

Grandes déviations en temps petit pour un processus de Markov à espace d'états dénombrable

Cécile ANÉ

Laboratoire de Statistique et Probabilités, Université Paul-Sabatier, 31062 Toulouse, France.
E-mail : ane@cict.fr

Résumé. Étant donné un processus de Markov X_t à valeurs dans un espace dénombrable, nous établissons un principe de grandes déviations pour la famille de processus $(X^\varepsilon(t)) = (X(\varepsilon t))$, $0 \leq t \leq 1$. La vitesse de convergence obtenue est beaucoup plus lente que dans le cas des diffusions.

*Large deviations in small time for a Markov process
with countable state space*

Abstract. *A large deviation principle for Markov processes on countable state spaces is proved. The speed is much slower than in the case of diffusions.*

1. Introduction et résultats de grandes déviations

Considérons un processus de Markov $(X_t)_{t \geq 0}$ à valeurs dans un espace dénombrable χ , avec, pour simplifier, un générateur L borné : pour tout $x \in \chi$, $\lambda(x) = |L(x, x)| \leq \Lambda$. Dans ce travail, nous établissons un principe de grandes déviations pour la famille de processus $(X_t^\varepsilon)_{0 \leq t \leq 1} = (X_{\varepsilon t})_{0 \leq t \leq 1}$. La fonction de taux est déterminée par la géométrie de χ pour le générateur L . La vitesse de convergence, quant à elle, est en $\log(1/\varepsilon)$. Elle est donc beaucoup plus lente que dans le cas des diffusions, qui ont d'ailleurs motivé notre étude (voir, par exemple, [1]). Ce comportement se détecte dès l'exemple de la probabilité de transition $p_\varepsilon(x, y)$ pour deux points x et y différents (mais reliés), puisque, pour un certain $\delta \in \mathbb{N}^*$,

$$\log p_\varepsilon(x, y) = \log \left(\sum_{n \geq 1} \frac{\varepsilon^n}{n!} L^n(x, y) \right) \sim \log(\varepsilon^\delta).$$

L'hypothèse de générateur borné (qu'il est possible de relaxer en une croissance contrôlée de λ autour du point de départ) est locale, et ceci s'explique par le fait que le comportement le plus probable pour le processus est celui du chemin identiquement égal à son point de départ.

Note présentée par Paul MALLIAVIN.

C. Ané

Un graphe orienté est associé au processus, avec χ comme ensemble de sommets, et avec comme arêtes les couples (x, y) tels que $L(x, y) > 0$. On note alors $\delta(x, y)$ (resp. $d(x, y)$) le nombre minimal d'arêtes orientées (resp. non orientées) nécessaires pour aller de x à y . Nous supposons, sans perte de généralité, que le graphe non orienté est connexe.

Rappelons que la trajectoire du processus X est une fonction de sauts, ayant des discontinuités aux instants T_n , $n \geq 1$, et prenant les positions successives Y_n , $n \geq 1$. La suite (Y_n) est une chaîne de Markov (avec des points cimetières éventuellement), dont la matrice de transition K est donnée par $K(x, y) = L(x, y)/\lambda(x)$ si $\lambda(x) > 0$. Sinon, x est un point cimetière (la chaîne s'arrête en ce point). Les instants de sauts sont déterminés par la propriété : $T_{n+1} - T_n$ suit la loi exponentielle de paramètre $\lambda(Y_n)$ conditionnellement à \mathcal{F}_{T_n} , \mathcal{F}_t étant la filtration naturelle associée au processus (voir [2] et [4]). Pour fixer les idées, on peut penser au processus de Poisson, ou bien au graphe \mathbb{Z}^d muni du générateur défini par $\lambda \equiv 1$ et par la matrice de transition K associée à la marche aléatoire simple sur \mathbb{Z}^d .

Fixons un point de départ $x_0 \in \chi$, et soit \mathcal{C} l'espace des trajectoires x càd-làg de $[0, 1]$ dans χ , issues de x_0 , à valeurs successives x_1, \dots, x_n telles que pour tout i , $\delta(x_{i-1}, x_i) = 1$. Pour deux chemins x et x' de \mathcal{C} , notons t_1, \dots, t_n (resp. $t'_1, \dots, t'_{n'}$) les instants de sauts de x (resp. x'), et x_1, \dots, x_n (resp. $x'_1, \dots, x'_{n'}$) leurs positions successives. On munit \mathcal{C} de la distance de Skorohod (voir [3], VI §5) :

$$\underline{d}(x, x') = \begin{cases} +\infty & \text{si } n \neq n' \text{ ou si } \exists i, x_i \neq x'_i, \\ \sup_{i \leq n=n'} |t_i - t'_i| & \text{sinon.} \end{cases}$$

Notons μ_ε la loi de X^ε .

THÉOREME 1. – *La famille $(\mu_\varepsilon)_{\varepsilon \rightarrow 0}$ vérifie un principe de grandes déviations de vitesse $\log(1/\varepsilon)$, et de fonction de taux S , où $S(x) = n$ si x admet exactement n sauts sur $[0, 1]$. Ainsi, pour toute partie mesurable A de \mathcal{C} ,*

$$\inf_A S \leq \liminf_{\varepsilon \rightarrow 0} \frac{\log \mu_\varepsilon(A)}{\log(1/\varepsilon)} \leq \limsup_{\varepsilon \rightarrow 0} \frac{\log \mu_\varepsilon(A)}{\log(1/\varepsilon)} \leq \inf_A S = \inf_{\bar{A}} S.$$

On en déduit facilement le « lemme de Varadhan » qui suit, mais que nous ne démontrerons pas. Le point clé n'est pas que la fonction de taux est bonne (c'est faux en général !), mais le fait qu'elle soit à valeurs dans un espace dénombrable.

COROLLAIRE 2. – *Soit $F : \mathcal{C} \rightarrow \mathbb{R}$ une fonction continue bornée. Alors la limite suivante existe :*

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\log(1/\varepsilon)} \log \left(\int_{\mathcal{C}} (1/\varepsilon)^{F(x)} \mu_\varepsilon(dx) \right) = \sup_{\mathcal{C}} (F - S).$$

Après la démonstration succincte du résultat de grandes déviations, quelques exemples d'ensembles de continuité et de discontinuité sont donnés. Ces ensembles sont souvent définis à partir des distances usuelles sur l'espace de chemins \mathcal{C} , qui sont d'ailleurs plus naturelles que \underline{d} .

2. Démonstrations

Le théorème principal s'obtient très facilement à partir de la proposition suivante qui étudie le comportement de probabilités de boules.

PROPOSITION 1. – Fixons des points x_0, \dots, x_n de \mathcal{X} , et notons $\lambda_i = \lambda(x_i)$.

(i) Fixons des instants $0 \leq t'_i < t''_i \leq 1$ pour $1 \leq i \leq n$. Soit D le domaine dans \mathbb{R}^n des points (v_1, \dots, v_n) vérifiant $v_1 + \dots + v_i \in]t'_i, t''_i[$ et $v_i \geq 0$ pour tout $i \leq n$. Supposons que D est de volume de Lebesgue V non nul. Alors,

$$\mathbb{P}^{x_0} \{ \forall i \leq n, Y_i = x_i \text{ et } T_i \in]\varepsilon t'_i, \varepsilon t''_i[\} \underset{\varepsilon \rightarrow 0}{\sim} V \lambda_0 \cdots \lambda_{n-1} \mathbb{P}^{x_0} \{ \forall i, Y_i = x_i \} \varepsilon^n.$$

De plus, $\mathbb{P}^{x_0} \{ \forall i \leq n, T_i \in]\varepsilon t'_i, \varepsilon t''_i[\} \underset{\varepsilon \rightarrow 0}{\sim} C \varepsilon^n$ pour une certaine constante $C = C(x_0, n)$.

(ii) Fixons un instant $t'_{n+1} \geq 0$ supplémentaire. Alors on a encore :

$$\mathbb{P}^{x_0} \{ \forall i \leq n, Y_i = x_i, T_i \in]\varepsilon t'_i, \varepsilon t''_i[\text{ et } T_{n+1} > \varepsilon t'_{n+1} \} \sim V \lambda_0 \cdots \lambda_{n-1} \mathbb{P}^{x_0} \{ \forall i, Y_i = x_i \} \varepsilon^n.$$

Démonstration. – (i) Par conditionnements successifs, il vient :

$$\mathbb{P}^{x_0} \{ \forall i \leq n, Y_i = x_i \text{ et } T_i \in]\varepsilon t'_i, \varepsilon t''_i[\} = K(x_0, x_1) \cdots K(x_{n-1}, x_n) \lambda_0 \cdots \lambda_{n-1} \varepsilon^n V(\varepsilon),$$

où $V(\varepsilon) = \int_D e^{-\varepsilon(\lambda_0 v_1 + \dots + \lambda_{n-1} v_n)} dv_1 \cdots dv_n$. On obtient ainsi, par convergence dominée, la première assertion de (i). Ensuite, $\mathbb{P}^{x_0} \{ \forall i \leq n, Y_i = x_i \text{ et } T_i \in]\varepsilon t'_i, \varepsilon t''_i[\}$ est la somme des probabilités précédentes, prise sur tous les (x_1, \dots, x_n) . Les équivalents peuvent s'additionner, car d'une part la série

$$\sum_{x_1, \dots, x_n \in \mathcal{X}} \lambda_0 \cdots \lambda_{n-1} K(x_0, x_1) \cdots K(x_{n-1}, x_n)$$

est convergente, et d'autre part $V(\varepsilon)$ converge uniformément vers V .

(ii) Par un calcul tout à fait similaire à celui de (i), on obtient :

$$\begin{aligned} \mathbb{P}^{x_0} \{ \forall i \leq n, Y_i = x_i, T_i \in]\varepsilon t'_i, \varepsilon t''_i[\text{ et } T_{n+1} > \varepsilon t'_{n+1} \} \\ = K(x_0, x_1) \cdots K(x_{n-1}, x_n) \lambda_0 \cdots \lambda_{n-1} \varepsilon^n \tilde{V}(\varepsilon), \end{aligned}$$

où $\tilde{V}(\varepsilon) = \int_D e^{-\varepsilon(\lambda_0 v_1 + \dots + \lambda_{n-1} v_n)} e^{-\varepsilon \lambda_n (0 \vee t'_{n+1} - (v_1 + \dots + v_n))} dv_1 \cdots dv_n$. On conclut encore par convergence dominée.

Démonstration du théorème principal. – La majoration s'obtient en écrivant que $A \subset \{S \geq \inf_A S\}$ et en utilisant la proposition (i) pour voir que

$$\frac{\log \mu_\varepsilon(\{S \geq \inf_A S\})}{\log(1/\varepsilon)} \xrightarrow{\varepsilon \rightarrow 0} -\inf_A S.$$

Pour la minoration, choisissons une boule $B = \{y; \underline{d}(x, y) < \eta\}$ incluse dans $\overset{\circ}{A}$ et centrée en un chemin x de longueur minimale. La proposition (ii) implique que $\log \mu_\varepsilon(B) / \log(1/\varepsilon)$ tend vers $-S(x) = -\inf_{\overset{\circ}{A}} S$ lorsque ε tend vers 0, et le résultat recherché en découle. Le théorème est établi.

3. Exemples

3.1. Ensembles de continuité

Ce sont les ensembles mesurables A pour lesquels il y a égalité entre $\inf_A S$ et $\inf_{\overset{\circ}{A}} S$. Afin d'en construire quelques uns, nous allons utiliser les distances usuelles sur \mathcal{C} , $d_\infty(x, y) =$

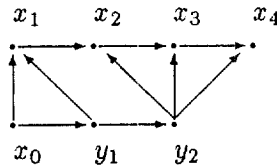
C. Ané

$\sup_{[0,1]} \{d(x(t), y(t))\}$, et $d_p(x, y) = \left(\int_0^1 d(x(t), y(t))^p dt \right)^{1/p}$ pour $1 \leq p < \infty$. Voici, sans démonstration, quelques exemples d'ensembles de continuité.

- Les ensembles cylindriques $A = \{x \in \mathcal{C}; (x(t_1), \dots, x(t_n)) \in E\}$, définis par des instants t_1, \dots, t_n et une partie E de χ^n .
- Les complémentaires de boules contenant le chemin x_0 , de la forme $A = \{y; d_p(x, y) > \varepsilon\}$ ou $A = \{y; d_p(x, y) \geq \varepsilon\}$ lorsque $d_p(x_0, x) < \varepsilon$. Ils sont de continuité si $p = \infty$, ou si $p \neq \infty$ et ε est assez petit ($\varepsilon^p < 1/2$). Dans ce dernier cas, ils vérifient $\inf_A S = 1$.
- Les boules autour de chemins « déviants » pour d_p , avec $p \neq \infty$. Si le chemin x vérifie $d_p(x_0, x) > \varepsilon$, alors les boules $\{y; d_p(x, y) < \varepsilon\}$ et $\{y; d_p(x, y) \leq \varepsilon\}$ sont de continuité pour ε petit, et avec $\inf_A S = S(x)$.

3.2. Contre-exemples

Pour la distance d_∞ , les boules autour de chemins « déviants » ne sont plus forcément de continuité. Cette propriété dépend de la structure du graphe. Pour le cube $\chi = \{-1, +1\}^N$ (arêtes orientées dans les deux sens), ces boules sont effectivement de continuité. C'est vrai aussi sur \mathbb{Z} . En revanche, pour le graphe suivant, ce n'est plus vrai. Le chemin x est à d_∞ -distance 3 de x_0 , et la partie $A = \{y; d_\infty(x, y) \leq 1\}$ vérifie $\inf_A S = 2$ et $\inf_\circ S = 3$.



Dans tous les cas, on a quand même l'existence de la limite de $\log \mu_\varepsilon(A) / \log(1/\varepsilon)$, lorsque $\varepsilon \rightarrow 0$, qui vaut $-\inf_\circ S$.

Remerciements. Je tiens à remercier vivement Michel Ledoux pour ses précieux conseils qui m'ont guidée tout au long de l'élaboration de ce travail.

Note remise le 26 mai 1997, acceptée après révision le 29 septembre 1997.

Références bibliographiques

- [1] Azencott R., 1980. Grandes déviations et applications, École d'Été de Probabilités de Saint-Flour VIII-1978, Lect. Notes in Math., 774, p. 1-176, Springer-Verlag.
- [2] Chung K. L., 1967. *Markov chains with stationary transition probabilities*, Springer-Verlag, 2^e ed..
- [3] Gihman I. I. et Skorohod A. V., 1974. *The theory of stochastic processes I*, Springer-Verlag.
- [4] Revuz D., 1975. *Markov chains*, North-Holland.