Available online at www.sciencedirect.com
SCIENCE DIRECT®

C. R. Acad. Sci. Paris, Ser. I 340 (2005) 49–54

COMPTES RENDUS MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Probabilités

Étude de la consistance d'un algorithme stochastique dans un cadre mélangeant

Ahmed Ait Saidi, Abdelnasser Dahmani

Laboratoire de mathématiques appliquées, université de Béjaia, 06000 Béjaia, Algerie

Reçu le 26 mai 2004 ; accepté après révision le 22 septembre 2004

Disponible sur Internet le 19 décembre 2004

Présenté par Paul Deheuvels

Résumé

Cette Note est consacrée à l'étude de la consistance d'un algorithme stochastique ayant pour objet l'estimation d'un paramètre sous une condition de α-mélange. ***Pour citer cet article : A. Ait Saidi, A. Dahmani, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

Abstract

Study of the consistency of a stochastic algorithm under mixing. This Note is devoted to the study of the consistency of a stochastic algorithm having for object the estimates of a parameter under strong mixing assumption. ***To cite this article: A. Ait Saidi, A. Dahmani, C. R. Acad. Sci. Paris, Ser. I 340 (2005).***

Abridged English version

The problem

This Note is set in the framework of the estimate of a real parameter θ introduced into the autoregressive model (1) where $(\xi_n)_n$ is a sequence of stationary random variable with zero mean, and where f_θ and g are numerical functions defined on $\mathbb{R}$, with f_θ derivable with respect to θ. This work concerns the study of the parameter θ when the error variable ξ_n is under strong mixing assumption.

Estimator and asymptotic study

The main goal is to estimate θ, specifying the rate of convergence. To this end, we introduced the stochastic algorithm θ_n defined by (2) and studied the pointwise almost complete convergence of the estimator θ_n under some regularity conditions. More precisely, under the assumptions (6)–(13) and (14) we obtain (15).

Adresse e-mail : dahmany@wissal.dz (A. Dahmani).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2004.11.001

Some remarks about asymptotic result

Firstly, we consider the sequence $(\xi_n)_n$ as a geometrically strong mixing sequence. In this case, the mixing coefficients $\alpha(n)$ satisfy the following condition:

$$\exists\tilde{c}>0,\ \exists t\in\,]0,1[,\quad \alpha(n)\leqslant\tilde{c}t^n\quad\forall n.$$

Note that the condition (11) is verified for any b. Therefore we can simplify the hypothesis (12) as (31).

Secondly, we consider $p=+\infty$. In this case, we may write the hypothesis (12) (respectively (14)) as follows $\frac{3-b}{b+1}+\delta\leqslant a\nu-1$ (respectively $b>\frac{2-a\nu}{\nu_0}$).

Application

To forecasting the evolution of the rates of exchange while using the model (1), the function $f_\theta(X_n)$ can be interpreted like the conditional mean (dependent on θ) of X_{n+1} given the backward shift and $g(X_n)$ like the conditional standard deviation.

1. Introduction

Cette Note est consacrée à l'étude de la consistance d'un algorithme stochastique ayant pour objet l'estimation d'un paramètre θ inconnu, réel et non nul, introduit dans un modèle autorégressif non linéaire

$$X_{n+1}=f_\theta(X_n)+g(X_n)\xi_{n+1} \tag{1}$$

où $(\xi_n)_n$ est une suite stationnaire de variables aléatoires réelles centrées et où f_θ et g sont deux fonctions numériques définies sur $\mathbb{R}$, avec f dérivable par rapport à θ.

Ce modèle peut être vu comme une extension du modèle autorégressif à indice fixé ($\theta=1$) (ou modèle ARCH, voir [3–5]). De fait, ce modèle hérite des potentialités des modèles ARCH en termes d'applications. L'interprétabilité du paramètre introduit dans ce modèle le rend particulièrement intéressant dans la pratique et notamment très populaire dans la communauté de l'économétrie de la finance. L'estimation du paramètre θ de la procédure (1), dans le cadre d'erreurs aléatoires indépendantes, a été amorcée par Bondarev et Dahmani [1]. Plus récemment, Dahmani et Ait Saidi [2] ont généralisé ce résultat au cadre fortement mélangeant lorsque les erreurs aléatoires ξ_n sont bornées. Ce cadre englobe de nombreux modèles utilisés en statistique mathématique, comme le montre Doukhan [5] dans son ouvrage sur le mélange. Dans un contexte analogue au notre et concernant les différentes modélisations de la dépendance du bruit d'un algorithme stochastique, on peut se reporter à la note de Bandrière et Doukhan [6].

Notre travail consiste à étendre les résultats cités en [2] au cas de variables aléatoires ξ_n non bornées en donnant un résultat sur la vitesse de convergence presque complète ponctuelle pour l'estimateur de θ.

2. Estimateur et étude asymptotique

L'objectif principal consiste à estimer le paramètre θ en précisant la vitesse de convergence presque complète pour son estimateur. Dans ce but, nous introduisons l'algorithme de Kholev [7] défini par :

$$\theta_{n+1}=\theta_n+\frac{a}{n}\big(X_{n+1}-f_{\theta_n}(X_n)\big)\frac{\partial f_{\theta_n}}{\partial\theta_n}(X_n),\quad n\in\mathbb{N}^*,\ a\in\mathbb{R}_+^*. \tag{2}$$

Si on pose

$$H_{\theta_n}(x)=\frac{\partial f_{\theta_n}}{\partial\theta_n}(x)g(x)\quad\text{et}\quad R_{\theta_n}(\theta,x)=\frac{f_{\theta_n}(x)-f_\theta(x)}{\theta_n-\theta}\frac{\partial f_{\theta_n}}{\partial\theta_n}(x) \tag{3}$$

on obtient

$$\theta_{n+1}-\theta=(\theta_n-\theta)\left(1-\frac{a}{n}R_{\theta_n}(\theta,X_n)\right)+\left(\frac{a}{n}\right)H_{\theta_n}(X_n)\xi_{n+1}; \tag{4}$$

par itérations successives, la relation (4) devient

$$|\theta_n-\theta| = \left|\prod_{i=1}^{n}\left(1-\frac{a}{i}R_{\theta_i}(\theta,X_i)\right)\right| \times\left(\left|(\theta_1-\theta)+\sum_{m=1}^{n}\left(\frac{a}{m}\prod_{i=1}^{n}\left(1-\frac{a}{i}R_{\theta_i}(\theta,X_i)\right)^{-1}H_{\theta_m}(X_m)\xi_{m+1}\right)\right|\right). \tag{5}$$

Introduisons maintenant les hypothèses suivantes :

$$\forall x\in\mathbb{R},\ \forall\theta\in\mathbb{R}^*,\quad 0<\nu\leqslant R_\theta(\theta,x)\leqslant\rho<\infty, \tag{6}$$

$$\forall\theta\in\mathbb{R}^*,\quad \left|H_\theta(x)\right|\leqslant B<\infty, \tag{7}$$

$$\theta \text{ est un paramètre qui vérifie a priori } |\theta_1-\theta|\leqslant M<\infty, \tag{8}$$

$$\Upsilon(\varepsilon,M)=\varepsilon\exp(a\nu\gamma)n^{a\nu}-M>0 \tag{9}$$

où γ est la constante d'Euler, $n\geqslant 1$, $\varepsilon>0$, $a>0$,

$$\exists p>2,\ \forall t>0,\quad P\left[|\xi_m|>t\right]<t^{-p}. \tag{10}$$

Par ailleurs, afin de simplifier les hypothèses ainsi que les aspects techniques, nous nous limiterons au cas où les ξ_i, sont algébriquement fortement mélangeants en ce sens que les coefficients de mélange sont tels que, pour des constantes $c\geqslant 1$ et b un réel positif,

$$\alpha(n)\leqslant cn^{-b} \tag{11}$$

et que la condition de décroissance algébrique (11) soit satisfaite pour une valeur de b vérifiant

$$\exists\delta>0,\quad \frac{4b+p(3-b)}{(b+1)p}+\delta\leqslant a\nu-1. \tag{12}$$

Enfin, notons que si les erreurs aléatoires ξ_{m+1} sont α-mélangeantes alors les transformées aléatoires $(\frac{a}{m}\prod_{i=1}^{m}(1-\frac{a}{i}R_{\theta_i}(X_i))^{-1}H_{\theta_m}(X_m)\xi_{m+1})$ sont aussi fortement mélangeantes de coefficients de mélange inférieurs ou égaux à ceux de la suite $(\xi_{m+1})_m$.

Nous pouvons maintenant énoncer le résultat suivant :

Théorème 2.1. *Sous les hypothèses* (6)–(12) *et si*

$$0<a<\frac{1}{2\rho}, \tag{13}$$

alors, pour tout réel positif b satisfaisant à

$$b>\frac{(2-a\nu)q}{\nu_0}\quad \textit{avec } \nu_0\in]0,1[\textit{ et } q \textit{ est tel que } \frac{2}{p}+\frac{1}{q}=1, \tag{14}$$

on a :

$$\theta_n-\theta=\mathrm{O}(\sqrt{n^{-a\nu}\log n})\quad p.co. \tag{15}$$

Démonstration. Nous avons

$$\log\prod_{i=1}^{n}\left(1-\frac{a}{i}R_{\theta i}(\theta,X_i)\right)\leqslant\sum_{i=1}^{n}-\frac{a\nu}{i}=-a\nu(\log n+\gamma_n)$$

où γ_n est définie par la relation $\gamma_n=\sum_{i=1}^{n}\frac{1}{i}-\log n=\gamma+(\psi(n+1)-\log n)$ où $\psi(\cdot)$ désigne la fonction digamma.

Il est facile de vérifier que $\gamma_n - \gamma_{n-1} = \log(1 - \frac{1}{n}) + \frac{1}{n} < 0$, pour tout $n > 1$. On aboutit alors au résultat bien connu que la suite γ_n décroît vers la constante d'Euler γ, soit

$$\gamma_n > \gamma = \lim_{n\to+\infty}\left\{\sum_{i=1}^{n}\frac{1}{i} - \log n\right\} = 0.577215\ldots.$$

A partir de cette relation, nous obtenons

$$\log\prod_{i=1}^{n}\left(1 - \frac{a}{i}R_{\theta i}(\theta, X_i)\right) \leqslant -a\nu(\log n + \gamma_n) \leqslant -a\nu(\log n + \gamma)$$

ce qui permet de conclure que

$$\prod_{i=1}^{n}\left(1 - \frac{a}{i}R_{\theta i}(\theta, X_i)\right) \leqslant \exp(-a\nu\gamma)n^{-a\nu} \tag{16}$$

et par suite, en utilisant (5) on déduit que

$$P\big[|\theta_n - \theta| > \varepsilon\big] \leqslant P\left[\left|\sum_{m=1}^{n}\frac{a}{m}\prod_{i=1}^{m}\left(1 - \frac{a}{i}R_{\theta i}(\theta, X_i)\right)^{-1}H_{\theta_m}(X_m)\xi_{m+1}\right| > \Upsilon(\varepsilon, M)\right].$$

Par ailleurs, pour un entier naturel n assez grand, on a $M/(n^{a\nu}\exp(a\nu\gamma)) < \varepsilon/2$ ce qui implique que

$$\Upsilon(\varepsilon, M) = n^{a\nu}\exp(a\nu\gamma)\left(\varepsilon - \frac{M}{n^{a\nu}\exp(a\nu\gamma)}\right) > \frac{\varepsilon}{2}n^{a\nu}, \tag{17}$$

d'où

$$P\big[|\theta_n - \theta| > \varepsilon\big] \leqslant P\left[\left|\sum_{m=1}^{n}\frac{a}{m}\prod_{i=1}^{m}\left(1 - \frac{a}{i}R_{\theta i}(\theta, X_i)\right)^{-1}H_{\theta_m}(X_m)\xi_{m+1}\right| > 4\lambda n^{a\nu}\right]$$

avec $\lambda = \frac{\varepsilon}{8}$. Une application directe de l'inégalité exponentielle de Fuk–Nagaev donnée par Rio (1999, formule (6.19a)), aux variables fortement mélangeantes

$$\zeta_m = \frac{a}{m}\prod_{i=1}^{m}\left(1 - \frac{a}{i}R_{\theta i}(X_i)\right)^{-1}H_{\theta_m}(X_m)\xi_{m+1} \tag{18}$$

amène alors en utilisant (10), que pour $\varepsilon > 0$ et pour $r \geqslant 1$ on a :

$$P\big[|\theta_n - \theta| > \varepsilon\big] \leqslant 4\left(1 + \frac{(\lambda n^{a\nu})^2}{rs_n^2}\right)^{-r/2} + 4Cnr^{-1}\left(\frac{r}{\lambda n^{a\nu}}\right)^{(b+1)p/(b+p)} \tag{19}$$

avec $C = 2p(2p-1)^{-1}(2^b c)^{(p-1)/(b+p)}$ et

$$s_n^2 = \sum_{m=1}^{n}\sum_{s=1}^{n}\left|\operatorname{cov}(\zeta_m, \zeta_s)\right| = \sum_{m=1}^{n}\operatorname{var}(\zeta_m) + \sum_{m=1}^{n}\sum_{s\neq m}^{n}\left|\operatorname{cov}(\zeta_m, \zeta_s)\right|. \tag{20}$$

D'une part, sous l'hypothèse (10), nous avons

$$\exists N < +\infty : E(\xi_m^2) \leqslant N \quad \text{et} \quad \operatorname{var}(\zeta_m) \leqslant C_1 m^{2(a\rho-1)}$$

avec $C_1 = B^2a^2N\exp(2a\rho\gamma + (a\rho)^2)$. Comme $a < \frac{1}{2\rho}$, on en déduit que

$$\sum_{m=1}^{n}\operatorname{var}(\zeta_m) = \sum_{m=1}^{n}\frac{C_1}{m^{2(1-a\rho)}} \leqslant DC_1 \tag{21}$$

car il s'agit d'une somme partielle d'une série à termes positifs convergente.

D'autre part, pour $m \neq s$,

$$\left|\operatorname{cov}(\zeta_m, \zeta_s)\right| \leqslant C_2 m^{a\rho-1} s^{a\rho-1} \left|E(\xi_{m+1}\xi_{s+1})\right|$$

avec $C_2 = B^2 a^2 \exp(2a\rho\gamma + (a\rho)^2)$. Sous la condition (5), nous pouvons utiliser l'inégalité de Davydov–Rio donnée par Rio (1999, formule (1.12c)) pour obtenir :

$$\left|E(\xi_{m+1}\xi_{s+1})\right| \leqslant 2q\bigl(\alpha\bigl(\left|(m+1)-(s+1)\right|\bigr)\bigr)^{1/q} = 2q\bigl(\alpha\bigl(|m-s|\bigr)\bigr)^{1/q}$$

et donc

$$\left|\operatorname{cov}(\zeta_m, \zeta_s)\right| \leqslant 2qC_2 m^{a\rho-1} s^{a\rho-1} \bigl(\alpha\bigl(|m-s|\bigr)\bigr)^{1/q}. \tag{22}$$

En appliquant une seconde fois l'inégalité de Davydov–Rio aux variables ζ_m, nous obtenons

$$\forall m \neq s, \quad \left|\operatorname{cov}(\zeta_m, \zeta_s)\right| \leqslant 2q\bigl(\alpha\bigl(|m-s|\bigr)\bigr)^{1/q} C_2 \tag{23}$$

car les coefficients de mélange de la suite $(\zeta_m)_m$ sont inférieurs ou égaux à ceux de la suite $(\xi_m)_m$.

En rassemblant (22) et (23), on a

$$\sum_{m=1}^{n} \sum_{m\neq s}^{n} \left|\operatorname{cov}(\zeta_m, \zeta_s)\right| \leqslant \sum_{m=1}^{n} \frac{1}{m^{2(1-a\rho)}} \sum_{k=1}^{n} 2qC_2\bigl(\alpha(k)\bigr)^{1/q} + 2n^2 C_2 q\bigl(\alpha(u_n)\bigr)^{1/q} \tag{24}$$

ou encore

$$\sum_{m=1}^{n} \sum_{m\neq s}^{n} \left|\operatorname{cov}(\zeta_m, \zeta_s)\right| \leqslant 2DC_2 D_1 + 2n^2 C_2 q\bigl(\alpha(u_n)\bigr)^{1/q} \tag{25}$$

avec $D_1 = \sum_{k=1}^{n} q(\alpha(k))^{1/q}$.

Ainsi, en prenant $u_n = [n^{\nu_0}]$, les crochets désignant la partie entière, et en utilisant (11), (14), (21) et (25) nous obtenons, pour n assez grand :

$$s_n^2 = \mathrm{o}(n^{a\nu}) \tag{26}$$

car

$$\frac{2n^2 C_2 q(\alpha(u_n))^{1/q}}{n^{a\nu}} \leqslant \frac{2C_2 q c^{1/q}}{n^{a\nu-2+\nu_0 b/q}} \longrightarrow 0.$$

Ainsi en tenant compte de (26), on arrive à

$$P\bigl[|\theta_n - \theta| > \varepsilon\bigr] \leqslant A_1 + A_2 \tag{27}$$

avec $A_1 = 4(1 + \lambda^2 n^{a\nu}/r)^{-r/2}$ et $A_2 = 4Cnr^{-1}(r/(\lambda n^{a\nu}))^{(b+1)p/(b+p)}$. Pour obtenir la vitesse de convergence, on prend $\lambda = \frac{\eta}{4}\sqrt{n^{-a\nu}\log n}$. On peut toujours choisir r sous la forme $r = A(\log n)^2$, et on obtient

$$A_1 = 4\left(1 + \frac{\eta^2 \log n}{16r}\right)^{-r/2} \leqslant A\exp\left(-\eta^2 \frac{\log n}{32}\right) = An^{-\eta^2/32} \tag{28}$$

où A désigne une constante positive générique. En ce qui concerne A_2, nous avons

$$A_2 = 4Cnr^{-1} r^{\frac{p(b+1)}{b+p}} \lambda^{-\frac{p(b+1)}{b+p}} n^{-\frac{a\nu p(b+1)}{b+p}} = 4Cnr^{-1+\frac{p(b+1)}{b+p}} \left(\frac{\eta}{4}\right)^{-\frac{p(b+1)}{b+p}} (\log n)^{\frac{-p(b+1)}{2(b+p)}} n^{-\frac{p(b+1)}{2(b+p)}} n^{-(a\nu-1)\frac{p(b+1)}{2(b+p)}}.$$

A cause de $r = A(\log n)^2$, on a

$$A_2 = 4C(\log n)^{\frac{b(3p-4)-p}{2(b+p)}} n^{\frac{b(2-p)+p}{2(b+p)}} n^{-(a\nu-1)\frac{p(b+1)}{2(b+p)}}.$$

Grâce à l'hypothèse (12), nous obtenons

$$A_2 \leqslant 4C(\log n)^{\frac{b(3p-4)-p}{2(b+p)}} n^{-1-\frac{\delta p(b+1)}{2(b+p)}}.$$

Donc il existe $d = \frac{\delta p(b+1)}{2(b+p)} > 0$ tel que

$$A_2 \leqslant An^{-1-d}, \tag{29}$$

d'où, pour $\varepsilon = 2\eta\sqrt{n^{-a\nu}\log n}$ suffisamment grand, on a

$$P\Big[|\theta_n - \theta| > 2\eta\sqrt{n^{-a\nu}\log n}\Big] \leqslant An^{-\eta^2/4} + An^{-1-d} \leqslant An^{-1-d}. \tag{30}$$

Le terme de droite de l'inégalité précédente est celui d'une série convergente. Ainsi (30) achève la démonstration.

Remarque 1. Si les variables aléatoires ξ_n sont géométriquement fortement mélangeantes en ce sens que les coefficients de mélange sont tels que, pour des constantes $\widetilde{c} \in \mathbb{R}_+^*$ et $t \in]0, 1[$, $\alpha(n) \leqslant \widetilde{c}t^n$ $\forall n \geqslant 1$, alors la condition (11) est vérifiée pour tout b. Dans ce cas, sans nuire à la vitesse de convergence, on peut simplifier la condition (12) comme suit

$$\exists \delta > 0, \quad \frac{4-p}{p} + \delta \leqslant a\nu - 1. \tag{31}$$

Remarque 2. Dans le cas limite $p = +\infty$, la condition (12) (respectivement (14)) devient $\frac{3-b}{b+1} + \delta \leqslant a\nu - 1$ (respectivement $b > \frac{2-a\nu}{\nu_0}$).

3. Application

Pour prévoir l'évolution des taux de change en utilisant le modèle (1), la fonction $f_\theta(X_n)$ peut s'interpréter comme la moyenne conditionnelle (dépendante de θ) de X_{n+1} sachant le passé et $g(X_n)$ comme l'écart type conditionnel.

Références

[1] B.V. Bondarev, A. Dahmani, Sur les estimations des paramètres inconnus dans des procédures stochastiques, in : Théorie des processus aléatoires, Kiev, 1990, pp. 25–34 (en russe).

[2] A. Dahmani, A. Ait Saidi, Intervalle de confiance pour un paramètre dans un processus stochastique α-mélangeant, Maghreb Math. Rev. 11 (1) (2002) 20–30.

[3] J. Diebolt, D. Guégan, Probabilistic properties of the general nonlinear markovian process of order one and application to time series modelling, Tech. report. 125, L.S.T.A., Jussieu, Paris, §2.4, 1990.

[4] J. Diebolt, D. Guégan, Le modèle de série chronologique autorégressive β-ARCH, C. R. Acad. Sci. Paris, Ser. I 312 (1991) 625–630.

[5] P. Doukhan, Mixing: Properties and Exemples, Lecture Notes in Statist., vol. 85, Springer-Verlag, Berlin, 1994.

[6] P. Doukhan, O. Bandrière, Dependent noise for stochastic algorithms, C. R. Acad. Sci. Paris, Ser. I 337 (7) (2003) 473–476.

[7] A.S. Kholev, Estimation du paramètre de démolition d'un processus de diffusion avec la méthode d'approximation stochastique, in : Etude sur la théorie de l'autoconstruction des systèmes, Centre de calcul de l'académie des sciences de l'URSS, Moscou, 1967, pp. 179–200 (en russe).