Int. Journal of Math. Analysis, Vol. 6, 2012, no. 25, 1205 - 1226

Maximum Likelihood Estimators in Nonlinear Autoregressive Processes

Meriem Henkouche

Département de mathématiques
USTOMB BP1505, El-m'naouar-Oran, Algérie
mhenk1094@yahoo.fr

Abstract

This paper is concerned with some asymptotic properties of maximum likelihood estimator of multivariate parameter for stable nonlinear autoregresssive .Under suitable assumptions,the consistency normality and the rate of convergence in distribution ($O(n^{-1/2})$) are settled. This rate is the same as in idd case.

Mathematics Subject Classification: 34H05, 35Q30

Keywords: MLE, Autoregressive process, Berry Essen Bound, Edgeworth, Speed of convergence

1 Introduction

Let $(X_n)_{n\in N}$ be a non linear autoregressive process defined by:

$$\begin{cases} X_0 = x_0, X_{-1} = x_{-1}., ..., X_{-d} = x_{-d+1} \\ X_n = f(X_{n-1}, X_{n-2}, ..., \theta) + \varepsilon_n \end{cases}$$

$x_0, x_{-1}, ..., x_{-d+1}$,are real numbers, the parameter θ lies on the compact parametric set $\Theta \in R^{d'}$.
$(\varepsilon_n)_{n\in N}$ is a sequence of unobserved real valued random variables on the probability space (Ω, A, P) . $(X_n)_{n\in N}$ are observed.

We are interested by asymptotic properties of the maximum likelihood estimators.
Under suitable assumptions, the consistency, the asymptotic normality and the rate of convergence in distribution $o(n^{-1/2})$ are settled. This rate is the same that in i-d-d case.

Let us denote $(Y_n)_{n\in N}$ the d.dimensional random vector
$Y_n =^t (X_n, X_{n-1}, ..., X_{n-d+1})$
$Y_0 = y_0 =^t (x_0, x_{-1}, ..., x_{-d+1})$
$e =^t (1, 0, ..., 0) \in R^d$

Therefore $\begin{cases} Y_n = F(Y_{n-1}, \theta) + \varepsilon_n \\ Y_0 = y_0 \end{cases}$

where $F(y, \theta) = (f(x_1, ..., x_d, \theta), x_1, ..., x_{d-1})$
$y =^t (x_1, ..., x_d)$

$P_{y,\theta}$ the distribution of $(Y_n)_{n\in N}$ on R^N of the chain $(Y_n)_{n\in N}$ with $Y_0 = y_0$.The transition probability of $(Y_n)_{n\in N}$ is defined by:

$$P(y, dz) = g(z_d - f(y, \theta))dz_d\delta_{\{y_d\}}dz_{d-1}...\delta_{\{y_2\}}(dz_1)$$

δ_x is the Dirac measure in x.

Part I

Probabilistic Study

2 Assumptions (HP)

(i) $(\varepsilon_n)_{n\in N}$ are real valued , i.d.d , centred random variables defined

on the the space probability (Ω, A, P) with distribution G admitting a density

g with respect to Lebesgue measure on R.

(ii) g is positive on R

(iii) A positive real number δ can be found such that

$$\int_R |\varepsilon|^\delta g(\varepsilon)d\varepsilon < \infty$$

(iv) $f : R^d \times \Theta \to R$ is continous , satisfies the following condition:

There exist two positive real numbers $a_1, a_2, ..., a_d$ verifying $\sum_{i=1}^d a_i < \rho^d$ such that for any vector θ
$f(x, \theta) \leq M + a_1 |x_1| + a_2 |x_2| + ... + a_d |x_d|$

(v) There exist two real numbers λ_0 *and* α_0 such that

$$\int_R \exp(\lambda_0 |\varepsilon|^{\alpha_0})g(\varepsilon)d\varepsilon < \infty$$

3 Assumptions1 (HP)

(i) $(\varepsilon_n)_{n\in N}$ are real valued , idd, centred random variables defined on

the probability space (Ω, A, P) with distribution G admitting a density g with respect to the Lebesgue measure on R.

(ii) g is positive on R.

(iii) $Lim_{|\varepsilon|\to\infty} g(\varepsilon) = 0$

(iv) *A positive number* δ can be found such that

$$\int |\varepsilon|^{\delta} g(\varepsilon) d\varepsilon < \infty$$

4 Theorem1

Under assumptions (HP), the markov chain $(Y_n)_{n\in N}$ is ergodic, aperiodic, Harris recurrent with an invariant measure of probability $\upsilon_\theta$equivalent to Lebesgue measure $l^{\otimes d}$of R^d.
Moreover this invariant measvre $\upsilon_\theta$has a finite moment of order δ.

4.1 Proof

Condition(C) $[Mok]$

For any compact in K in $R^d$and any borelian set B in R^d
with $l^{\otimes d}(B) > 0$,there exists an integer n_0 such that
$\inf_{y\in K} P^{n_0}(y, B) > 0$
We use the framework of $[Mokh]$.

The transition probability satisfies $[Mok]$condition C.
Let k a positive integer and $Z_n = Y_{nd}$

The transition probability $P^{kd}(.,dz)$of thischain is equivalent to the Lebesgue measure $l^{\otimes d}$.

If k is such that $(kd+1)^{\delta-1}\rho^{kd\delta} < 1$.
$(Z_n)_{n\in N}$ *is* an Harris recurrent, ergodic and geometrically ergodic chain with stationnary distribution equivalent to the Lebesgue measure and admitting a finite moment of order δ.

4.2 Definition

Let υ *a* positive $\sigma - finite$ measure on (E, B_E) *a* Markov chain $(Z_n)_{n\in N}$ on

(E, B_E) *is* said to

(i) be $\upsilon - irreductible\ if$:

$$\forall B \in B_E\ \upsilon(B) < 0 \rightarrow P(\bigcup_{n\in N}(Z_n \in B)/Z_0 = x) > 0, \forall x \in E$$

(ii) be Harris recurrent if

$$\upsilon_\theta(B) > 0 \rightarrow P(\bigcup_{n\in N}(Zn \in B)/Z_0 = x) = 1, \forall x \in E$$

$Then(Y_n)_{n\in N}$ can be shown to be an harris recurrent aperiodic chain.
Using the Orey theorem [Revuz] ,the geometric ergodicity of the chain can be proved in the following way:
Given any positive integer n, let us set:
$n = qd + r; 0 \leq r \leq d - 1$
$\left\| P^n_{(x,.)} - \upsilon_\theta \right\| \leq \left\| P^{qd+r}_{(x,.)} - \upsilon_\theta \right\| = \left\| P^r P^{qd}_{(x,.)} - \rho^r \upsilon_\theta \right\| \leq$
$\|P^r\| \, \|P - \upsilon_\theta\| \leq$
$\leq \|P^r\| \left\| P^{qd}_{(x,.)} - \upsilon_\theta \right\| \leq K\rho^q \leq (\frac{K}{\rho})\rho^{n/d}$.

5 Proposition

(i) Under assumptions (HP), for all $\alpha (0 < \alpha \leq \delta)$,

there exist positive real constants $M_\alpha, \rho, C_\alpha (0 < \rho < 1)$ such that :
$E\left[\|Y_n(y)\|^\alpha\right] \leq M_\alpha + C_\alpha \rho^{n\alpha} \|y\|^\alpha$

(ii) If there exist two positive constants α, β such that

$E\left[\exp \beta \left|\epsilon_1\right|^\alpha\right] < \infty$, then there exists a constant $\lambda (0 < \lambda < \beta)$ such that :

$E\left[\exp \lambda \|Y_n(y)\|^\alpha\right] < B(\exp \lambda C_\alpha \rho^n \|y\|^\alpha)$

B is a constant depending on ε_j

5.1 The proof is technical

6 Assumptions (HS)

(i) $(\varepsilon_n)_{n \in N}$ are , i.i.d real valued , centred variables admitting

a density strictly, positive on R

(ii) $\lim_{|\varepsilon| \to \infty} g(\varepsilon) = 0$

(iii) $\int |\varepsilon|^{2r} g(\varepsilon) < \infty \qquad (2r \geq d+1)$

(iv) The density g is continously derivable with derivative g′.

We suppose that $\quad I^\delta = \int_R \left|\frac{g'(\varepsilon)}{g(\varepsilon)}\right|^\delta g(\varepsilon) d\varepsilon \quad$ is finite
We set : $\quad I^{(2)} = \int_R \left|\frac{g'(\varepsilon)}{g(\varepsilon)}\right|^2 g(\varepsilon) d\varepsilon$

(v) There exist two positive real constants C and ρ such that

$$\left|\frac{g'}{g}(u) - \frac{g'}{g}(v)\right| < C\,|u - v|\,(1 + |u|^{p+1} + |v|^{p+1})$$
$$\left|\frac{g''}{g}(u) - \frac{g''}{g}(v)\right| < C\,|u - v|\,(1 + |u|^{p} + |v|^{p})$$

For any x in R^d the function f of the vector parameter $\theta = (\theta_1, ..., \theta_{d'})$ has continous and bounded partial

derivatives in $\otimes$

(vi) *For any x in* R^d*, the function f of the vector parameter* $\theta = (\theta_1, ..., \theta_{d'})$ *has continous and bounded partial derivatives*

in $\otimes$

(vii) Let $Df(x, \theta)$ be the gradient vector ${}^t(\frac{\partial f}{\partial \theta_1}(x, \theta), ..., \frac{\partial f}{\partial \theta_{d'}}(x, \theta))$

$\int Df(x, \theta)^t Df(x, \theta) dx$ is not singular.

(vii) There exist two positive real numbers C and $q (2q < \delta)$ *such* that

a-$\|Df(x, \theta)\| \leq C(1 + \|x\|^q)$
b-$\|Df(x, \theta) - Df(x, \theta')\| \leq C\,\|\theta - \theta'\|\,(1 + \|x\|^q)$

$\|.\|$ denote the euclidian norm in $R^{d\prime}$.

(ix) For any x in R^d , $f(x, \theta)$ has a function of the vector parameter θ

has continous and partial derivatives .

Let $D^2 f$ the hessian of f. We assume the following conditions:

$$\|D^2 f(x, \theta)\| \leq C(1 + \|x\|^q)$$
$$\|D^2 f(x, \theta) - D^2 f(x, \theta^*)\| \leq C\,\|\theta - \theta^*\|\,(1 + \|x\|^q)$$

(x) For any θ in $\otimes$, $\exists M > 0, f(\theta, M)$ the function of the vector

$x =^t (x_1, x_2, ..., x_d)$ has partial derivatives in R^d, There exists a positive real number $\rho(0 < \rho < 1)$, such that for any (x, θ) *in* $R^d \times \Theta$ the following relation holds:

$$\sum_{i=1}^{d} \left| \frac{\partial f}{\partial x_i}(x, \theta) \right| \leq \rho^d$$

(xi) There exist two real numbers $\rho_1, and\ C(\rho_1)$ such that

for any $\theta \in \Theta, \lambda$, positive real number and τ vector in $R^{d\prime}$ with $\|\tau\| = 1$, we have:

$$E\left[\Delta^2(\theta, \tau)1_{(0,1)}(\Delta(\theta, \lambda\tau)) + (1 - \rho_1)1_{(1,\infty)}(\Delta(\theta, \lambda\tau))/Y_1 = y\right] \geq \lambda^2 C(\rho_1)$$

where $\Delta(\theta, \lambda\tau) = |f(Y_{d^*+1}, \theta + \lambda\tau) - f(Y_{d^*+1}, \theta)|$
and $d^* = sup(d, d')$.

7 Theorem2

Under assumptions **(HP),(HS)**, the sequence of maximum

likelihood estimators (MLE) $(\hat{\theta}_n)_{n \in N}$ verifies :

(i) $(\hat{\theta}_n)_{n \in N}$ is consistent

(ii) $L_{y,\theta}(\sqrt{n}(\hat{\theta}_n - \theta)) \rightarrow N(0, I_\theta^{-1})$.

$$I_\theta = I^{(2)} E_{\upsilon_\theta}\left[Df(.,\theta)^t Df(.,\theta)\right]$$

(iii) Large deviation inequality:

For all real positif number γ any compact set in R^d,
there exist three positive real constants A_1,A_2 ,
such that:

$$\sup_{\otimes \times K} P_{\theta,y_0}(\left\|\hat{\theta}_n - \theta\right\| > \gamma) < A_1 \exp -A_2\gamma^2 n$$

(iv) There exists a positive real number B such that:

$$\sup{}_{\otimes \times K} P_{\theta,y_0}\left[\sqrt{n} I_\theta^{1/2}\left\|\hat{\theta}_n - \theta\right\| > B\sqrt{Logn}\right] = 0(\tfrac{1}{\sqrt{n}}).$$

7.1 Proof

Let $Z_n(t) = \prod_{i=1}^n \frac{g(Xi - f(Yi-1,\theta+t/\sqrt{n}))}{g(Xi-f(Yi-1,\theta))}$ be the random

likelihood ratio for t such that $\theta + t \in \Theta$.

We achieved the goals using the well known method
of Ibragimov and Khas'minski.

We need the convergence of finite dimensional distributions
of the process Z_n(t).(Lemma 8.2)

Then a majoration of the trajectories of process (lemma 8.3)
and their modulus of continuity (Lemma 3).
The Theorem is then a consequence of th 10.1 p104
[I.A,Ibragimov and Khas′minski] .

7.2 Lemma

Under the assumptions **(HS),**

$$LogZn(t)\text{-}\frac{1}{\sqrt{n}}\sum_{j=1}^{n} < Df(Y_{j-1},\theta),t)\frac{g'}{g}(\varepsilon_j) + \frac{1}{2} < I_\theta t,t > \rightarrow_{P_{y,\theta}} 0$$

uniformly in $\theta \in \Theta$, *and* $y_0 \in K_0$

7.2.1 Proof

It is sufficient to show that uniformly on $\Theta \times K_0 \times N$, *the* rand*om function*:

$\tau^* \rightarrow \phi j(\theta,\tau^*) = \frac{g^{1/2}(X_j - f(Y_{j-1},\theta+\tau^*))}{g^{1/2}(X_j - f(Y_{j-1},\theta))}$ is differentiable in quadratic mean at the point $\tau^* = 0$ for $P_{y,\theta}$.

[*G.Roussas*.Contiguity of probability measures] .

7.3 Lemma

Under the assumptions **(HS)**, there exists a positive real constant C such for any θ in Θ, y *in* K_0 and t such that $\theta + \frac{t}{\sqrt{n}} \in \Theta$,

$$E_{\theta,y_0}\left[Z_n^{1/2}(t)\right] \leq \exp -C\left\|t\right\|^2$$

7.3.1 Proof

The proof is technical.

Part II

Speed of convergence

8 Speed of M.L.E Convergence

8.1 Notations

Let l(.,.,.) the function defined by

$l(.,.,.) : R \times R^d \times \Theta \to R$

$$l(x,y,\theta) = Logg(x - f(y,\theta))$$

For i=1,2,... , we set $l_i(\omega,\theta) = l(X_i(\omega), Y_{i-1}(\omega), \theta))$

$$l'(x,y,\theta) =^t (\frac{\partial l}{\partial \theta_1}(x,y,\theta), ..., \frac{\partial l}{\partial \theta_k}(x,y,\theta))$$
$$l"(x,y,\theta) =^t (\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}(x,y,\theta))_{i,j=1,...,k}$$

$$l_i'(\theta) = l'(X_i, Y_{j-1}, \theta)$$
$$l_i"(\theta) = (\frac{\partial^2 l}{\partial \theta_i \partial \theta_j}(X_i, Y_{j-1}, \theta))_{i,j=1,...,k}$$

$$Df(y,\theta) =^t (\frac{\partial f}{\partial \theta_1}(y,\theta), ..., \frac{\partial f}{\partial \theta_k}(y,\theta)).$$

We have :$l'(x,y,\theta)^t l'(x,y,\theta) = \frac{g'^2}{g^2}(x - f(y,\theta))Df(y,\theta)^t Df(y,\theta)$

and $l'(x,y,\theta) = \frac{g'}{g}(x - f(y,\theta))Df(y,\theta)$

$$I_\theta = E_{\upsilon_\theta}[l'(.,\theta)^t l'(.,\theta))] = -E_{\upsilon_\theta}[l"(.,\theta)]$$

We denote

$$m(x,y,\theta) = I_\theta^{-1/2} l'(x,y,\theta)$$
$$h(x,y,\theta) = I_\theta^{-1/2} l''(x,y,\theta) I_\theta^{-1/2} + I$$

I is the identity operator in R^d

$m_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n I_\theta^{-1/2} l_i'(\theta) h_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n (I_\theta^{-1/2} l'' i(\theta) I_\theta^{-1/2} + I)$

$hn = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (I_\theta^{-1/2} l_i''(\theta) I_\theta^{-1/2} + I)$

Let $S_{d\prime}$ the set of symetric matrix on $\mathrm{R}^{d\prime}$ and
π the bijection : $R^{d\prime} \times S_{d\prime} \rightarrow R^{d\prime(d\prime+3)/2} = R^{d^*}$

$$(u,v) \rightarrow \pi(u,v) = \begin{cases} u_i & i = 1,..d' \\ w_{j(i),l(i)} & i = d'+1, ...d^* \end{cases}$$

where $(j(i), l(i))$ is the solution of $d^* + j + \frac{l(\ l\ -1)}{2} = i, \qquad j \leq l.$

Conversly , given the R^{d^*} $vector\ r, u(r) = (r_1, ..., r_{d'})$
W(r) is the symetric matrix , where te term of the j-th
row and l .th colomn $(j \leq l)$ is

$w_{j,l} = r_i,\ and\ \ i = d^* + j + \frac{l(l-1)}{2}.$
We denote $\qquad \pi_n(\omega, \theta) = \pi(m_n(\omega,\theta), h_n(\omega,\theta)).$
$C(\theta) = E_{\theta \otimes v_\theta} \left[\pi(m(.,\theta), h(.,\theta))^t \pi(m(.,\theta), h(.,\theta) \right].$

Given a set A in $R^{d\prime}$ and a positive real number B,

$A^\beta = \{ r / \|r - s\| < \beta \text{ for some s } \in A \}$,
$A^{-\beta} = C(CA)^\beta.$

Given A in $R^{d\prime}$, Z(A) is the cylinder $\mathrm{A} \times R^{d^*-d} \subset R^{d^*}$,
For $a > 0, A \subset R^k, L_n(a) = \left\{ r \in R^k / \|r\| \leq a^{-1/2} n^{1/4} \right\}$

$L\prime_n(a) = \left\{ r \in R^{d^*} / \|r\| \leq a^{-1/2} n^{1/4} \right\}$
We set :

$An(a) = A \bigcap Ln(a)$
$A_n^*(a) = \left\{ r \in R^k, \|r - s\| \leq a^* n^{-1/2} \|s\|^2 \ \ for\ s\ \in A_n(a) \right\}$

8.2 Theorem

Under assumptions **(HP)**,**(HS)**, the sequence of maximum likelihood estimators $(\hat{\theta}_n)_{n\in N}$ satisfies the following property:
For any compact set K in R^d , there exists a positive number C(K) such that for every convex set in R^d

$$sup_{\theta\in\Theta,y\in K}\left|P_{\theta,y}(I_\theta^{1/2}(\hat{\theta}n-\theta)\in E)-\Phi(E)\right|\leq\frac{C(K)}{\sqrt{n}}$$

Where Φ is the standard normal distribution on R^d , $N(0,I_{d'})$.
I$_{d'}$ is the standard identity operator on $R^{d'}$.

8.3 Proof of the theorem

The result of the theorem derives mainly from Berry Essen Bound

of suitably normalized random vector sums and this following expansion:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n l_i'(\hat{\theta}_n)=\frac{1}{\sqrt{n}}\sum_{i=1}^n l_i'(\theta)+\frac{1}{\sqrt{n}}\sum_{i=1}^n l_i''(\theta)\sqrt{n}(\hat{\theta}_n-\theta)+$$
$$+\frac{1}{\sqrt{n}}\sum_{i=1}^n\int_0^1(l''i(\theta+(\hat{\theta}_n-\theta))-l''(\theta))\sqrt{n}(\hat{\theta}_n-\theta)du.$$

With the notations , we have :

$$m_n(\theta)+\frac{1}{\sqrt{n}}(h_n-I)t_n=K_n(.,\theta)\frac{\|t_n\|^2}{(\sqrt{n})}$$

Under assumptions(**HP**)(**v**), (**HS**)(**v**)(**viii**)(**ix**),
there exist two positive real constants$C$ and $\bar{k}$ such that :

$$sup\left\{P_{\theta,y}^N(K_n(\theta))>\bar{k},\theta\in\Theta,n\in N,y\in K\right\}\leq\frac{C}{\sqrt{n}}$$

Let $\tilde{A}_n=\tilde{A}_n(\theta)=\left\{\omega/\ \|K_n(\omega,\theta)\|\leq\bar{k}\right\}$

when $\omega\in\tilde{A}_n$, we have :

$$\left\|m_n(\omega,\theta)+\frac{1}{\sqrt{n}}\left\{h_n(\omega,\theta)-I\right\}t_n\right\|\leq\bar{k}\frac{\|t_n\|^2}{\sqrt{n}}$$

Given a convex set E,

$$\{t_n\in E\}=\left\{(t_n,W(\pi_n))\in E\times R^{d^*-d'}\right\}$$

$et\ C_n(A,b) = \left\{ r \in R^{d^*} / \left\| u(r) + (\frac{W(r)}{\sqrt{n}} - I)t \right\| \leq \frac{b}{\sqrt{n}} \left\| t \right\|^2, t \in A \cap L_n(b) \right\}.$

We have , $\forall n \geq n_0, \theta \in \Theta, \omega \in A_{n,\theta} \cap \tilde{A}_n$

$$\left\| m_n(\omega,\theta) + (\tfrac{1}{\sqrt{n}} h_n(\omega,\theta) - I)\sqrt{n} I_\theta^{1/2}(\hat{\theta}_n - \theta) \right\| \leq \frac{b}{\sqrt{n}} \left\| \sqrt{n} I_\theta^{1/2}(\hat{\theta}_n(\omega) - \theta) \right\|^2.$$

$b = \bar{K}_{sup\theta\in\Theta} \left\| I_\theta^{-1/2} \right\|^2$

$\forall n \geq n_0, \theta \in \Theta, \omega \in A_{n,\theta} \cap \tilde{A}_n$

$\sqrt{n} I_\theta^{1/2}(\hat{\theta}_n - \theta) \in E \Rightarrow \sqrt{n} I_\theta^{1/2}(\hat{\theta}_n(\omega) - \theta) \in E \cap L_n(b)$
$\Rightarrow \pi(m_n(\omega,\theta), h_n(\omega,\theta)) \in C_n(E,b)$

We have : $C_n(E,b) \subset (ZE)_n^*(\sqrt{2}b) \cup \overline{L', (2b)}$ $\qquad [Pf.7.1]$

Therefore $P_{\theta,y}^N \left| [\omega \in \Omega / \pi_n(\omega,\theta) \in C_n(,b)] - \Phi(E) \right| \leq |P_1| + |P_2|.$

$P_1 = \left| P_{\theta,y} \left[\omega \in \Omega / \pi_n(\omega,\theta) \in (ZE)_n^*(\sqrt{2}b) \right] - \Phi(E) \right|$

$P_2 = P_{y,\theta} \left[\omega \in \Omega / \pi_n(\omega,\theta) \in \overline{L'_n(2b)} \right].$

$P_2 = o(\frac{1}{\sqrt{n}})$

Let $S_n(y) = \frac{1}{\sqrt{n}} \sum_{i=1}^n V(\varepsilon_k, Y_{k-1}(y))$
We are interested by the speed of convergence in law of such fonctionnals:

The quantities such as will be evalued are
$P_{\theta,y}\left[\| S_n(y) \| \geq \gamma \right], \qquad \gamma > 0$
$\Psi_n(t) = E \left[\exp i < t, \frac{1}{\sqrt{n}} \sum_{i=1}^n V(\varepsilon_k, Y_{k-1}(y)) > \right]$

8.4 Assumptions(HT)

Let beV(ε, y) a function $R \times R^d \to R^r$,

η is a function R$\to [1, \infty[, \gamma, \delta, \tau (\delta < \alpha_0)$, *three* positive real numbers such that:

(i) $\|V(\varepsilon, y)\| \leq \eta(\varepsilon)(1 + \|y\|^{1+\tau})$

(ii) $\|V(\varepsilon, y) - V(\varepsilon, y')\| \leq \eta(\varepsilon)(1 + \|y\|^{\tau} + \|y'\|^{\tau}) \|y - y'\|$

(iii) $\int \eta^6(\varepsilon) \exp \gamma(|\varepsilon|^{\delta}) g(\varepsilon) d\varepsilon < \infty$

(iv) The vectors $V(\varepsilon, y)$ are centred and reduced under $g(\varepsilon) d\varepsilon \upsilon_\theta(dy)$.

First we introduce a Banach space and a linear operator linked to the

conditionnal expectation of the chain .

Let $B = B_{\alpha,\lambda}$ be the set of the functions $h : R^d \to R^r$ such that :

$M(h) = sup_{x \in R^d} \frac{|h(x)|}{(1+\|x\|^{1+\tau+c}}$,

$L(h) = sup_{x,y \in R^d} \frac{|h(x)-h(y)|}{\|x-y\|(1+\|x\|^{\tau}+\|y\|^{\tau})}$

$\tau > 0, c > 0.$

The operator U$_\theta$ is defined on B

$U_\theta h(y) = \int h(\varepsilon e + F(y, \theta) g(\varepsilon) d\varepsilon$

8.5 Lemma1:

Under assumptions **(HS)**, $\mathbf{U}_\theta$ is quasi-compact and admits the following decomposition

$$U_\theta = \pi_\theta + Q_\theta$$

where π_θ is the projector, Q$_\theta$ is an operator with spectral radius ρ_θ less than 1.

Moreover, there exist two real positive constants **C*** and ρ $(0<\rho < 1)$ such for any $\theta\ in\ \Theta, we\ have\ \|Q_\theta^n\| \leq C^* \rho^n$

where $\|.\|$is the usual operator norm.

8.6 Proposition

Under the assumptions **(HS)**, for every $\gamma \in K$ (compact in R^d), *there* exist a positive constant C such that :
$\forall s\ (2 \leq s \leq 6),\ P^N_{\theta,y}\left[\|S_n(y)\| > \gamma\right] \leq \frac{C}{\gamma^s}$

8.6.1 Proof

For $s \geq 2, by$ Markov inequality,

$$\mathrm{P}_{\theta,y}\left[\|S_n(y)\| > s\right] \leq \frac{1}{n^{1/2}\gamma^s} E\left[\|\sum_{i=1}^n V(\varepsilon_k, Y_{k-1}(y)\|^s\right] = \frac{B_n}{\gamma^s n^{s/2}}$$

Let $H(y) = \int V(\varepsilon, y)\mu(d\epsilon)$.
$B_n = E_{\theta,y}\left[\|\sum V(\epsilon_k, Y_{k-1}) - H(Y_{k-1}(y) + H(Y_{k-1}(y)\|\right] \leq$

$$2^{s-1}(E\left[\|B_n^1\|^s\right] + E\left[\|B_n^2\|^s\right])$$

$B_n^1 = \sum_{i=1}^n V(\varepsilon_k, Y_{k-1}) - H(Y_{k-1}(y))$
$B_n^2 = \sum_{i=1}^n H(Y_{k-1}(y))$
B_n' is a norm of increments of martingales .
$E\left[\|B_n'\|^s\right] \leq C_s E_{\theta,y}\left[\|B_n'\|\ 2\right]^{s/2}$

$$\leq C_5 n^{s/2-1} 2^s \sum_{i=1}^n E_{\theta,y}\left[\|V(\varepsilon_k, Y_{k-1})\|^s\right].$$

Using assumptions about $\|V(\varepsilon, y)\|$,we conclude that
$E_{\theta,y}\left[\|B'n\|^s\right]$ is bounded for all n

8.7 A representation of the characteristic function of S_n

Let us set, for a function $h : R^d \to R$

$U_\theta h(y) = \int \exp i < t, S_n > h(\varepsilon e + F(y,\theta))g(\varepsilon)d\varepsilon$

Then the characteristic function of S_n can be written

$\Psi n(t) = E\left[\exp i < t, S_n >\right] = U_{\theta + \frac{t}{\sqrt{n}}} 1.$

where 1 is the constant function :y→ 1.

- $U_{\theta,t}$ is a linear operator on $B_{\lambda,\alpha}$

- Using the perturbation theory of linear, we will derive an

Edgeworth'expansion of the characteristic function of S_n

For α and λ *two* positive $cons\tan ts$ $0 < \alpha < 1), (0 < \lambda < \beta),$
For $h : R^d \to C, let$
$|h|_\lambda = \sup_{\times \in R^d} \frac{|h(x)|}{(1+\|x\|)^{1+\lambda} \exp \lambda \|x\|^\alpha}$
and $m_\lambda(h) = \sup_{x,y \in R^d} \frac{|h(x)-h(y)|}{\|x-y\| \exp \lambda [\|x\|^\alpha + \|y\|^\alpha]}$

$B_{\lambda,\alpha}$ *with the norm* $(|.|_\lambda + m_\lambda)$ is a Banach space.

8.8 Proposition

Under $(HS), for\ h \in L_\lambda,$ $\quad U_\theta^n h = \int h d\nu_\theta + Q_n^0 h$

Q_0 is an operator with spectral radius less than 1 and $Q_0 1=0$.

8.8.1 Proof

By Marinescu Tulcea theorem.

8.9 Proposition

Under $(HS), there$ exist a positive constant on such that :

$$\|t\| < \eta, h \in B_\lambda,\ n \geq 1$$
$$U_t^n h = s^n(t)\pi(t)h + Q^n(t)h$$

where s(t) is an eigenvalue of U_t, its eigenspace E(t) has its dimension equal to 1, $\pi(t)$ is the projector of E(t)
$Q(t)$ has its spectral radius <1.
$Q(t)\pi(t) = \pi(t)Q(t) = 0$

t$\rightarrow s(t), t \rightarrow \pi(t)$ are C^3.

THE PERTURBATION THEORY OF LINEAR YIELDS

8.10 Proposition

For $\|t\|$small enough ($\|t\| < T$),the operator U$_{\theta,t}$ has the following representation:
U$_{\theta,t} = s_\theta(t)\pi_\theta(t) + Q_{\theta,t}$.

where $s_\theta(t)$ is an simple eigenvalue,$\pi_\theta(t)$ the projector on the one dimensionnal eigenspace of s$_\theta(t), Q_\theta(t)$, an operator with spectral radius less than 1(s$_\theta(0) = 1, \pi_\theta(0) = \pi_\theta, Q_{\theta,0} = Q_\theta$)

The functions $s_\theta(.), \pi(.), Q(.)$, have five derivatives in a neighborhood of the point $t = 0$.

These derivatives are uniformly bounded on Θ, and there exist two positive real constants C$_1^*$ and ρ_1 such that for any θ in Θ.

$$\|Q_\theta^n\| \leq C_1^* \rho_1^n \qquad \text{when} \qquad \left\|\frac{t}{\sqrt{n}}\right\| \leq T_1$$

on a:$\Psi_n(\frac{t}{\sqrt{n}}) = U^n_{\theta,\frac{t}{\sqrt{n}}}1 = s_\theta^n(\frac{t}{\sqrt{n}})\pi_\theta(\frac{t}{\sqrt{n}})1 + Q^n_{\theta,\frac{t}{\sqrt{n}}}1.$
$= exp(nLogs_\theta(\frac{t}{\sqrt{n}}))\pi_\theta(\frac{t}{\sqrt{n}})1 + Q^n_{\theta,\frac{t}{\sqrt{n}}}1.$

8.11 Notations

$s_\theta^{(k)} = s_k, \pi_\theta^{(k)} = \pi^k$ is the derivative k-th derivative of the functions .These are multilinear on R^d.

$s^{(k)}(t) = s^{(k)}(t_1,...,t_k);$

$\pi^{(k)}(t) = \pi^{(k)}(t_1, ..., t_k)$

8.12 Lemma

For any $\theta \in \Theta, t \in R^d, s_\theta^{(1)}(0) = 0, s_\theta^{(2)}(t) = -\|t\|^2$

8.12.1 Proof

For all t near zéro , we have:

$\frac{1}{n}E\left[< S_n(y), t >\right] \to < \int\int V(\varepsilon, y)\mu(d\varepsilon)\nu(dy), t >= 0$

and $\frac{1}{n}E\left[< S_n(y)^t Sy(y)t, t >\right] \to \|t\|^2$

$t = (e_i)_{i=1}^p$

We take the derivative in the left and the right at the point zero , $n \to \infty$

$U^n_{\frac{t}{\sqrt{n}}} = s_n(\frac{t}{\sqrt{n}})\pi(\frac{t}{\sqrt{n}}) + Q_n(\frac{t}{\sqrt{n}})$

8.13 Proposition

There exist two positive constants , C, and C_4 such that for $\left\|\frac{t}{\sqrt{n}}\right\| < A$

$$\left|\hat{H}_n(\frac{t}{\sqrt{n}})\right| = \left|U^n_{\frac{t}{\sqrt{n}}}1 - \exp -\frac{\|t\|2}{2}(1 + \frac{s^{(3)}(0)t^{(3)}}{3!\sqrt{n}} + \pi'(0)1\frac{t}{\sqrt{n}}\right| \leq$$
$$\frac{C}{n}\exp -\frac{\|t\|^2}{4}(\|t\|^2 + \|t\|^8) + C_4\tau^n.$$

8.14 Lemma1

For $\|t\| < T\sqrt{n}$, there exists a real positive constant C such that :

$$\left| s^n(\frac{t}{\sqrt{n}}) - \exp(-\frac{\|t\|^2}{2})(1 + n^{-1/2}\frac{s^{(3)}(t)}{6}) \right| \leq$$
$$C((\frac{\|t\|^4}{n} + \frac{\|t\|^6}{n}) \exp -\frac{\|t\|^2}{4})$$

8.14.1 Proof

Using Taylor's expansion of $s(\frac{t}{\sqrt{n}})$ about zero

$$s(\frac{t}{\sqrt{n}}) = 1 - n^{-1} \left\| \frac{t}{2} \right\|^2 + \frac{n^{-3/2}}{3!} s^{(3)}(t) + n^{-2}\eta_1(\frac{t}{\sqrt{n}}) \|t\|^4 .$$

$$\eta_{1,\infty} = \sup \{ |\eta_1(t)| , \theta \in \Theta, \|t\| < T_1 \} < \infty$$

After,

$$\pi(\frac{t}{\sqrt{n}}) = 1 + n^{-1/2}\pi^{(1)}(t)1 + \frac{\|t\|2}{n}\eta_2(\frac{t}{\sqrt{n}}).$$

Then

$$\left| s^n(\frac{t}{\sqrt{n}})\pi(\frac{t}{\sqrt{n}}) - \exp -\frac{\|t\|^2}{2}(1 + n^{-1/2}(\frac{s^{(3)}(t)}{6} + \pi^{(1)}(t)1)) \right|$$
$$\leq C\frac{\|t\|^4 + \|t\|^8}{\eta} \exp -\frac{\|t\|^2}{4}.$$

The norm of the term $Q^n_{\frac{t}{\sqrt{n}}}1$ will be majored by C_1^*, ρ_1^n.

8.15 Acknowledgement

The author is indebted to Professor Xavier Milhaud
(Département de statistique et de probabilités, Toulouse, France)
for his help during the prepartion of this work.

References

[1] R.H Battacharya and Ranga Rao, Normal approximation and asymptotic expansions , Wiley, New York, 1976.

[2] P.Doukhan, M.Ghindès,

a.Etude du processus X_n=f(X_{n-1})+ϵ_n,

Thèse de 3 ème cycle , Université Paris Sud, Département de mathématiques , Orsay-France, 1980.

b.Etude du processus: X_{n+1}=f(X_n).ϵ_n, $CRAS\ IT$ 290 serie$A$, (1980), p921 − 923.

[3] I.A.Ibragimov et Khas'minki, Asymptotic behaviour of statistical estimators in the smooth case

I.Study of the likelihood ratio , Theor Probab Appl , vol17, no=3 (1972), p445-462.

II.Statistical estimation , Asymptotic theory, Springer Verlag, 1981.

[4] T.Kato, Perturbation thory for linear operators, Springer Verlag, Berlin, Heidelberg, 1966.

[5] X.Milhaud et A.Raugi, Etude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autoregressif: convergence , normalité symptotique et vitesse de convergence, Annales de l'institut Poincarré no=4, 1989.

[6] X.Milhaud , Problèmes asymptotiques dans l'étude des chaînes de Markov, Thèse de doctorat , Université Paul Sabatier, Toulouse,1981.

[7] A.Mokkadem, Thèse de doctorat d'état ès sciences mathématiques 25 septembre, Paris Sud-Centre Orsay, 1987.

[8] J.pfanzagl,

1-On the accuracy of normal approximation of minimum contrast estimates, Z.Warhr.Verw.Gebiete, 25 (1973), p171-180.

2-The berry Essen Bound for minimum contrast estimate, Metrika, 17 (1971), p82-91.

[9] B.M.Pötscher, Convergence results for maximum likelihood type estimators in multivariate ARMA models, J.Multivariate Anal, 21 (1987), p29-52.

[10] Revuz, Markov chains, North Holland, American, Elsevier, 1976.

[11] G.Roussas, Contiguity of probability mesures, Cambridge, University Press, 1971.

[12] D.Tjøstein, Estimation in nonlinear times series, Stochastic processes and their applications, 21 (1986), p251-273.

[13] R/L Tweedie, Criteria for classifying general markov chains, Adv.Appl.Prob, 8 (1976), p737-771.

Received: November, 2011