Convergence uniforme presque sûre d'une classe d'M-estimateurs par noyau pour des observations dépendantes

Carlos TENREIRO

Departamento de Matemática

Universidade de Coimbra

Apartado 3008

3000 COIMBRA

PORTUGAL

fax: (351) 39.32568

e-mail: tenreiro@mat.uc.pt

décembre 1995

Convergence uniforme presque sûre d'une classe d'M-estimateurs par noyau pour des observations dépendantes

Carlos TENREIRO

Departamento de Matemática, Universidade de Coimbra

Mots clés: M-estimateurs par noyau, stationnarité, α-mélangeance, convergence uniforme presque sûre.

Key words and phrases: Kernel M-estimators, stationarity, α-mixing, strong uniform convergence.

Classification AMS 1991: 62G07, 62G05.

RÉSUMÉ

Nous établissons la convergence uniforme presque sûre sur un compact d'une classe d'M-estimateurs par noyau, lorsque le processus des observations est fortement stationnaire et satisfait certaines conditions de mélange. Ces résultats généralisent ceux de Collomb et Härdle [10] et de Gouriéroux, Monfort et Tenreiro [14].

ABSTRACT

We establish the strong uniform convergence over a compact of a classe of kernel M-estimators, when the processus of observations is strictly stationary and satisfies some mixing conditions. These results improve previous ones of Collomb and Härdle [10] and Gouriéroux, Monfort and Tenreiro [14].

1 Introduction

Dans ce travail nous considérons une classe d'estimateurs fonctionnels obtenus par composition d'une fonction réelle φ définie sur $C \times \Theta$, où C et Θ sont des parties compactes de $\mathbb{R}^d$ et $\mathbb{R}^p$ respectivement, et d'un M-estimateur à noyau. Nous nous intéressons ainsi à l'étude des estimateurs fonctionnels définis, pour $s \in C$, par

$$\widehat{\varphi}_n(s) = \varphi(s; \widehat{\theta}_n(s)),$$

avec

$$\widehat{\theta}_n(s) = arg\ min_{\theta \in \Theta} \sum_{i=1}^{n} K\left(\frac{s - S(Z_i)}{h_n}\right) \psi(Z_i; s; \theta),$$

où $(Z_i, i \in \mathbb{Z})$ est un processus stochastique fortement stationnaire à valeurs dans $\mathbb{R}^q$, S une application connue de $\mathbb{R}^q$ dans $\mathbb{R}^d$, K un noyau sur $\mathbb{R}^d$, c'est-à-dire, une application réelle mesurable sur $\mathbb{R}^d$ telle que $\int_{\mathbb{R}^d} K(u)du = 1$, (h_n) une suite de nombres réels strictement positifs qui converge vers zéro, lorsque n tend vers l'infini, et ψ une application réelle connue que l'on désigne par fonction critère.

La convergence presque sûre, uniformément par rapport à $s \in C$, de la suite d'estimateurs $\widehat{\theta}_n(\cdot)$ vers la solution supposée unique (condition d'identifiabilité asymptotique) du problème d'optimisation

$$min_{\theta \in \Theta} \, E[\psi(Z; s; \theta) \mid S(Z) = s],$$

a été obtenue par Collomb et Härdle [10] dans le cadre des processus ϕ-mélangeants. En employant une technique plus proche de celle utilisée dans le cadre des M-estimateurs paramétriques, Gouriéroux, Monfort et Tenreiro [13] et [14] ont dérivé, dans le cadre des processus géométriquement α-mélangeants, la convergence presque sûre, ponctuelle et uniforme, de l'estimateur $\widehat{\theta}_n(\cdot)$. En particulier, la convergence uniforme a été obtenue en admettant que la fonction critère est indépendante de s.

Sous des conditions analogues à celles utilisées dans ces dernières références, nous nous proposons dans cet article d'établir la convergence uniforme presque sûre de la suite d'estimateurs fonctionnels $\widehat{\varphi}_n(\cdot)$ lorsque le processus des observations satisfait certaines conditions de mélange. Nous généralisons de cette façon, à la classe des estimateurs introduits ci-dessus, les résultats de Collomb et Härdle [10] et Gouriéroux, Monfort et Tenreiro [14]. On le fera en remplaçant l'habituelle condition d'identifiabilité par une condition plus faible, qui n'impose pas la singularité de l'ensemble des solutions du problème d'optimisation précédent, mais qu'assure l'existence d'une application réelle $\varphi_\infty(\cdot)$ définie dans C telle que pour toute la solution $\theta(s), s \in C$ d'un tel problème d'optimisation on ait

$$\varphi(s; \theta(s)) = \varphi_\infty(s), \text{ pour tout } s \in C.$$

On montrera alors que

$$\sup_{s \in C} \mid \widehat{\varphi}_n(s) - \varphi_\infty(s) \mid \xrightarrow[n \to +\infty]{p.s.} 0.$$

La littérature est riche en exemples d'estimateurs qui appartiennent à la classe définie ci-dessus. Robinson [25], Härdle [17] et Boente et Fraiman [2] considèrent des estimateurs robustes de paramètres conditionnels de position en prenant des fonctions critère du type $\psi(Z; s; \theta) = \rho(Y - \theta)$ où ρ est une fonction convexe et $\varphi(s; \theta) = \theta$. La convergence ponctuelle presque sûre de tels estimateurs a été dérivée par Härdle [17] dans un cadre d'échantillonnage et par Boente et Fraiman [2] dans le cas dépendant (en effet ces derniers auteurs ont considéré une classe plus vaste d'estimateurs). Si $\rho(y) = y^2$, $\widehat{\varphi}_n(\cdot)$ est l'estimateur à noyau usuel de la régression $E[Y|S(Z) = \cdot]$ introduit par Nadaraya [24] et Watson [28] et dont la convergence uniforme presque sûre a été établie par Mack et Silverman [23] dans un cadre d'échantillonnage. Si $\rho(y) = |y|$, $\widehat{\varphi}_n(s)$ est l'estimateur de la médianne de la distribution conditionnelle $Y|S(Z) = s$. Un estimateur robuste au sens de Huber [21] d'un tel paramètre fonctionnel peut être obtenu, dans le cas où la densité de probabilité de $Y|S(Z) = s$ est symétrique, en considérant $\rho(y) = y^2/2$ si $|y| \leq c$, et $\rho(y) = c|y| - c^2/2$ si $|y| > c$ (cf. Robinson [25]). Lorsque $\psi(Z; s; \theta) = \big(Y - \theta_0 - \theta_1 X - \ldots - \theta_p X^p\big)^2$ et $\varphi(s; \theta) = \theta_0 + \theta_1 s + \ldots + \theta_p s^p$, $\widehat{\varphi}_n(\cdot)$ est un M-estimateur par noyau de la régression $E[Y|X = s]$ correspondant à l'estimateur pondéré des k_n-points les plus proches considéré par Cleveland [7]. Plus généralement si $\psi(Z; s; \theta) = \big(Y - m(X; \theta)\big)^2$ et $\varphi(s; \theta) = m(s; \theta)$, $\widehat{\varphi}_n(\cdot)$ a été introduit par Gozalo et Linton [15] qui ont aussi dérivé sa convergence uniforme presque sûre dans un cadre d'échantillonnage. Si maintenant $\psi(Z; s; \theta) = \big(Y - \theta_0 - \theta_1(X - s) - \ldots - \theta_p(X - s)^p\big)^2$ et $\varphi(s; \theta) = \theta_0$, $\widehat{\varphi}_n(\cdot)$ est l'estimateur par polynômes locaux de la régression, considéré par exemple dans Fan [12] et Härdle et Tsybakov [20]. Des estimateurs par noyau de la moyenne et de la variance conditionnelles peuvent être obtenus en définissant $\psi(Z; s; \theta) = \ln \theta_1 + \frac{(Y - \theta_0)^2}{\theta_1}$ et en considérant $\varphi(s; \theta) = \theta_0$ et $\varphi(s; \theta) = \theta_1$, respectivement. Finalement, la classe d'estimateurs considérée ci-avant apparaît naturellement lorsqu'on développe des procédures de diagnostique sur l'hypothèse de constance de certains paramètres fonctionnels implicites (cf. Gouriéroux, Monfort et Tenreiro [13] et [14] et Gozalo et Linton [15]).

Le résultat le plus important de ce papier concernant la convergence uniforme sur un compact d'un estimateur fonctionnel de la classe introduite ci-dessus est présenté dans la Section 3. Il est dérivé à l'aide d'un théorème général de convergence uniforme que nous obtenons dans la Section 2. Des applications à l'estimation de la régression sont considérées dans la Section 4. Les démonstrations des divers résultats sont rassemblées dans la Section 5.

2 Convergence d'un M-estimateur fonctionnel

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé, Θ une partie compacte de $\mathbb{R}^p$ muni de sa tribu borélienne $\mathcal{B}(\Theta)$ et C un sous-ensemble de $\mathbb{R}^d$. Soit de plus $L_n, n \in \mathbb{N}^*$ une suite d'applications réelles définies sur $\Omega \times C \times \Theta$, telle que pour tout $n \in \mathbb{N}^*$:

A$_1$) l'application $L_n(\cdot; s; \theta) : \omega \to L_n(\omega; s; \theta)$ est mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

pour tout $s \in C$ et $\theta \in \Theta$;

A$_2$) l'application $L_n(\omega; s; \cdot) : \theta \rightarrow L_n(\omega; s; \theta)$ est continue pour tout $\omega \in \Omega$ et $s \in C$.

Dans les conditions précédentes il existe, pour chaque $s \in C$ et $n \in \mathbb{N}^*$, une application mesurable $\widehat{\theta}_n(\cdot; s)$ de $(\Omega, \mathcal{A})$ dans $(\Theta, \mathcal{B}(\Theta))$ telle que pour tout $\omega \in \Omega$ (cf. Jennrich [22], Lemme 2)

$$\widehat{\theta}_n(\omega; s) = arg\ min_{\theta \in \Theta}\ L_n(\omega; s; \theta).$$

$\mathcal{M}$ étant un mode stochastique (probabilité, presque sûre ou presque complète), le théorème suivant établit des conditions suffisantes pour la convergence uniforme, suivant le mode $\mathcal{M}$ de l'estimateur fonctionnel défini, pour $s \in C$, par

$$\widehat{\varphi}_n(\omega; s) = \varphi(s, \widehat{\theta}_n(\omega; s)),$$

où φ est une application connue de $C \times \Theta$ dans $\mathbb{R}$. Par simplicité on écrit $\widehat{\theta}_n(s)$ et $\widehat{\varphi}_n(s)$ à la place de $\widehat{\theta}_n(\omega; s)$ et $\widehat{\varphi}_n(\omega; s)$.

Théorème 2.1 *Soit φ une application continue dans $C \times \Theta$ où C et Θ sont des parties compactes de $\mathbb{R}^d$ et $\mathbb{R}^p$ respectivement, et admettons satisfaites les hypothèses* A$_1$) *et* A$_2$) *précédentes. On suppose de plus:*

B$_1$) *il existe une application continue $L_\infty : C \times \Theta \rightarrow \mathbb{R}$ telle que*

$$\sup_{s \in C} \sup_{\theta \in \Theta} \left| L_n(s; \theta) - L_\infty(s; \theta) \right| \xrightarrow[n \rightarrow +\infty]{\mathcal{M}} 0;$$

B$_2$) *il existe une application continue $\theta_\infty : C \rightarrow \Theta$ telle que pour tout $s \in C$, $\theta_\infty(s)$ est solution du problème d'optimisation*

$$min_{\theta \in \Theta}\ L_\infty(s; \theta); \tag{2.2}$$

B$_3$) *si $\theta^* : C \rightarrow \Theta$ est une solution de* (2.2) *on a*

$$\varphi(s; \theta^*(s)) = \varphi(s; \theta_\infty(s)) \equiv \varphi_\infty(s), \ \ \textit{pour tout } s \in C.$$

Alors,

$$\sup_{s \in C} \left| \widehat{\varphi}_n(s) - \varphi_\infty(s) \right| \xrightarrow[n \rightarrow +\infty]{\mathcal{M}} 0.$$

Remarque 2.3 Lorsque la fonction φ est définie par $\varphi(s; \theta) = \theta$ les hypothèses B$_2$) et B$_3$) se réduisent à l'habituelle condition d'identifiabilité: il existe une solution unique et continue $\theta_\infty : C \rightarrow \Theta$ du problème d'optimisation (2.2).

3 Convergence d'un M-estimateur à noyau

Soient $(Z_i, i \in \mathbb{Z})$ un processus stochastique fortement stationnaire à valeurs dans $\mathbb{R}^q$ et S une application mesurable de $\mathbb{R}^q$ dans $\mathbb{R}^d$ telle que la variable aléatoire $S(Z_0)$ est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. On note f_S sa densité de probabilité.

De façon à mesurer la dépendance entre le passé $F_{-\infty}^0$ et le futur $F_i^{+\infty}$, où, pour $n, m \in \mathbb{Z} \cup \{-\infty, +\infty\}$, F_n^m est la tribu engendrée par les variables $\{Z_i, n \le i \le m\}$, nous considérons les coefficients de mélange d'ordre i de type ϕ et α définis par

$$\phi(i) = \sup_{\substack{A \in F_{-\infty}^0 : P(A)>0 \\ B \in F_i^{+\infty}}} \left| P(B|A) - P(B) \right| \quad \text{et}$$

$$\alpha(i) = \sup_{\substack{A \in F_{-\infty}^0 \\ B \in F_i^{+\infty}}} \left| P(A \cap B) - P(A)P(B) \right|,$$

(cf. Billingsley [1] et Rosenblatt [26], respectivement). Le rapport entre ces deux mésures de dépendence se déduit de l'inégalité $2\alpha(i) \le \phi(i)$, valable pour $i \in \mathbb{N}^*$ (cf. Bradley [5]).

Etant données φ une application réelle définie dans $C \times \Theta$ où C et Θ sont des parties compactes de $\mathbb{R}^d$ et $\mathbb{R}^p$, respectivement, considérons l'estimateur $\widehat{\varphi}_n(\cdot)$ défini pour $s \in C$, par

$$\widehat{\varphi}_n(s) = \varphi(s; \widehat{\theta}_n(s)),$$

où $\widehat{\theta}_n(s)$ est une des solutions du problème d'optimisation

$$min_{\theta \in \Theta} \, L_n(s; \theta),$$

avec

$$L_n(s;\theta) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_n^d} K\left(\frac{s - S(Z_i)}{h_n} \right) \psi(Z_i; s; \theta), \tag{3.1}$$

où (h_n) est une suite de nombres réels strictement positifs convergeant vers zéro lorsque n tend vers l'infini, K est un noyau borné à support borné satisfaisant la condition de Lipschitz

$$\exists\, \gamma > 0,\ c_K > 0\ \forall\, u, v \in \mathbb{R}^d \ : \ \left| K(u) - K(v) \right| \le c_K \|u - v\|^\gamma,$$

et ψ est une application réelle vérifiant

$$\begin{aligned} &\exists\, \eta > 0,\ c_\psi > 0 \text{ avec } Ec_\psi(Z_0) < \infty : \ \forall\, s, s' \in C\ \forall\, \theta, \theta' \in \Theta : \\ &\qquad \left| \psi(z; s; \theta) - \psi(z; s'; \theta') \right| \le c_\psi(z) \big\|(s, \theta) - (s', \theta')\big\|^\eta, \end{aligned}$$

et dont l'application partielle $\psi(\cdot;s;\theta) : z \to \psi(z;s;\theta)$ est mesurable de $\mathbb{R}^q$ dans $\mathbb{R}$ pour tout $s \in C$ et $\theta \in \Theta$.

Le résultat suivant établit des conditions suffisantes pour la convergence uniforme presque sûre de l'estimateur $\widehat{\varphi}_n(\cdot)$. Remarquons qu'aucune condition de positivité sur K ou de convexité sur ψ est admise.

Théorème 3.2 *Dans les conditions générales précédentes on suppose que:*

H$_1$) *il existe $\beta > 1$ tel que*

$$E\left|\sup_{s\in C}\sup_{\theta\in\Theta}|\psi(Z_0;s;\theta)|\right|^{\beta} < \infty;$$

H$_2$) *il existe une suite entière (s_n) satisfaisant une des conditions suivantes:*

ϕ) $\sum_{i=1}^{s_n}\phi(i)/s_n \underset{n\to+\infty}{\longrightarrow} 0$ *et* $\exists A > 0 : n\phi(s_n)/s_n \leq A$, *avec* $s_n < n$;

α) $\exists A > 0 : n\sum_{i=1}^{s_n}\alpha(i)/s_n^2 \leq A$ *et* $n\big[\alpha(s_n)\big]^{2s_n/3n}/s_n \leq A$, *avec* $s_n < n/2$;

H$_3$) *la suite (h_n) vérifie conjointement avec (s_n) et β l'hypothèse*

$$\frac{n\,h_n^d}{M_n\,s_n\,\log n}\underset{n\to+\infty}{\longrightarrow}+\infty,$$

où $M_n \uparrow +\infty$ satisfait $\sum_{n=1}^{+\infty}M_n^{-\beta} < \infty$;

H$_4$) *il existe $\beta^* \in]1,\beta]$ tel que l'application*

$$z \longrightarrow E\Big[\,\Big|\sup_{s\in C}\sup_{\theta\in\Theta}|\psi(Z_0;s;\theta)|\Big|^{\beta^*}\mid S(Z_0) = z\Big]f_S(z),$$

est bornée dans un ouvert contenant C;

H$_5$) *l'application $L(r;s;\theta)$ définie pour $r \in \mathbb{R}^d$, $s \in C$ et $\theta \in \Theta$ par*

$$L(r;s;\theta) = E[\psi(Z_0;s;\theta)\mid S(Z_0) = r]f_S(r),$$

satisfait: $\quad \forall\,\epsilon > 0\ \exists\,\alpha > 0 : \sup_{s\in C}\sup_{\theta\in\Theta}\sup_{u:\|u-s\|\leq\alpha}\big|\,L(u;s;\theta) - L(s;s;\theta)\,\big| < \epsilon;$

H$_6$) *les hypothèses* B$_2$) *et* B$_3$) *du Théorème 2.1 sont satisfaites avec*

$$L_\infty(s;\theta) = L(s;s;\theta).$$

Alors, la suite d'estimateurs $\widehat{\varphi}_n(\cdot)$ converge presque sûrement et uniformément sur C, vers $\varphi_\infty(\cdot)$:

$$\sup_{s\in C}\big|\,\widehat{\varphi}_n(s) - \varphi_\infty(s)\,\big| \underset{n\to+\infty}{\overset{p.s.}{\longrightarrow}} 0.$$

Remarque 3.3 Les hypothèses H_2) et H_3) concernant la vitesse de convergence vers zéro des coefficients de mélange et de la suite (h_n) sont liées à l'utilisation dans le théorème de convergence précédent de deux inégalités distinctes de type Bernstein dues respectivement à Collomb [9], celle qui concerne les coefficients de mélange ϕ, et à Carbon [6], celle qui concerne les coefficients de mélange α (voir la Section 5). Telles hypothèses sont satisfaites dès que l'une des conditions ci-dessous est vérifiée, où β est donné dans l'hypothèse H_1):

M1) le processus $(Z_i, i \in \mathbb{Z})$ est géométriquement ϕ-mélangeant, i.e., $\exists\, C > 0$, $\rho \in]0,1[$, $\forall i \in \mathbb{N}^*$: $\phi(i) \leq C\rho^i$, et $\exists\, \eta \in]0, 1-\frac{1}{\beta}[$: $n^\eta\, h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$ ($M_n = n^\gamma$ avec $\gamma \in]1/\beta, 1-\eta[$ et $s_n = [n^\delta]$ avec $\delta \in]0, 1-\gamma-\eta[$).

M2) le processus $(Z_i, i \in \mathbb{Z})$ est algébriquement ϕ-mélangeant, i.e., $\exists\, C > 0$, $\omega > 0$, $\forall i \in \mathbb{N}^*$: $\phi(i) \leq C i^{-\omega}$, et $\exists\, \eta \in]0, 1-\frac{1}{\beta}-\frac{1}{1+\omega}[$: $n^\eta\, h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$ où $\beta > (1+\omega)/\omega$.

M3) le processus $(Z_i, i \in \mathbb{Z})$ est géométriquement α-mélangeant, i.e., $\exists\, C > 0$, $\rho \in]0,1[$, $\forall i \in \mathbb{N}^*$: $\alpha(i) \leq C\rho^i$, et $\exists\, \eta \in]0, \frac{1}{2}-\frac{1}{\beta}[$: $n^\eta\, h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$ où $\beta > 2$.

Remarque 3.4 Lorsqu'il existe $M > 0$ tel que $\sup_{s\in C} \sup_{\theta\in\Theta} |\psi(Z_0; s; \theta)| \leq M$, on conclut facilement que l'hypothèse H_3) peut être remplacée par l'hypothèse H_3') suivante:

H_3') la suite (h_n) vérifie conjointement avec (s_n) la condition $\dfrac{n\, h_n^d}{s_n\, \log n} \underset{n\to+\infty}{\longrightarrow} +\infty$.

Comme dans la remarque précédente, les hypothèses H_2) et H_3') sont satisfaites dès que l'une des hypothèses ci-dessous est vérifiée:

N1) le processus $(Z_i, i \in \mathbb{Z})$ est m-dépendant, i.e., les tribus $F_{-\infty}^0$ et $F_m^{+\infty}$ sont indépendantes, et $n\, h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$.

N2) le processus $(Z_i, i \in \mathbb{Z})$ est géométriquement ϕ-mélangeant et $n\, h_n^d / (\log n)^2 \underset{n\to+\infty}{\longrightarrow} +\infty$.

N3) le processus $(Z_i, i \in \mathbb{Z})$ est algébriquement ϕ-mélangeant et $n\, h_n^d / (n^{1/(1+\omega)} \log n) \underset{n\to+\infty}{\longrightarrow} +\infty$.

N4) le processus $(Z_i, i \in \mathbb{Z})$ est géométriquement α-mélangeant, et $\exists\, \eta \in]0, \frac{1}{2}[$: $n^\eta\, h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$.

4 Application à l'estimation de la régression

Dans cette section nous désignons par $(X_i, Y_i), i \in \mathbb{Z}$ un processus stochastique fortement stationnaire à valeurs dans $\mathbb{R}^d \times \mathbb{R}$, où X_0 est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$ de densité de probabilité f_X. Nous nous intéressons à l'estimation de la fonction de régression $r(\cdot) = E[Y_0|X_0 = \cdot]$ dans un sous-ensemble compact C de $\mathbb{R}^d$, où $r(\cdot)$ est supposée continue. Par simplicité d'exposition on considère uniquement le cas où le processus des observations est géométriquement α-mélangeant.

En reprenant les notations de la section précédente, (h_n) est une suite de nombres réels strictement positifs convergeant vers zéro lorsque n tend vers l'infini et K est un noyau sur $\mathbb{R}^d$ satisfaisant la condition de Lipschitz qui y a été introduite.

Considérons dans un premier temps l'estimateur défini pour $x \in C$, par

$$r_{1n}(x) = \arg\min_{\theta \in r(C)} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)(Y_i - \theta)^2.$$

Pour $x \in C$ et $r(x)$ à l'intérieur de $r(C)$, $r_{1n}(x)$ est, pour n assez grand, l'estimateur de Nadaraya-Watson.

Le Théorème 3.2 et la Remarque 3.3 permettent d'énoncer le résultat de convergence suivant (pour des résultats voisins voir Bosq [4], Györfi et al. [16] et Roussas [27]).

Théorème 4.1 *Soient $E|Y_0|^{4+\delta} < \infty$ pour un $\delta > 0$, $f_X(\cdot)$ strictement positive dans C et $r(\cdot)$, $E[Y_0^2|X_0 = \cdot]$ et $f_X(\cdot)$ continues dans un ouvert contenant C. S'il existe $\eta \in]0, \frac{1}{2} - \frac{2}{4+\delta}[$ tel que $n^\eta h_n^d / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$, alors*

$$\sup_{x \in C} \left| r_{1n}(x) - r(x) \right| \xrightarrow[n\to+\infty]{p.s.} 0.$$

Dans le cas réel $d = 1$, un autre estimateur de la régression a été, dans les dernières années, l'objet de nombreux travaux de recherche. On parle de l'estimateur par polynômes locaux considéré par exemple dans Fan [12]. Il est défini pour $x \in C$ par

$$r_{2n}(x) = \theta_{0n}(x), \text{ où}$$

$$\big(\theta_{0n}(x), \theta_{1n}(x)\big) = \arg\min_{(\theta_0,\theta_1) \in r(C)\times B} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)\big(Y_i - \theta_0 - \theta_1(X_i - x)\big)^2,$$

et B est un sous-ensemble compact de $\mathbb{R}$.

Le théorème suivant sur la convergence presque sûre de l'estimateur $r_{2n}(\cdot)$ est, à notre connaissance, nouveau.

Théorème 4.2 *Soient $E|X_0|^{4+\delta} < \infty$ et $E|Y_0|^{4+\delta} < \infty$ pour un $\delta > 0$, $f_X(\cdot)$ strictement positive dans C et $r(\cdot)$, $E[Y_0^2|X_0 = \cdot]$ et $f_X(\cdot)$ continues dans un ouvert contenant C. S'il existe $\eta \in]0, \frac{1}{2} - \frac{2}{4+\delta}[$ tel que $n^\eta h_n / \log n \underset{n\to+\infty}{\longrightarrow} +\infty$, alors*

$$\sup_{x \in C} \left| r_{2n}(x) - r(x) \right| \xrightarrow[n\to+\infty]{p.s.} 0.$$

5 Démonstrations

Le résultat suivant joue un rôle fondamental dans la démonstration du Théorème 2.1.

Lemme 5.1 *Sous les conditions générales du Théorème 2.1, admettons satisfaites les hypothèses* B_1) *et* B_2)*. Alors l'hypothèse* B_3) *est équivalente à*

$$\forall \delta > 0\ \exists \epsilon > 0\ \forall s \in C\ \forall \theta \in \Theta :$$

$$\left| \varphi(s;\theta) - \varphi(s;\theta_\infty(s)) \right| \geq \delta \Rightarrow L_\infty(s;\theta) - L_\infty(s;\theta_\infty(s)) > \epsilon. \tag{5.2}$$

Démonstration: Supposons que la condition (5.2) n'est pas satisfaite. Il existe alors $\delta > 0$ et deux suites $(s_n) \subset C$ et $(\theta_n) \subset \Theta$ tels que pour $n \in \mathbb{N}^*$:

$$\left| \varphi(s_n;\theta_n) - \varphi(s_n;\theta_\infty(s_n)) \right| \geq \delta \text{ et } L_\infty(s_n;\theta_n) - L_\infty(s_n;\theta_\infty(s_n)) \leq 1/n. \tag{5.3}$$

Comme $C \times \Theta$ est une partie compacte de $\mathbb{R}^d \times \mathbb{R}^p$, la suite (s_n, θ_n) admette une sous-suite (s_{k_n}, θ_{k_n}) convergeant vers un point (s^*, θ^*) de $C \times \Theta$. Alors la condition B_3) n'est pas vérifiée puisque l'application $\theta^* : C \to \Theta$ définie par $\theta^*(s) = \theta_\infty(s)$ si $s \neq s^*$ et $\theta^*(s^*) = \theta^*$ est une solution de (2.2) distincte de θ_∞. En effet, d'après (5.3) et tenant compte de la continuité de L_∞ et de θ_∞, nous avons

$$\begin{aligned} L_\infty(s^*;\theta^*) &= \lim_{n\to+\infty} L_\infty(s_{k_n};\theta_{k_n}) \\ &= \lim_{n\to+\infty} L_\infty(s_{k_n};\theta_\infty(s_{k_n})) \\ &= L_\infty(s^*;\theta_\infty(s^*)), \text{ et} \end{aligned}$$

$$\begin{aligned} \left| \varphi(s^*;\theta^*(s^*)) - \varphi(s^*;\theta_\infty(s^*)) \right| &= \left| \varphi(s^*;\theta^*) - \varphi(s^*;\theta_\infty(s^*)) \right| \\ &= \lim_{n\to+\infty} \left| \varphi(s_{k_n};\theta_{k_n}) - \varphi(s_{k_n};\theta_\infty(s_{k_n})) \right| \\ &\geq \delta. \end{aligned}$$

Afin d'établir l'implication réciproque, admettons qu'il existe $\theta^* : C \to \Theta$ solution de (2.2) et $s^* \in C$ tels que

$$\varphi(s^*;\theta^*(s^*)) \neq \varphi(s^*;\theta_\infty(s^*)).$$

D'après la condition (5.2) et pour $\delta = \left| \theta^*(s^*) - \theta_\infty(s^*) \right|$, il existe $\epsilon > 0$ tel que

$$\forall s \in C\, \forall \theta \in \Theta : \left| \varphi(s;\theta) - \varphi(s;\theta_\infty(s)) \right| \geq \delta \Rightarrow L_\infty(s;\theta) - L_\infty(s;\theta_\infty(s)) > \epsilon.$$

En particulier, pour $s^* \in C$ et $\theta^*(s^*) \in \Theta$, nous avons

$$L_\infty(s^*;\theta^*(s^*)) - L_\infty(s^*;\theta_\infty(s^*)) > \epsilon,$$

ou encore

$$L_\infty(s^*;\theta^*(s^*)) > L_\infty(s^*;\theta_\infty(s^*)),$$

ce qui permet de conclure. $\square$

Démonstration du Théorème 2.1: D'après le lemme précédent, on conclut que pour tout $\delta > 0$ il existe $\epsilon > 0$ tel que

$$\begin{aligned}
&\Big\{ \sup_{s\in C} \big| \varphi_n(s) - \varphi_\infty(s) \big| \geq \delta \Big\} \\
&\qquad\qquad = \Big\{ \sup_{s\in C} \big| \varphi(s;\widehat{\theta}_n(s)) - \varphi(s;\theta_\infty(s)) \big| \geq \delta \Big\} \\
&\qquad\qquad \subset \Big\{ \exists s \in C \,:\, L_\infty(s;\widehat{\theta}_n(s)) - L_\infty(s;\theta_\infty(s)) > \epsilon \Big\} \\
&\qquad\qquad \subset \Big\{ \sup_{s\in C} \sup_{\theta\in\Theta} \big| L_n(s;\theta) - L_\infty(s;\theta) \big| > \epsilon/2 \Big\}.
\end{aligned}$$

En effet, si $\omega \in \Omega$ est tel que $\sup_{s\in C} \big| \varphi_n(s) - \varphi_\infty(s) \big| \geq \delta$, il existe $s_o \in C$ tel que $\big| \varphi(s_o;\widehat{\theta}_n(s_o)) - \varphi(s_o;\theta_\infty(s_o)) \big| \geq \delta$. Ainsi, d'après le Lemme 5.1 on déduit qu'il existe $\epsilon > 0$ indépendant de s_o tel que

$$L_\infty(s_o;\widehat{\theta}_n(s_o)) - L_\infty(s_o;\theta_\infty(s_o)) > \epsilon.$$

De plus, si pour tout $s \in C$ et $\theta \in \Theta$, $\big| L_n(s;\theta) - L_\infty(s;\theta) \big| \leq \epsilon/2$, nous avons

$$\begin{aligned}
L_\infty(s;\widehat{\theta}_n(s)) - L_\infty(s;\theta_\infty(s)) &\leq \epsilon/2 + L_n(s;\widehat{\theta}_n(s)) - L_\infty(s;\theta_\infty(s)) \\
&\leq \epsilon/2 + L_n(s;\theta_\infty(s)) - L_\infty(s;\theta_\infty(s)) \\
&\leq \epsilon,
\end{aligned}$$

pour tout $s \in C$. La conclusion découle maintenant de l'hypothèse B_1). $\square$

Démonstration du Théorème 3.2: En tenant compte du Théorème 2.1 et l'hypothèse H_6), il nous suffit d'établir la convergence

$$\sup_{s\in C} \sup_{\theta\in\Theta} \big| L_n(s;\theta) - L_\infty(s;\theta) \big| \xrightarrow[n\to+\infty]{p.s.} 0, \tag{5.4}$$

où $L_n(s;\theta)$ est donné par (3.1) et $L_\infty(s;\theta) = E[\psi(Z_0;s;\theta) \mid S(Z_0) = s] f_S(s)$. Pour cela, considérons pour $s \in C$ et $\theta \in \Theta$ la décomposition

$$L_n(s;\theta) - L_\infty(s;\theta) = [L_n(s;\theta) - EL_n(s;\theta)] + [EL_n(s;\theta) - L_\infty(s;\theta)].$$

Chaqu'un des termes précédents est étudié dans la suite.

A) Convergence uniforme de $EL_n(s;\theta) - L_\infty(s;\theta)$ vers zéro

Comme $h_n \to 0, n \to +\infty$, et K est à support borné, on conclut que pour tout $\delta > 0$ il existe $n_0 \in \mathbb{N}^*$ tel que, pour $n \geq n_0$, on a

$$\begin{aligned}
&\sup_{s\in C}\sup_{\theta\in\Theta} \left| EL_n(s;\theta) - L_\infty(s;\theta) \right| \\
&\quad \leq \sup_{c\in C}\sup_{\theta\in\Theta} \int_{\mathbb{R}^d} \left| L(s-uh_n;s;\theta) - L(s;s;\theta) \right| |K(u)| du \\
&\quad \leq \sup_{s\in C}\sup_{\theta\in\Theta} \sup_{u:\|u-s\|\leq\delta} \left| L(u;s;\theta) - L(s;s;\theta) \right| \int_{\mathbb{R}^d} |K(u)| du,
\end{aligned}$$

où $L(r;s;\theta) = E[\psi(Z_0;s;\theta)|S(Z_0) = r]f_S(r)$. L'hypothèse H_5) permet maintenant de conclure.

B) Décomposition de $L_n(s;\theta) - EL_n(s;\theta)$

Pour $s \in C$ et $\theta \in \Theta$ fixés, considérons la décomposition

$$\begin{aligned}
L_n(s;\theta) &= L_n'(s;\theta) + L_n''(s;\theta) \\
&= \frac{1}{n}\sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s-S(Z_i)}{h_n}\right) \psi(Z_i;s;\theta) \mathbb{I}_{|\psi(Z_i;s;\theta)|\leq M_n} \\
&+ \frac{1}{n}\sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s-S(Z_i)}{h_n}\right) \psi(Z_i;s;\theta) \mathbb{I}_{|\psi(Z_i;s;\theta)|> M_n},
\end{aligned}$$

où la suite (M_n) est donnée dans l'hypothèse H_3).

Nous avons

$$L_n'(s;\theta) - EL_n'(s;\theta) = \frac{1}{n}\sum_{i=1}^n X_{n,i},$$

où d'après les hypothèses H_1) et H_4) il existe des constantes $a_0, b_0, c_0 > 0$ telles que pour $i = 1, \ldots, n$, et n assez grand on a

$$(5.5) \qquad EX_{n,i} = 0, |X_{n,i}| \leq \frac{a_0 M_n}{h_n^d}, E|X_{n,i}| \leq b_0 \text{ et } E(X_{n,i})^2 \leq \frac{c_0 M_n^{2-\beta^*}}{h_n^d}.$$

En conséquence de l'hypothèse H_2) nous établirons dans la suite que pour tout $\epsilon > 0$ il existe $\mathbf{a}, \mathbf{b} > 0$ tels que pour n assez grand

$$\sup_{s\in C}\sup_{\theta\in\Theta} P\Big(\left| L_n'(s;\theta) - EL_n'(s;\theta) \right| > \epsilon \Big) \leq \mathbf{a} \exp\left(-\mathbf{b} \frac{n\, h_n^d}{M_n\, s_n} \right).$$

On le fera en considérant, dans un premier temps, valable la condition ϕ) de l'hypothèse H_2), et dans un deuxième temps valable la condition α) de la même hypothèse. Les coefficients $\phi(i)$ et $\alpha(i)$ considérés dans la suite sont ceux définis dans la Section 3.

i) La condition ϕ) est vérifiée

Nous utiliserons l'inégalité exponentielle suivante due à Collomb [9] (Lemme 1, pg. 449) qui améliore une inégalité précédente de Bosq [3]:

Lemme 5.6 *Soit, pour tout $n \in \mathbb{N}^*$, une suite de v.a.r. $X_{n,i} = g_n(Z_i), i \in \mathbb{N}^*$ vérifiant*

$$EX_{n,i} = 0, |X_{n,i}| \leq d_n, E|X_{n,i}| \leq \delta_n \text{ et } E(X_{n,i})^2 \leq D_n,$$

pour $i = 1, \ldots, n$ et $n \geq n_0$ avec $n_0 \in \mathbb{N}^*$ fixe.

Alors $\forall\, \epsilon > 0\, \forall\, n \geq n_0$ on a

$$P\left(\left|\sum_{i=1}^{n} X_{n,i}\right| > n\epsilon\right) \leq 2\exp\left(-\frac{n}{k}\left[\frac{\epsilon}{4d_n} - \frac{3\big(D_n + 4\delta_n d_n \sum_{i=1}^{k}\phi(i)\big)}{8kd_n^2} - 3\sqrt{e}\,\phi(k)\right]\right),$$

pour tout $k \in \mathbb{N}^*$ avec $k < n$.

En tenant compte des majorations (5.5) et du lemme précédent en choisissant $k = s_n$, on obtient pour tout $\epsilon > 0$, n assez grand et $s \in C$ et $\theta \in \Theta$

$$\begin{aligned} & P\Big(\,\big|\,L_n'(s;\theta) - EL_n'(s;\theta)\,\big| > \epsilon\Big) \\ \leq\quad & 2\exp\left(-\frac{n\,h_n^d}{M_n\,s_n}\left[\frac{\epsilon}{4a_0} - \frac{3c_0}{8a_0^2 M_n^{\beta^*-1} s_n} - \frac{3b_0\sum_{i=1}^{k}\phi(i)}{2a_0 s_n}\right] + 3\sqrt{e}\,\phi(s_n)/s_n\right). \end{aligned}$$

L'hypothèse ϕ) permet maintenant de conclure.

ii) La condition α) est vérifiée

Nous utiliserons l'inégalité exponentielle suivante due à Carbon [6] (Proposition 23, pg. 138):

Lemme 5.7 *Soit, pour tout $n \in \mathbb{N}^*$, une suite de v.a.r. $X_{n,i} = g_n(Z_i), i \in \mathbb{N}^*$ vérifiant*

$$EX_{n,i} = 0, |X_{n,i}| \leq d_n \text{ et } E(X_{n,i})^2 \leq D_n,$$

pour $i = 1, \ldots, n$ et $n \geq n_0$ avec $n_0 \in \mathbb{N}^*$ et $n_0 \geq 3$ fixe.

Alors $\forall\, \epsilon > 0\, \forall\, n \geq n_0$ on a

$$P\left(\left|\sum_{i=1}^{n} X_{n,i}\right| > n\epsilon\right) \leq 2\exp\left(-\frac{n}{k}\left[\frac{\epsilon}{4d_n e} - \frac{D_n + 8d_n^2\sum_{i=1}^{k}\alpha(i)}{4kd_n^2 e} - 2\sqrt{e}\,\alpha(k)^{2k/3n}\right]\right),$$

pour tout $k \in \mathbb{N}^*$ avec $k < n/2$.

Les majorations (5.5) et l'application du lemme précédent en choisissant $k = s_n$ nous conduisent à l'inégalité suivante valable pour tout $\epsilon > 0$, n assez grand et $s \in C$ et $\theta \in \Theta$

$$P\Big(\,|\, L'_n(s;\theta) - EL'_n(s;\theta) \,|\, > \epsilon \Big)$$

$$\leq \; 2\exp\bigg(-\frac{n\, h_n^d}{M_n\; s_n}\Big[\frac{\epsilon}{4a_0 e} - \frac{c_0}{4a_0^2 M_n^{\beta^*-1}\, s_n\, e}\Big] + \frac{2n\sum_{i=1}^k \alpha(i)}{e s_n^2} + 2\sqrt{e}\, n\big(\alpha(s_n)\big)^{2s_n/3n}/s_n \bigg).$$

L'hypothèse α) permet maintenant de conclure.

C) <u>Convergence uniforme presque sûre de $L''_n(s;\theta) - EL''_n(s;\theta)$ vers zéro</u>

D'après l'inégalité de Bienaymé-Chebyshev et l'hypothèse H$_1$), on a, pour $\epsilon > 0$,

$$P\big(\Psi(Z_n) > \epsilon M_n\big) \leq \epsilon^{-\beta} M_n^{-\beta} E|\Psi(Z_0)|^{\beta},$$

où $\Psi(z) = \sup_{s\in C}\sup_{\theta\in\Theta} |\psi(z;s;\theta)|$. La convergence de la série $\sum_{n=1}^{+\infty} M_n^{-\beta}$ et le théorème de Borel-Cantelli, permettent d'écrire

$$\frac{\Psi(Z_n)}{M_n} \underset{n\to+\infty}{\overset{p.s.}{\longrightarrow}} 0.$$

La suite M_n étant croissante on obtient (cf. Mack et Silverman [23] pg. 408)

$$\overline{\lim}\, \frac{\sup_{j=1,\ldots,n} \Psi(Z_j)}{M_n} \leq 1, \text{ p.s.},$$

et donc

$$\sup_{s\in C}\sup_{\theta\in\Theta} \big|\, L''_n(s;\theta) \,\big| \overset{p.s.}{=} 0, \text{ pour } n \text{ assez grand.}$$

D'autre part,

$$EL''_n(s;\theta) = \frac{1}{h_n^d}\int_{\mathbb{R}^d} K\Big(\frac{s-z}{h_n}\Big)\, E\big[\psi(Z_0;s;\theta) \mathbb{1}_{|\psi(Z_0;s;\theta)|>M_n} \mid S(Z_0) = z\big] f_S(z) dz,$$

et donc l'inégalité de Hölder et une nouvelle application de l'inégalité de Bienaymé-Chebyshev permettent d'obtenir, pour $s \in C$, $\theta \in \Theta$ et $\beta^* \in]1,\beta]$ donné dans l'hypothèse H$_4$), la majoration

$$\begin{aligned} \big|\, EL''_n(s;\theta) \,\big| \;\; &\leq \;\; \frac{1}{h_n^d}\int_{\mathbb{R}^d} \Big| K\Big(\frac{s-u}{h_n}\Big)\Big| E^{1/\beta^*}\big[|\Psi(Z_0)|^{\beta^*} \mid S(Z_0) = z\big] \\ &\qquad\qquad \times E^{1-1/\beta^*}\big[\mathbb{1}_{|\Psi(Z_0)|>M_n} \mid S(Z_0) = z\big] f_S(z) dz \\ &\leq \;\; \Big(\frac{1}{M_n}\Big)^{\beta^*-1} \int_{\mathbb{R}^d} |K(u)| E\big[|\Psi(Z_0)|^{\beta^*} \mid S(Z_0) = s - uh_n\big] f_S(s-uh_n) du. \end{aligned}$$

D'après l'hypothèse H$_4$) et du fait que K est à support borné on a alors

$$\sup_{s\in C}\sup_{\theta\in\Theta} \big|\, EL''_n(s;\theta) \,\big| = O\Big(\frac{1}{M_n}\Big)^{\beta^*-1},$$

ce qui permet d'obtenir la convergence souhaitée.

D) Equicontinuité de $L_n(s;\theta)$

D'après les conditions de Lipschitz sur K et ψ, on a pour $s, s' \in C$ et $\theta, \theta' \in \Theta$,

$$\begin{aligned}
&\left| L_n(s;\theta) - L_n(s';\theta') \right| \\
= & \left| \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_n^d}\left[K\left(\frac{s-S(Z_i)}{h_n}\right) - K\left(\frac{s'-S(Z_i)}{h_n}\right)\right]\psi(Z_i;s;\theta) \right. \\
& \left. + \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_n^d} K\left(\frac{s'-S(Z_i)}{h_n}\right)\left[\psi(Z_i;s;\theta) - \psi(Z_i;s';\theta')\right] \right| \\
\leq & \; C_K \frac{\|s-s'\|^\gamma}{h_n^{d+1}} \frac{1}{n}\sum_{i=1}^{n}\Psi(Z_i) + \sup_{u\in\mathbb{R}^d}|K(u)| \frac{\|(s,\theta)-(s',\theta')\|^\eta}{h_n^d} \frac{1}{n}\sum_{i=1}^{n} C_\psi(Z_i),
\end{aligned}$$

et donc de l'application du théorème ergodique, on déduit

$$\left| L_n(s;\theta) - L_n(s';\theta') \right| = O\left(\frac{\|s-s'\|^\gamma}{h_n^{d+1}} + \frac{\|(s,\theta)-(s',\theta')\|^\eta}{h_n^d} \right).$$

E) Convergence uniforme presque sûre de $L_n(s;\theta) - EL_n(s;\theta)$ vers zéro

Considérons un recouvrement fini $C_\ell \times \Theta_{\ell'}$, $\ell = 1, \ldots, N$, $\ell' = 1, \ldots, N'$, du compact $C \times \Theta$, dont les éléments sont de diamètre (par rapport à la norme du maximum) inférieur ou égal à $\left(h_n n^{-1}\right)^{1/\min(\gamma,\eta)}$. Ce recouvrement peut être choisi de façon que $NN' = O\left(\left(h_n n^{-1}\right)^{(d+p)/\min(\gamma,\eta)}\right)$.

En choisissant $s_\ell \in C_\ell$ et $\theta_{\ell'} \in \Theta_\ell$, pour $\ell = 1, \ldots, N$ et $\ell' = 1, \ldots, N'$, on a, d'après les points C) et D) précédents

$$\begin{aligned}
&\sup_{s\in C}\sup_{\theta\in\Theta} \left| L_n(s;\theta) - EL_n(s;\theta) \right| \\
= & \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \sup_{s\in C_\ell}\sup_{\theta\in\Theta_{\ell'}} \left| L_n(s;\theta) - EL_n(s;\theta) \right| \\
\leq & \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \sup_{s\in C_\ell}\sup_{\theta\in\Theta_\ell} \left| L_n(s;\theta) - L_n(s_\ell;\theta_{\ell'}) - E\left[L_n(s;\theta) - L_n(s_\ell;\theta_{\ell'})\right] \right| \\
& + \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \left| L'_n(s_\ell;\theta_{\ell'}) - EL'_n(s_\ell;\theta_{\ell'}) \right| + o(1) \\
\leq & \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \sup_{s\in C_\ell}\sup_{\theta\in\Theta_{\ell'}} O\left(\frac{\|s-s_\ell\|^\gamma}{h_n^{d+1}} + \frac{\|(s,\theta)-(s_\ell,\theta_{\ell'})\|^\eta}{h_n^d} \right) \\
& + \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \left| L'_n(s_\ell;\theta_{\ell'}) - EL'_n(s_\ell;\theta_{\ell'}) \right| + o(1) \\
= & \; O\left(\frac{1}{nh_n^d}\right) + \max_{\substack{\ell=1,\ldots,N \\ \ell'=1,\ldots,N'}} \left| L'_n(s_\ell;\theta_{\ell'}) - EL'_n(s_\ell;\theta_{\ell'}) \right| + o(1).
\end{aligned}$$

Comme $nh_n^d \to +\infty$, $n \to +\infty$, il suffit d'établir la convergence presque sûre vers zéro du second terme. D'après le point B), pour $\epsilon > 0$ quelconque, et pour n assez grand, nous avons

$$
\begin{aligned}
P\Big(\max_{\substack{\ell=1,\dots,N \\ \ell'=1,\dots,N'}} \big| L_n'(s_\ell;\theta_{\ell'}) - EL_n'(s_\ell;\theta_{\ell'}) \big| > \epsilon \Big) & \\
\leq & \sum_{\ell=1}^{N} \sum_{\ell'=1}^{N'} P\Big(\big| L_n'(s_\ell;\theta_{\ell'}) - EL_n'(s_\ell;\theta_{\ell'}) \big| > \epsilon \Big) \\
\leq & \ \mathbf{a} N N' \exp\Big(-\mathbf{b} \frac{n\, h_n^d}{M_n\, s_n} \Big).
\end{aligned}
$$

D'après l'hypothèse $\lim_{n\to+\infty} \frac{n\, h_n^d}{M_n\, s_n\, \log n} = +\infty$, et en tenant en compte que $NN' = O(n^{t(d,p,\gamma,\eta)})$ avec $t(d,p,\gamma,\eta) = \frac{(d+1)(d+p)}{d\min(\gamma,\eta)}$, nous obtenons

$$
\sum_{n=1}^{\infty} P\Big(\max_{\substack{\ell=1,\dots,N \\ \ell'=1,\dots,N'}} \big| L_n'(s_\ell;\theta_{\ell'}) - EL_n'(s_\ell;\theta_{\ell'}) \big| > \epsilon \Big) < +\infty,
$$

ce qui entraîne la convergence presque sûre souhaitée.

La convergence (5.4) découle maintenant des points précédents. □

Démonstration des Théorèmes 4.1 et 4.2: Conséquence immédiate du Théorème 3.2 en faisant pour $i \in \mathbb{Z}$, $Z_i = (X_i, Y_i)$, $S(Z_i) = X_i$, et en considérant $\psi(Z_i; x; \theta) = (Y_i - \theta)^2$ et $\varphi(x;\theta) = \theta$ dans le cas de l'estimateur $r_{1n}(\cdot)$ et $\psi(Z_i; x; \theta) = (Y_i - \theta_0 - \theta_1(X_i - x))^2$ et $\varphi(x;\theta) = \theta_0$ dans le cas de l'estimateur $r_{2n}(\cdot)$. □

Bibliographie

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[2] Boente, G. et Freiman, R. (1989). Robust nonparametric regression estimation for dependent observations. *Ann. Statist.*, **17**, 1242-1256.

[3] Bosq, D. (1975). Inégalité de Bernstein pour les processus stationnaires et mélangeants. Applications. *C. R. Acad. Sci. Paris*, **281**, 1095-1098.

[4] Bosq, D. (1988). Prévision nonparamétrique d'un processus stationnaire non borné. DP. LSTA 81, Université Paris VI.

[5] Bradley, R.C. (1986). Basic properties of strong mixing conditions. *In: Dependence in Probability and Statistics, a survey of recent results*, ed. Eberlein, E. et Taqqu, M., Birkhäuser, Boston, 165-192.

[6] Carbon, M. (1988) *Inégalités de Grandes Deviations dans les Processus. Applications à l'Estimation Fonctionnelle.* Thèse de Doctorat de l'Université de Paris VI.

[7] Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.*, **74**, 829-836.

[8] Collomb, G. (1979). Conditions nécessaires et suffisantes de convergence uniforme d'un estimateur de la régression, estimation des dérivées de la régression. *C. R. Acad. Sci. Paris*, **288**, 161-164.

[9] Collomb, G. (1984). Propriétés de convergence presque complète du prédicteur à noyau. *Z. Wahrsch. Verw. Gebiete*, **66**, 441-460.

[10] Collomb, G. et Härdle, W. (1986). Strong uniform convergence rates in robust nonparametric time series analysis and prediction: kernel regression estimation from dependent observations. *Stochastic Process. Appl.*, **23**, 77-89.

[11] Devroye, L.P. (1978). The uniform convergence of the Nadaraya-Watson regression function estimate. *Canad. J. Statist.*, **6**, 179-191.

[12] Fan, J. (1992). Design-adaptative Nonparametric Regression. *J. Amer. Statist. Assoc.*, **87**, 998-1004.

[13] Gouriéroux, C., Monfort, A. et Tenreiro, C. (1994). Kernel M-estimators: nonparametric diagnostics for structural models. Document de travail 9405, CEPREMAP, Paris.

[14] Gouriéroux, C., Monfort, A. et Tenreiro, C. (1995). Kernel M-estimators and functional residuals plots. Document de travail 9546, CREST, Paris.

[15] Gozalo, P. et Linton, O. (1994). Local nonlinear least squares estimation: using parametric information nonparametrically. Cowles foundation discussion paper 1075, Yale University.

[16] Györfi, L., Härdle, W., Sarde, P. et Vieu, P. (1989). *Nonparametric Curve Estimation from Time Series*. Lect. Notes in Statist., 60, Springer-Verlag, New York.

[17] Härdle, W. (1984). Robust regression function estimation. *J. Multivariate Anal.*, **14**, 169-180.

[18] Härdle, W. et Luckhaus, S. (1984). Uniform consistency of a class of regression function estimators. *Ann. Math. Statist.*, **12**, 612-623.

[19] Härdle, W. et Tsybakov, A.B. (1988). Robust nonparametric regression with simultaneous scale curve estimation. *J. Multivariate Anal.*, **14**, 169-180.

[20] Härdle, W. et Tsybakov, A.B. (1995). Local polynomial estimators of the volatility function in nonparametric autoregression. Manuscrit non publié.

[21] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **33**, 73-101.

[22] Jennrich, R. (1969). Asymptotic properties of nonlinear least squares estimators. *Ann. Math. Statist.*, **40**, 633-643.

[23] Mack, Y.P. et Silverman, B. (1982). Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrsch. Verw. Gebiete*, **61**, 405-415.

[24] Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.*, **9**, 141-142.

[25] Robinson, P.M. (1984). Robust nonparametric autoregression. *In: Robust and Nonlinear Time Series Analysis*, Lect. Notes in Statist., 26, 247-255.

[26] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA.*, **42**, 43-47.

[27] Roussas, G.G. (1990). Nonparametric regression estimation under mixing conditions. *Stochastic Process. Appl.*, **36**, 107-116.

[28] Watson, G.S. (1964). Smooth regression analysis. *Sankhya A*, **26**, 359-372.

Departamento de Matemática, Universidade de Coimbra

Apartado 3008, 3000 Coimbra, Portugal

e-mail: tenreiro@mat.uc.pt