EMILE LE PAGE

Calcul des probabilités - Théorème des grands écarts et théorème de la limite centrale pour certains produits de matrices aléatoires

Publications des séminaires de mathématiques et informatique de Rennes, 1979, fascicule 1
« Séminaire de probabilités », , exp. nº 5, p. 1-8
<http://www.numdam.org/item?id=PSMIR_1979___1_A5_0>

CALCUL DES PROBABILITES. - *Théorème des grands écarts et théorème de la limite centrale pour certains produits de matrices aléatoires.*
Note(*) d'Emile LE PAGE transmise par

Soit $(g_n)_{n\geq 1}$ une suite de variables aléatoires à valeurs dans $SL(2,\mathbb{R})$, indépendantes et de même loi p à support compact ; on suppose de plus que le sous groupe fermé G_p engendré par le support de p est non compact et a ses sous groupes d'indice fini irréductibles. Notant $\|\ \|$ une mesure quelconque sur $\mathbb{R}^2$, on étudie pour $x \in \mathbb{R}^2-\{0\}$ la suite de variables aléatoires $(\mathrm{Log}\|g_n\, g_{n-1}\dots g_1 x\|)_{n\geq 1}$. On établit un théorème des grands écarts puis un théorème de la limite centrale.

Let $(g_n)_{n\geq 1}$ be a sequence of independent, identically distributed random variables with values in $SL(2,\mathbb{R})$, having a distribution with compact support ; moreover we suppose that the closed group G_p generated by the support of p is non compact and such that no subgroup of G_p of finite index is reductible. If $\|\ \|$ is some norm on $\mathbb{R}^2$ we study for all $x \in \mathbb{R}^2-\{0\}$ the sequence of random variables $(\mathrm{Log}\|g_n\, g_{n-1}\dots g_1 x\|)_{n\geq 1}$. We prove a theorem of large deviations, and a central limit theorem.

§1 - Résultats préliminaires

1-1 Considérons une probabilité p à support compact S_p portée par $SL(d,\mathbb{R})$ $d\geq 1$. Nous supposerons de plus que p admet une unique probabilité invariante ν sur l'espace projectif $P(\mathbb{R}^d)$ (condition (U)). Des conditions suffisantes sur le support de p pour que (U) soit réalisé sont données dans [3] et [5]. En particulier si le semi groupe fermé T_p engendré par le support de p est égal à $SL(d,\mathbb{R})$ ou si T_p est un réseau (U) est vérifiée.

Soit maintenant $(g_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi p et soit $(P(x,.))_{x\in P(\mathbb{R}^d)}$ la probabilité de transition définie de $P(\mathbb{R}^d)$ dans les boréliens de $P(\mathbb{R}^d)$ par

$$P(x,A) = \int 1_A\ (g\ x)\ p(dg)$$

Notons de plus σ l'application de $SL(d,\mathbb{R}) \times P(\mathbb{R}^d)$ dans $\mathbb{R}$ définie par

$$\sigma(g,\bar{u}) = \frac{\|g\ u\|}{\|u\|}$$

$g \in SL(d,\mathbb{R})$, $\bar{u} \in P(\mathbb{R}^d)$ étant l'image dans $P(\mathbb{R}^d)$ de $u \in \mathbb{R}^d-\{0\}$et $\|\ \|$ étant une norme quelconque sur $\mathbb{R}^d$.

Avec ces notations nous pouvons énoncer la

Proposition 1

Soit p une probabilité à support compact portée par $SL(d,\mathbb{R})$ et satisfaisant à la condition (U), alors la suite de fonctions $f_n(x) = \frac{1}{n} E \operatorname{Log}\|g_n \dots g_1 x\|$ converge uniformément sur $S_{d-1} = \{x/\|x\|=1\}$ vers $\gamma = \iint_{SL(d,\mathbb{R}) \times P(\mathbb{R}^d)} \operatorname{Log} \sigma(g,\bar{u})\ p(dg\ \nu(d\bar{u})$

Démonstration

La proposition 1 résulte du fait que pour toute suite $(u_n)_{n\geq 1} \in P(\mathbb{R}^d)$ la suite de probabilités $(\frac{1}{n} \sum_{k=1}^{n} P^k(u_n,.))_{n\geq 1}$ converge vaguement vers ν , car toute valeur d'adhérence de cette suite est une probabilité p invariante et est donc égale à ν .

1-2 On suppose désormais que d=2. On définit sur $P(\mathbb{R}^2)$ la distance δ compatible avec la topologie de $P(\mathbb{R}^2)$

$$\delta(\bar{x},\bar{y}) = |\sin \theta(x,y)| \qquad \bar{x},\ \bar{y} \in P(\mathbb{R}^2) \qquad x,y \in \mathbb{R}^2-\{0\}$$

où $\theta(x,y) = |\arg x - \arg y|$ $\quad 0 \leq \arg x < 2\pi$

$\qquad 0 \leq \arg y < 2\pi$

On considère également l'opérateur P_λ, $\lambda \geq 0$ de l'espace vectoriel $\mathcal{C}(P(\mathbb{R}^2))$ des fonctions continues sur $P(\mathbb{R}^2)$ muni de la topologie de la convergence uniforme sur $P(\mathbb{R}^2)$, défini par

$$P_\lambda f(\bar{x}) = \int \frac{1}{|\sigma(g,\bar{x})|^\lambda} f(g\ \bar{x})\ p(dg) \qquad \bar{x} \in P(\mathbb{R}^2)$$

$$f \in \mathcal{C}(P(\mathbb{R}^2))$$

Alors

THEOREME 1

Soit p une probabilité à support compact portée par $SL(2,\mathbb{R})$ telle que le groupe fermé G_p engendré par le support de p soit non compact et ait ses sous groupes d'indice fini irréductibles, et telle que la condition (U) soit satisfaite ; et soit $(g_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de même loi p, alors

1°) Il existe un $\lambda_0 > 0$ telle que pour $0<\lambda\leq\lambda_0$ la norme spectrale de l'opérateur P_λ soit strictement inférieure à 1

2°) $\overline{\lim_n}\,[\sup_{\bar{x},\bar{y}\,P(\mathbb{R}^2)} E\delta(g_n\cdots g_1\bar{x},\ g_n\cdots g_1\bar{y})]^{1/n} = \rho < 1$

Démonstration

Le 1°) se déduit de ce que pour $n\geq 0$ et $\bar{x} \in P(\mathbb{R}^2)$ on a

$$\frac{d}{d\lambda}\,[P^n\, 1(\bar{x})]\Big|_{\lambda=0} = -E\, \mathrm{Log}\| g_n\cdots g_1\, x\|$$

et ce que pour n assez grand $\inf_{x\in S_1} E\,\mathrm{Log}\| g_n\, g_{n-1}\cdots g_1\, x\| > 0$ ceci résultant de la proposition 1 et du fait que d'après le théorème (8-6) de [2]

$$\gamma = \iint_{SL(2,\mathbb{R})\times P(\mathbb{R}^2)} \mathrm{Log}\,\sigma(g,\bar{u})\ p(dg)\ \nu(d\bar{u}) > 0$$

Le 2°) se déduit du 1°) à l'aide des inégalités

$$\delta(g\bar{x},g\bar{y}) \leq [\delta(gx,gy)]^\lambda \leq \frac{[\delta(\bar{x},\bar{y})]^\lambda}{\|gx\|^\lambda\|gy\|^\lambda}$$

$$\bar{x},\bar{y} \in P(\mathbb{R}^2) \qquad x,y \in S_1 \qquad 0<\lambda<1$$

et de l'inégalité de Schwartz.

Avant d'énoncer 2 corollaires donnons quelques notations. Pour $0<\lambda\leq 1$ et pour toute fonction f bornée sur $P(\mathbb{R}^2)$ nous définissons $m_\lambda(f)$ par :

$$m_\lambda(f) = \sup_{\bar{x},\bar{y}\in P(\mathbb{R}^2)} \frac{f(\bar{x})-f(\bar{y})}{[\delta(\bar{x},\bar{y})]^\lambda}$$

et $\mathcal{L}_\lambda = \{y/\|f\|_\lambda = \sup_{x\in P(\mathbb{R}^2)} |f(x)| + m_\lambda(f) < +\infty\}$

$\mathcal{L}_\lambda$ est une algèbre de Banach unitaire muni de la norme $\|\ \|_\lambda$.

On a alors

Corollaire 1

Sous les hypothèses du théorème 1 il existe un $0<\lambda_0\leq 1$ tel que pour tout $0<\lambda\leq\lambda_0$ il existe un réel $0<\rho_1<1$ et un entier N_1 tels que $\forall f\in \mathscr{C}_\lambda$ $\forall n\geq N_1$ $\forall m\geq 0$ on ait les inégalités :

$$m_\lambda(P^n f) \leq m_\lambda(f)\ \rho_1^n$$

et $\sup_{\bar{x}\in P(\mathbb{R}^2)} |P^{n+m} f(\bar{x}) - P^n f(\bar{x})| \leq m_\lambda(f)\ \rho_1^n$

Corollaire 2

Sous les hypothèses du théorème 1 pour toute fonction $f \in \mathscr{C}(P(\mathbb{R}^2$

$$\lim_n \sup_{\bar{x}\in P(\mathbb{R}^2)} |P^n f(x) - \nu(f)| = 0$$

§2 - Un théorème des grands écarts

On énonce un théorème de grands écarts pour une chaine semi-markovienne puis nous l'appliquons au cas d'une marche aléatoire sur $SL(2,\mathbb{R})$.

2-1 Soient un espace compact X à base dénombrable, P une probabilité de transition de X dans les boréliens $\mathcal{B}_X$ de X, F une probabilité de transition de X x X dans les boréliens $\mathcal{B}_\mathbb{R}$ de $\mathbb{R}$. On considère une chaine semi markovienne (X_n,U_n) à valeurs dans $X \times \mathbb{R}$ de probabilité de transition $\tilde{Q}$ définie par

$$\tilde{Q}(x,u,A,B) = \tilde{Q}(x,0,A,B) = Q(x,A,B) = \int_X P(x,dy)\ 1_A(y) \int_\mathbb{R} 1_B(\lambda)\ F(d\lambda,x,y$$

$x \in X \quad u \in \mathbb{R} \quad A \in \mathcal{B}_X \quad , B \in \mathcal{B}_\mathbb{R}$

Notons $(Q_x)_{x\in X}$ la famille de probabilités associées sur l'espace des trajectoires $\Omega = (X \times \mathbb{R})^\mathbb{N}$

Faisons les hypothèses suivantes :

(H_1) P admet une probabilité invariante π et pour toute fonction ϕ continue sur x on a

$$\lim_n \sup_{x\in X} |P^n\phi(x) - \pi(\phi)| = 0$$

(H_2) Il existe un compact C de $\mathbb{R}$ tel que pour tous $x,y \in X$ le support de la probabilité $F(d\lambda,x,y)$ soit contenu dans C

(H_3) L'application $\alpha(x) = \int_X P(x,dy) \int_{\mathbb{R}} \lambda F(d\lambda,x,y)$ est continue sur X

On a alors le

THEOREME 2

Sous les hypothèses (H_1), (H_2), (H_3)

1°) $\forall x \in X$

$$Q_x \text{ p.s. } \lim_n \frac{1}{n} (U_1+U_2+\dots+U_n) = \int_X \alpha(x)\, \pi(dx)$$

2°) Il existe une constante $0<a<1$ telle que

$\forall \varepsilon > 0$

$$\overline{\lim_n} \left[\sup_{x\in X} Q_x\left\{\left|\frac{U_1+U_2+\dots+U_n}{n} - \int_X \pi(dx)\, \alpha(x)\right| > \varepsilon\right\}\right]^{1/n} \leq a^{\varepsilon^2} < 1$$

Démonstration

Elle utilise la

Proposition 2

Soit $(\Omega,\mathcal{F},P)$ un espace probabilisé et $(Y_n)_{n\geq 1}$ une suite de variables aléatoires réelles définies sur cet espace telles que

a) $\forall n\geq 1 \quad |Y_n| \leq L$ Pps

b) $\forall n\geq 1 \quad E(Y_n|\mathcal{F}_{n-1}) = 0$ Pps

où $\mathcal{F}_0 = \{\phi,\Omega\}$ et $\mathcal{F}_n$ $n\geq 1$ est la tribu engendrée par les variables aléatoires $(Y_k)_{1\geq k\geq n}$

alors

1°) Pps $\lim_n \frac{1}{n} (Y_1+Y_2+\dots+Y_n) = 0$

2°) $\forall \varepsilon>0$ et $\forall n\geq 1$ $P\left(\frac{|Y_1+Y_2+\dots+Y_n|}{n} > \varepsilon\right) \leq 2e^{-\frac{n\varepsilon^2}{2L}}$

et la décomposition suivante par une méthode apparue dans [1] et reprise dans [6] :

Notons $(\mathcal{F}_n)_{n\geq 1}$ la tribu $(X_k,U_k)\ k\leq n)$ et $\mathcal{F}_0 = \{\Omega,\phi\}$ pour $x \in X$ on définit alors Q_x ps les variables aléatoires

$$U_n^{(1)} = \begin{cases} U_n - E_x(U_n|\mathcal{F}_{n-1}) & \text{si } n > 1 \\ 0 & \text{si } n = 1 \end{cases}$$

et pour $j \geq 2$ $U_n^{(j)} = \begin{cases} E_x (U_n | \mathcal{F}_{n-j+1}) - E_x(U_n|\mathcal{F}_{n-j}) & \text{si} \quad n > j \\ 0 & \text{si} \quad 1 \leq n \leq j \end{cases}$

On a alors pour $n > j \geq 1$

$$\frac{S_n}{n} - \alpha = \frac{U_1 + U_2 + \ldots + U_n}{n} - \int_X \pi(dx)\alpha(x) = \sum_{p=1}^{j} \frac{1}{n} \Big(\sum_{k=j+1}^{n} U_k^{(p)} \Big) + \frac{1}{n} \sum_{k=j+1}^{n} [E_x(U_k | \mathcal{F}_{k-j}) - \alpha]$$

$$+ \frac{1}{n} \sum_{k=1}^{j} U_k - \frac{j}{n} \alpha = T_1(n,j) + T_2(n,j) + T_3(n,j)$$

où pour tout $p \geq 1$ $\Big(\sum_{k=j+1}^{n} U_k^{(p)} \Big)_{n \geq 1}$ est une martingale.

2-2 Compte tenu du corollaire 2 la chaine semi markovienne (X_n, U_n) à valeurs dans $P(\mathbb{R}^2) \times \mathbb{R}$ définie par

$$(X_n, U_n) = (g_n \ldots g_1 \bar{x} \;,\; \operatorname{Log} \sigma(g_{n+1}, g_n \ldots g_1 \bar{x})) \qquad \bar{x} \in P(\mathbb{R}^2)$$

satisfait aux hypothèses du théorème 2 d'où le

THEOREME 3

Sous les hypothèses du théorème 1 il existe $0 < a < 1$ tel que $\forall \varepsilon > 0$

$$\lim_n \Big[\sup_{\|x\|=1} P\Big(\Big|\frac{1}{n} \operatorname{Log} \|g_n\, g_{n-1} \ldots g_1 x\| - \gamma\Big| > \varepsilon\Big)\Big]^{1/n} \leq a^{\varepsilon^2} < 1$$

§3 - Un théorème de la limite centrale

Le corollaire 1 montre que pour $0 < \lambda \leq \lambda_0$ l'opérateur P est un opérateur de Doeblin-Fortet régulier [7] sur $\mathcal{L}_\lambda$. Il en est de même pour l'opérateur de transition associé à la chaine semi markovienne $(g_n \ldots g_1 \bar{x},\, g_{n+1})$ à valeurs dans $P(\mathbb{R}^2) \times S_p$, agissant sur l'espace L_λ , analogue de $\mathcal{L}_\lambda$ pour l'espace $P(\mathbb{R}^2) \times S_p$ muni de la métrique $d((\bar{x},g)(\bar{y},g')) = \operatorname{Sup} (\delta(\bar{x},\bar{y}), \|g-g'\|)$.

En utilisant un théorème de la limite centrale énoncé dans ce cadre dans [7], ou en appliquant un théorème de [6] établi dans le cadre des processus d'apprentissages nous pouvons prouver le

THEOREME 4

Sous les hypothèses du théorème 1

1) $\forall x \in \mathbb{R}^2-\{0\}$ la suite $\sigma_n^2(x) = \frac{1}{n}$ (E Log $\|g_n \ldots g_1 x\| - n\gamma)^2$ converge vers une constante $\sigma^2 > 0$ indépendante de x

2) $\forall x \in \mathbb{R}^2-\{0\}$ la suite de variables aléatoires
$$Z_n(x) = \frac{1}{\sigma\sqrt{n}} (\text{Log } \|g_n \ldots g_1 x\| - n\gamma)$$
converge en loi vers une loi normale $N(0,1)$

3) Il existe une constante $C > 0$ et un réel $\alpha \geq 0$ tels que
$$\forall t \in \mathbb{R} \quad \forall n \geq 1 \quad \sup_{\|x\|=1} \left[P(Z_n(x) \leq t) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{u^2}{2}} du\right] \leq \frac{C}{n^\alpha}$$

§4 - Remarques

1) Les théorèmes 3 et 4 s'étendent aux groupes semi-simples de rang 1 en particulier aux groupes du type SO(n,1) et SU(n,1).

2) Le résultat du théorème 4 est analogue à ceux établis dans [8] pour d quelconque lorsque la probabilité p admet une densité et également dans [4] dans le cas où la suite de matrices aléatoires est une suite de matrices positives.

(*) séance du

[1] BREIMAN L. : The strong law of large numbers for a class of Marko chains - Ann. Math. Stat vol 31 (1960), p.801-803.

[2] FURSTENBERG H. : Non Commuting random products - TAMS vol 108 (1963) p. 377-428.

[3] FURSTENBERG H. : Amer Math Soc Summer Inst on Harmonic Analysis on homogeneous spaces, Williamstown, Massachussets (1972).

[4] FURSTENBERG H. et KESTEN H. : Products of random matrices - Ann Math stat 31 (1960), p. 457-469.

[5] GUIVARC'H Y. : Etude des produits de matrices alèatoires - Cours de Saint Flour (1978).

[6] KAIJSER T. : Some limit theorems for Markov chains with applications to learning models and products of random matrices - Report Institut Mittag Leffler (1972).

[7] NORMAN F. : Markov Processes and Learning models - Academic Press New York (1972).

[8] TUTUBALIN V.N. : On limit theorems for the product of random matrices - Theory of Probability and its applications, volume 10 number 1 (1965) p. 15-27.

E. LE PAGE
Laboratoire dè Probabilit
L.A. 305 du C.N.R.S.,
B.P. n° 25 A
35031 Rennes Cédex.