Quelques résultats sur les Processus de Markov Agrégés

Christiane COCOZZA-THIVENT et Xavier GUYON

Mars 1998

Résumé

Soit $X = \{X_t; t \in \mathbb{R}^+\}$ un processus markovien de sauts, irréductible, défini sur un espace d'états E fini. Soit $P = \{A_\alpha; \alpha = 1; N\}$ une partition de E. Le processus $\tilde{X}$ agrégé associé est défini par : $\tilde{X} = \alpha \iff X_t \in A_\alpha$. Nous résumons ici certaines propriétés du processus agrégé : temps de séjour et temps de séjours successifs, loi du processus agrégé $\tilde{X}$, identifiabilité de X à partir de $\tilde{X}$, c'est à dire que peut-on apprendre dur X à partir de $\tilde{X}$? On examine ensuite le cas où X est une chaîne de Markov. Ce travail présente, généralise et simplifie les résultats de Fredkin et altri ([?],[?]), Ball et altri [?], Rubino et Sericola [?] et Csenki [?].

Mots clefs : Processus markovien de sauts ; chaîne de Markov ; agrégation d'états ; temps de séjour ; identifiabilité paramétrique.

Table des matières

Soit $X = (X_t; t \geq 0)$ un processus de Markov indexé par $\mathbb{R}^+$, à valeurs dans un espace d'états fini $E = \{1; 2; \cdots; n\}$ (on parle aussi de processus markovien de sauts, et on notera p.m.s.). On supposera le processus homogène dans le temps et on notera Q la matrice $n \times n$ générateur infinitésimal de X (pour les résultats généraux sur les p.m.s., on consultera [?], §8). On supposera toujours le processus X irréductible (c.a.d. Q est une matrice irréductible). Puisque E est fini, X est ergodique, admettant une unique loi invariante π vérifiant $\pi Q = 0$. Si $P = \{A_\alpha; \alpha = 1; N\}$ est une partition de E, le processus agrégé $\tilde{X}$ associé est défnit par :

$$\tilde{X}_t = \alpha \iff \{X_t \in A_\alpha\}$$

On s'intéressera ici aux propriétés du processus $\tilde{X}$. La majeure partie des résultats présentés ici sont issus de [?], [?], [?], [?] et [?].

1 Temps de séjour dans un sous ensemble A

Soit $E = A \cup B$ une partition de E en deux parties. A cette partition, on associe la décomposition suivante de Q :

$$Q = \begin{pmatrix} Q_{AA} & Q_{AB} \\ Q_{BA} & Q_{BB} \end{pmatrix}$$

où, si on note n_F le cardinal de F, Q_{FG} est un bloc de dimension $n_F \times n_G$ canoniquement associé à $F \times G$. En $t = 0$, on suppose que $X_0 \in A$, et ceci avec les probabilités $\pi_A = (\pi_A(i); i \in A)$, vecteur ligne $n_A \times 1$ (on notera $X_0 \sim \pi_A$, et on a : $\sum_{i \in A} \pi_A(i) = 1$). Si $F \subset E$, 1_F est le vecteur colonne constitué de n_F fois la valeur 1. On note :

- T_A le (premier) temps de séjour en A

$$T_A = \inf\{t > 0; X_t \in B\}$$

- soit $i \in A$ et $j \in B$: la densité conditionnelle à $\{X_0 = i\}$, $f_{i;j}(t)$, de T_A en t pour une entrée en B par j,

$$P(T_A \leq t; X_{T_A} = j \mid X_0 = i) = \int_0^t f_{i;j}(u)du$$

- $f(t)$ la densité de T_A si $X_0 \sim \pi_A$ (bien que cela n'apparaisse pas, f dépend de π_A),

$$f = {}^t\pi_A f\, 1_B$$

On a le résultat suivant,

Proposition 1 Loi du temps de séjour dans A

$$f_{ij}(t) = \left(e^{Q_{AA}t}Q_{AB}\right)_{i;j} \quad \text{et} \quad f(t) = \pi_A e^{Q_{AA}t}Q_{AB}1_B$$

Démonstration :
Soit X^* le p.m.s. absorbant associée à X, ayant pour états absorbants chaque état de B : dès que l'entrée en B se fait par j, la chaîne reste en j. X^* a pour générateur infinitésimal et pour transition en t :

$$Q^* = \begin{pmatrix} Q_{AA} & Q_{AB} \\ 0 & 0 \end{pmatrix}, \quad P^*(t) = e^{Q^*t} = \begin{pmatrix} P^*_{AA}(t) & P^*_{AB}(t) \\ 0 & I \end{pmatrix}$$

avec $P^*_{AA}(t) = e^{Q_{AA}t}$. Pour $i \in A$ et $j \in B$, définissons

$$F_{ij}(t) = P\{T_A \leq t \text{ et entrée dans } B \text{ par } j \mid X_0 = i\}$$

D'après la définition de X^*,

$$F_{ij}(t) = P\{X^*_t = j \mid X^*_0 = i\} = (P^*_{AB}(t))_{ij}$$

Puisque P^* est liée à Q^* par la relation $P^{*\prime}(t) = P^*(t)Q^*$, on obtient

$$P^{*\prime}_{AB}(t) = P^*_{AA}(t)Q_{AB} = e^{Q_{AA}t}Q_{AB}$$

ce qui est le résultat annoncé puisque $f_{ij}(t) = F'_{ij}(t)$

On remarquera que dans la formule donnant $f(t)$, on peut remplacer $Q_{AB}1_B$ par $-Q_{AA}1_A$ puisque, en effet, chaque ligne de Q est de somme 0. On notera encore f_{ij} pour f_{ij}.

2 Temps de séjours successifs

Plus généralement, soit P la partition $E = A_1 \cup A_2 \cup \cdots \cup A_N$ de E en N sous ensembles, ou agrégats. On notera A_α un tel agrégat, les indices α, β repèreront les d'agrégats. Dans le contexte du paragraphe précédent, $N = 2$, $A = A_1$ et $B = A_2 = A^c$.

On suppose que $X_0 \in A_{\alpha_0}$ suivant la loi π_{α_0}. Supposons que les agrégats successivement visités sont $\alpha_0, \alpha_1, \alpha_2, \cdots, \alpha_r$, avec des temps de séjours successifs $T_{\alpha_k} \equiv T_{A_{\alpha_k}}$; $k = 0, r$. Notons $f_{\alpha_0,\alpha_1,\cdots,\alpha_r}(t_0, t_1, \cdots, t_r)$ la densité jointe sur $(R^{*+})^{r+1}$ des temps de séjours successifs pour la suite d'agrégats visités d'indices $(\alpha_0, \alpha_1, \cdots, \alpha_r)$

$$\begin{aligned} &P_{\pi_{\alpha_0}}(T_{\alpha_0} \leq t_0; X_{T_{\alpha_0}} \in A_{\alpha_1}; T_{\alpha_1} \leq t_1; X_{T_{\alpha_1}} \in A_{\alpha_2}; \cdots; T_{\alpha_r} \leq t_r) \\ &= \int_0^{t_0} \cdots \int_0^{t_r} f_{\alpha_0,\alpha_1,\cdots,\alpha_r}(u_0, u_1, \cdots, u_r)\, du_0 du_1 \cdots du_r \end{aligned}$$

On a le résultat suivant,

Proposition 2 Loi des temps de séjours successifs

$$f_{\alpha_0,\alpha_1,\cdots,\alpha_r}(t_0, t_1, \cdots, t_r) = \pi_{\alpha_0} e^{Q_{\alpha_0\alpha_0}t_0} Q_{\alpha_0\alpha_1} e^{Q_{\alpha_1\alpha_1}t_1} \cdots Q_{\alpha_{r-1}\alpha_r} e^{Q_{\alpha_r\alpha_r}t_r}(-Q_{\alpha_r\alpha_r}1_{\alpha_r})$$

Démonstration :

D'après la proposition précédente, $\pi_{\alpha_0} e^{Q_{\alpha_0\alpha_0}t_0} Q_{\alpha_0\alpha_1}$ est le n_{α_1} vecteur des probabilités d'entrée en A_{α_1} pour une durée de séjour en A_{α_0} de t_0. Le résultat annoncé s'obtient alors directement par récurrence sur r en appliquant la propriété de Markov forte à l'instant $T_{A_{\alpha_r}}$.

3 La chaîne de Markov des états d'entrée en A

3.1 Chaîne Z_P des états d'entrée dans les différents agrégats

Relativement à la partition P de E, on peut définir le processus à temps discret Z_P des états d'entrée de X dans les agrégats. C'est une chaîne de Markov dont la matrice de transition P_P sur E s'obtient ainsi : si $i \in A_\alpha$ et $j \in A_\beta, \alpha \neq \beta$, alors, par marginalisation par rapport au temps de séjour en A_α, la transition de i vers j vaut :

$$P_P(i,j) = \left[\int_0^{+\infty} e^{Q_{\alpha\alpha}t} Q_{\alpha\beta} dt\right]_{ij} = [-Q_{\alpha\alpha}^{-1} Q_{\alpha\beta}]_{ij} \text{ (cf également[?], proposition8.21)}$$

Relativement à la partition, les matrices blocs de P_P sont : $P_P(\alpha,\alpha) = 0$ pour tout α, et pour $\alpha \neq \beta$, $P_P(\alpha,\beta) = -Q_{\alpha\alpha}^{-1} Q_{\alpha\beta}$.

Ce résultat est une conséquence directe du résultat préliminaire suivant : soit A un sous ensemble de E, $A \neq E$. Alors (cf [?], lemme 2.2.1.; [?], propositions 9.13 et 9.14) :

Lemma 3 (1) La classe A est transiente pour le p.m.s. X^* (de la proposition 1) si et seulement si Q_{AA} est inversible.

(2) Si X est irréductible, Q_{AA} est inversible et ses valeurs propres sont à partie réelle strictement négatives. De plus, pour tout $k \in \mathbb{N}^*$

$$\int_0^{+\infty} t^k e^{Q_{AA}t} dt = (-1)^{k+1} k! (Q_{AA})^{-(k+1)}$$

3.2 Chaîne Z_A des états d'entrée en A

Si A est un sous ensemble propre de E, $A \neq \emptyset$, il définit la partition $E = A \cup B$ avec $B = A^c$. La suite Z_A des états d'entrée successifs dans A forme une chaîne de Markov dont la transition s'obtient par marginalisation en t et s de la loi des temps de séjours successifs en A puis en B, $f_{AB}(t,s)$. Plus précisement, notons, pour $i, j \in A$,

$$\{f_{AB}\}_{ij}(t,s) = [e^{Q_{AA}s} Q_{AB} e^{Q_{BB}t} Q_{BA}]_{ij}$$

cette probabilité pour une entrée en A par j, sachant $X_0 = i$. Soit P_A la matrice de transition de Z_A;

$$P_A = \int_0^{+\infty}\int_0^{+\infty} f_{AB}(t,s) dt ds$$

Puisque X est irréductible, on a, d'après la proposition 2 et le lemme précédent

Proposition 4 La matrice de transition des états d'entrée en A est

$$P_A = Q_{AA}^{-1} Q_{AB} Q_{BB}^{-1} Q_{BA}$$

Démonstration : Si $X_0 = i \in A$, la densité jointe des deux premiers temps de séjours $(T_A; T_B)$ en $(s; t)$ est, si la deuxième entrée en A se fait en j,

$$[e^{Q_{AA}s} Q_{AB} e^{Q_{BB}t} Q_{BA}]_{i;j}$$

Le résultat annoncé s'obtient alors à partir du lemme précédent par marginalisation en s et t.

Les éléments récurrents pour la chaîne Z_A sont ceux atteignables en un pas depuis B, définissant la "porte d'entrée en A", $R(A) = \{j \in A \mid \exists i \in A$ t.q. $Q_{ij} > 0\}$.

Loi invariante de P_A. Puisque E est fini et Q irréductible, X est ergodique (cf.[?], prop. 8.29). Ceci n'entraine pas automatiquement que la chaîne Z_A le soit. Par exemple, pour $A = \{1; 2\}; B = \{3; 4; 5\}$, pour ce classement, et $*$ notant un coéfficient non nul,

$$Q = \begin{pmatrix} * & 0 & * & 0 & * \\ 0 & * & 0 & * & 0 \\ 0 & * & * & 0 & * \\ * & 0 & 0 & * & 0 \\ 0 & * & * & 0 & * \end{pmatrix}; \quad P_A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Q est irréductible, donc ergodique, mais P_A est $2-$périodique, donc non ergodique.

Supposons que P_A soit ergodique, et notons $\pi = (\pi_A; \pi_B)$ la loi invariante de $(X; Q)$. En régime stationnaire, et sachant $X_0 \in B$, la loi de X_0 est $\nu_B = \pi_B / \pi_B 1_B$. La loi stationnaire de P_A qui est la loi d'entrée en A est (cf [?], lemme 9.16) :

$$\nu_{P_A} = \frac{\pi_B Q_{BA}}{\pi_B Q_{BA} 1_A}$$

Autres chaînes.

On peut définir d'autres chaînes : par exemple, si $E = A \cup B \cup C$ est une partition en trois agrégats, on peut étudier la chaîne des états d'entrée en A, sans s'autoriser à passer par C, de transition

$${}_C P_A(i; j) = P(\text{ retour en } A \text{ par } j \text{ sans jamais passer par } C \mid X_0 = i)$$

La propriété de Markov forte appliquée à T_A donne :

$$_CP_A(i;j) = \sum_{k \in B} P_i(X_{T_A} = k)P_k(X_{T_B} = j)$$

On en déduit, utilisant la chaine Z_P des états d'entrée dans les agrégats :

$$_CP_A = Q_{AA}^{-1}Q_{AB}Q_{BB}^{-1}Q_{BA}$$

4 Moments et corrélations des temps d'entrée

4.1 Moments du temps de séjour T_A

Si $X_0 \in A$ avec la loi π_A, la densité de T_A est $f(t) = -\pi_A e^{Q_{AA}t}Q_{AA}1_A$. On obtient donc par application directe du lemme 3,

$$\begin{cases} E(T_A) = -\pi_A Q_{AA}^{-1}1_A \\ E(T_A^k) = \pi_A(-1)^k k! Q_{AA}^{-k}1_A \\ Var(T_A) = \pi_A Q_{AA}^{-1}(2I - 1_A\pi_A)Q_{AA}^{-1}1_A \end{cases}$$

4.2 l-ième retour en A et corrélation

Si $B = A^c$; la densité des temps de séjours successifs en $A; B; A$ est $f_{A;B;A}(s;t;u) = -\pi_A e^{Q_{AA}s}Q_{AB}e^{Q_{BB}t}Q_{BA}e^{Q_{AA}u}Q_{AA}1_A$. On en déduit que la loi jointe des deux premiers temps de séjours en A est

$$f_{A(0);A(1)}(s;u) = \pi_A e^{Q_{AA}s}Q_{AA}P_A e^{Q_{AA}u}Q_{AA}1_A$$

Une extension immédiate donne, pour le l-ième temps de séjour en $A; l \geq 1$;

$$f_{A(0);A(l)}(s;u) = \pi_A e^{Q_{AA}s}Q_{AA}P_A^l e^{Q_{AA}u}Q_{AA}1_A$$

Si P_A est ergodique, et $\pi_A = \pi_{P_A}$, $T_{A(0)}$ et $T_{A(l)}$ ont même espérance et leur covariance vaut

$$cov_A(l) = \pi_A Q_{AA}^{-1}[P_A^l - 1_A\pi_A]Q_{AA}^{-1}1_A$$

Soit $M = \inf\{rang(Q_{AB}); rang(Q_{BA})\}$. P_A étant ergodique, $P_A^l \xrightarrow[l \to \infty]{} 1_A\pi_A$, et ceci avec une vitesse exponentielle. Si P_A est diagonalisable, de valeurs propres autres que 1, (λ_i);

$$P_A^l - 1_A\pi_A = \sum_{i=1}^{M-1} \alpha_i^A \lambda_i^l; \text{ avec } |\lambda_i| < 1; \ i = 1; M - 1$$

La covariance est un mélange d'au plus $(M-1)$ exponentielles. Sans hypothèse de diagonalisation sur P_A, la décomposition de Jordan donne, pour $K \le M-1$ et des polynomes P_i :

$$P_A^l - 1_A\pi_A = \sum_{i=1}^{K} \alpha_i^A \lambda_i^l P_i(l); \text{ avec } \sum_{i=1}^{K} \partial P_i = M - K - 1$$

5 Problèmes d'identifiabilité paramétrique

Soit $\tilde{X}$ le processus agrégé associé à la partition $E = A_1 \cup A_2 \cup \cdots \cup A_N$. Notons $n_\alpha = n_{A_\alpha}$. $\tilde{X}$ est à état dans $I = \{1, 2, \cdots, N\}$, caractérisé par :

$$\tilde{X}_t = \alpha \iff \{X_t \in A_\alpha\}$$

Supposons que la loi $X = (X_t)_{t \ge 0}$ (caractérisée par Q en situation stationnaire) dépende d'un paramètre $\theta \in \mathbb{R}^q$; avec $q \le n(n-1)$. Si $X_0 \sim \pi_{\alpha_0}$, si $(S_n)_{n \ge 1}$ est la suite des instants de saut de $\tilde{X}$; la loi $\tilde{P}$ de $\tilde{X}$ est définie par ses lois finies dimensionnelles de densités $f_{\alpha_0;\alpha_1;\cdots;\alpha_r}$: :

$$\begin{aligned} P_{\pi_{\alpha_0}}(S_1 \le t_0; \tilde{X}_{S_1} = \alpha_1; S_2 - S_1 \le t_1; \tilde{X}_{S_2} = \alpha_2; \cdots; \tilde{X}_{S_r} = \alpha_r; S_{r+1} - S_r \le t_r) \\ = \int_0^{t_0} \cdots \int_0^{t_r} f_{\alpha_0;\alpha_1;\cdots;\alpha_r}(u_0; u_1; \cdots; u_r) du_1 \cdots du_r \end{aligned}$$

$$f_{\alpha_0;\alpha_1;\cdots;\alpha_r}(t_0; t_1; \cdots; t_r) = \pi_{\alpha_0} e^{Q_{\alpha_0\alpha_0} t_0} Q_{\alpha_0\alpha_1} \cdots Q_{\alpha_{r-1}\alpha_r} e^{Q_{\alpha_r\alpha_r} t_r} (-Q_{\alpha_r\alpha_r} 1_{\alpha_r})$$

Supposons que cette loi dépende d'un paramètre $\varphi \in \mathbb{R}^k$, la loi de X dépendant d'un paramètre $\theta \in \mathbb{R}^q$ ($Q = Q(\theta)$ et $\varphi = \varphi(\theta)$) et que seul $\tilde{X}$ est observé.

La question préliminaire à toute étude statistique est la suivante :

"Quelle information $\tilde{X}$ apporte t-elle sur le processus X ?"

Ou encore : qu'est-ce que $\tilde{X}$ nous apprend sur le processus latent X ?

(1) Information sur la taille n_α des agrégats $A_\alpha; \alpha = 1, N$?

(2) Quelle majoration peut-on donner de la dimension k du modèle agrégé ?

(3) $\varphi = \varphi(\theta)$ permet-elle l'identification de θ ? Si non, quels paramètres de θ sont identifiables ?

Nous examinerons ces questions sous l'hypothèse simplificatrice suivante :

(D) : Chacunes des matrices $Q_{\alpha\alpha}$ est diagonalisable

On notera $((-\lambda_i^\alpha; V_i^\alpha); i = 1; n_\alpha)$ les valeurs/vecteurs propres de Q_α et on omettra l'incice α quand il n'y aura pas d'ambiguité.

5.1 Minoration de la taille n_α d'un agrégat

5.1.1 A partir de la densité f_α du temps de séjour en A_α

La décomposition spectrale de $Q_{\alpha\alpha}$ est, notant P_i^α la projection sur V_i, parrallèlement à $\oplus_{j\neq i}(V_j)$

$$Q_{\alpha\alpha} = -\sum_{i=1}^{n_\alpha} \lambda_i^\alpha P_i^\alpha$$

Il est alors aisé d'établir que

$$e^{Q_{\alpha\alpha}t} = \sum_{i=1}^{n_\alpha} e^{-\lambda_i^\alpha t} P_i^\alpha$$

Supposons alors que $X_0 \in A_\alpha$ avec une loi π_α fixée. Cette décomposition de $e^{Q_{\alpha\alpha}t}$ conduit à la formule suivante pour la densité de probabilité $f_\alpha(t)$ de $T_\alpha = T_{A_\alpha}$,

Proposition 5 Posons : $a_i^\alpha = \pi_\alpha P_i^\alpha q_\alpha$, avec $q_\alpha = -Q_{\alpha\alpha} 1_\alpha$. Alors :

$$f_\alpha(t) = \sum_{i=1}^{n_\alpha} a_i^\alpha e^{-\lambda_i^\alpha t}$$

Minoration de n_α. $t \to f_\alpha(t)$ étant un mélange de n_α-exponentielles, f_α est solution d'une équation différentielle linéaire et à coefficients constants d'ordre n_α. Soit p un entier positif, $t(p) = \{t_l; l = 1, p\}$; p réels non négatifs distincts et

$$W(f_\alpha; t(p)) = \det(f_\alpha^{(j-1)}(t_i))_{i,j=1,p}$$

Si $p < n_\alpha$, il existera un $t(p)$ t.q. $W(f_\alpha; t(p)) \neq 0$, alors que pour tout $p \geq n_\alpha$, $W(f_\alpha; t(p)) = 0$. On en déduit donc une méthode de minoration de n_α : si pour un p, on trouve un $t(p)$ t.q. $W(f_\alpha; t(p)) \neq 0$, alors $n_\alpha \geq p$. Remarquons que si $p < n_\alpha$, et si on choisit $t(p)$ au hasard (pour une mesure absolument continue par rapport à la mesure de Lebesgue), alors $W(f_\alpha; t(p)) \neq 0$ p.s.. Cette méthode permet donc d'identifier n_α p.s..

5.1.2 Evaluation de n_α à partir des covariances $l \to c_\alpha(l)$

Si P_{A_α} est diagonalisable, $l \to c_\alpha(l)$, fonction connue, est un mélange d'au plus $M_\alpha = \inf\{n_\alpha; n - n_\alpha\}$ exponentielles. On en déduit donc une autre procédure d'évaluation de n_α.

5.2 Majoration de la dimension paramétrique k de la chaîne cachée

5.2.1 Les densités bidimensionnelles $f_{\alpha\beta}$ engendrent la loi de $\tilde{X}$

Utilisant les décompositions spectrales de $\exp Q_{\alpha\alpha}$ et de $\exp Q_{\beta\beta}$, il est facile de voir que la densité des temps de séjours succésifs en A_α, puis en A_β, $\alpha \neq \beta$, est

$$f_{\alpha,\beta}(s,t) = \sum_{i=1}^{n_\alpha}\sum_{j=1}^{n_\beta} a_{i,j}^{\alpha,\beta} \exp -\{\lambda_i^\alpha s + \lambda_j^\beta t\}, \text{ avec } a_{i,j}^{\alpha,\beta} = \pi_\alpha P_i^\alpha Q_{\alpha\beta} P_j^\beta q_\beta$$

Plus généralement, la loi jointe $\tilde{\mathbb{P}}$ admet pour densités finie dimensionnelles, pour $\alpha_k \neq \alpha_{k+1}$, $k = 0, r-1$:

$$f_{\alpha_0,\alpha_1,\cdots,\alpha_r}(t_0,t_1,\cdots,t_r) = \sum_{\alpha_0=1}^{n_{\alpha_0}} \cdots \sum_{\alpha_r=1}^{n_{\alpha_r}} a_{i_0\cdots i_r}^{\alpha_0\cdots\alpha_r}\exp -\{\sum_{k=0} \lambda_{i_k}^{\alpha_k} t_k\}$$

avec $a_{i_0\cdots i_r}^{\alpha_0\cdots\alpha_r} = \pi_{\alpha_0} P_{i_0}^{\alpha_0} Q_{\alpha_0\alpha_1}\cdots Q_{\alpha_{r-1}\alpha_r} P_{i_r}^{\alpha_r} q_{\alpha_r}$.

Proposition 6 Supposons que $X_0 \in A_{\alpha_0}$ avec la probabilité π_{α_0}, et que pour tout $\{(\alpha,i); i=1,n_\alpha; \alpha=1,N\}$, $a_i^\alpha \neq 0$. Alors, les lois $f_{\alpha_0,\alpha_1,\cdots,\alpha_r}$, $r \geq 3$, s'obtiennent toutes à partir des densités bidimensionnelles $f_{\alpha\beta}$.

Démonstration : Elle repose sur le résultat suivant

Lemma 7 Soit $B = \{V_1, V_2, \cdots, V_n\}$ une base de R^n, P la projection sur V_1 parrallèlement à $\oplus_{i:i\neq 1}(V_i)$. Soit p et q deux vecteurs arbitraires de R^n. Alors :

$$Pq\,{}^tpP = ({}^tpPq)P$$

Démonstration du lemme : Il suffit d'expliciter P dans la base B et de constater l'égalités des deux menbres de l'égalité.

Notant alors que $P_{i_1}^{\alpha_1} q_{\alpha_1} \pi_{\alpha_1} P_{i_1}^{\alpha_1} = P_{i_1}^{\alpha_1}(\pi_{\alpha_1} P_{i_1}^{\alpha_1} q_{\alpha_1})$, on a :

$$a_{i_0\cdots i_r}^{\alpha_0\cdots\alpha_r} = \frac{[\pi_{\alpha_0}P_{i_0}^{\alpha_0}Q_{\alpha_0\alpha_1}(P_{i_1}^{\alpha_1}\{q_{\alpha_1}][\pi_{\alpha_1}P_{i_1}^{\alpha_1}\})Q_{\alpha_1\alpha_2}\cdots P_{i_r}^{\alpha_r}q_{\alpha_r}]}{\pi_{\alpha_1}P_{i_1}^{\alpha_1}q_{\alpha_1}}$$

c'est à dire

$$a_{i_0\cdots i_r}^{\alpha_0\cdots\alpha_r} = \frac{a_{i_0 i_1}^{\alpha_0\alpha_1} \times a_{i_1\cdots i_r}^{\alpha_1\cdots\alpha_r}}{a_{i_1}^{\alpha_1}}$$

Le résultat annoncé s'obtient par récurrence sur r.

Comme la densité $f_\alpha(s)$ s'obtient par marginalisation de $f_{\alpha\beta}(s,t)$ en t et

$$f_\alpha(s) = \sum_{\beta:\beta\neq\alpha}\int_0^{+\infty} f_{\alpha\beta}(s,t)\,dt$$

la loi de $\tilde{X}$ est complétement caractérisée par les densités bidimensionnelles $f_{\alpha\beta}$ et la loi initiale de X_0.

5.2.2 Une majoration de la dimension paramétrique de $\tilde{X}$

Si X est en régime stationnaire (de loi invariante π), $\tilde{X}$ l'est également, de loi stationnaire

$$\rho_\alpha = P(\tilde{X} = \alpha) = P(X \in A_\alpha) = \sum_{i \in A_\alpha} \pi_i; \quad \alpha = 1; N$$

Le lemme suivant montre que les densités bidimensionnelles $f_{\alpha\beta}$ caractérisent cette loi stationnaire ρ : en quelque sorte, si $\tilde{X}$ est en régime stationnaire, les densités bidimensionnelles sont caractéristiques de $\tilde{X}$. C'est la propriété que nous utiliserons pour l'évalation de la dimension de $\tilde{X}$.

Lemma 8 La loi ρ est caractérisée par les densités $f_{\alpha\beta}$.

Démonstration : Soit Z_P la chaîne des états d'entrée dans les agrégats de P, P_P sa transition. Sans condition sur P_P, il éxiste une loi π, déterminée de façon unique par la condition $\pi = \pi P_P$, limite de Cesaro des itérées de P_P (cf [?])

$$1_E \times \pi = \lim_{K \to \infty} \frac{1}{K} \sum_{k=0}^{K} P_P^k$$

Notons $F(\alpha_0; \alpha_1; \cdots; \alpha_r) = \int_{[0; \infty[^r} f_{\alpha_0\alpha_1\cdots\alpha_r}(t_0; t_1; \cdots; t_r) dt_0 dt_1 \cdots dt_r$ la probabilité de passages succésifs en $\alpha_1; \cdots \alpha_r$, si $X_0 \in \alpha_0$ suivant une loi ν_{α_0}. Utilisant la forme analytique de la densité $f_{\alpha_0\alpha_1\cdots\alpha_r}$, on obtient

$$F(\alpha_0; \alpha_1; \cdots; \alpha_r) = \nu_{\alpha_0}[-Q_{\alpha_0\alpha_0}]^{-1} Q_{\alpha_0\alpha_1} \cdots [-Q_{\alpha_{r-1}\alpha_{r-1}}]^{-1} Q_{\alpha_{r-1}\alpha_r} 1_{\alpha_r}$$

Comparant avec l'expression de la transition P_P, on obtient, pour la loi du $r^{ième}$ état d'entrée :

$$\rho_r(\alpha_r \mid \nu_{\alpha_0}) = \sum_{\alpha_1; \cdots; \alpha_{r-1}} F(\alpha_0; \alpha_1; \cdots; \alpha_r) = \sum_{i \in A_{\alpha_0}} \sum_{j \in \alpha_r} \nu_{\alpha_0}(i) P_P^r(i; j)$$

On obtient donc pour la somme de Césaro des $\rho_r(\alpha_r \mid \nu_{\alpha_0})$:

$$\rho(\alpha_r) = \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \rho_r(\alpha_r \mid \nu_{\alpha_0}) = \sum_{i \in A_{\alpha_0}} \sum_{j \in \alpha_r} \nu_{\alpha_0}(i) \pi_j$$

Il suffit alors de constater que cette limite ne dépend, comme F, que des lois bidimensionnelles $(f_{\alpha\beta})_{\alpha \neq \beta}$

Notons, pour $\alpha \neq \beta$, $r_{\alpha\beta} = \mathrm{rang}(Q_{\alpha\beta})$; on a $r_{\alpha\beta} \leq \inf\{n_\alpha; n_\beta\}$.

Proposition 9 On suppose que le processus X est en régime stationnaire, que (D) est vérifiée ainsi que l'hypothèse de non nullité des a_i^{α}. Alors, on a la majoration suivante de la dimension paramétrique de la chaîne agrégée $\tilde{X}$

$$k = \dim \tilde{X} \le \sum_{(\alpha,\beta):\alpha\neq\beta} r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta})$$

Démonstration : Puique les lois $f_{\alpha\beta}$ caractérisent $\tilde{X}$, il faut compter les paramètres de ces lois, et retrancher le nombre de contraintes libres existant entre ces densités.

• Il y a $n = \sum_\alpha n_\alpha$ paramètres exponentiels λ_i^α

•• Il y a au plus $\sum_{(\alpha,\beta):\alpha\neq\beta} r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta})$ paramètres $a_{ij}^{\alpha\beta}$.

En effet, $Q_{\alpha\beta}$ étant de rang $r_{\alpha\beta}$ s'écrit, pour des vecteurs $u_l \in \mathbb{R}^{n_\alpha}$, $v_l \in \mathbb{R}^{n_\beta}$, $l = 1, r_{\alpha\beta}$: $Q_{\alpha\beta} = \sum_{l=1,r_{\alpha\beta}} u_l\, {}^t v_l$. Revenant à la forme analytique des $a_{ij}^{\alpha\beta}$,

$$a_{ij}^{\alpha\beta} = \sum_{l=1}^{r_{\alpha\beta}} (\pi_\alpha P_i^\alpha u_l)({}^t v_l P_j^\beta q_\beta)$$

on obtient que pour la matrice A $n_\alpha \times n_\beta$, $A = (a_{ij}^{\alpha\beta}) = CD$ où C est une matrice $n_\alpha \times r_{\alpha\beta}$ et D est $r_{\alpha\beta} \times n_\beta$. On a donc $\mathrm{rang}(A^{\alpha\beta}) \le r_{\alpha\beta}$. L'évaluation annoncée résulte alors du fait q'une matrice $n \times m$ de rang r dépend d'au plus $r(n + m - r)$ paramètres.

Examinons maintenant les contraintes existant entre les différentes $f_{\alpha\beta}$.

◦ Pour tout α, $\sum_{\beta:\beta\neq\alpha} \int_{[0,+\infty[^2} f_{\alpha\beta}(s,t)\,ds\,dt = 1$: cela fait N contraintes.

◦◦ Contraintes entre les $f_{\alpha\beta}$ et les $f_{\beta\alpha}$. Puique l'on est en régime stationnaire, il y a deux façon d'obtenir f_α : l'une par marginalisation depuis le futur, l'autre depuis le passé,

$$f_\alpha(s) = \sum_{\beta:\beta\neq\alpha} \int_0^{+\infty} f_{\alpha\beta}(s,t)\,dt \quad \text{et} \quad f_\alpha(s) = \sum_{\beta:\beta\neq\alpha} \rho_\beta \int_0^{+\infty} f_{\beta\alpha}(t,s)\,dt$$

Comme f_α est un mélange de n_α exponentielles indépendantes, cela conduit pour chaque α à $(n_\alpha - 1)$ contraintes, étant donnée qu'il y a une contrainte de densité $\int_0^{+\infty} f_\alpha(s)\,ds = 1$, redondante avec celles déjà comptabilisée antérieurement. On obtient alors $\sum_\alpha (n_\alpha - 1) = (n - N)$ contraintes supplémentaires.

Le bilan est donc le suivant : il y a au plus

$$n + \sum_{(\alpha,\beta):\alpha\neq\beta} r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta}) - N - (n - N)$$

paramètres, ce qui est le résultat annoncé.

Condition nécessaire d'identifiabilité :

Une condition nécéssaire pour que l'observation de $\tilde{X}$ rende $(X;Q)$ identifiable est donc que $\dim Q \leq \sum_{(\alpha;\beta):\alpha\neq\beta} r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta})$. Puisque $r_{\alpha\beta} \leq \inf\{n_\alpha; n_\beta\}$, $r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta}) \leq n_\alpha n_\beta$. On a donc la majoration plus grossière :

$$\dim(\tilde{X}) \leq 2 \sum_{1\leq\alpha<\beta\leq N} n_\alpha n_\beta$$

et la condition nécessaire moins stricte : $\dim Q \leq 2\sum_{1\leq\alpha<\beta\leq N} n_\alpha n_\beta$. La quantité majorante s'interprète facilement : c'est le nombre de termes de la matrice Q de taille $n \times n$ à laquelle on a enlevé les blocs diagonaux $Q_{\alpha\alpha}$, $\alpha = 1;N$. Si chaque agrégat est à un seul point, c'est le nombre de termes hors diagonale de Q, soit encore $\dim X$ si il n'y a pas de contrainte sur Q.

6 Quelques exemples

6.1 Modèle de canaux ioniques

Si par exemple il y a deux agrégats $E = A_0 \cup A_1$ (c'est le cas des modèles de canaux ioniques, cf [?], [?]), l'observation du processus agrégé permet d'estimer au plus $2n_0n_1$ paramètres. A l'autre extrême, si chaque agrégat est réduit à un élément ($N = n$), $X \equiv \tilde{X}$ et on retrouve que $k = \dim Q \leq n(n-1)$.

6.2 Processus markovien de sauts à régime caché.

Soit X un processus markovien de saut à r-états, $E = \{1;2;\cdots;r\}$, de générateur $Q^X = (q_{ij})$. Soit le processus markovien de saut, conditionnel à X, $(Y \mid X)$, à k-états $F = \{1;2;\cdots;k\}$, de transitions définies par :

$$P(Y_t = m \mid (X_u = x_u)_{0\leq u<t}; X_t = j; Y_0 = l) = a^j_{l;m} \times t + o(t)$$

Il est facile de vérifier que $(X;Y)$ est un processus de saut markovien à rk états $E \times F$; de générateur Q^{XY}

$$\begin{cases} \text{Si } i \neq j : q_{XY}((i;l);(j;l)) = q_{ij} \\ \text{Si } l \neq m : q_{XY}((i;l);(i;m)) = a^i_{lm} \\ \text{Si } i \neq j \text{ et } l \neq m : q_{XY}((i;l);(j;m)) = 0 \\ q_{XY}((i;l);(i;l)) = -\sum_{j:j\neq i} q_{ij} + \sum_{m:m\neq l} a^i_{lm} \end{cases}$$

La dimension paramétrique de X est $r(r-1)$; celle de $(Y \mid X)$ est $rk(k-1)$; enfin, celle de $(X;Y)$ est $r[(r-1) + k(k-1)]$.

Observer seulement Y (on parle de processus de Markov caché), équivaut à observer, pour le processus $(X;Y)$; l'appartenance a l'un des agrégats A_l de $E \times F$:

$$A_l = \{(i;l); i \in E\} = E \times \{l\}; \quad l = 1;k$$

La dimension paramétrique du processus caché est au plus $r^2k(k-1)$. Puisque $\dim(Q_{XY}) < r^2k(k-1)$, le modèle $(X;Y)$ est peut être identifiable à partir de la seule observation Y.
Condition suffisante d'identifiabilité : Il n'y a pas de condition suffisante générale assurant que l'observation de $\tilde{X}$ rend X identifiable : l'identifiabilité de X à partir de $\tilde{X}$ s'examinera au cas par cas.

6.3 Sous-systèmes en interaction

6.3.1 Description et réduction de l'espace

Considérons deux processus markoviens de sauts (p.m.s.), $X^1 = (X_t^1; t \geq 0)$ et $X^2 = (X_t^2; t \geq 0)$ indépendants. Alors le processus $X = ((X_t^1; X_t^2); t \geq 0)$ est un p.m.s.. Supposons que ces deux processus soient à valeurs respectivement dans $E_1 = \{1; \cdots; r_1\}$ et $E_2 = \{1; \cdots r_2\}$ et aient pour matrices génératrices respectives Q_1 et Q_2, alors la matrice génératrice Q de X est donnée, pour $(i_1; i_2) \neq (j_1; j_2)$, par ([?], proposition 9.1) :

$$Q((i_1; i_2); (j_1; j_2)) = \begin{cases} Q_1(i_1; j_1) & \text{si } i_2 = j_2; \\ Q_2(i_2; j_2) & \text{si } i_1 = j_1; \\ 0 & \text{sinon.} \end{cases}$$

Si $r_2 = 2$, c'est-à-dire $E_2 = \{1; 2\}$, la matrice Q_2 s'écrit :

$$Q_2 = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

Supposons que les éléments de $E_1 \times E_2$ soient rangés dans l'ordre suivant :

$$(1;1);\ (2;1);\ \cdots;\ (r_1;1); \ldots (1;2);\ (2;2);\ \cdots;\ (r_1;2)$$

et posons :

$$Q' = \begin{pmatrix} Q_1 & \alpha I_{r_1} \\ \beta I_{r_1} & Q_1 \end{pmatrix}:$$

Alors la matrice Q s'obtient en modifiant les termes diagonaux de Q' pour que la somme des lignes soit égale à 0.

Supposons maintenant que $X^1 = (X_t^1; t \geq 0); \ldots; X^N = (X_t^N; t \geq 0)$ sont N p.m.s. indépendants qui représentent les évolutions au cours du temps de N sous-systèmes, alors $X = (X_t^1, \cdots, X_t^N)$ est un p.m.s.. Si chacun des processus X^k est à valeurs dans $\{1, 2\}$, la matrice génératrice Q de X s'obtient par induction, à partir de la construction ci-dessus (cette construction se généralise d'ailleurs sans difficulté au cas où $r_2 > 2$, voir par exemple [?], p. 285). Cela permet, à partir d'un logiciel tel que matlab, de construire Q de manière automatique.

Si tous les processus X^k ont même matrice génératrice, on peut agréger l'espace d'états $E = \{1, \cdots, r\}^N$, en regroupant les points de E qui ne diffèrent que par l'ordre de leurs composantes. L'espace ainsi réduit est noté $\bar{E}$. Notons que dans la situation présente, le processus agrégé $\bar{X}$, à valeurs dans $\bar{E}$, reste markovien : en effet, la condition d'agrégation forte (voir par exemple [?], chapitre 8, §7) est satisfaite.

Autorisons maintenant les interactions entre processus. Les X^k ne sont plus nécessairement indépendants : le taux de transition du processus X^k entre un état i_k et un état j_k ($1 \le i_k, j_k \le r$) peut dépendre des états des autres processus X^j. Si on suppose que les différents sous-systèmes se comportent de manière identique, c'est-à-dire que la transition de l'état $(i_1, \cdots, i_{k-1}, i_k, i_{k+1}, \cdots, i_N) \in E$ vers l'état $(i_1, \cdots, i_{k-1}, j_k, i_{k+1}, \cdots, i_N)$, ne dépend pas de k et de l'ordre dans lequel sont les i_j, alors on peut, comme précédemment, agréger les N-uples qui ne diffèrent que par l'ordre de leurs composantes et le processus agrégé $\bar{X}$, construit sur l'espace réduit $\bar{E}$ reste markovien.

Condition nécessaire d'identifiabilité. Examinons la condition d'identifiabilité de la proposition 9. Sans contraintes sur le comportement des sous systèmes, la dimension de X est $N r^N (r - 1)$. Cette dimension diminue si les sous systèmes se comportent identiquement, passant enfin à $r(r - 1)$ pour des systèmes indépendants si il n'y a pas de contraintes sur la transition initiale. Spécifions le cas $N = 2$. L'espace réduit $\tilde{E}$ est à $\frac{r(r+1)}{2}$ éléments :

$$\tilde{E} = \{(i, i); i = 1, r; \{(i, j), (j, i)\}; 1 \le i < j \le r\}$$

Si α et β sont réduits à un point, $Q_{\alpha\beta} = 0$. Si $\alpha = (i, i)$ et $\beta = \{(k, j), (j, k)\}$, $j \neq k$, $Q_{\alpha\beta}$ est de rang 1 si $k = i$, 0 sinon. Enfin, si α et β sont l'une et l'autre à deux éléments, $Q_{\alpha\beta} = 0$ sauf pour $\alpha = \{(i, j), (j, i)\}$ et $\beta = \{(i, k), (k, i)\}, k \neq i$, auquel cas $Q_{\alpha\beta}$ est de rang 2. On vérifie facilement que $\dim \tilde{X} \le 4r(r - 1) + 4r(r - 1)(r - 2)$. La réduction de E à $\tilde{E}$ laisse possible l'identifiabilité de X.

6.3.2 Processus de superposition

Nous considérons désormais N sous-systèmes identiques (au sens de ci-dessus) éventuellement en interaction. Supposons que (A, B) est une partition de $\{1,\cdots,r\}$. A chaque élément de la partition est associé un niveau (qui par exemple décrit la performance du sous-sytème lorsque son état appartient à l'élément de la partition considéré). Le niveau du système global est, par définition, égal à la somme des niveaux des sous-systèmes. Pour fixer les idées, le niveau correspondant à A sera pris égal à 1 et celui correspondant à B sera 0. Le niveau du système appartient donc à $\{0,1,\cdots,N\}$.

Notons D_m le sous-ensemble de l'espace dont les éléments correspondent à un même niveau m pour le système. Alors $(D_0, D_1, \cdots, D_N)$ est une partition de $\bar{E}$ à laquelle on peut appliquer les résultats précédemment décrits :

– soit en travaillant avec toute la partition, la matrice Q se décompose alors en :

$$Q = \begin{pmatrix} Q_{0,0} & Q_{0,1} & \cdots & Q_{0,N} \\ Q_{1,0} & Q_{1,1} & \cdots & Q_{1,N} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{N,0} & Q_{N,1} & \cdots & Q_{N,N} \end{pmatrix},$$

la matrice bloc $Q_{i,j}$ étant nulle si $|i-j| > 1$, car en une transition, seul un des sous sytèmes évolue, le niveau général, s'il est modifié, l'est donc de $+1$ ou de -1.

– soit en considérant la partition formée d'un ensemble D_m et de son complémentaire. Dans ce cas, par exemple, si ν_m est la restriction à D_m de la loi stationnaire du processus $\tilde{X}$, alors la mesure π_m, définie par

$$\pi_0 = \nu_1 Q_{10}, \quad \pi_N = \nu_{N-1} Q_{N-1,N},$$

$$\pi_m = \nu_{m-1} Q_{m-1,m} + \nu_{m+1} Q_{m+1,m} \qquad m = 1, 2, \cdots, N-1,$$

est stationnaire pour la chaine des états successifs d'entrée dans D_m, lorsque celle-ci est apériodique.

– soit en s'intéressant à deux niveaux m et n ($m \neq n$) et en prenant la partition $(D_m, D_n, (D_m \cup D_n)^c)$.

Vérifions que la condition d'identifiabilité de la proposition 9 est bien satisfaite dans le cas $N = 2$ et de la partition en trois parties. Notons a, b les cardinaux de A et B. Seuls le bloc $Q_{B^2, A\times B\cup B\times A}$ et ses trois analogues interviennent dans la majoration de dim $\tilde{X}$. Il est facile de voir que le majorant de cette dimension est $4ab(a^2+b^2)$, valeur toujours supérieure à la dimension de X (qui vaut $2(a+b)^2(a+b-1)$). Le processus après superposition laisse possible l'identification du processus initial X.

On peut aussi regarder, par exemple, le temps de séjour dans l'ensemble D_m, lorsque l'entrée dans cet ensemble correspond à une transition d'un sous sytème vers la classe A et que la sortie de cet ensemble correspond à une transition d'un sous-sytème vers la classe B (cf Yeo et al. [?]). Sa densité en régime stationnaire est :

$$\frac{\alpha_{m-1} Q_{m-1,m}}{\alpha_{m-1} Q_{m-1,m} \mathbb{1}_m} e^{Q_{m,m}t} Q_{m,m-1} \mathbb{1}_{m-1}.$$

Remarquons enfin que si le système est constitué de deux sous-systèmes et qu'on s'intéresse à l'un deux, qui est observable, alors que l'autre ne l'est pas, on doit alors travailler sur l'espace complet E (et non sur l'espace réduit $\bar{E}$). Si le système est constitué de $N > 2$ sous-systèmes identiques en interaction, dont uniquement l'un est observable, on peut réduire l'espace correspondant aux $N - 1$ sous-systèmes non observables et travailler ainsi sur un espace dont la taille est intermédiaire entre celle de E et celle de $\bar{E}$.

7 Agrégation : le cas de Chaînes de Markov

Nous allons examiner les questions étudiées précedemment dans le cas où le processus X est à temps discret : $X = \{X_l; l \in \mathbb{N}\}$ est une chaîne de Markov homogène sur un espace d'état fini E, de transition P, $\mathcal{P}$ une partition de E et $\tilde{X}$ est le processus à temps discret associé à $(X; \mathcal{P})$. On supposera X irréductible.

7.1 Chaîne absorbante

Soit $E = A \cup B$ une partition de E en deux sous ensembles, et la décomposition de la transition associée,

$$P = \begin{pmatrix} P_{AA} & P_{AB} \\ P_{BA} & P_{BB} \end{pmatrix}$$

On a le résultat suivant :

Proposition 10 (1) Si la classe A est transitoire (par exemple si P est irréductible), alors $(I - P_{AA})^{-1}$ existe, est à termes ≥ 0, valant

$$(I - P_{AA})^{-1} = \sum_{k \geq 0} P_{AA}^k$$

(2) Moments factoriels : pour tout entier $p \geq 1$,

$$\sum_{k \geq 0} k(k-1) \cdots (k-p+1) P_{AA}^k = p! P_{AA}^p (I - P_{AA})^{-(p+1)}$$

Démonstration : Soit X^* la chaîne de Markov absorbante de transition

$$P^* = \begin{pmatrix} P_{AA} & P_{AB} \\ 0 & I_B \end{pmatrix}$$

A est la classe transiente de X^*, et B sa classe absorbante. Supposons que $X^*_0 = i \in A$, et notons : m_i le nombre minimun de pas pour atteindre B depuis i ($m_i < \infty$), p_i la probabilité d'être encore dans A après m_i pas ($p_i < 1$), $p = \max\{p_i; i \in A\}$ ($p < 1$) et $m = \max\{m_i; i \in A\}$ ($m < \infty$).

Pour tout $i; j \in A$; $P^k_{AA}(i;j) \leq p^{[\frac{k}{m}]}$; ce qui garantit la convergence normale de la série $\sum_{k \geq 0} P^k_{AA}$. Il suffit alors de vérifier que $(I - P_{AA})(\sum_{k \geq 0} P^k_{AA}) = I$.

De même, la série (2) converge et l'égalité $\frac{\partial^p}{\partial P^p}\{I - P\}^{-1} = p!(I - P)^{-(p+1)}$ conduit au résultat.

7.2 Temps de séjours successifs

Soit A un sous ensemble de E ; si $X_0 \in A$, le temps de séjour en A est

$$T_A = \inf\{k \in \mathbb{N}^* \text{ t.q. } X_k \notin A\}$$

Notons que l'instant $t = 0$ est compté dans ce temps. Pour une partition $\mathcal{P}$ de E, et si $X_0 \in A_{\alpha_0}$, on notera $(T_{\alpha_0}; T_{\alpha_1}; \cdots; T_{\alpha_r})$ les $(r + 1)$ premiers temps de séjours successifs dans les agrégats $\alpha_0; \alpha_1; \cdots; \alpha_r$, $\alpha_l \neq \alpha_{l+1}$ pour $l = 0; r - 1$. Si $X_0 \in A_{\alpha_0}$, et ceci avec la densité $\pi_{\alpha_0} = (\pi(i); i \in A_{\alpha_0})$, on notera $f_{\alpha_0\alpha_1\cdots\alpha_r}(k_0; k_1; \cdots; k_r)$ la densité sur $(\mathbb{N}^*)^{r+1}$ de ces $(r + 1)$ temps. Pour $r = 0$, on notera $A_{\alpha_0} = A$, $T_{\alpha_0} = T_A$, $\pi_{\alpha_0} = \pi_A$. Notons :

$$\begin{aligned} f_{\alpha_0\alpha_1\cdots\alpha_r}(k_0; k_1; \cdots; k_r) &= P_{\pi_{\alpha_0}}(T_{\alpha_0} = k_{\alpha_0}; X_{k_0} \in A_{\alpha_0}; T_{\alpha_1} = k_1; X_{k_0+k_1} \in A_{\alpha_1}; \\ \cdots; T_{\alpha_{r-1}} &= k_{r-1}; X_{k_0+\cdots+k_{r-1}} \in A_{\alpha_r}; T_{\alpha_r} = k_r) \end{aligned}$$

On a le résultat suivant :

Proposition 11 Loi des temps de séjours.

(1) Supposons que $E = A \cup B$, $k \in \mathbb{N}^*$, $i \in A$ et $j \in B$. On a :

$$P(T_A = k; \text{ et entrée en } B \text{ par } j \mid X_0 = i) = [P^{k-1}_{AA} P_{AB}]_{ij}$$

En particulier, $P(T_A = k) = \pi_A P^{k-1}_{AA} P_{AB} 1_B$ (la loi de T_A est un mélange de n_α exponentielles) et le vecteur des probabilités d'entrée en B est

$$[P(\text{entrée en } B \text{ par } j)]_{j \in B} = \pi_A (I - P_{AA})^{-1} P_{AB}$$

(2) Si $X_0 \in A_{\alpha_0}$ avec la loi π_{α_0}

$$f_{\alpha_0\alpha_1\cdots\alpha_r}(k_0;k_1;\cdots;k_r) = \pi_{\alpha_0}P_{\alpha_0\alpha_0}^{k_0-1}P_{\alpha_0\alpha_1}\cdots P_{\alpha_{r-1}\alpha_r}P_{\alpha_r\alpha_r}^{k_r-1}(I - P_{\alpha_r\alpha_r})1_{\alpha_r}$$

(3) Si $E = A \cup B$, la chaîne de Markov des états d'entrée dans A admet pour matrice de transition

$$P_A = (I - P_{AA})^{-1}P_{AB}(I - P_{BB})^{-1}P_{BA}$$

(4) Si $E = A \cup B$, la loi jointe du premier et du $(l+1)^{ième}$ temps de séjour en A est

$$f_{T_0;T_l}(m;n) = \pi_A P_{AA}^{m-1}(I - P_{AA})P_A^l P_{AA}^{n-1}(I - P_{AA})1_A$$

Démonstration :

(1) résulte de l'égalité

$$P(X_l \in A \text{ pour } l = 1;k-1 \text{ et } X_k = j \mid X_0 = i) = \sum_{i_l \in A; l=1;k-1} P_{ii_1}P_{i_1i_2}\cdots P_{i_{k-2}i_{k-1}}P_{i_{k-1};j}$$

Les probabilités d'entrée en B s'obtiennent en notant que $\sum_{k\geq 1} P_{AA}^{k-1} = (I - P_{AA})^{-1}$.

(2) s'obtient directement par récurrence en remarquant que, pour $T_{\alpha_0} = k_0$, la probabilité d'entrée en A_{α_1} est $\pi_{\alpha_0}P_{\alpha_0\alpha_0}^{k_0-1}P_{\alpha_0\alpha_1}$, et que d'autre part, $\sum_{\beta:\beta\neq\alpha} P_{\beta\alpha}1_\alpha = (I - P_{\beta\beta})1_\beta$.

(3) Soit $(T_A;T_B)$ les deux premiers temps de séjour. On a pour i et $j \in A$:

$$P(T_A = k; T_B = l;\ 2^{ième} \text{ entrée en A par } j \mid X_0 = i) = [P_{AA}^{k-1}P_{AB}P_{BB}^{l-1}P_{BA}]_{ij}$$

Le résultat annoncé s'obtient par marginalisation en k et en l.

(4) s'obtient en utilisant (2) pour les temps de séjours successifs en A et B, et en marginalisant par rapport aux temps de séjours en B et au temps de séjour dans A autres que le premier et le dernier.

Remarques.

(1) Notons $A' = R(A) = \{j \in A \mid \exists i \notin A \text{ t.q. } P_{ij} \neq 0\}$ l'ensemble des états récurrents de P_A, $A'' = A \setminus A'$ les états transients. Supposons que $n_{A'} > 0$ et $n_{A''} > 0$, $P_A = \begin{pmatrix} P_{A'} & 0 \\ P_{A''} & 0 \end{pmatrix}$. [?] établissent la formule :

$$P_{A'} = \{I - P_{AA'} - P_{A'A''}(I - P_{A''A''})^{-1}P_{A''A'}\}^{-1}\{P_{A'B} + P_{A'A''}(I - P_{A''A''})^{-1}\}(I - P_{BB})^{-1}P_{BA}$$

Là où P_A demandait une inversion $n_A \times n_A$, $P_{A'}$ s'obtient à partir de deux inversions de tailles réduites, l'une $n_{A'} \times n_{A'}$ et l'autre $n_{A''} \times n_{A''}$. D'autre part, l'ergodicité de P_A est assurée par l'ergodicité de $P_{A'}$.

(2) La chaîne des états d'entrés dans les agrégats de la partition $\mathcal{P}$ admet pour transition,

$$\text{Pour } \alpha \neq \beta,\ i \in \alpha \text{ et } j \in \beta,\ P_{\mathcal{P}}(i;j) = [(I - P_{\alpha\alpha})^{-1}P_{\alpha\beta}]_{i;j}$$

les blocs $\alpha \times \alpha$ étant nuls.

(3) Supposons que $X_0 \in A$; et notons T_{l+1} le $(l+1)-$ième temps de séjour en A. On a :

$$P(T_0 = m; T_{l+1} = n) = \pi_A P_{AA}^{m-1}(I - P_{AA})P_A^l P_{AA}^{n-1}(I - P_{AA})1_A$$

Supposons que $\pi_A P_A = \pi_A$. Alors, T_0 et T_{l+1} ont même loi et leur covariance vaut

$$\mathrm{Cov}(T_0; T_{l+1}) = \pi_A(I - P_{AA})^{-1}[P_A^l - 1_A\pi_A](I - P_{AA})^{-1}1_A$$

Si P_A est ergodique, la covariance est un mélange de M exponentielles, où $M = \inf\{\mathrm{rang}(P_{AB}),\mathrm{rang}(P_{BA})\}$. De même, l'espérance, la variance et les différents moments de T_A s'obtiennent directement à partir des moments factoriels d'ordre p de la suite des $(P_{AA}^k)_k$.

7.3 Dimension paramétrique de la chaîne agrégée

L'ensemble des résultats établis pour le processus de Markov agrégé se transpose directement à la chaîne agrégée $\tilde{X}$, que ce soit pour l'identification des n_α, ou pour la dimension du modèle agrégé.

Dans une marginalisation temporelle, $\int_0^{+\infty}$ est remplacée par $\sum_{k \geq 0}$, $Q_{\alpha\beta}$ par $P_{\alpha\beta}$, les fonctions de base $e^{-\lambda_i^\alpha t}$ par $(\lambda_i^\alpha)^n$ (et réel$(\lambda_i^\alpha) > 0$ par $|\lambda_i^\alpha| < 1$), $e^{Q_{\alpha\alpha}t}$ par $P_{\alpha\alpha}^n$. Les définitions des coefficients a se correspondent alors, et sous l'hypothèse (D) de diagonalisation des différents $P_{\alpha\alpha}$, ainsi que la non nullité des $\{(a_i^\alpha)_{i;\alpha}\}$, on montre que les lois bidimensionnelles $\{(f_{\alpha\beta})_{\alpha;\beta}\}$ caractérisent la loi agrégée. En particulier, en régime stationnaire pour X, la loi marginale de $\tilde{X}$ est caractérisée par les $\{(f_{\alpha\beta})_{\alpha;\beta}\}$; et la dimension du modèle agrégé est

$$\dim \tilde{X} = \sum_{\alpha \neq \beta} r_{\alpha\beta}(n_\alpha + n_\beta - r_{\alpha\beta}) \text{ si } r_{\alpha\beta} = \mathrm{rang}(P_{\alpha\beta})$$

Références

[1] Ball F.G. et Rice J.A., 1992, Stochastic models for ion channels : introduction and bibliography. Math. Biosci. 112, p.189-206.

[2] Ball F., Milne R.K., Tame I.A. et Yeo G.F., 1997, Superposition of interacting aggregated continuous-time Markov chains, Adv. Appl. Prob. 29, p.56-91.

[3] Cocozza-Thivent Ch., 1997, Processus de Markov et …abilité des systèmes, Springer.

[4] Csenki A., 1992,The joint distribution of sejourn time in …nite Markov processes, Adv. Appl. Prob. 24, p.141-160.

[5] Fredkin D.R., Montal M. et Rice J., 1985, Identi…cation of aggregated markovian models : application to the nicotinic acetylcholine receptor. Proc. Berkeley Conf. Honor of J. Neyman and J. Kiefer, 1, ed. L. Le Cam et R. Olshen, p.269-290

[6] Fredkin D.R. et Rice J., 1986, On aggregated Markov processes, J. Appl. Prob. 23, p.208-214.

[7] Kemeny D. et Snell L.J., 1960, Finite Markov chains, Van Nostrand.

[8] Neuts M.F., 1981, Matrix-Geometric solutions in stochastic models, J. Hopkins Univ. Press, Baltimore.

[9] Rubino G. et Sericola B., 1989, Sojourn time in …nite Markov processes, J. Appl. Prob. 27, p.744-756.

[10] Yeo G.F., Edeson R.O., Milne R.K. et Madsen B.W., 1989, Superposition properties of independant ion channels, Proc. R. Soc. London B 238, 155-170