



Statistique

Développement d'Edgeworth d'ordre 1 pour des M -estimateurs dans le cas de chaînes V -géométriquement ergodiques

Edgeworth expansion for M -estimators of V -geometrically ergodic Markov chains

Déborah Ferré

Université Européenne de Bretagne, INSA IRMAR, UMR-CNRS 6625, institut national des sciences appliquées de Rennes, 20, avenue des Buttes de Coësmes, CS 70839, 35708 Rennes cedex 7, France

INFO ARTICLE

Historique de l'article :

Reçu le 11 janvier 2010

Accepté le 4 février 2010

Disponible sur Internet le 20 février 2010

Présenté par Paul Malliavin

RÉSUMÉ

Nous donnons des conditions assurant l'existence du développement d'Edgeworth d'ordre 1 pour des M -estimateurs dans le cas des chaînes de Markov V -géométriquement ergodiques. Ce développement est explicite.

© 2010 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

ABSTRACT

Conditions under which a first-order Edgeworth expansion for M -estimators of V -geometrically ergodic Markov chains exists are proposed in this Note. The expansion is explicit.

© 2010 Académie des sciences. Publié par Elsevier Masson SAS. Tous droits réservés.

1. Énoncés des hypothèses et du théorème

Soit $(Q_\theta)_{\theta \in \Theta}$ une famille paramétrique de noyaux de transition sur un espace d'états quelconque (E, \mathcal{E}) , où Θ est un compact de \mathbb{R} . On considère une chaîne de Markov homogène $(X_n)_{n \geq 0}$ à valeurs dans E , de noyau de transition $\{Q_\theta(x, \cdot); x \in E\}$, et on suppose que $(X_n)_{n \geq 0}$ est V -géométriquement ergodique uniformément en θ , à savoir :

Condition (\mathcal{M}) . Pour tout $\theta \in \Theta$, il existe une probabilité Q_θ -invariante notée π_θ , une fonction non bornée $V : E \rightarrow [1, +\infty[$ telles que $\sup_{\theta \in \Theta} \pi_\theta(V) < +\infty$. Il existe des constantes $C > 0$ et $0 < \kappa < 1$ telles que l'on ait, pour toute fonction $f : E \rightarrow \mathbb{R}$ mesurable vérifiant $|f| \leq V$:

$$\forall \theta \in \Theta, \forall n \in \mathbb{N}, \forall x \in E, \quad |Q_\theta^n f(x) - \pi_\theta(f)| \leq C\kappa^n V(x).$$

Sous des conditions d'irréductibilité et d'apériodicité, cette hypothèse à θ fixé est équivalente à la condition classique de dérive $Q_\theta^{n_\theta} V \leq \rho_\theta V + M_\theta 1_{C_\theta}$, avec $n_\theta \in \mathbb{N}^*$, $\rho_\theta < 1$, $M_\theta \in \mathbb{R}$, et C_θ un "petit" ensemble [4]. Pour obtenir (\mathcal{M}) , il suffit que n_θ , ρ_θ , M_θ et C_θ puissent être choisis indépendants de θ .

La loi de X_0 est notée μ_θ , et on suppose dans ce qui suit que $\sup_{\theta \in \Theta} \mu_\theta(V) < +\infty$. La loi de notre modèle statistique associé à Q_θ et μ_θ est notée $\mathbb{P}_{\theta, \mu_\theta}$ et $\mathbb{E}_{\theta, \mu_\theta}$ l'espérance associée. Soit $\alpha \equiv \alpha(\theta) \in I$ le paramètre d'intérêt, où I est un intervalle ouvert de \mathbb{R} . Pour une fonction mesurable réelle F sur $I \times E^2$, on pose, pour $n \geq 1$,

Adresse e-mail : Deborah.Ferre@insa-rennes.fr.

$$M_n(\alpha) := \frac{1}{n} \sum_{k=1}^n F(\alpha, X_{k-1}, X_k).$$

On suppose que $F^{(1)}$ existe et est intégrable, où $F^{(j)} := \partial^j F / \partial \alpha^j$, et que pour tout $\theta \in \Theta$, il existe un unique $\alpha_0 \equiv \alpha_0(\theta) \in I$ tel que $\mathbb{E}_{\theta, \pi_\theta}[F^{(1)}(\alpha_0, X_0, X_1)] = 0$. Enfin, on suppose que la v.a.r. $\hat{\alpha}_n = \operatorname{argmin} M_n(\alpha)$ est définie pour tout $n \geq 1$.

Nous dirons que $F : I \times E^2 \rightarrow \mathbb{R}$ vérifie la condition de Lipschitz (L) s'il existe $\eta \in]0, 1/2[$ et $C > 0$ tels que :

$$\forall (\alpha, \alpha') \in I^2, \forall (x, y) \in E^2, \quad |F(\alpha, x, y) - F(\alpha', x, y)| \leq C |\alpha - \alpha'| (V(x) + V(y))^\eta, \quad (L)$$

et que F vérifie la condition de domination (D_{m_0}) ($m_0 \in \mathbb{N}^*$) s'il existe $\varepsilon > 0$ et $C > 0$ tels que :

$$\forall \alpha \in I, \forall (x, y) \in E^2, \quad |F(\alpha, x, y)|^{m_0 + \varepsilon} \leq C(V(x) + V(y)). \quad (D_{m_0})$$

Introduisons maintenant les conditions d'application du théorème à venir.

- (a) $M_n^{(1)}(\hat{\alpha}_n) = 0$ pour tout $n > 1$.
- (b) $\forall n \geq 1, \forall d > 0, \sup_{\theta \in \Theta} \mathbb{P}_{\theta, \mu_\theta} \{|\hat{\alpha}_n - \alpha_0| \geq d\} = o(n^{-\frac{1}{2}})$ (consistance en $o(n^{-\frac{1}{2}})$).
- (c) Les dérivées partielles $F^{(2)}$ et $F^{(3)}$ existent, sont intégrables, et vérifient (L).
- (d) $F^{(1)}, F^{(2)}, F^{(3)}$ vérifient la condition de domination (D_3). Soient $m_j(\theta) := \mathbb{E}_{\theta, \pi_\theta}[F^{(j)}(\alpha_0, X_0, X_1)]$ pour $1 \leq j \leq 3$. On rappelle que par définition de α_0 , $m_1(\theta) = 0$.
- (e) $\inf_{\theta \in \Theta} m_2(\theta) > 0$. Comme $F^{(1)}, F^{(2)}, F^{(3)}$ vérifient (D_2), les variances asymptotiques suivantes sont bien définies (cf. [3])

$$\sigma_j(\theta)^2 := \lim_n \frac{1}{n} \mathbb{E}_{\theta, \mu_\theta} \left[\left(\sum_{k=1}^n F^{(j)}(\alpha_0, X_{k-1}, X_k) - n m_j(\theta) \right)^2 \right] \quad (1 \leq j \leq 3);$$

en outre, on sait qu'elles sont indépendantes de μ_θ et telles que $\sup_{\theta \in \Theta} \sigma_j(\theta) < +\infty$.

(f) $\inf_{\theta \in \Theta} \sigma_j(\theta) > 0$ pour $j = 1, 2$.

(g) Condition de non-arithméticité. Nous précisons cette condition dans la remarque 1.

Soit $\sigma(\theta) := \frac{\sigma_1(\theta)}{m_2(\theta)}$, et pour $j = 1, 2 : S_n^{(j)}(\theta) := \sum_{k=1}^n F^{(j)}(\alpha_0, X_{k-1}, X_k) - n m_j(\theta)$. On note \mathcal{N} la fonction de répartition de la loi $\mathcal{N}(0, 1)$, et η sa densité.

Théorème. *Sous les hypothèses (M) et (a)–(g), il existe un polynôme $A(\cdot, \theta)$ tel que l'on ait :*

$$\forall n \geq 1, \quad \sup_{\theta \in \Theta, u \in \mathbb{R}} \left| \mathbb{P}_{\theta, \mu_\theta} \left\{ \frac{\sqrt{n}}{\sigma(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \right\} - \mathcal{N}(u) - \eta(u) n^{-\frac{1}{2}} A(u, \theta) \right| = o(n^{-\frac{1}{2}}).$$

En outre, $A(\cdot, \theta)$ admet des coefficients uniformément bornés en θ , et il s'écrit

$$A(u, \theta) := \left[-\frac{1}{6} \frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} + \frac{b_1(\theta)}{\sigma_1(\theta)} \right] + \left[\frac{1}{6} \frac{m_{3,1}(\theta)^3}{\sigma_1(\theta)^3} - \frac{\sigma_{12}(\theta)}{\sigma_1(\theta)m_2(\theta)} + \frac{\sigma_1(\theta)}{2m_2(\theta)^2} m_3(\theta) \right] u^2$$

où on définit $b_1(\theta) := \lim_n \mathbb{E}_{\theta, \mu_\theta}[S_n^{(1)}(\theta)]$, $\sigma_{12}(\theta) := \lim_n \mathbb{E}_{\theta, \mu_\theta}[S_n^{(1)}(\theta) S_n^{(2)}(\theta)]/n$ et $m_{3,1}(\theta)^3 := \lim_n \mathbb{E}_{\theta, \pi_\theta}[(S_n^{(1)}(\theta))^3]/n$.

Le théorème précédent a été démontré dans le cas i.i.d. par Pfanzagl [6] au changement suivant près : les conditions de moment d'ordre 3 du cas i.i.d. sont remplacées par la condition (D_3). Cette condition de domination est naturelle dans le cadre d'ergodicité V -géométrique. La condition (D_2) a déjà été utilisée pour établir le théorème central limite dans [1] pour $\hat{\alpha}_n$, tandis que (D_3) fournit le théorème de Berry–Esseen dans [3].

2. Idée de la preuve

Le résumé ci-dessous de l'extension de [6] au cas markovien montrera que la difficulté principale réside dans le fait de disposer, sous les conditions (M), de développements d'Edgeworth d'ordre 1 pour $S_n(\xi) := \sum_{k=1}^n \xi(X_{k-1}, X_k)$ pour ξ vérifiant (D_3). Plus précisément, ces propriétés devront être établies pour les fonctionnelles $\xi_\theta(x, y) = F^{(1)}(\alpha_0, x, y)$, puis $\xi_\theta(x, y) = F^{(2)}(\alpha_0, x, y)$, et enfin pour $\xi_{\theta, (p, v)}(x, y) = F^{(1)}(\alpha_0, x, y) + \zeta_\theta(p, v)(F^{(2)}(\alpha_0, x, y) - m_2(\theta)) + \frac{\zeta_\theta(p, v)^2}{2} (F^{(3)}(\alpha_0, x, y) - m_3(\theta))$ avec $(p, v) \in J := \{(p, v) \in \mathbb{N}^* \times \mathbb{R}; |v| < 2\sqrt{\ln p}\}$, et $\zeta_\theta(p, v)$ telle que $\sup_{\theta \in \Theta, |v| < 2\sqrt{\ln p}} \zeta_\theta(p, v)^2 = o(p^{-\frac{1}{2}})$. En outre, les constantes intervenant dans ces développements d'Edgeworth devront être contrôlées uniformément en $(\theta, (p, v)) \in \Theta \times J$ via les constantes de (D_3). Dans [6], ces résultats sont obtenus grâce aux développements d'Edgeworth classiques pour v.a.i.i.d. Pour les modèles markoviens vérifiant la condition (M), ces propriétés résultent du travail récent [2] qui doit être appliqué ici aux noyaux de Fourier $Q_{\theta, (p, v)}(t)(x, dy) := e^{it\xi_{\theta, (p, v)}(x, y)} Q_\theta(x, dy)$ (qui permettent l'étude de fonctions de (X_{k-1}, X_k)).

Grâce à ces résultats probabilistes, on peut comme dans [6] étudier $\mathbb{P}_{\theta, \mu_\theta} \{ \frac{\sqrt{n}}{\sigma(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \}$ selon que $|u| \geq 2\sqrt{\ln n}$ ou $|u| < 2\sqrt{\ln n}$. Dans un premier temps, on désigne par $A(\cdot, \theta)$ un polynôme quelconque, et l'on suppose que les coefficients de $A(\cdot, \theta)$ sont uniformément bornés en θ . Posons $P_{n, \theta}(u) := |\mathbb{P}_{\theta, \mu_\theta} \{ \frac{\sqrt{n}}{\sigma(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \} - \mathcal{N}(u) - \eta(u)n^{-\frac{1}{2}} A(u, \theta)|$.

- Cas $|u| \geq 2\sqrt{\ln n}$. Comme les coefficients de $A(\cdot, \theta)$ sont uniformément bornés en θ , on montre facilement que $\sup_{\theta \in \Theta, |u| \geq 2\sqrt{\ln n}} P_{n, \theta}(u) \leq o(n^{-\frac{1}{2}}) + \sup_{\theta \in \Theta} \mathbb{P}_{\theta, \mu_\theta} \{ \frac{\sqrt{n}}{\sigma(\theta)} |\hat{\alpha}_n - \alpha_0| \geq 2\sqrt{\ln n} \}$. Par l'égalité des accroissements finis, il existe une v.a.r. $\tilde{\alpha}_n$ vérifiant $|\tilde{\alpha}_n - \alpha_0| < |\hat{\alpha}_n - \alpha_0|$ et

$$0 = M_n^{(1)}(\hat{\alpha}_n) = M_n^{(1)}(\alpha_0) + (\hat{\alpha}_n - \alpha_0)M_n^{(2)}(\tilde{\alpha}_n).$$

On utilise ensuite le développement d'Edgeworth pour $F^{(1)}$ pour montrer que $\sup_{\theta \in \Theta, |u| \geq 2\sqrt{\ln n}} P_{n, \theta}(u) \leq o(n^{-\frac{1}{2}}) + \sup_{\theta \in \Theta} \mathbb{P}_{\theta, \mu_\theta} \{ \frac{3}{2} |M_n^{(2)}(\tilde{\alpha}_n)| \leq m_2(\theta) \}$. Enfin, l'hypothèse de consistance et le développement d'Edgeworth pour $F^{(2)}$ permettent de conclure à $\sup_{\theta \in \Theta, |u| \geq 2\sqrt{\ln n}} P_{n, \theta}(u) = o(n^{-\frac{1}{2}})$.

- Cas $|u| < 2\sqrt{\ln n}$. On a $\{ \frac{\sqrt{n}}{\sigma(\theta)} (\hat{\alpha}_n - \alpha_0) \leq u \} = \{ \hat{\alpha}_n \leq \tau \}$, avec $\tau := \alpha_0 + \varsigma_\theta(n, u)$ et $\varsigma_\theta(n, u) := \frac{\sigma(\theta)}{\sqrt{n}} u$. En outre, il existe une v.a.r. $\tilde{\alpha}'_n$ vérifiant $|\tilde{\alpha}'_n - \tau| < |\hat{\alpha}_n - \tau|$ et

$$0 = M_n^{(1)}(\hat{\alpha}_n) = M_n^{(1)}(\tau) + (\hat{\alpha}_n - \tau)M_n^{(2)}(\tilde{\alpha}'_n).$$

On montre par les mêmes arguments que précédemment que $\sup_{\theta \in \Theta, |u| < 2\sqrt{\ln n}} \mathbb{P}_{\theta, \mu_\theta} \{ M_n^{(2)}(\tilde{\alpha}'_n) \leq 0 \} = o(n^{-\frac{1}{2}})$, ce qui ramène l'étude au terme $\sup_{\theta \in \Theta, |u| < 2\sqrt{\ln n}} \mathbb{P}_{\theta, \mu_\theta} \{ M_n^{(1)}(\tau) \geq 0 \}$. Par ailleurs, il existe une v.a.r. $\tilde{\alpha}''_n$ vérifiant $|\tilde{\alpha}''_n - \alpha_0| < |\tau - \alpha_0|$ et

$$M_n^{(1)}(\tau) = M_n^{(1)}(\alpha_0) + \varsigma_\theta(n, u)M_n^{(2)}(\alpha_0) + \frac{\varsigma_\theta(n, u)^2}{2} M_n^{(3)}(\tilde{\alpha}''_n).$$

On considère alors l'événement $B_{n, \theta, u}^\pm := \{ M_n^{(1)}(\alpha_0) + \varsigma_\theta(n, u)M_n^{(2)}(\alpha_0) + \frac{\varsigma_\theta(n, u)^2}{2} M_n^{(3)}(\alpha_0) \pm c \geq 0 \}$ avec la constante $c = c_n(u, \theta) > 0$ choisie telle qu'il suffise de montrer l'existence d'un polynôme $A(\cdot, \theta)$ vérifiant $\sup_{\theta \in \Theta, |u| < 2\sqrt{\ln n}} |\mathbb{P}_{\theta, \mu_\theta} \times \{ B_{n, \theta, u}^\pm \} - \mathcal{N}(u) - \eta(u)n^{-\frac{1}{2}} A(u, \theta)| = o(n^{-\frac{1}{2}})$ pour terminer la preuve du théorème par encadrement. Pour cela, on considère la somme $S_n(\xi)$ associée à la fonctionnelle $\xi_{\theta, (p, v)}(x, y)$ définie plus haut pour trouver l'expression donnée dans l'énoncé du théorème.

Remarque 1. La condition de non-arithméticité (g) s'écrit comme suit : pour tout compact $K_0 \subset \mathbb{R}^*$, il existe $0 < \rho < 1$ tel que l'on ait : $\forall n \geq 1, \forall t \in K_0, \sup_{\theta \in \Theta, (p, v) \in J} \| Q_{\theta, p, v}(t)^n \|_V \leq O(\rho^n)$, où $\| \cdot \|_V$ désigne la norme d'opérateur sur l'espace $\mathcal{B}_V := \{ f : E \rightarrow \mathbb{C} \text{ mesurables telles que } \|f\|_V := \sup_{x \in E} |f(x)|/V(x) < +\infty \}$. À (θ, p, v) fixé, cette condition se réduit à une condition de type "nonlattice" (cf. [2]). Ici, la difficulté provient du fait qu'on requiert un contrôle uniforme en (θ, p, v) (cf. l'exemple ci-dessous).

Remarque 2. La méthode résumée ci-dessus dans le cas $|u| < 2\sqrt{\ln n}$, légèrement différente de [6], simplifie la mise en place des arguments de dérivées et des techniques spectrales de [2]. La condition (g) est plus forte que celle de [6] qui ne fait intervenir que $F^{(1)}$. Cependant, sous des hypothèses assez générales, on montre que la condition (g) se réduit aussi à une condition portant uniquement sur $F^{(1)}$, comme le suggère l'exemple suivant.

Exemple. Dans le cas des processus vectoriels autorégressifs, un théorème de Berry–Esseen avec une vitesse en $(\ln n)^{1/2}/\sqrt{n}$ est obtenu dans [5] pour l'estimateur du maximum de vraisemblance d'un paramètre scalaire. Leur hypothèse de type moment exponentiel est toutefois assez restrictive. À titre d'exemple, considérons le cas simple du système linéaire scalaire $X_n = \theta X_{n-1} + W_n$, avec $(W_n)_{n \geq 0}$ une suite de v.a.i.i.d. réelles de densité $f_W(x) := \sqrt{2}/[\pi(1+x^4)]$ et indépendantes de X_0 , et Θ un compact inclus dans $\{ \theta \in \mathbb{R}; |\theta| < 1 \}$. Le noyau de transition associé est $Q_\theta(x, dy) := f_W(y - \theta x) dy$, et on introduit $V(x) := (1 + |x|)^\gamma$ pour $0 < \gamma < 3$. Alors (\mathcal{M}) est satisfaite. Si de plus on considère l'estimateur du maximum de vraisemblance de θ ($\alpha(\theta) := \theta$), alors $F(\theta, x, y) := -\ln f_W(y - \theta x)$ et les hypothèses (a)–(g) sont vérifiées. En particulier, F ne s'écrivant pas comme la somme de deux fonctions de x et y , on montre que la condition de non-arithméticité ci-dessus est bien satisfaite.

Références

[1] D. Dehay, J.F. Yao, On likelihood estimation for discretely observed Markov jump processes, Aust. N. Z. J. Stat. 49 (1) (2007).
 [2] L. Hervé, F. Pène, The Nagaev method via the Keller–Liverani theorem, Bull. Soc. Math. France (2009), Hal-00203408, in press.
 [3] L. Hervé, J. Ledoux, V. Patilea, A uniform Berry–Esseen theorem on M -estimators for geometrically ergodic Markov chains, 2009, submitted for publication.
 [4] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Springer-Verlag, New York, Heidelberg, Berlin, 1993.

- [5] X. Milhaud, A. Raugi, Etude de l'estimateur du maximum de vraisemblance dans le cas d'un processus auto-régressif : convergence, normalité asymptotique, vitesse de convergence, *Ann. Inst. H. Poincaré* 25 (4) (1989) 383–428.
- [6] J. Pfanzagl, Asymptotic expansions related to minimum contrast estimators, *Ann. Statist.* 1 (6) (1973) 993–1026.