XAVIER MILHAUD

Un M estimateur pour un processus autorégressif non explosif. Quelques propriétés asymptotiques

Publications de l'Institut de recherche mathématiques de Rennes, 1988, fascicule 1
« Probabilités », , p. 65-90
<http://www.numdam.org/item?id=PSMIR_1988___1_65_0>

UN M ESTIMATEUR POUR UN PROCESSUS AUTOREGRESSIF NON EXPLOSIF. QUELQUES PROPRIETES ASYMPTOTIQUES.

X. MILHAUD
Institut de Mathématiques
Place Eugène Bataillon
34060 MONTPELLIER CEDEX 1

0. Introduction.

Nous étudions ici le comportement asymptotique d'un M estimateur dans un processus autoregressif non explosif d'ordre d où les innovations sont independantes et de loi symetrique inconnue.

L'idée d'un tel estimateur est issue d'une monographie [1] où l'auteur H.J. Bierens l'emploie dans des problèmes de regression.

Nous étudions d'abord sa consistance et sa normalité asymptotique, les innovations ayant un moment d'ordre $\delta > 0$ [Th. 2.1 et 2.2]. La démonstration en est classique.

Ensuite nous évaluons de façon explicite, avec calcul de constantes, une vitesse de convergence de l'estimateur vers le paramètre. Nous ne passons pas par des coefficients de melange [6]. Nous considérons plutôt l'espérance conditionnelle comme un opérateur agissant sur un espace fonctionnel convenablement choisi. Cette méthode a déjà été introduite dans [5]. Dans cette dernière publication les propriétés spectrales d'un tel opérateur se déduisent du théorème de Ionescu-Tulcea et Marinescu [7]. Les renseignements obtenus sur le spectre sont plutôt qualitatifs et valables pour chaque valeur θ du paramètre.

Dans le présent article une évaluation plus précise du spectre se fait grâce a un autre choix des espaces fonctionnels et par une méthode tout à fait différente. Nous obtenons alors une convergence uniforme sur l'espace paramétrique et sur un ensemble convenablement choisi de lois de l'innovation.

Les constantes obtenues sont cependant très grandes, cela est dû à l'utilisation des martingales mais aussi aux majorations statistiques employées.

Le plan de l'article est le suivant :

1- Présentation générale du processus et M-estimateur.
2- Théorèmes principaux.
3- Hypothèses H.
4- Quelques résultats probabilistes.
5- Démonstration des théorèmes 2.1 et 2.2.
6- Operateurs associés à la chaîne.
7- Vitesse de convergence (démonstration du théorème 2.3).

1. Présentation générale du processus et M estimateur.

1. Le processus.

On considère un processus autoregressif d'ordre d AR(d) défini par les équations suivantes

$$X_0 = x_0 \dots X_{-d+1} = x_{-d+1}$$

(1.1)

$$X_n = \theta_1 X_{n-1} + \dots + \theta_d X_{n-d+1} + \varepsilon_n$$

où $(x_0,\dots,x_{-d+1})$ sont des constantes réelles arbitraires mais fixées. Les X_i $(i=1,2,\dots)$ sont observés.

$(\varepsilon_n,\ n = 1,2,\dots)$ sont des variables aléatoires indépendantes identiquement distribuées de densité g relativement à la mesure de Lebesgue, symétriques, unimodales [cf : Hypothèses 3.H.1.1.].

$\theta_1,\dots,\theta_d$ sont des paramètes réels tels que le processus autoregressif ne soit pas explosif [cf. Hypothèses 3 H.2]. Θ désignera l'espace paramétrique.

Pour $n \geq 0$ on notera :

$$Y_n = {}^t(X_n,\dots,X_{n-d+1})\ ;\ y_0 = {}^t(x_0,\dots,x_{-d+1})\ ;\ \theta = {}^t(\theta_1,\dots,\theta_d)$$

Pour chaque $\theta \in \Theta$ la chaîne $(Y_n,\ n \geq 1)$ admettra une loi stationnaire ν_θ et $P^{\mathbb{N}}_{\theta,y_0}$ désignera la loi sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\otimes \mathbb{N}})$ induite par la chaîne.

2. Méthode d'estimation.

Soient ρ et s deux fonctions reelles positives paires de la variable reelle. ρ est supposee unimodale (cf. Hypothèses 3 H.3.)

Soit T un cube compact de $\mathbb{R}^d$ contenant l'espace paramétrique Θ et centré à l'origine.

Etant donné les observations $X_1(\omega),\ldots,X_n(\omega)$ et les fonctions ρ et s définies par les hypothèses H3 et H4 on désigne par $\hat{\theta}_n(\omega)$ l'une des solutions (qui existera toujours) de l'équation

$$\sum_{i=1}^{n} \rho\left(\frac{X_i(\omega) - <\hat{\theta}_n(\omega), Y_{i-1}(\omega)>}{s\left(\| Y_{i-1}(\omega)\|\right)}\right) = \sup_{\theta^* \in T} \sum_{i=1}^{n} \rho\left(\frac{X_i(\omega) - <\theta^*, Y_{i-1}(\omega)>}{s\left(\| Y_{i-1}(\omega)\|\right)}\right)$$

$\hat{\theta}_n(.)$ est alors un M-estimateur du paramètre θ.

2. Théorèmes principaux.

1. Consistance.

Théorème.

Sous les hypothèses H.1.1.,2,3.1,3.2,4.1,4.2 et pour chaque $\theta \in \Theta$ la suite des M estimateurs$(\hat{\theta}_n(.), n \in \mathbb{N})$ est $\mathbb{P}_{\theta,y}^{\mathbb{N}}$ presque sûrement convergente.

Démonstration en 5.1.

2. Convergence en loi.

Théorème.

Sous les hypothèses H.1.1,2,3,4.1,4.2 pour chaque θ dans Θ la suite normalisée d'estimateurs $(\sqrt{n}\ (\hat{\theta}_n-\theta), n \in \mathbb{N})$ converge en loi vers une loi normale $N(0,\Gamma(\theta))$ centrée de covariance $\Gamma(\theta)$ sur $\mathbb{R}^d$, définie par

$$\Gamma = M_\theta^{-1}\, \Lambda_\theta\, {}^tM_\theta^{-1}$$

avec

$$M(\theta) = \int_{\mathbb{R}\times\mathbb{R}^d} \psi'\left(\frac{\varepsilon}{s^2(\|y\|)}\right) \frac{y^t y}{s^2(\|y\|)}\, g(\varepsilon)\, d\varepsilon\, \nu_\theta(dy)$$

$$\Lambda(\theta) = \int_{\mathbb{R}\times\mathbb{R}^d} \psi^2\left(\frac{\varepsilon}{s(\|y\|)}\right) \frac{y^t y}{s^2(\|y\|)}\, g(\varepsilon)\, d\varepsilon\, \nu_\theta(dy)$$

où ψ et ψ' sont définies en H.3.

Démonstration du théorème en 5.3.

3. Vitesses de convergence.

Théorème.

Sous les hypothèses H. étant donné deux nombres réels positifs γ et k il existe 2 constantes positives (J_k^i, i=1,2) ne dependant que de γ et k telles que

$$\sup_{\theta\in\Theta} P^{\mathbb{N}}_{\theta,y_0}\left(\|\hat{\theta}_n - \theta\| > \gamma\right) < \left(\frac{J_k^1}{n^{k/2}} + \frac{J_k^2}{n^k}\right)\frac{2p}{(\Delta(\gamma))^k} \times R^k$$

où $$p = \left(\left[\frac{3s_1\psi_1\sqrt{d}\,|T|}{2\Delta(\gamma)}\right] + 1\right)^d$$

(avec [a] = partie entière du nombre réel a, | T | longueur de l'arête de T)

$$\Delta(\gamma) = \frac{1}{2}\rho^*(1-\alpha)^2\left((\gamma s_0)^2 \wedge v_0^2\right)$$

(avec $a \wedge b = \inf(a,b)$)

$$J_k^1 = 2^{3k-2}\left(2^{k-2}\left(\frac{18k^{3/2}}{(k-1)^{1/2}}\right)^k c^\delta \mu^*_k\left(\frac{k}{\delta(1-\tau)}\right)^k\left(1+C^\delta(1-\tau)^{-\delta}\left(\mu_\delta + \|y_0\|^\delta\right)\right) + \left(\frac{18k^{3/2}}{(k-1)^{1/2}}\right)^k\right)$$

$$J_k^2 = 2^{3k-2}\, c^\delta \mu^*_k\left(\frac{k}{\delta(1-\tau)}\right)^k\left(1+C^\delta\left(\mu_\delta + \|y_0\|^\delta(1-\tau)\right)(1-\tau)^{-\delta}\right)$$

où $\mu^*_k = 2^k$ si $\mu_\delta^{1/k} < 1 \cdot \tau^{\delta/k}$

$= 2^{k-1} \left(1+\left(\frac{k}{\delta(1-\tau)}\right)^k \mu_\delta\right)$ sinon

et $R = 3\,(\rho(0) + \psi_1(T + T\, s_1 \|s'\|_\infty + \mu_\delta^{1/\delta} \|s'\|_\infty)$.

Les diverses quantités sont introduites dans les hypothèses.

Démonstration en 7.

3. Hypothèses H :

Nous exposons ici toutes les hypothèses dont nous avons besoin pour les théorèmes énoncés ci-dessus.

H.1.1. (ε_n, $n \in \mathbb{N}$) est une suite de variable aléatoires réelles, définies sur un espace de probabilité($\Omega, \mathcal{A}, \mathbb{P}$), indépendantes, de même loi admettant une densité g relativement à la mesure de Lebesgue, paire, unimodale. Il existe un nombre réel positif δ vérifiant

(3.1) $\displaystyle\int_{\mathbb{R}} |\varepsilon|^\delta\, g(\varepsilon)\, d\,\varepsilon = \mu(\delta) = \mu_\delta < \infty$

1.2. Il existe des constantes positives β_1, β_2, α $(0 < \alpha < 1)$ telles que

(3.2) $P(\beta_1 < |\varepsilon| < \beta_2) \geq 1\text{-}\alpha$

H.2. L'espace paramétrique Θ est d'intérieur non vide dans $\mathbb{R}^d$ compact et égal à la fermeture de son intérieur ($\Theta = \overline{\mathring{\Theta}}$). Il est tel que pour tout $\theta = {}^t(\theta_1,\ldots,\theta_d)$ dans Θ l'équation

(3.3) $z^d - \theta_1 z^{d-1} - \theta_2 z^{d-2} - \ldots - \theta_d = 0$

ait toutes ses racines réelles ou complexes de module inférieur ou égal à un nombre réel τ^* fixe $(0 < \tau^* < \tau < 1)$.

Nous savons alors qu'il existe une constante réelle positive C telle que les matrices

$$A_\theta = \begin{pmatrix} \theta_1 & \dots & & \theta_d \\ 1 & 0 & & \dots 0 \\ 0 & 1 & 0 & \dots 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}$$

vérifient pour tout n

$$\sup_{\theta \in \Theta} \|A_\theta^n\| \leq C\, \tau^n$$

nous ne calculerons pas C dans cet article.

H.3.1. ρ est une fonction réelle de la variable réelle, positive, continue, bornée, paire, unimodale, intégrable pour la mesure de Lebesgue. On pose

$$\rho_0 = \rho(0) = \max_x \rho(x)$$

3.2. ρ est continuement dérivable, sa dérivée ψ est bornée. On pose

$$\psi_1 = \max_{x \in \mathbb{R}} |\psi(x)|$$

3.3. ρ admet une dérivée seconde ψ' bornée, absolument continue sur $\mathbb{R}_+$ négative en zéro positive pour x grand, ne s'annulant qu'en en seul point de $\mathbb{R}_+$

$$\psi_2 = \sup_{x \in \mathbb{R}} \{|\psi'(x)|, x \in \mathbb{R}\}$$

$$\psi_3 = \sup_{x \neq y} \left\{ \frac{|\psi'(x) - \psi'(y)|}{|x - y|} \right\} < \infty$$

3.4. Il existe deux constantes positives ρ^* et v_0 telles que les relations $0 < \varepsilon < \beta_2$ et $0 < v < v_0$ impliquent

$$\rho(\varepsilon + v) \leq \rho(\varepsilon) - \rho^*\, v^2.$$

H.4.1. s est une fonction réelle de la variable réelle, continue, paire, strictement positive, croissante sur $\mathbb{R}_+$ bornée inférieurement par s_0 nombre réel positif supérieur ou égal à 1

(3.4.) $1 \leq s_0 = s(0) = \inf \{ s(x) \mid x \in \mathbb{R} \}$

la fonction s et donc s_0 est choisie telle que

(3.5) $\displaystyle\int_{\mathbb{R}} \psi'\left(\frac{\varepsilon}{s_0}\right) g(\varepsilon)\, d\varepsilon \leq \tilde{\psi}' < 0$

4.2. On suppose de plus l'existence d'un réel positif s_1

(3.6) $\displaystyle s_1 = \sup_{x \in \mathbb{R}} \frac{|x|}{s(x)}$

4.3. il existe un reel positif s_2 tel que

(3.7) $s_2 = \min \{ \frac{x}{s(x)} \mid x > \beta_1 \kappa(d,\tau) \}$

avec

(3.8) $\displaystyle \kappa(d,\tau) = \frac{(d-1)^{\frac{d-1}{2}} \left(1-\tau^2\right)^{\frac{d-1}{2}}}{\sqrt{d}\, C^{d-1}}$

4.4. s' est supposé bornée sur $\mathbb{R}$. On pose

$$\|s'\|_\infty = \sup_{x \in \mathbb{R}} (|s'(x)| \; |x \in \mathbb{R})$$

Commentaire:

1) Les hypothèses H.3 et H.4 dépendent du statisticien. Il a entièrement le choix des fonctions ρ et s. Par contre les hypotheses H.1 et H.2 sont plus difficilement controlables. Elles dependent de la "Nature".

2) Exemples de fonctions ρ :

$$x \to \rho(x) \qquad = e^{-\frac{x^2}{2}}$$
$$= \frac{1}{1+x^2}$$

Exemple de fonctions s :

$$x \to (1+x^2)\, \mathbb{1}_{|x|\le 1} + 2x\, \mathbb{1}_{|x|>1}$$

3) Le choix de s(0) supérieur à un est simplement dû à des raisons techniques.

4. Quelques résultats probabilistes.

4.1. Pour tout θ dans Θ, nous notons par A_θ la matrice $d \times d$

$$A_\theta = \begin{pmatrix} \theta_1 \theta_2 & & \dots \theta_d \\ 1 \;\; 0 & & 0 \\ 0 \;\; 1 \;\; 0 & & 0 \\ 0 \;\dots & 1 & 0 \end{pmatrix}$$

et e le vecteur de $\mathbb{R}^d$

$$e = {}^t(1,0,\dots,0)$$

D'après H.2 le rayon spectral de A_θ est inférieur ou égal à $\tau' < \tau$ pour tout $\theta \in \Theta$. Pour toute norme compatible avec la structure d'algèbre des matrices nous avons pour chaque $\theta \in \Theta$

$$\|A_\theta^n\| \le C_\theta\, \tau^n$$

Comme cela a été montré [5] nous pouvons trouver une constante C ne dépendant que de l'espace paramétrique Θ telle que pour tout θ dans Θ et tout entier n positif

(4.1) $\quad \|A_\theta^n\| \le C\, \tau^n$

C peut être choisi supérieur ou égal à $\mathbb{1}$.

Les relations de recurrence (1.1) s'ecrivent vectoriellement :

$Y_{n+1} = A_\theta\, Y_n + \varepsilon_n\, e$

$Y_0 = y_0$

La chaîne vectorielle $(Y_n, n \in \mathbb{N})$ est une chaîne de Markov, récurrente au sens de Harris et admet une loi limite stationnaire ν_θ. Cette loi ν_θ est aussi celle de la variable aléatoire

$$\tilde{Y}(\theta) = \varepsilon_1 e + \varepsilon_2 A_\theta e + \ldots + \varepsilon_j A_\theta^{j-1} e + \ldots$$

Des hypothèses H. 1 faites sur la densité de g on déduit que le support S de ν_θ contient un ouvert de $\mathbb{R}^d$. ν_θ est équivalente à la mesure de Lebesgue sur S.

Nous noterons

$$(4.2) \qquad \begin{cases} Y_n = Y_n(\theta,y) = A_\theta^n y + \varepsilon_n e + \varepsilon_{n-1} A_\theta e + \ldots + \varepsilon_1 A_\theta^{n-1} e \\ \tilde{Y}_n(\theta,y) = A_\theta^n y + \varepsilon_1 e + \varepsilon_2 A_\theta e + \ldots + \varepsilon_n A_\theta^{n-1} e \\ \tilde{Y}(\theta) = \varepsilon_1 \varepsilon + \varepsilon_2 A_\theta e + \ldots + \varepsilon_n A^{n-1} e + \ldots \\ Z = C\left(|\varepsilon_1| + \tau |\varepsilon_2| + \ldots + \tau^n |\varepsilon_{n+1}| + \ldots \right) \end{cases}$$

4.2. Proposition.

Si ε_1 admet un moment absolu d'ordre δ avec $\delta > 1$, $E(|\varepsilon_1|^\delta) = \mu(\delta)$, alors

$$E(Z^\delta) \leq \frac{C^\delta}{(1-\tau)^{\delta-1}} \mu(\delta)$$

$$E_{\theta,y}^{1/\delta}(\|Y_n\|^\delta) \leq (C \tau^n \|y\|) + \frac{C}{(1-\tau)^{(\delta-1)/\delta}} \times (\mu(\delta))^{1/\delta}$$

Démonstration. Les résultats proviennent de ce que Y_n et $\tilde{Y}_n$ ont même loi, et de l'application d'inégalités de convexité.

On utilise aussi l'inégalité

$$\tilde{Y}_n(\theta,y) \leq \|A_\theta^n y\| + Z$$

5. Démonstration des théorèmes 2.1 et 2.2.

1. Demonstration du theoreme 2.1.

Nous notons pour tout $n \in \mathbb{N}$, $t \in \mathbb{R}^d$

$$\overset{*}{Q}_n(.,t) = \overset{*}{Q}_n(t) = \sum_{j=0}^{n} \rho\left(\frac{X_{j+1} - <\theta + t, Y_j>}{s(\|Y_j\|)}\right)$$

$\hat{\theta}_n$ est une solution de :

$$\overset{*}{Q}_n(\hat{\theta}_n - \theta) = \sup\{Q_n(t) \,;\, t \in T\}$$

Pour chaque $t \in T$ $\left(\frac{1}{n}\overset{*}{Q}_n(t),\ n \in \mathbb{N}\right)$ est une suite de variables aléatoires réelles bornées.

Le théorème ergodique nous permet d'affirmer que pour chaque t dans T et θ dans Θ

$$\frac{1}{n}\overset{*}{Q}_n(t) \to \overset{*}{Q}_\theta(t) \qquad \mathbb{P}^{\mathbb{N}}_{\theta,y_\theta} \text{ presque sûrement}$$

où

$$\overset{*}{Q}_\theta(t) = \int_{\mathbb{R}^d}\int_{\mathbb{R}} \rho\left(\frac{\varepsilon - <t,y>}{s(\|y\|)}\right) g(\varepsilon)\, d\varepsilon\, \nu_\theta(dy)$$

Etant donne un ensemble T* dénombrable partout dense de T il existe donc un ensemble mesurable N_θ de Ω, $\mathbb{P}$-nul tel que pour tout $\omega \notin N_\theta$

$$\frac{1}{n}\overset{*}{Q}_n(\omega,t) \to \overset{*}{Q}_\theta(t)$$

des hypothèses H.3.2, 4.1,4.2 on déduit que pout tout ω dans Ω, la fonction de t $\overset{*}{Q}_n(\omega,t)$ admet un gradient borné sur T. T étant compact $\overset{*}{Q}_n(\omega,.)$ converge uniformément vers $\overset{*}{Q}_\theta(.)$.

D'après le lemme 2 ci-dessous la fonction $\overset{*}{Q}_\theta(.)$ atteint son maximum en $t = 0$.

La convergence uniforme de $\frac{1}{n}\overset{*}{Q}_n(\omega,.)$ tend vers $\overset{*}{Q}_\theta(.)$ et la continuité de $\overset{*}{Q}_n(\omega,.)$ impliquent que pour $\omega \notin N_\theta$, $\hat{\theta}_n(\omega)$ converge vers θ quand n tend vers l'infini.

2. Lemme.

Avec les hypotheses du theoreme 2.1, pour chaque θ dans Θ, la fonction de t, $Q_\theta(t)$ definie ci-dessus atteint son maximum en $t = 0$.

Démonstration.

g et ρ sont deux fonctions positives paires, unimodales, et intégrables pour la mesure de Lebesgue. Par application du lemme de T.W. Anderson [4] p. 155 la fonction de t

$$r_\theta(t,x) = \int_{\mathbb{R}} \rho\left(\frac{\varepsilon - \langle t,x\rangle}{s(\|x\|)}\right) g(\varepsilon)\, d\varepsilon$$

atteint son maximum en $\langle t,x\rangle = 0$, ν_θ étant équivalente à la mesure de Lebesgue sur S le maximum de $Q^*(t)$ est atteint en $t = 0$.

3. Démonstration du théorème 2.2.

Notation. Si R est une fonction numérique "régulière" de la variable vectorielle t de $\mathbb{R}^d$, $t = (t_1,\ldots,t_d)$. Nous identifions R avec la fonction notée encore R des variables reelles $t_1,\ldots,t_d$

$$R(t) = R(t_1,\ldots,t_d).$$

Nous notons :

$R^i(t)$ la valeur en t de la dérivée partielle de R relativement à la ième variable

$$R^i(t), = \frac{\partial R(t)}{\partial t_i}$$

Demême

$$R^{i,j}(t) = (R^i)^j(t)$$

D R(t) désignera la valeur du gradient de R en t. C'est un vecteur de $\mathbb{R}^d$.

$D^2 R(t)$ désignera la matrice : $(R^{i,j}(t))_{i,j=1,d}$.

Démonstration. Donnons une idée de la démonstration qui est standard. Posons $Q_n(\theta') = Q^*_n(\theta'-\theta)$. Pour chaque ω, en $\hat\theta_n(\omega)$ nous avons :

$$D\, Q_n(\hat\theta_n(\omega)) = 0.$$

Faisons un développement de Taylor de chaque composante du gradient au voisinage de θ.

$$0 = Q^i_n(\hat\theta_n) = Q^i_n(\theta) + \sum_{j=1}^{d} Q^{ij}_n(\theta)\,(\hat\theta_{n,j}-\theta_j) + \eta_{i,n}(\theta,\hat\theta_n)$$

avec

$$\eta_{i,n}(\theta,\hat\theta_n) = \sum_{j=1}^{d} (Q^{ij}_n(\theta^{*i}) - Q^{ij}_n(\theta))\,(\hat\theta_{n,j}-\theta_j)$$

où

θ^{*i} est tel que $\|\theta^{*i}-\theta\| < \|\hat{\theta}_n-\theta\|$

$\hat{\theta}_{n,j}$ est la $j^{ème}$ composante de $\hat{\theta}_n$.

Toutes ces relations s'écrivent matriciellement

$$(1) \qquad \frac{1}{\sqrt{n}} D\, Q_n(\theta) - \frac{1}{n} D^2\, Q_n(\theta)\, \sqrt{n}\,(\hat{\theta}_n-\theta) + \frac{1}{\sqrt{n}}\, \eta_n(\theta,\hat{\theta}_n) = 0$$

où

$$\eta_n(\theta,\hat{\theta}_n) = {}^t(\eta_{1,n}(\theta,\hat{\theta}_n),\dots,\eta_{d,n}(\theta,\hat{\theta}_n))$$

La consistance de $\hat{\theta}_n$ et les hypothèses H.3.3 et H.4.1,4.2 impliquent que la convergence en probabilité de $\eta_n(\theta,\hat{\theta}_n)$.

Le théorème ergodique implique la convergence presque sûre de $\frac{1}{n} D^2\, Q_n(\theta)$ vers $M(\theta)$.

$M(\theta)$ est inversible car le support S de ν_θ contient un ouvert de $\mathbb{R}^d$. ν_θ est équivalente à la mesure de Lebesgue sur S, l'intégrale de ψ' est non nulle, négative, majorée par $\tilde{\psi}'$

Enfin par [3]

$$\mathcal{L}_\theta\left(\frac{1}{\sqrt{n}} D\, Q_n(\theta)\right) \to N(0,\Lambda(\theta))$$

où $\Lambda(\theta)$ est défini comme dans le théorème.

D'où la conclusion.

6. Opérateurs associés à la chaîne.

6.1. Soit $\mathcal{C}_n$ la tribu engendrée par $\varepsilon_1,\dots,\varepsilon_n$.

Etant donné une fonction f définie sur $\mathbb{R}^d$ à valeur dans $\mathbb{R}^{d'}$ nous considérons pour chaque θ (quand elle existe) l'espérance conditionnelle $E_\theta(f(Y_n)/\mathcal{C}_n)$. La chaîne est markovienne, un representant de cette espérance conditionnelle nous est donne par la fonction de $\mathbb{R}^d$ dans $\mathbb{R}^{d'}$

$$(6.1) \qquad y \to U_\theta f(y) = \int_{\mathbb{R}} f(\varepsilon e + A_\theta y)\, g(\varepsilon)\, d\varepsilon$$

Nous allons étudier les propriétés de l'opérateur U_θ opérant sur certains espaces fonctionnels convenablement choisis. Pour alléger les notations dans ce paragraphe nous supprimons l'indice θ. ν désignera la loi stationnaire associée à la chaîne.

6.2 Rappelons que $E(|\varepsilon_1|^\delta) = \mu(\delta) = \mu_\delta < \infty$

$$\|A^n\| \leq C\, r^n$$

Pour chaque réel β, $0 < \beta \leq 1$, soit B_β l'espace des fonctions de $\mathbb{R}^d$ dans $\mathbb{R}^{d'}$ vérifiant

$$M_\beta(f) = \sup_{x \in \mathbb{R}^d} \frac{\|f(x)\|}{1+\|x\|^\beta} < \infty$$

$$\ell_\beta(f) = \sup_{x,y \in \mathbb{R}^d} \frac{\|f(x)-f(y)\|}{\|x-y\|^\beta} < \infty$$

où $\|.\|$ désigne indifféremment tant qu'il n'y a pas confusion la norme dans $\mathbb{R}^d$ ou $\mathbb{R}^{d'}$. On pose

$$|f|_\beta = M_\beta(f) + \ell_\beta(f)$$

Alors B_β est un espace de Banach et on a

$$\int_{\mathbb{R}^d} |f(y)|\ \nu(dy) < \infty$$

6.3. Proposition

L'opérateur U défini en (6.1) de B_β dans lui-même admet la décomposition:

$$U = \pi + Q$$

où π est le projecteur de B_β sur l'espace des constantes

$$f \to \pi f = \int f\, d\nu$$

Q est un opérateur de B_β dans lui même tel que

1) $\pi Q = Q\pi = 0$

2) pour tout entier n positif

$$\|Q^n\|_\beta \leq C_\beta\, r^{\beta n}$$

Pour tout $\beta < 1$

$$C_\beta = C^\beta \, [\max (1, \frac{\mu(\delta)^{\beta/\delta}}{1-\tau^\beta}) + 1].$$

Démonstration.

Notons par $\mathscr{C}_\beta$ le sous espace vectoriel des fonctions f de B_β centrées pour la loi stationnaire ν de la chaine.

La stationnarité de ν implique

$$U(\mathscr{C}_\beta) \subset \mathscr{C}_\beta$$

Notons par V la restriction de U à $\mathscr{C}_\beta$. Soit f une fonction de $\mathscr{C}_\beta$ évaluons $\|V^n f\|_\beta$. Nous remarquons que

$$V^n f(x) = E(f(\tilde{Y}_n(x)) - f(\tilde{Y}))$$

car $E(f(\tilde{Y})) = \int f \, d\nu = 0$

f appartenant à B_β nous avons :

$$|f(\tilde{Y}_n(x)) - f(\tilde{Y})| \le \ell_\beta(f) \, \|A^n x + A^n e\varepsilon_{n+1} + A^{n+1} e\varepsilon_{n+2} + \dots\|^\beta$$

$$|f(\tilde{Y}_n(x)) - f(\tilde{Y})| \le \ell_\beta(f) (\|A^n\|^\beta \|x\|^\beta + \sum_{p=n}^{\infty} \|A^p\|^\beta |\varepsilon_{p+1}|^\beta)$$

Tenant compte de la majoration de la norme de A^p nous obtenons :

$$\|V^n f(x)\| \le \ell_\beta(f) \; C^\beta \, \tau^{\beta n} (\|x\|^\beta + \frac{1}{1-\tau^\beta} \, (\mu(\delta))^{\beta/\delta})$$

D'autre part, un calcul aisé nous permet d'obtenir pour tout $\beta > 0$

$$\|V^n f(x) - V^n f(y)\| \le \ell_\beta(f) \; C^\beta \, \tau^{\beta n} \|x-y\|^\beta$$

Sommant M_β et ℓ_β nous obtenons

$$\|V^n f\|_\beta \le C_\beta \, \tau^{\beta n} \, \ell_\beta(f)$$

Une fonction quelconque f de B_β se décompose en : $\pi(f) + (f - \pi(f))$

$f - \pi(f)$ appartient à $\mathscr{C}_\beta$

Alors U admet la décomposition

$$U = \pi + Q$$

où $Q = V \circ (I - \pi)$.

6.4. Soit φ une fonction bornée lipschitzienne défnie sur $\mathbb{R}^d$ a valeurs dans $\mathbb{R}^{d'}$ nous notons par

$$\|\varphi\|_\infty = \sup_x \frac{\|\varphi(x)\|}{\|x\|}$$

$$\ell(\varphi) = \sup_{x \neq y} \frac{\|\varphi(x)-\varphi(y)\|}{\|x-y\|}$$

Nous posons $|\varphi|_1 = \|\varphi\|_\infty + \ell(\varphi)$.

Le lemme suivant nous permet d'évaluer la norme de φ dans B_β.

Lemme.

Soit β un nombre réel compris entre 0 et 1. $(0 < \beta \leq 1)$

Soit φ une fonction bornée lipschitzienne.

Alors avec les notations ci-dessus **6.2.**

1) $M_\beta(\varphi) \leq \|\varphi\|_\infty$

2) $\ell_\beta(\varphi) \leq 2^{1-\beta} \|\varphi\|_\infty^{1-\beta} \ell^\beta(\varphi) \leq (1-\beta)2\|\varphi\|_\infty + \beta\, \ell(\varphi)$

3) $\|\varphi\|_\beta \leq 3|\varphi|_1$.

Démonstration.

3) est manifestement une conséquence de 1) et 2).

1) est évident

2) $\ell(\varphi) = 0$ implique $\ell_\beta(\varphi) = 0$

si $\ell(\varphi) > 0$

* Soit $\|x-y\| \leq 2\|\varphi\|_\infty \ell^{-1}(\varphi)$.

Nous avons alors

$$\frac{\|\varphi(x)-\varphi(y)\|}{\|x-y\|^\beta} \leq \ell(\varphi)\|x-y\|^{1-\beta} \leq 2^{1-\beta} \|\varphi\|_\infty^{1-\beta} \ell^\beta(\varphi)$$

* Soit $\|x-y\| \geq 2\|\varphi\|_\infty \ell^{-1}(\varphi)$

$$\frac{\|\varphi(x)-\varphi(y)\|}{\|x-y\|^\beta} \leq 2\|\varphi\|_\infty (2\|\varphi\|_\infty \ell^{-1}(\varphi))^{-\beta} \leq 2^{1-\beta} \|\varphi\|_\infty^{1-\beta} \ell^\beta(\varphi)$$

d'ou la premiere inegalite.

La deuxième inégalité est due à la concavité de la fonction logarithme.

6.5. Nous nous intéressons maintenant à l'évaluation d'expressions du type

$$J_{n,k}(t) = E(\langle t, \sum_{i=1}^{n} F(\varepsilon_i, Y_{i-1}) \rangle^k)$$

où k est un entier positif, t un vecteur de $\mathbb{R}^r$, F une fonction de $\mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^r$ telles que :

1) F est borné, $\sup_{(\varepsilon,y)} |F(\varepsilon,y)| = \|F\|_\infty$

2) Pour tout y et z de $\mathbb{R}^d$, $\|F(\varepsilon,y) - F(\varepsilon,z)\| \leq (D_0|\varepsilon| + D_1)\|y-z\|$

3) F est centré pour la loi $g(\varepsilon)\, d\varepsilon\, \nu(dy)$.

Posons $\overline{F}(y)$ la fonction de y définie par :

$$\overline{F}(y) = \int F(\varepsilon,y)\, g(\varepsilon)\, d\varepsilon$$

Alors : $\|\overline{F}\|_\infty \leq \|F\|_\infty$

$\ell(\overline{F}) \leq (D_0\, \mu^{1/\delta}(\delta) + D_1)$

Nous avons une majoration de $|\overline{F}|_1$

(1) $\quad |\overline{F}|_1 \leq \|F\|_\infty + (D_0\, \mu_\delta^{1/\delta} + D_1)$

6.6. Proposition.

Supposons qu'il existe $\delta > 1$ tel que $E|\varepsilon_1|^\delta$ soit fini, soit F soumis aux trois conditions ci-dessus et la condition H.2. réalisée.

1- Alors pour chaque entier n positif et chaque k tel que $\beta = \frac{\delta}{k} < 1$ nous avons

$$J_{nk}(t) \leq 4^{k-1} (2^{k-1} C_\beta^k (1-\tau^\beta)^{-k} [(2+2^{k-1} C_k(n-1)^{k/2})(1+C^\delta \mu(\delta)(1-\tau)^{-\delta})$$
$$+ C^\delta \|y_0\|^\delta (1-\tau)^{1-\delta} (1+\tau^{n-1}+2^{k-1} C_k(n-1)^{[k/2]-1} \sum_{i=1}^{n-1} \tau^i)] \times \|t\|^k \times |3\overline{F}|_1^k$$
$$+ 2^k C_k n^{k/2} . \|t\|^k . \|F\|_\infty^k)$$

où $|\overline{F}|_1$ est donné en (6.5(1)). C_β en proposition 6.3., $C_k = (18 \frac{k^{3/2}}{(k-1)^{1/2}})^k$.

2- Nous avons les majorations suivantes

$$(1-\tau^{\beta})^{-k} \leq (k/\delta)^{k} (1-\tau)^{-k}$$

$$C_\beta \leq 2C^{\delta/k} \qquad \text{pour } (\mu(\delta))^{1/k} < 1-\tau^{\delta/k}$$

$$C_\beta \leq C^{\delta/k} \left(1 + \frac{\mu(\delta)^{1/k}}{1-\tau^{\delta/k}}\right) \leq C^{\delta/k} \left(\frac{k}{\delta} \frac{(\mu(\delta))^{1/k}}{1-\tau} + 1\right) \qquad \text{sinon.}$$

Démonstration.

Notons par $\mathcal{C}_{i-1}$ la tribu engendrée par $\varepsilon_1 \ldots \varepsilon_{i-1}$, et posons :

$$\overline{F}(Y_{i-1}) = E(F(\varepsilon_i, Y_{i-1}) / \mathcal{C}_{i-1})$$

$$A = \sum_{i=1}^{n} F(\varepsilon_i, Y_{i-1}) - \overline{F}(Y_{i-1})$$

$$B = \sum_{i=1}^{n} \overline{F}(Y_{i-1})$$

En vertu du lemme 6.7 nous pouvons écrire :

$$\begin{aligned} B &= \sum_{i=1}^{n} (I-U)\, h(Y_{i-1}) \\ &= \sum_{i=2}^{n} (h(Y_{i-1}) - Uh(Y_{i-2})) - Uh(Y_{n-1}) + h(Y_0) \\ &= \qquad B' \qquad - Uh(Y_{n-1}) + h(Y_0) \end{aligned}$$

Nous obtenons alors

$$(1) \qquad \begin{aligned} J_{n,k}(t) &= E(|\langle t,A\rangle + \langle t,B\rangle|^k) \\ &\leq 4^{k-1}(E(|\langle t,A\rangle|^k) + E(|\langle t,B'\rangle|^k) + E(|\langle t,Uh(Y_{n-1})\rangle|^k) + E(|\langle t,h(Y_0)\rangle|^k)) \end{aligned}$$

$\langle t,A\rangle$ est une somme d'accroissement de martingales.

Par l'inégalité de B.D.G [2,p.22]

$$\begin{aligned} E(|\langle t,A\rangle|^k) &\leq C_k E\left\{ \sum_{i=1}^{n} |\langle t, F(\varepsilon_i, Y_{i-1}) - \overline{F}(Y_{i-1})\rangle|^2 \right\}^{k/2} \\ &\leq C_k\, n^{(k-2)/2} \sum_{i=1}^{n} E\{ |\langle t, F(\varepsilon_i, Y_{i-1}) - \overline{F}(Y_{i-1})\rangle|^k \} \end{aligned}$$

$$(2) \qquad E(|\langle t,A\rangle|^k) \leq C_k\, n^{k/2} \|t\|^k (2\|F\|_\infty)^k .$$

D'après le lemme 6.7, nous avons :

(3) $E(|<t,h(Y_{n-1})>|^k) \le$

$\le 2^{k-1} C_\beta^k (1-\tau^\beta)^{-k} (1+C^\delta \mu(\delta)(1-\tau)^{-\delta} + C^\delta \tau^{n-1} \|y_0\|^\delta (1-\tau)^{1-\delta}) \|t\|^k |3\overline{F}|_1^k$

(4) $E(|<t,h(Y_0)>|^k) \le$

$\le 2^{k-1} C_\beta^k (1-\tau^\beta)^{-k} (1+C^\delta \mu(\delta)(1-\tau)^{-\delta} + C^\delta \|y_0\|^\beta (1-\tau)^{1-\delta}) \|t\|^k |3\overline{F}|_1^k$

D'autre part $<t,B'>$ est une somme d'accroissements de martingales, nous avons donc la majorationvial'inégalitéde[2,p.22].

$$E(|<t,B'>|^k) \le C_k (n-1)^{(k-2)/2} \sum_{i=2}^{n} E(|<t,h(Y_{i-1})-Uh(Y_{i-2})>|^k) \le$$

$$\le 2^{k-1} C_k (n-1)^{(k-2)/2} \sum_{i=2}^{n} E(|<t,h(Y_{i-1})>|^k)$$

Maintenant grâce au lemme 6.7 ci-dessous nous obtenons :

(5) $E(|<t,B'>|^k) \le 2^{k-1} C_\beta^k (1-\tau^\beta)^{-k} 2^{k-1} C_k (n-1)^{(k-2)/2} \times$

$\times \quad ((n-1)(1+C^\delta \mu(\delta)(1-\tau)^{-\delta} + C^\delta \|y_0\|^\beta (1-\tau)^{1-\delta} \sum_{i=2}^{n} \tau^{i-1}) \|t\|^k |3\overline{F}|_1^k$

(1) et les inégalités (2),(3),(4),(5) donnent le résultat.

La majoration de C_k s'obtient à partir de [2, p. 22], celle de $(1-\tau^\beta)^{-k}$ de l'inégalité

$1-\tau^\alpha \ge \alpha(1-\tau)$ pour $0 < \tau < 1$ et $0 < \alpha < 1$.

6.7. Lemme.

Soit $\overline{F}$ une fonction de $\mathbb{R}^d$ dans $\mathbb{R}^{d'}$ borné lipschitzienne centrée pour la loi stationnaire v.

$$|\overline{F}|_1 = \|\overline{F}\|_\infty + \ell(\overline{F})$$

Notons par h la fonction définie formellement par la série

$$h(y) = \overline{F}(y) + \dots + U^m \overline{F}(y) + \dots$$

Alors :

1) Pour $\beta \in (0,1)$ la série définissant h est convergente dans B_β et nous avons

$$\|h\|_\beta \le C_\beta (1-\tau^\beta)^{-1} 3|\overline{F}|_1$$

2) Si $\delta > 1$, $\beta k = \delta$, $\beta < 1$, pour tout entier positif i nous avons

$$E(\|h(Y_i)\|^k) \leq 2^{k-1} C_\beta^k (1-\tau^\beta)^{-k} (1+C^\delta \mu(\delta)(1-\tau)^{-\delta} + C^\delta \tau^i \|y_0\|^\delta (1-\tau)^{1-\delta}) |3\overline{F}|_1^k$$

Démonstration.

1) $\overline{F}$ est un élément de B_β centré pour ν c'est donc un élément de $\mathscr{C}_\beta$. D'après la proposition 6.3 et le lemme 6.4

$$(1) \qquad \|U^m \overline{F}\|_\beta \leq \|Q^m\|_\beta \|\overline{F}\|_\beta \leq C_\beta \tau^{\beta m} 3|\overline{F}|_1$$

la série définissant h converge donc dans B_β et l'évaluation de $\|h\|_\beta$ se déduit aisément de la majoration précédente (1).

2) Si $\beta k = \delta$ avec $\delta > 1$.

Pour tout entier positif i nous avons :

$$(2) \qquad \|h(Y_i(y))\|^k \leq M_\beta^k(h)(1+\|Y_i(y)\|^\beta)^k$$

De la définition de h et de (1) et du lemme 6.4 3)

$$(3) \qquad M_\beta(h) \leq C_\beta(1-\tau^\beta)^{-1} . |3\overline{F}|_1$$

D'autre part :

$$(4) \qquad (1+\|Y_i(y_0)\|^\beta)^k \leq 2^{k-1}[1+(\|A^i y_0\|+\sum_{j=1}^{i} |\varepsilon_j| \|A^{i-j} e\|)^\delta]$$

$$E[(1+\|Y_i(y_0)\|^\beta)^k] \leq 2^{k-1}[1+(C\tau^i\|y_0\| + \sum_{j=1}^{i} C\tau^{i-j} |\varepsilon_j|)^\delta]$$

Grâce à l'inégalité de convexité

$$E[(1+\|Y_i(y_0)\|^\beta)^k] \leq 2^{k-1}[1+C^\delta (\frac{1-\tau^{\delta-1}}{1-\tau})(\tau^i \|y_0\|^\delta + \sum_{j=1}^{i} \tau^{i-j} \mu(\delta)]$$

$$(5) \qquad \leq 2^{k-1}[1+C^\delta((1-\tau)^{1-\delta} \tau^i \|y_0\|^\delta + (1-\tau)^{-\delta} \mu(\delta))]$$

(3) et (5) nous permettent d'obtenir l'inégalité.

7. Vitesse de convergence (Démonstration du théorème 2.3)

7.1. Lemme.

Posons

$$r(\theta,\theta')=\int_{\mathbb{R}^d}\int_{\mathbb{R}}\rho\left(\frac{\varepsilon-<\theta'-\theta,y>}{s(\|y\|)}\right)g(\varepsilon)\,d\varepsilon\,v_\theta(dy)$$

Sous les conditions H

$$\Delta(\|\theta'-\theta\|)=r(\theta,\theta)-r(\theta,\theta')\geq\frac{1}{2}\rho^*\left(\left(\|\theta'-\theta\|^2 . s_2^2\right)\wedge v_0^2\right)(1-\alpha)^2$$

ou ρ^*, v_0, et s_2 sont définis en H.3.4 et H.4.3.

Démonstration.

Nous noterons ici $\kappa(d,\tau)$ par κ, g étant une densité, paire, unimodale, il en est de même de la densité de la loi stationnaire v_θ relativement à la mesure de Lebesgue sur $\mathbb{R}^d$. ρ étant aussi unimodale, en posant $t=\theta'-\theta$, et utilisant l'hypothèse H.3.4. nous avons :

$$r(\theta,\theta')=r(\theta,\theta+t)\leq A+B$$

avec

$$A=\int_{\mathbb{R}^d}\int_{\mathbb{R}}\rho\left(\frac{\varepsilon}{s(\|y\|)}\right)\left(1_{\frac{|\varepsilon|}{s(\|y\|)}>\beta_2}+1_{\left[\varepsilon\,.\,<t,y>\,\leq 0\right]}+1_{\left[|<\frac{t}{\|t\|},y>|\leq\beta_1\kappa\right]}\right)g(\varepsilon)\,d\varepsilon\,v_\theta(dy)$$

$$B=\int_{\mathbb{R}^d}\int_{\mathbb{R}}\left[\rho\left(\frac{\varepsilon}{s(\|y\|)}\right)-\rho^*\left(<t,\frac{y}{s(\|y\|)}>^2\wedge v_0^2\right)\right]\times$$

$$\times\left[1_{\left[0\leq\frac{\varepsilon}{s(\|y\|)}<\beta_2\right]}.1_{\left[<\frac{t}{\|t\|},y>>\beta_1\kappa\right]}+1_{\left[-\beta_2\leq\frac{\varepsilon}{s(\|y\|)}<0\right]}1_{\left[<\frac{t}{\|t\|},y><-\beta_1\kappa\right]}\right]g(\varepsilon)\,d\varepsilon\,v_\theta(dy)$$

D'après H.4.1 $s(\|y\|)\geq s_0>1$, d'autre part g et $\frac{dv_\theta}{dy}$ sont des densités symetriques, nous avons alors

$$r(\theta,\theta)-r(\theta,\theta+t)\geq \rho^{*}\left[\left(\|t\|^{2}<\frac{t}{\|t\|},\frac{y}{s\|y\|}>^{2}\right)\wedge v_{0}^{2}\right]\times P$$

avec

$$P=2\int_{\mathbb{R}^{d}}\int_{\mathbb{R}} 1_{0<\varepsilon<\beta_2} 1_{\left[<\frac{t}{\|t\|},y>\,>\beta_1\kappa\right]}$$

La parité de g et la symétrie de v_θ à l'origine et l'hypothese H.1.3 impliquent que

$$P\geq\frac{1}{2}(1-\alpha)\int_{\mathbb{R}^{d}} 1_{\left\{\left|<\frac{t}{\|t\|},y>\right|>\beta_1\kappa\right\}}v_\theta(dy)$$

$$P\geq\frac{1}{2}(1-\alpha)\qquad v_\theta\left(\left|<\frac{t}{\|t\|},\tilde{Y}_\theta>\right|>\beta_1\kappa\right)$$

De la définition de $\tilde{Y}(\theta)$ dont la loi est v_θ nous déduisons :

$$<\frac{t}{\|t\|},\tilde{Y}(\theta)>=\sum_{i=1}^{\infty}\varepsilon_i<\frac{t}{\|t\|},A_\theta^{i-1}e>$$

Du lemme 7.2. ci-dessous nous avons que

$$\max_{i=1,d}\left|<\frac{t}{\|t\|},A^{i-1}e>\right|\geq\kappa$$

Notant i_0 l'un des indices i tel que

$$\left|<\frac{t}{\|t\|},A^{i_0-1}e>\right|=\max_{i=1,d}<\frac{t}{\|t\|},A^{i-1}e_i>\ \geq\kappa$$

les variables aléatoires étant indépendantes de loi unimodale symétrique on a :

$$v_\theta\left(\left|<\frac{t}{\|t\|},\tilde{Y}>\right|\geq\beta_1\kappa\right)\geq P\left(\left|<\frac{t}{\|t\|},A^{i_0-1}e>\varepsilon_{i_0}\right|>\beta_1\kappa\right)\geq P\left(\left|\varepsilon_{i_0}\right|>\beta_1\right)\geq 1-\alpha$$

D'ou

$$P\geq\frac{1}{2}(1-\alpha)^2.$$

7.2. Lemme.

Etant donné le système de vecteurs de $\mathbb{R}^d$

$v_1=e={}^t(1,0\ldots0)$

$v_i={}^t(\alpha_{i-1},\alpha_{i-2},\ldots\alpha_1,1,0\ldots0)$

$v_d={}^t(\alpha_{d-1},\ldots\alpha_1,1)$

Désignant par $<\,,\,>$ et $\|.\|$ le produit scalaire et la norme euclidienne de $\mathbb{R}^d$ nous avons

1) $$\min_{\|t\|=1}\max_{i=1,d}\left|<t,v_i>\right|\geq\frac{1}{\sqrt{d}}\;\frac{(d-1)^{\frac{d-1}{2}}}{\left(\sum_{i=1}^{d}\|v_i\|^2\right)^{\frac{d-1}{2}}}$$

Pour $v_i = A_\theta^{i-1} e$, nous avons

$$2)\quad \min_{\|t\|=1} \max_{i=1,d} |\langle t, A^{i-1} e\rangle| \geq \frac{1}{\sqrt{d}} \frac{\left(\left(1-\tau^2\right)(d-1)\right)^{\frac{d-1}{2}}}{C^{d-1}} = \kappa(d,\tau)$$

Démonstration.

Pour tout vecteur t de $\mathbb{R}^d$ nous avons

$$\max_{i=1,d} |\langle t, v_i\rangle|^2 \geq \frac{1}{d} \sum_{i=1}^{d} \langle t, v_i\rangle^2 = \frac{1}{d} Q(t)$$

Notons par $e_1 = e, e_2, \ldots e_d$ la base canonique de $\mathbb{R}^d$, $\overline{v}_i$ (i=1,d) la matrice colonne des composantes de v_i dans cette base. Alors, toujours dans cette base, la forme quadratique Q(t) admet pour matrice

$$B = \overline{v}_1 {}^t\overline{v}_1 + \ldots + \overline{v}_d {}^t\overline{v}_d .$$

B s'écrit aussi :

$$B = \left(\sum_{j=1}^{d} \langle e_i, v_j\rangle \overline{v}_j\right)_{i=1,d}$$

min { Q(t)| ||t||=1 } est aussi la plus petite valeur propre de B.

Or la somme des valeurs propres est : s = trace(B).

Nous avons :

$$s = \mathrm{trace}(B) = \sum_{i=1}^{d}\sum_{j=1}^{d} \langle e_i, v_j\rangle^2 = \sum_{j=1}^{d}\sum_{i=1}^{d} \langle e_i, v_j\rangle^2$$

$$s = \mathrm{trace}(B) = \sum_{j=1}^{d} \|v_j\|^2 .$$

Le produit des valeurs propres est : det B.

Or pour tout vecteur v_j = (j = i,d)

$$\langle e_i, v_j\rangle \quad \begin{aligned} &= 1 \quad \text{si } i = j \\ &= 0 \quad \text{si } i > j \end{aligned}$$

La matrice B s'écrit donc

$$B = \left(\overline{v}_1 - \sum_{j=2}^{d} \langle e_1, v_j\rangle \overline{v}_j,\ \overline{v}_2 + \sum_{3}^{d} \langle e_2, v_j\rangle, \ldots, \overline{v}_d\right)$$

On voit alors que : det B = 1, et toutes les valeurs propres sont donc strictement positives.

Soient $\lambda_1,\ldots,\lambda_d$ les valeurs propres de B et λ_d la plus petite d'entre elles

(1) $$\lambda_d = \frac{1}{\prod_{i=1}^{d-1} \lambda_i}$$

Pour λ_d donné $\sum_{i=1}^{d-1} \lambda_i = s - \lambda_d$ le deuxième membre de (1) se minore en prenant pour tout $i = 1 \ldots d-1$

$$\lambda_i = \frac{1}{d-1}(s - \lambda_d)$$

Nous avons donc à partir de (1)

(2) $$\lambda_d \geq \frac{(d-1)^{d-1}}{\left(s-\lambda_d\right)^{d-1}} \geq \frac{(d-1)^{d-1}}{s^{d-1}}$$

nous en déduisons alors la première inégalité du lemme. La deuxième inégalité provient de l'évaluation $\|v_i\| = \|A_\theta^{i-1} e\| \leq C\,\tau^{i-1}$.

7.3. Démonstration du théorème 2.3.

Soit θ la valeur du paramètre, T est l'ensemble défini en 1.2. Etant donné $\gamma > 0$ posons $T^* = T(\theta,\gamma) = \{\theta^* \in T \mid \|\theta^*-\theta\| \geq \gamma\}$
cet ensemble est compact. L'événement $\{\|\hat{\theta}^*_n-\theta\| \geq \gamma\}$ équivaut à $\hat{\theta}_n \in T(\theta,\gamma) = T^*$.

Il résulte du lemme 7.1 que pour tout $\theta' \in T^*$

(1) $$r_{\theta,\theta} - r_{\theta,\theta'} \geq \frac{1}{2} \rho^*((\gamma \,.\, s_2)^2 \wedge v_0^2)\,(1-\alpha)^2 = \Delta(\gamma) = \Delta$$

Etant donné θ', θ^* quelconque dans T^*, il résulte des hypothèses H.3.2 et H.4.1 que pour (x,y) dans $R \times R^d$ on a :

$$\rho\left(\frac{x - <\theta^*, y>}{s(\|y\|)}\right) \leq \rho\left(\frac{x - <\theta', y>}{s(\|y\|)}\right) + s_1\,\psi_1\,\|\theta^*-\theta\|$$

avec $\psi_1 = \sup_{u\in R} |\psi(u)|$ et $s_1 = \sup_{x\in R} \frac{|x|}{s(|x|)}$.

Notons $D(\theta')$ l'ensemble

(2) $$D\left(\theta'\right) = \left\{ \theta^* \in T \mid \|\theta^*-\theta'\| \leq \frac{\Delta}{3 s_1 \psi_1} \right\}$$

en posant

(3) $$m\left(x,y,\theta'\right)=\rho\left(\frac{x-<\theta',y>}{s\left(\|y\|\right)}\right)+\frac{\Delta}{3}$$

nous avons pour tout θ^* dans $D(\theta')$ et (x,y) dans $\mathbb{R}\times\mathbb{R}^d$

(3') $$r\left(x,y,\theta^*\right)=\rho\left(\frac{x-<\theta^*,y>}{s\left(\|y\|\right)}\right)\leq m\left(x,y,\theta'\right)$$

T^* étant compact il existe un ensemble fini $\theta_1,\ldots,\theta_p$ d'éléments de T^* tels que les voisinages $D(\theta_1),\ldots,D(\theta_p)$ définis comme en (2) recouvrent T^*. Un calcul élémentaire mène à une majoration de p

(4) $$p\leq\left(\left[\frac{3s_1\psi_1\sqrt{d}\,|T|}{2\Delta(\gamma)}\right]+1\right)^d$$

où [a] est la partie entière du réel [a] et $|T|$ la longueur du côté du cube T.

Pour une observation ω posons :

$x_j=X_j(\omega)\qquad y_j=Y_j(\omega)\qquad j=1,2\ldots$

Soit $\hat{\theta}_n(\omega)$ tel que $\|\hat{\theta}_n(\omega)-\theta\|\geq\gamma$

Nous avons alors :

$$\frac{1}{n}\sum_{i=1}^{n}r\left(x_j,y_{j-1},\hat{\theta}_n(\omega)\right)\geq\frac{1}{n}\sum_{j=1}^{n}r\left(x_j,y_{j-1},\theta\right)$$

$\hat{\theta}_n$ appartient à l'un des $D(\theta_i)$ (i=1,p), et donc pour l'un des i convenablement choisi

$$\frac{1}{n}\sum_{j=1}^{n}m\left(x_j,y_{j-1},\theta_i\right)\geq\frac{1}{n}\sum_{j=1}^{n}r\left(x_j,y_{j-1},\theta\right)$$

Nous en déduisons alors :

$$(5)\qquad P^{\mathbb{N}}_{\theta,y_0}\left(\|\hat{\theta}_n-\theta\| \geq \gamma\right) \leq \sum_{i=1}^{p} P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} m\left(X_j, Y_{j-1}, \theta_i\right) \geq \frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta\right)\right) = \sum_{i=1}^{p} P_i$$

Evaluons chaque terme de la somme du 2e membre de (5). La relation (3) nous permet de remplacer la fonction m par r, introduisant de plus l'espérance sous la loi stationnaire en θ de m et r nous obtenons :

$$P_i = P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta_i\right) + \frac{\Delta}{3} - r\left(\theta,\theta_i\right) \geq \frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta\right) - r\left(\theta,\theta\right) + r\left(\theta,\theta\right) - r\left(\theta,\theta_i\right)\right)$$

Majorons cette probabilité :

$$(7)\qquad \begin{aligned} P_i \leq\ & P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta_i\right) - r\left(\theta,\theta_i\right) \geq \frac{\Delta}{3}\right) \\ & + P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta\right) - r\left(\theta,\theta\right) \leq -\frac{\Delta}{3}\right) \end{aligned}$$

Nous obtenons alors la majoration à partir de (5) et (7)

$$(8)\qquad \begin{aligned} P^{\mathbb{N}}_{\theta,y_0}\left(\|\hat{\theta}_n-\theta\| \geq \gamma\right) \leq\ & \sum_{i=1}^{p} P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta\right) - r\left(\theta,\theta_i\right) \geq \frac{\Delta}{3}\right) \\ & + p\, P^{\mathbb{N}}_{\theta,y_0}\left(\frac{1}{n}\sum_{j=1}^{n} r\left(X_j, Y_{j-1}, \theta\right) - r\left(\theta,\theta\right) \leq -\frac{\Delta}{3}\right) \end{aligned}$$

La conclusion vient en appliquant l'inégalité de Bernstein :

$$P(Z > a) \leq \frac{E|Z|^k}{a^k} \qquad \text{avec } a > 0$$

et rassemblant les résultats antérieurs.

Bibliographie.

[1] H.J. Bierens, 1981
Robust Methods and Asymptotic theory in non linear Econometrics.
Lecture notes in Economics and mathematical Systems, Springer-Verlag, 192.

[2] D.L. Burkholder, 1973
Distribution function inequalities for martingales.
Ann Probability, vol. 1, n°1, pp. 19-42.

[3] A. Dvoretsky, 1970
Central limit Theorems for dependent random variables.
Actes du Congrès International des Mathématiciens, p. 565.

[4] I.A. Ibragimov, R.Z. Has'minsky, 1981
Statistical Estimation Asymptotic Theory.
Applications of Mathematics n° 16, Springer-Verlag.

[5] X. Milhaud, A. Raugi
Etude de l'estimateur du maximum de vraisemblance dans le cas d'un processus autoregressif. Convergence Normalité asymptotique, vitesse de convergence.
Annales de l'Institut Henri Poincaré. A paraître.

[6] A. Mokkadem, 1988
Mixing properties of ARMA Processes.
Stochastic Processes Appl. 29, 1988, pp. 309-315.

[7] M.F. Norman, 1972
Markov Processes and Learning Models.
Mathematics in Sciences and Engeneering, Vol. 84, Academic Press.