Séminaire Mathématique de Béjaïa (LaMOS)
Volume 12, 2013, pp. 115 - 120

14

Itérations des Fonctions aléatoires et application à la simulation

B. BESSAD[a], F. LADJIMI[b] et M. A. BOUDIBA[c]

[c] Faculté de Sciences, Campus de BASTOS, Université Mouloud Mammeri, Tizi-Ouzou, Algérie,
[c] email : `m.arezki boudiba @ yahoo.fr`

Résumé The purpose is some observations on the Markov Chains $(X_n^x)_n$, on the model $X_n^x = F_n \circ F_{n-1} \cdots \circ F_1(x)$, where $(F_n)_n$ is a sequence of i.i.d. random functions. We discribe a motivating example related to the growth of a population. We state a result on the limits of $(X_n^x)_n$ in the particular case where the F_n are generated by the functions $f_{Y_n}(x) = k(Y_n) + g(x)$ with $(Y_n)_n$ a sequence of i.i.d. random variables and where k and g are some suitable functions. We end our exposition by a survey of general theory and an application of such models to exact simulation algorithm of Propp-Wilson. We notice in the conclusion the relationsheep with the recurrency class of the backward chain functions, for the "coalescence" time to be finite.

Key words : Markov chain ; Dynamic systems ; Stationary measure.

14.1 La chaîne de Markov $X_{n+1} = Y_{n+1}X_n(1 - X_n)$

Dans [7], Devaney rapporte que la taille X_t, à l'instant t, de certaines populations de quelque espèce biologique est gouvernée par l'équation différentielle ordinaire,

$$\frac{dX_t}{dt} = KX_t(1 - X_t);$$

où K est un coefficient de proportionnalité propre au milieu.
Un modèle plus adéquat est de considérer que le milieu agit par un coefficient aléatoire $K = Y_t$, à chaque instant t. En discrétisant le temps nous pouvons alors considérer que la taille X_n de la population à l'instant n est modélisée par l'équation

$$X_n = Y_nX_{n-1}(1 - X_{n-1});$$

où on suppose que X_0 est une variable aléatoire donnée et indépendante des Y_n, et $(Y_n)_n$ est une suite de variables aléatoires i.i.d. de loi μ.
Pour x et $y \in \mathbb{R}$ soit alors $x \mapsto f_y(x) = yx(1 - x)$. La chaîne de Markov définie ci-dessus s'écrit alors, si $X_0 = x$

$$X_{n+1}^x = f_{Y_{n+1}}(X_n^x).$$

Cette chaîne de Markov a fait l'objet de nombreux travaux (Cf. [1, 13], Athreya et Dai, ...). En particulier Rabi Bhattacharaya dans [1] montre, sous des conditions assez restrictives

sur le noyau P de la chaîne, que si la loi μ des Y_n est concentrée dans $[0,\ 4]$ et admet une composante absolument continue par rapport à la mesure de Lebesgue, la chaîne de Markov $(X_n^x)_n$ admet une loi stationnaire et est récurrente au sens de Harris, en adoptant la terminologie de Revuz dans [12].
De nombreux auteurs se sont intéressés à ce type de chaînes de Markov et la littérature est riche de nombreux travaux dont certains ont fait progresser de façon notable la théorie générale (Cf. [13, 5, 14],...). Un moyen d'étude par exemple est de savoir dans quelles conditions le processus $(X_n^x)_n$ converge. Le travail consiste aussi à déterminer dans quelles conditions la chaîne de Markov admet une mesure stationnaire.

14.2 Etude de systèmes dynamiques

Nous rassemblons ici quelques outils de base. Dans le cas où $f_y(x) = y + g(x)$ pour x et $y \in \mathbb{R}$, nous avons, la proposition suivante.

Proposition 1 *Soit $(Y_n)_n$ une suite de variables aléatoires réelles i.i.d. S'il existe deux fonctions réelles k et g de la variable réelle x ; k mesurable et g dérivable et de dérivée $g' < 1$ sur $\mathbb{R}$ tel que $f_y(x) = k(y) + g(x)$ pour x et $y \in \mathbb{R}$, alors le processus $(X_n^x)_n$ défini par*

$$X_0 = x,\ X_{n+1}^x = f_{Y_n}(X_n^x);$$

converge en loi.

Démonstration. Définissons le processus $(H_n^x)_n$, pour $x \in \mathbb{R}$, en posant

$$H_0^x = x \text{ et pour } n > 0, H_n^x = f_{Y_1} \circ f_{Y_2} \circ \cdots \circ f_{Y_n}(x).$$

Remarquons que le processus $(H_n^x)_n$ n'est pas une chaîne de Markov en général. Si g est dérivable et de dérivée $g' < 1$, alors $(H_n^x)_n$ est une suite de Cauchy p.s. Par suite $(H_n^x)_n$ converge p.s. Il s'en suit qu'il converge en loi. Comme les processus $(H_n^x)_n$ et $(X_n^x)_n$ sont de même loi, la proposition, est donc établie.

Dans le cas contractant, le résultat de base est le principe de contraction de Letac (Cf. [13])qui s'énonce ainsi :

Théorème 14.1 (Letac(1985)-Principe de contraction) *Soient $(E, \mathcal{B}_E)$ un espace mesurable avec E localement compact et $\mathcal{B}_E$ sa tribu borélienne, $(F, \mathcal{F}, Q)$ un espace de probabilité et $f : E \times F \to E$ une application mesurable par rapport à la tribu $\mathcal{B}_E \otimes \mathcal{F}$ tel que pour tout y fixé dans F, $x \mapsto f_y(x) = f(x, y)$ est contractante. Soit alors $(Y_n)_n$ une* *suite de v.a. i.i.d à valeurs dans F de loi Q, et X_0 une variable aléatoire à valeurs dans E indépendante de $(Y_n)_n$. La suite de v.a. $(X_n)_n$ définie par X_0 et pour $n > 0$ par*

$$X_{n+1} = f(X_n, Y_{n+1})$$

est une chaîne de Markov d'espace d'états E, de loi initiale $\mu = \mathcal{L}(X_0)$ et de noyau de transition P, la loi image de Q par f_x. De plus si $(Z_n)_n$ est la suite de variables aléatoires définies par $Z_n = f_{Y_1} \circ f_{Y_2} \circ \cdots \circ f_{Y_n}(X_0)$, si $Z = \lim_n Z_n(x)$ existe p.s. , et si π est la loi de Z, alors elle est unique et $(X_n)_n$ admet comme mesure stationnaire π.

Dans le cas de la chaîne de Markov $(X_n)_n$ sur $\mathbb{N}$, définie par

$$X_0, \text{ et pour } n > 0, X_n = |Y_n - X_{n-1}|,$$

où $(Y_n)_n$ est une suite de variables aléatoires i.i.d, à valeurs dans $\mathbb{N}$ et tel que $\mathbb{E}[Y_n] < \infty$, le Principe de Contraction s'applique et nous montrons(Cf. [2] que $\lim H_n^x$ existe p.s. et ne dépend pas de x. Ce qui assure l'existence et l'unicité de la mesure stationnaire pour la chaîne de Markov $(X_n)_n$. Celle-çi est bornée et par suite les classes essentielles de la chaîne sont récurrentes positives. Pour une étude détaille de cette chaîne voir [2, 10, 15, 12].
Dans le cas de la chaîne de Markov $X_{n+1} = Y_{n+1}X_n(1 - X_n)$, les fonctions f_y ne sont pas contractantes et alors une notion de contraction plus large semble nécessaire pour montrer l'existence et l'unicité d'une mesure stationnaire, c'est la contraction en moyenne (Cf. [14, 8])
De façon générale, soit (E, d) un espace métrique et soit $(f_y)_{y\in\Theta}$ une famille de fonctions de E dans E lipschitziennes de rapport K_y i.e. :

$$\forall x, x' \in E,\ d(f_y(x), f_y(x') \leq K_y d(x, x').$$

Soit $(Y_n)_n$ une suite de variables aléatoires i.i.d. de loi μ, à valeurs dans Θ . Considérons alors la chaîne de Markov $(X_n^x)_n$ définie par X_0 une variable aléatoire indépendante des Y_n et à valeurs dans E et pour $n > 0$

$$X_n = f_{Y_n}(X_{n-1}).$$

Supposons que les conditions suivantes sont satisfaites :

1. K_{Y_n} a une loi a queue lourde i.e. il existe α et β tel que pour u assez grand

$$\mu\{y,\ K_y \geq u\} \leq \frac{\beta}{u^\alpha};$$

2. Il existe $x_0 \in E$ tel que la variable aléatoire A_Y définie par $A_Y = d(f_Y(x_0), x_0$ est aussi de loi à queue lourde ;
3. les fonctions f_y sont contractantes en moyenne i.e. :

$$\int \ln(K_y)\mu(dy) < 0.$$

Alors d'après Diaconis in [14], il en résulte en particulier que :

1. La chaîne de Markov $(X_n^x)_n$ admet une probabilité stationnaire ν unique ;

2. Le processus (H_n^x) associé converge p.s. et sa limite ne dépend pas de x.

Nous conjecturons alors la proposition suivante :

Proposition 2 *Si $(Y_n)_n$ une suite de variables aléatoires i.i.d. de loi μ tel que supp$(\mu) \subset [0,\ a]$ avec $a > 1$, soit $(X_n^x)_n$ la chaîne de Markov définie pour $x \in [0,\ a]$, par*

$$X_0 = x, \quad \text{pour } n > 0,\ X_n^x = Y_n X_{n-1}^x (1 - X_{n-1}^x).$$

Alors pour certaines lois μ, la chaîne de Markov $(X_n^x)_n$ admet une probabilité stationnaire unique.

Cette conjecture est motivée par de nombreuses simulations effectuées qui montre que les trajectoires de la chaîne finissent par se stabiliser dans certaines conditions (Cf. [4]) et les conditions 1., 2., 3., de Diaconis in [7] semblent vérifiées pour quelques cas de lois des Y_n considérées dans ces simulations.

14.3 Algorithme de Propp-Wilson de simulation exacte

Pour échantillonner une loi de probabilités π sur un ensemble fini S, différents algorithmes de simulation sont disponibles : Métropolie, MCMC,...(Cf. [20] pour un excellent ouvrage sur la simulation). La plupart de ces algorithmes permettent de construire un échantillon distribué approximativement suivant la loi π. La rupture par rapport à ces méthodes est l'algorithme de simulation exacte, proposé par Propp et Wilson (Cf. [16, 17]). Il est particulièrement utile si S est grand et π difficile à approcher par les méthodes de Monté Carlo classiques.

L'idée est la suivante. Si on veut construire un échantillon d'une loi donnée π sur S, on se donne une suite $(Y_n)_n$ de variables aléatoires à valeurs dans un ensemble Θ de loi μ. Pour tout $y \in \Theta$, on se donne une fonction $x \mapsto f_y(x)$ de S dans S. On construit alors la chaîne de Markov $(X_n^x)_n$, en posant $X_0 = x$ et pour $n > 0$,

$$X_n^x = f_{Y_n}(X_{n-1}^x),$$

d'une telle manière que son noyau de transition P admet π comme loi invariante. On s'assure aussi que $(X_n^x)_n$ est irréductible et apériodique.
Définissons le processus associé $(H_n^x)_n$, pour $x \in S$, par $H_n^x = x$ et pour $n > 0$:

$$H_n^x = f_{Y_1} \circ f_{Y_2} \circ \cdots \circ f_{Y_n}(x).$$

L'algorithme de Propp-Wilson se fonde sur le fait essentiel que $lim_n H_n^x$ existe p.s. et de plus $\lim_n H_n^x = Constante$. Si $\lim_n H_n^x = x_1$, comme $\lim_n H_n^x$ converge en loi vers une variable aléatoire de loi π, l'échantillon $(x_1, x_2, \cdots x_N)$ construit en répétant la procédure est distribué exactement suivant π. Les x_i sont générés en produisant un échantillon $(y_1, y_2, \cdots y_n)$ de la loi μ des Y_n . Le pas d'arrêt n est le nombre d'itérations nécessaire pour que $\lim_n H_n^x = Constante$.
Propp et Wilson dans [20, 16], optent pour un choix assez intuitif des f_Y, qui assure d'abord que $lim_n H_n^x$ existe et d'autre part que cette limite est atteinte au bout d'un temps fini. Diaconis et Freedman généralisent dans [8] et établissent que sous les conditions 1., 2. et 3. de la section 2, on a effectivement que $lim_n H_n^x$ existe et que cette limite est constante. Benaïm montre, dans certains cas, avec une démonstration élégante dans [17], que le temps de coalescence i.e. le temps mis par le processus H_n^x, pour atteindre sa limite est fini p.s.
Le défaut de cet algorithme est bien sûr que le test d'arrêt reste assez intuitif. Pour y remédier Propp et Wilson considèrent des système monotones, dans le sens où en munissant S d'un ordre partiel noté $\leq$, les f_y vérifient

$$x, x' \in S,\ x \leq x',\ \Rightarrow f_y(x) \leq f_y(x') \forall y \in \Theta.$$

Si a est le plus petit élément et b le plus grand élément dans S pour $\leq$, le test d'arrêt pour l'algorithme est alors $H_n^a = H_n^b$. Une fois que cette condition est vérifiée, on prend la valeur x_i produite par H_n^x.
Nous terminons cet exposé en signalant que de nombreux auteurs se sont intéressés à cet algorithme et ont y apportés des améliorations (Cf. [18], [19], [20], ...).

14.4 Conclusion

L'algorithme de Propp-Wilson de simulation exacte, réalise une rupture par rapport aux méthodes de simulation classiques basées en général sur des approximations de la loi cible. Cela replace dans l'actualité mathématique, l'étude des chaînes de Markov obtenues par des itérations aléatoires. L'algorithme de Propp-Wilson, exploite la convergence du processus $(H_n^x)_n$ vers des constantes, dans les conditions de Diaconis (Cf. [8]). Cela signifie que la chaîne de Markov $(H_n)_n$ sur le semi-groupe engendré par les produits finis $f_{y_1} \circ f_{y_2} \cdots \circ f_{y_n}$ avec les y_i dans le support de la loi de Y_1, admet pour seules classes récurrentes les fonctions constantes. L'exemple de la chaîne de Feller $(X_n^x)_n$ avec $f_y(x) = |y - x|$ est intéressant, car dans ce cas nous savons (Cf.[2]) que les constantes sont des classes récurrentes pour la chaîne associée $(H_n)_n$, si $\mathbb{E}[Y_1] < \infty$. Il reste à trouver des conditions moins restrictives (cas contractant) pour que ce soit les seules classes récurrentes.

Références

1. R. Bhattacharaya and M. Majumdar, *Stability in distribution of randomly perturbed random maps as Markov Process.*
2. M. A. Boudiba, *La chaîne de Markov* $X_n = |X_{n-1} - Y_n|$ *et les chaînes associées, avec* $(Y_n)_n$ *une suite i.i.d. de variables aléatoires,* $\mathbb{E}[Y_n] < \infty$. Compt. Rend. Acad. Sciences, Paris, 1985.
3. M. Benaïm et N. El Karaoui, *Promenades aléatoires, Chaines de Marlov et simulation ; Martingales et stratégies.* Editions de l'Ecole Polytechnique, Paris 2007.
4. B. Bessad, *Chaînes de Markov et Application à la Cinétique des réactions chimiques.* Mémoire de magister, Option Probabilités et Statistique, Faculté de Sciences-UMMTO, 2011.
5. F. K. R. Chung, P. Diaconis and R.L. Graham, *Random walks arising in random number generation.* The Annals of propbability, Vol. 15,No 3, 1987.
6. J. N. Corcoran and R. L. Tweedie, *Perfect samling from independant Metropolis-Hastings Chains.* Journal of Statistical Planing and Inference, 2002.
7. R. Devaney, *Chaotic Dynamical Systems.* Willey, N-Y, 1985.
8. P. Diaconis and D. Freedman,*Iterated andom Functions.* SIAM review,Vol. 41, No 1,1999.
9. P. Diaconis and M. Shahshahani, *Products of random matrices as they arise in the study of random walks on groups.* Technical report no. 229, Department of Statistics, Stanford Univrsity, Novembre 1984.
10. W. Feller, *Introduction to Probability Theory and Stochastic Process with Applications*, 2nd Edition Willey and Sons, N-Y , 1966.
11. O. Häggström, *Finite Markov Chains and Algorithmic Applications.* Cambridge University Press, 2007.
12. F. B. Knight, *The absolute difference chaîn.* Wah. Gebiete, Berlin, 1978.
13. G. Letac, *A contraction principle for certain Markov Chains and its applications.* Contemporary Mathematics, Volume 50, 1986. American Mathematecal Society, Rhodes Island, N-Y.
14. M. Mirek, *Heavy tail phenomenom and convergence to stable laws for iterated Lipschitz random maps.* Probability Theory and Related Fields, 151 (2011)
15. M. Peigné, *Marches aléatoires sur le semi-groupe des contractions de* $\mathbb{R}^d$ *avec chocs elastiques.* Annales de l'IHP, 1992.
16. J. Propp and D. Wilson, *Exact sampling with coupled Markov chains, Random Structures.* Algorithms, 9, pp. 223-252, 1996.
17. J. Propp and D. Wilson, *How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph.* Journal of Algorithms, No 27, 1998.
18. A. Revuz, *Markov Chains.* North Holland, 1975.
19. D. Wilson, *How to couple from the past using a read-once source of randomness.* Random Strucures and Algorithms, 16, 2000.
20. B. Ycart, *Modèles et Algorithmes markoviens*, Mathématiques et Applications, Springer, 2002.