m+1}\}$. Or, $B(t) \cap \{T < J_{m+1}\} \subset B(t) \cap \{t < J_{m+1}\}$, d'où $B(t) \cap \{T < J_{m+1}\} = B_m(t) \cap \{T < J_{m+1}\}$. Alors la v.a. $T_m = \sup_{t \in \mathbb{Q}_+} t \mathbb{1}_{B_m(t)}$ est $\mathcal{G}_m$-mesurable et vérifie

$$\begin{aligned} T_m \mathbb{1}_{T < J_{m+1}} &= \sup_{t \in \mathbb{Q}_+} (t \mathbb{1}_{B(t)}) \mathbb{1}_{T < J_{m+1}} \\ &= \sup_{t \in \mathbb{Q}_+} (t \mathbb{1}_{t < T}) \mathbb{1}_{T < J_{m+1}} = T \mathbb{1}_{T < J_{m+1}}, \end{aligned}$$

ce qu'il fallait démontrer pour T.

Soit maintenant $A \in \mathcal{F}_T$. Pour tout $0 < h < t$, on pose

$$A(t, h) := A \cap \{t - h < T \leq t\} \in \mathcal{F}_t.$$

Il suit de (8) qu'il existe $A_m(t, h) \in \mathcal{G}_m$ tel que $A_m(t, h) = A(t, h)$ sur $\{T < J_{m+1}\}$, et donc sur $\{T + h < J_{m+1}\}$. Soit

$$A_m(h) := \bigcup_{t \in \mathbb{Q},\ t > h} A_m(t, h) \quad \text{et} \quad A_m := \liminf_{n \to +\infty} A_m(\tfrac{1}{n}) = \bigcap_{i \geq 1} \left(\bigcup_{j \geq i} A_m(\tfrac{1}{j}) \right).$$

Alors pour tout $h' \geq h$,

$$A_m(h) \cap \{T + h' < J_{m+1}\} = \left(\bigcup_{t \in \mathbb{Q}, t > h} A(t, h) \right) \cap \{T + h' < J_{m+1}\} = A \cap \{T + h' < J_{m+1}\}.$$

Il en découle que

$$\begin{aligned} A_m \cap \{T < J_{m+1}\} &= \bigcup_{k \geq 1} \left(A_m \cap \{T + \tfrac{1}{k} < J_{m+1}\} \right) \\ &= \bigcup_{k \geq 1} \left\{ \bigcap_{i \geq k} \left[\left(\bigcup_{j \geq i} A_m(\tfrac{1}{j}) \right) \cap \{T + \tfrac{1}{k} < J_{m+1}\} \right] \right\} \\ &= \bigcup_{k \geq 1} \left\{ \bigcap_{i \geq k} \left(A \cap \{T + \tfrac{1}{k} < J_{m+1}\} \right) \right\} \\ &= A \cap \{T + \tfrac{1}{k} < J_{m+1}\}. \end{aligned}$$

D'où $A_m \cap \{T < J_{m+1}\} = A \cap \{T < J_{m+1}\}$. □

8 Équations forward et backward et autres caractérisations des CMTC

Le but de cette section est de donner plusieurs caractérisations alternatives des CMTC, qui peuvent être plus faciles à manipuler suivant les propriétés que l'on cherche à établir.

8.1 Cas I fini

On commence par supposer que I. On rappelle qu'un candidat pour la loi de la CMTC a été proposé dans la section 2 dans ce cas, à partir du semigroupe défini dans la section 1.

Théorème 8.1 *On suppose I fini. Soit $(X_t, t \geq 0)$ un processus dont les trajectoires sont dans Ω telles que $\zeta = \infty$ p.s. Soit Q une matrice de taux de transition sur I. On a équivalence entre les trois conditions suivantes.*

(a) *Pour tout $i \in I$, conditionnellement à $\{X_0 = i\}$, X est une CMTC de générateur Q et de loi initiale δ_i, au sens où sa chaîne incluse et sa suite des temps de séjour satisfont les propriétés de la définition 5.2.*

(b) **(définition infinitésimale)** *Pour tout $i \in I$ et $t, h \geq 0$, conditionnellement à $\{X_t = i\}$, X_{t+h} est indépendante de $(X_s, s \leq t)$ et, quand $h \to 0$, pour tout $j \in I$*

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = \delta_{i,j} + q_{i,j}h + o(h), \tag{9}$$

où le $o(h)$ vaut $h\varepsilon(i, j, t)$ où la fonction ω peut dépendre de (i, j, t) mais converge vers 0 lorsque $h \to 0$ uniformément en (i, j, t).

(c) **(probabilités de transition)** *Pour tout $n \geq 0$, $0 \leq t_0 \leq \ldots \leq t_{n+1}$ et $i_0, \ldots, i_{n+1} \in I$,*

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n,i_{n+1}}(t_{n+1} - t_n), \tag{10}$$

où $(p_{i,j}(t), i, j \in I, t \geq 0)$ est l'unique solution de l'équation forward

$$\begin{cases} P'(t) = P(t)Q & \text{pour tout } t \geq 0, \\ P(0) = Id. \end{cases}$$

Rappelons que lorsque I est fini, l'unique solution de l'équation forward est également l'unique solution de l'équation backward (théorème 1.4). Remarquons que la caractérisation (c) est la même que celle de (1). Elle donne donc la même construction de la loi de la CMTC que dans la section 2, avec la différence que les trajectoires du processus sont dans le sous-ensemble Ω de Ω_0. Remarquon que la caractérisation (b) justifie le nom de *générateur infinitésimal* pour la matrice Q, puisque pour tout vecteur $(v_i)_{i\in I}$, et pour tout $i \in I$,

$$(Qv)_i = \lim_{t\to 0} \frac{\mathbb{E}_i[v(X_t)] - v_i}{t},$$

où l'on note $v(i)$ pour v_i dans l'espérance. Rappelons enfin que pour une CMTC à valeurs dans I fini, $\zeta = \infty$ p.s. d'après le théorème 6.1.

Démonstration Montrons d'abord (a)⇒(b). D'après la propriété de Markov, (a) implique la propriété d'indépendance énoncée dans (b), et il suffit de montrer (9) pour $t = 0$ (et la propriété d'uniformité des $o(h)$ est également claire, puisqu'ils dépendent seulement de i et j qui parcourent un ensemble fini). Pour tout $i \in I$, on a

$$\mathbb{P}_i(X_h = i) \geq \mathbb{P}_i(J_1 > h) = e^{-q_i h} = 1 + q_{ii}h + o(h),$$

et pour tout $j \neq i$ dans I,

$$\mathbb{P}_i(X_h = j) \geq \mathbb{P}_i(J_1 < h, Y_1 = j, S_2 > h) = (1 - e^{-q_i h})\pi_{i,j}e^{-q_j h} = q_{i,j}h + o(h),$$

par définition de la matrice $\Pi = (\pi_{i,j})_{i,j\in I}$ dans (4). Pour obtenir les inégalités en sens inverse, on fait la somme de ces inégalités sur $j' \neq j$ et on utilise le fait que

$$\sum_{j' \neq j} \mathbb{P}_i(X_h = j) = 1 - \mathbb{P}_i(X_h = j').$$

Remarquons qu'on a utilisé pour cette dernière relation le fait que $\zeta = \infty$ p.s.

Montrons maintenant que (b)⇒(c). Soit $p_{i,j}(t) := \mathbb{P}_i(X_t = j)$. Pour tout $t, h \geq 0$, on a

$$\begin{aligned} p_{i,j}(t+h) &= \sum_{k\in I} \mathbb{P}_i(X_t = k)\mathbb{P}_k(X_{t+h} = j \mid X_t = k) \\ &= \sum_{k\in I} p_{i,k}(t)(\delta_{k,j} + q_{k,j}h + o(h)) = p_{i,j}(t) + \left(\sum_{k\in I} p_{i,k}(t)q_{k,j}\right) h + o(h). \end{aligned}$$

Donc $p_{i,j}$ est dérivable à droite. De plus, l'uniformité en t des $o(h)$ permet de remplacer t par $t - h$ dans la relation précédente, et on obtient de même la dérivabilité à gauche. Donc $p_{i,j}(t)$ satisfait l'équation forward, et le théorème 1.4 assure l'unicité.

Il nous reste à démontrer (10). Pour cela, on remarque que la propriété d'indépendance de (b) implique que, pour tout $i_0, \ldots, i_n \in I$ et $0 \leq t_0 \leq \ldots \leq t_n$, et pour tout $j \in I$ et $h \geq 0$,

$$\mathbb{P}(X_{t_n+h} = j \mid X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = \mathbb{P}(X_{t_n+h} = j \mid X_{t_n} = i_n). \qquad (11)$$

On pose alors pour tout $h, t \geq 0$ et tout $i, j \in I$ $\hat{p}^{(t)}_{i,j}(h) := \mathbb{P}(X_{t+h} = j \mid X_t = i)$. De la même manière que ci-dessus, on a

$$\hat{p}^{(t)}_{i,j}(s+h) = \hat{p}^{(t)}_{i,j}(s) + \left(\sum_{k\in I} \hat{p}_{i,k}(s)q_{k,j}\right) h + o(h),$$

et on conlut de la même manière que ci-dessus que $(\hat{p}^{(t)}_{i,j}(s), s \geq 0)_{i,j\in I}$ est solution de l'équation forward. L'unicité implique que $\hat{p}^{(t)}_{i,j}(s) = p_{i,j}(s)$, et on déduit (10) de (11).

Montrons enfin que (c)$\Rightarrow$(a). On sait qu'il existe un processus qui satisfait (a), construit dans les sections 3 et 5. On vient de démontrer que ce processus satisfait également (c). Or, comme remarqué en section 2, la propriété (c) caractérise les lois marginales fini-dimensionnelles de X (c.-à-d. les lois jointes de $X_{t_0}, \ldots, X_{t_n}$ pour tout $0 \leq t_0 \leq \ldots \leq t_n$), et donc la loi de X sur Ω muni de sa tribu cylindrique. En particulier, un processus X' qui satisfait (c) a la même loi que la CMTC X de (a), et donc sa chaîne incluse et ses temps de séjour ont même loi jointe que ceux de X. En particulier, cette loi satisfait la définition 5.2, et donc X' est bien une CMTC. $\square$

8.2 Cas I dénombrable

Le cas où I est dénombrable est plus délicat. La difficulté principale est de caractériser convenablement le semigroupe $P(t)$ solution des équations forward et backward. En effet, lorsque I est dénombrable, ces équations prennent la forme d'un système infini d'EDO, et on ne peut plus définir l'exponentielle de la matrice Q en toute généralité.

Théorème 8.2 *Soit I dénombrable et Q une matrice de taux de transition sur I. L'équation backward*

$$\begin{cases} P'(t) = QP(t) & \textit{si } t \geq 0, \\ P(0) = Id \end{cases}$$

admet une unique solution positive minimale[3] $(P(t), t \geq 0)$*. Cette solution form un semigroupe, c.-à-d. satisfait la relation matricielle $P(t+s) = P(t)P(s)$ pour tout $s, t \geq 0$. On l'appelle le* ***semigroupe associé à*** *Q. Ce semigroupe est également la solution positive minimale de l'équation forward*

$$\begin{cases} P'(t) = P(t)Q & \textit{si } t \geq 0, \\ P(0) = Id. \end{cases}$$

Une fois cette difficulté résolue, on obtient une caractérisation des probabilités de transition similaire à celle du théorème 8.1 (c).

Théorème 8.3 *On suppose I dénombrable. Soit $(X_t, t \geq 0)$ un processus dont les trajectoires sont dans Ω. Soit Q une matrice de taux de transition sur I, et $(P(t), t \geq 0)$ son semigroupe donné par le théorème précédent. On a équivalence entre les deux conditions suivantes.*

(a) *Pour tout $i \in I$, conditionnellement à $\{X_0 = i\}$, X est une CMTC de générateur Q et de loi initiale δ_i, au sens où sa chaîne incluse et sa suite des temps de séjour satisfont les propriétés de la définition 5.2.*

3. Au sens où toute autre solution positive lui est supérieure ou égale pour tout $t \geq 0$.

(b) **(probabilités de transition)** *Pour tout* $n \geq 0$, $0 \leq t_0 \leq \ldots \leq t_{n+1}$ *et* $i_0, \ldots, i_{n+1} \in I$,

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n, i_{n+1}}(t_{n+1} - t_n), \quad (12)$$

où $(p_{i,j}(t))_{i,j \in I} := P(t)$ *est l'unique solution minimale de l'équation backward.*

Nous donnons une preuve partielle de ces deux théorèmes à la fin de cette section.

Remarquons que rien ne dit que les matrices $P(t)$ soient stochastiques. Si ce n'est pas le cas, la CMTC peut exploser en temps fini, et d'après (b), pour tout $i \in I$,

$$\mathbb{P}_i(\zeta < t) = 1 - \sum_{j \in I} p_{i,j}(t).$$

La propriété du théorème 8.1 (b) est vraie sous certaines hypothèses pour une CMTC X sur I dénombrable (voir la proposition suivante), mais sans la propriété d'uniformité des $o(h)$. Cette propriété est indispensable pour pouvoir en déduire la formule des probabilités de transition, et donc on ne peut pas caractériser en général une CMTC sur I dénombrable par le développement limité des probabilités de transition.

Proposition 8.4 *Supposons* I *dénombrable. Si l'une des propriétés (a) ou (b) du théorème précédent est satisfaite et si pour tout* $i \in I$,

$$\sum_{k \in I} q_{i,k} q_k < \infty, \quad (13)$$

alors pour tout $t, h \geq 0$ *et* $i, j \in I$,

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) = \delta_{i,j} + q_{i,j} h + o(h).$$

Démonstration D'après la propriété de Markov forte, il suffit de le démontrer pour $t = 0$. En utilisant la même méthode que dans la preuve du théorème 8.1, on obtient des bornes inférieures sur $\mathbb{P}_i(X_h = j)$ qui ont le bon développement limité. Il suffit donc de montrer une borne inférieure, mais du fait que I est dénombrable, on ne peut pas utiliser la même astuce que dans la preuve du théorème 8.1. Soit donc $i, j \in I$ et $h \geq 0$,

$$\begin{aligned}
\mathbb{P}_i(X_h = j) &\geq \mathbb{P}_i(J_1 > h)\delta_{i,j} + \mathbb{P}_i(J_1 \leq h, Y_1 = j) + \mathbb{P}_i(J_1 \leq h, Y_1 \neq j, J_2 \leq h) \\
&= \delta_{i,j} e^{-q_i h} + (1 - e^{-q_i h})\pi_{i,j} + (1 - e^{-q_i h}) \sum_{k \neq j} \pi_{i,k}(1 - e^{-q_k h}) \\
&\leq \delta_{i,j}(1 + q_{i,i} h + o(h)) + h q_{i,j}(1 - \delta_{i,j}) + h^2 \sum_{k \notin \{i,j\}} q_{i,k} q_k,
\end{aligned}$$

où nous avons utilisé à plusieurs reprises l'inégalité $1 - e^{-x} \leq x$ dans la dernière ligne. L'hypothèse (13) permet de conclure. $\square$

Démonstration (partielle) des théorèmes 8.2 et 8.3 Supposons que X vérifie (a). On pose $p_{i,j}(t) := \mathbb{P}_i(X_t = j)$ et $P(t) := (p_{i,j}(t))_{i,j \in I}$.

Étape 1 : $P(t)$ satisfait l'équation backward. Conditionnellement à $\{X_0 = i\}$, $\mathbb{P}_i(X_t = j, \, t < J_1) = e^{-q_i t}\delta_{i,j}$ et pour tout $i, j, k \in I$,

$$\mathbb{P}_i(J_1 \leq t, \, X_{J_1} = k, \, X_t = j) = \int_0^t q_i e^{-q_i s} \pi_{i,k} p_{k,j}(t-s) \, ds.$$

On en déduit que $P(t)$ satisfait l'équation intégrale suivante :

$$\begin{aligned} e^{q_i t} p_{i,j}(t) &= \delta_{i,j} + e^{q_i t} \sum_{k \neq i} \int_0^t e^{-q_i s} q_{i,k} p_{k,j}(t-s) \, ds \\ &= \delta_{i,j} + \sum_{k \neq i} \int_0^t e^{-q_i u} q_{i,k} p_{k,j}(u) \, du, \end{aligned}$$

où la dernière ligne s'obtient par changement de variable $u = t - s$. Puisque les $p_{k,j}(u)$ sont bornés par 1, on en déduit par dérivation sous le signe somme que $p_{i,j}(t)$ est $\mathcal{C}^1$ et que

$$e^{q_i t}(q_i p_{i,j}(t) + p'_{i,j}(t)) = \sum_{k \neq i} e^{q_i t} q_{i,k} p_{k,j}(t). \tag{14}$$

D'où $p'_{i,j}(t) = \sum_{k \in I} q_{i,k} p_{k,j}(t)$.

Étape 2 : $P(t)$ est la solution positive minimale de l'équation backward. Soit $\widetilde{P}(t) = (\widetilde{p}_{i,j}(t))_{i,j \in I}$ une autre solution positive de l'équation backward. En particulier, $\widetilde{P}(t)$ est solution de (14), et donc

$$\widetilde{p}_{i,j}(t) = e^{-q_i t}\delta_{i,j} + \sum_{k \neq i} \int_0^t e^{-q_i s} q_{i,k} \widetilde{p}_{k,j}(t-s) \, ds.$$

D'autre part, un calcul similaire à celui de l'étape 1 montre que, pour tout $n \leq 0$,

$$\mathbb{P}_i(X_t = j, \, t < J_{n+1}) = e^{-q_i t}\delta_{i,j} + \sum_{k \neq i} \int_0^t e^{-q_i s} q_{i,k} \mathbb{P}_k(X_{t-s} = j, \, t - s < J_n) \, ds.$$

Puisque $\mathbb{P}_i(X_t = j, \, t < J_0) = 0 \leq \widetilde{p}_{i,j}(t)$, on montre par récurrence en comparant ces deux équations que, pour tout $n \geq 0$,

$$\mathbb{P}_i(X_t = j, t < J_n) \leq \widetilde{p}_{i,j}(t).$$

En faisant tendre n vers l'infini, on en déduit

$$p_{i,j}(t) = \mathbb{P}_i(X_t = j) = \mathbb{P}_i(X_t = j, t < \zeta) \leq \widetilde{p}_{i,j}(t).$$

Étape 3 : $(P(t), t \geq 0)$ est un semigroupe. Pour tout $s, t \geq 0$ et $i, j \in I$,

$$p_{i,j}(s+t) = \sum_{k \in I} \mathbb{P}_i(X_{t+s} = j \mid X_s = k)\mathbb{P}_i(X_s = k) = \sum_{k \in I} p_{i,j}(t) p_{j,k}(s),$$

où l'on a utilisé la propriété de Markov pour la dernière égalité.

Étape 4 : preuve de (12). Si $(X_t, t \geq 0)$ satisfait (a), la propriété de Markov appliquée au temps t_n implique que

$$\begin{aligned}\mathbb{P}(X_{t_{n+1}} = i_{n+1} \mid X_{t_0} = i_0, \ldots, X_{t_n} = i_n) &= \mathbb{P}_{i_n}(X_{t_{n+1}-t_n} = i_{n+1}) \\ &= p_{i_n, i_{n+1}}(t_{n+1} - t_n).\end{aligned}$$

Étape 5 : preuve de (b)⇒(a). Comme pour le théorème 8.1. [Remarquons que la caractérisation de $P(t)$ comme solution positive minimale de l'équation backward est suffisante pour conclure. On n'a donc pas besoin d'avoir démontré la dernière phrase du 8.2 pour conclure ici.]

Étape 6 : $P(t)$ est la solution positive minimale de l'équation forward. Admis. □

9 Conclusion

Nous n'abordons pas dans ce cours les propriétés asymptotiques des CMTC qui généralisent les propriétés classiques d'ergodicité et de convergence en temps long des CMTD. Cette théorie est très similaire au cas à temps discret, et passe par les notions d'irréductibilité, de récurrence nulle, récurrence positive et transience. Les résultats en temps continu sont globalement très similaires au cas discret, excepté pour tout ce qui concerne l'éventuelle périodicité de la chaîne, qui n'a plus d'influence sur les résultats de comportement en temps long en temps continu. On renvoie le lecteur au chapitre 3 de [2].

Références

[1] W. J. Anderson : *Continuous-time Markov chains.* Springer Series in Statistics : Probability and its Applications. Springer-Verlag, New York, 1991. An applications-oriented approach.

[2] J. R. Norris : *Markov chains*, vol. 2 de *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.