Une nouvelle description des possibilités d'agrégation d'une chaîne de Markov

James Ledoux

Une nouvelle description des possibilités d'agrégation exacte d'une chaîne de Markov finie

James Ledoux

INSA Rennes
20 Avenue des Buttes de Cöesmes
35043 Rennes cedex, FRANCE

Le modèle markovien reste l'outil analytique de base pour le calcul d'indicateurs de fiabilité et plus généralement de sûreté de fonctionnement. Cependant, le modèle développé peut présenter un très grand nombre d'états, ce qui le rend difficilement exploitable. Diverses techniques proposent d'agréger l'espace d'état selon des règles qui essaient de conserver au maximum l'information pertinente. Une question, au moins d'ordre théorique, est de déterminer les conditions initiales d'un modèle markovien homogène pour lesquelles, une fois certains états regroupés, le modèle agrégé possède toujours le caractère markovien homogène. Si une telle possibilité existe, on dit alors que le modèle original est faiblement agrégeable. Cette question est relativement ancienne et elle n'a connue de réponse précise que relativement récemment par les travaux de Rubino et Sericola [1] pour un modèle original irréductible. Pour un modèle d'un système dont certains états sont non-réparables, le problème est traité dans Ledoux et al. [2]. Dans tous ces travaux, l'objectif essentiel est de parvenir à calculer l'ensemble, noté $\mathcal{A}_{\mathcal{M}}$, des conditions initiales du modèle original qui permettent d'affirmer que le modèle déduit de l'agrégation d'états est encore markovien homogène. On montre alors que l'ensemble $\mathcal{A}_{\mathcal{M}}$ représente les vecteurs de probabilité du cône polyhédrique

$$\mathcal{C}_{\mathcal{M}} = \bigcap_{k=1}^{N} \mathcal{C}^k = \bigcap_{k=1}^{N} \{\alpha \geq 0 \ / \ \alpha H^{[k]} = 0\}$$

où N est le nombre initial d'états. Les éléments extrêmes du cône $\mathcal{C}_{\mathcal{M}}$ sont alors construits de manière incrémentale à partir du calcul des éléments extrêmes du cône $\mathcal{C}^1$. Dans Ledoux [3], certains invariants géométriques d'un modèle faiblement agrégeable sont mis en évidence. Ces propriétés permettent aussi bien de généraliser des résultats de [1], [2] ou [4] à des chaînes de Markov réductibles que de dériver des caractéristiques spectrales de la matrice des probabilités de transition du modèle original. L'invariant géométrique important est contenu dans le théorème suivant.

Théorème 1. *Soit $\mathcal{P} = \{C(1), \ldots, C(M)\}$ une partition en M classes de l'espace d'état du modèle de matrice des probabilités de transition P. Pour tout $l \in F = \{1, \ldots, M\}$, on définit la matrice $N \times M$ diagonale $I_l = diag(c_i)$ où $c_i = 1$ si $i \in C(l)$ et 0 autrement.*

L'ensemble $\mathcal{A}_\mathcal{M} \neq \emptyset$ ou $\mathcal{C}_\mathcal{M} \neq \{0\}$ ssi il existe une famille de M cônes polyhédriques $(\mathcal{C}_l)_{l\in F}$, distincts de $\{0\}$, tels que $\forall l, m \in F$

$$\begin{cases} \mathcal{C}_l \subseteq \mathcal{C}^1 I_l \\ \mathcal{C}_l P I_m \subseteq \mathcal{C}_m. \end{cases}$$

Lorsque la matrice des probabilités de transition est irréductible, quelque soit la distribution initiale, le modèle markovien "atteindra" un régime stationnaire caractérisé par la distribution de probabilité, notée π, invariante par P. Ainsi, s'il existe une certaine distribution initiale α donnant un processus agrégé $agr(\alpha, P, \mathcal{P})$ markovien homogène, alors π, choisie comme condition initiale, offre également cette possibilité. Ceci montre que tous les processus $agr(\alpha, P, \mathcal{P})$ avec $\alpha \in \mathcal{A}_\mathcal{M}$ partagent la même matrice des probabilité de transition $\widehat{P}$ et que l'ensemble $\mathcal{A}_\mathcal{M}$ est convexe. En résumé, si $\mathcal{A}_\mathcal{M} \neq \emptyset$ (resp. $\mathcal{C}_\mathcal{M} \neq \{0\}$) alors $\pi \in \mathcal{A}_\mathcal{M}$ et plus précisément $\mathcal{A}_\mathcal{M}$ (resp. $\mathcal{C}_\mathcal{M}$) contient un polytope noté $\mathcal{A}_\pi$ (resp. un cône $\mathcal{C}_\pi$) qui ne dépend que de π et de la partition choisie. L'objet principal de cette communication est de décrire de manière précise en quoi le cône $\mathcal{C}_\mathcal{M}$ peut différer du cône $\mathcal{C}_\pi$. Pour obtenir ce résultat, on utilise directement le fait que $agr(\alpha, P, \mathcal{P})$ est markovien homogène si et seulement si on a l'équivalence du processus $agr(\alpha, P, \mathcal{P})$ et de la chaîne de Markov $(\widehat{\alpha}, \widehat{P})$ sur l'espace d'état F. Si F^* désigne l'ensemble de toutes les suites finies d'éléments de F, on peut définir la matrice $N \times N$ suivante :

$$P^x = \begin{cases} I_{x_0} & \text{si } x = x_0 \\ I_{x_0} P I_{x_1} \cdots I_{x_{k-1}} P I_{x_k} & \text{si } x = x_0 x_1 \ldots x_k. \end{cases}$$

Maintenant, considérons le sous-espace vectoriel de R^N

$$\mathcal{N} = \{ \ v \in R^N \ / \ vP^x 1^t = 0, \forall x \in F^* \ \} = \oplus_{l\in F} \ [\mathcal{N} I_l].$$

Théorème 2. *Soit P une matrice irréductible et $\mathcal{P}$ une partition de l'espace d'état. Pour tout $l \in F$, on pose $k(l) = \mathrm{Dim}(\mathcal{N} I_l)$. Si $\mathcal{A}_{\mathcal{M}} \neq \emptyset$ (i.e $\mathcal{C}_{\mathcal{M}} \neq \{0\}$) alors*

$$\begin{aligned} \mathcal{C}_{\mathcal{M}} I_l &= (\mathrm{Vect}(\pi I_l) \oplus \mathcal{N} I_l) \cap R_+^N && (0.1) \\ \mathcal{C}_{\mathcal{M}} &= \oplus_{l \in F} \mathcal{C}_{\mathcal{M}} I_l && (0.2) \\ \mathrm{Dim}(\mathcal{C}_{\mathcal{M}} I_l) &= k(l) + 1. \end{aligned}$$

Ainsi, le cône $\mathcal{C}_{\mathcal{M}}$ est nécessairement donné par les formules (0.1) et (0.2). On construit le cône candidat à partir du calcul des bases des différents sous-epaces vectoriels $\mathcal{N} I_l$. Il ne reste plus qu'à vérifier que le cône obtenu satisfait les propriétés d'invariance du Théorème 1. Ce dernier point ne pose aucun problème une fois les rayons extrêmes du cône $\mathcal{C}_{\mathcal{M}}$ déterminés. Si tel est le cas, la famille de chaînes originales de matrice P est faiblement agrégeable et l'ensemble $\mathcal{A}_{\mathcal{M}}$ est la trace du cône sur le simplexe $\mathcal{A} = \{\alpha \in R^N \ / \ \alpha \geq 0 \text{ et } \alpha 1^t = 1\}$. Le Théorème 2 nous fournit les degrés de liberté autour du régime stationnaire, pour le choix d'une condition initiale autorisant une agrégation exacte d'une chaîne de Markov. En particulier, il caractérise le cas où $\mathcal{A}_{\mathcal{M}}$ se réduit à $\mathcal{A}_\pi = \mathrm{Vect}(\pi I_l, l \in F) \cap \mathcal{A}$: si $\mathcal{A}_{\mathcal{M}} \neq \emptyset$ alors

$$\mathcal{A}_{\mathcal{M}} = \mathcal{A}_\pi \text{ ssi } \mathcal{N} = \{0\}.$$

Cela précise totalement un résultat donné par Peng [4] (dans le cas irréductible) qui décrit la situation où l'agrégé est markovien si et seulement si il satisfait les équations de Chapman-Kolmogorov.

Bibliographie

[1] G. Rubino et B. Sericola (1991) A finite characterization of weak lumpable Markov processes. Part I : The discrete time case, Stoch. Proc. and Appl. vol 38 pp 195–204

[2] J. Ledoux, G. Rubino et B. Sericola (1994) Exact aggregation of absorbing of Markov processes using the quasi-stationary distribution, J. Appl. Probab. vol 31 pp 626–634.

[3] J. Ledoux (1995) Weak lumpability of general finite Markov chains and positive invariance of cones, Sousmis à J. Appl. Probab.

[4] N.P. Peng (1995) On Weak lumpability of a finite Markov chain, A paraître dans Statist. and Proba. Letters