



HAL
open science

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

Carlos Tenreiro

► **To cite this version:**

Carlos Tenreiro. Convergence uniforme presque sûre d'une classe d'M-estimateurs par noyau pour des observations dépendantes. *Annales de l'ISUP*, 1999, XXXXIII (1), pp.43-58. hal-03653437

HAL Id: hal-03653437

<https://hal.science/hal-03653437v1>

Submitted on 27 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pub. Inst. Stat. Univ. Paris
 XXXXIII, fasc. 1, 1999, 43 à 58

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

Carlos TENREIRO

Departamento de Matemática, Universidade de Coimbra

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], Ioffe et Katkovnik [21] et Gouriéroux *et al.* [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], Ioffe et Katkovnik [21] et Gouriéroux *et al.* [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 applica-

* Ce travail a bénéficié de l'appui du CMUC/FCT.

Mots clés: M-estimateurs par noyau, stationnarité, mélangeance, convergence uniforme presque sûre. *Key words and phrases:* Kernel M-estimators, stationarity, mixing, strong uniform convergence. *Classification AMS 1991:* 62G07, 62G05.

tion réelle connue continue en θ que l'on désigne par fonction critère.

Dans des études précédentes (voir Collomb et Härdle [10], Ioffe et Katkovnik [21]), la fonction ψ est supposée satisfaire certaines conditions de convexité. Or, comme l'avons marqué Gouriéroux *et al.* [13] et [14], il y a diverses situations où une telle condition n'est pas vérifiée. En employant une technique proche de celle utilisée dans le cadre des M-estimateurs paramétriques par Jenrich [22] et White [32], Gouriéroux *et al.* [14] ont dérivé, lorsque la fonction critère est indépendante de s et dans le cadre des processus géométriquement α -mélangeants, la convergence presque sûre, uniformément par rapport à $s \in C$, de la suite d'estimateurs $\hat{\theta}_n(\cdot)$ vers la solution supposée unique (condition d'identifiabilité asymptotique) du problème d'optimisation $\min_{\theta \in \Theta} E[\psi(Z; s; \theta) \mid S(Z) = s]$.

Sous des conditions analogues à celles utilisées dans cette dernière référence, nous nous proposons dans cet article d'établir la convergence uniforme presque sûre de la suite d'estimateurs fonctionnels $\hat{\varphi}_n(\cdot)$ 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], Ioffe et Katkovnik [21] et Gouriéroux *et al.* [14]. On le fera sans aucune hypothèse de positivité sur K ou de convexité ou différentiabilité sur ψ , 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 qui 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 $\varphi(s; \theta(s)) = \varphi_\infty(s)$, pour tout $s \in C$. On montrera alors que

$$\sup_{s \in C} |\hat{\varphi}_n(s) - \varphi_\infty(s)| \xrightarrow[n \rightarrow +\infty]{P.S.} 0.$$

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 $\rho(y) = y^2$, $\hat{\varphi}_n(\cdot)$ est l'estimateur à noyau usuel de la régression $E[Y|S(Z) = \cdot]$ introduit par Nadaraya [24] et Watson [31] et dont la convergence uniforme presque sûre a été établie par Mack et Silverman [23] dans un cadre d'échantillonnage. Si $\rho(y) = |y|$, $\hat{\varphi}_n(s)$ est l'estimateur de la médiane de la distribution conditionnelle $Y|S(Z) = s$. Un estimateur robuste au sens de Huber [20] 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 $\psi(Z; s; \theta) = (Y - \sum_{j=0}^p \theta_j X^j)^2$ et $\varphi(s; \theta) = \sum_{j=0}^p \theta_j s^j$, $\widehat{\varphi}_n(\cdot)$ est un M-estimateur par noyau de la régression $E[Y|X = s]$ correspondant à l'estimateur local pondéré considéré par Stone [29] (voir aussi Cleveland [7]). Plus généralement si $\psi(Z; s; \theta) = (Y - m(X; \theta))^2$ 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 $\psi(Z; s; \theta) = \Phi(Y - \sum_{j=0}^p \theta_j \varphi_j(X - s))$, où Φ est une fonction convexe dans \mathbb{R} avec un minimum à l'origine et φ_j , $j = 0, 1, \dots, p$ sont des fonctions non toutes nulles à l'origine, et $\varphi(s; \theta) = \sum_{j=0}^p \theta_j \varphi_j(0)$, $\widehat{\varphi}_n(\cdot)$ est l'estimateur de la régression considéré par Tsybakov [30] (avec $\varphi_j(u) = u^j$, Ioffe et Katkovnik [21] et Fan [12] (avec $\varphi_j(u) = u^j$ et $\Phi(u) = u^2$). 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 diagnostic sur l'hypothèse de constance de certains paramètres fonctionnels implicites (cf. Gouriéroux *et al.* [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 d'un M-estimateur fonctionnel que nous obtenons dans la Section 2. Une application à l'estimation de la régression est considérée dans la Section 4, où on dérive la convergence uniforme presque sûre sur un compact de l'estimateur par polynômes locaux considéré par Ruppert et Wand [28]. Les démonstrations des divers résultats sont rassemblées dans la Section 5.

2 Convergence uniforme d'un M-estimateur fonctionnel

Soient (Ω, \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 $n \in \mathbb{N}^*$:

A₁) l'application $L_n(\cdot; s; \theta) : \omega \rightarrow L_n(\omega; s; \theta)$ est mesurable de (Ω, \mathcal{A}) dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ pour tout $s \in C$ et $\theta \in \Theta$;

A₂) 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 (Ω, \mathcal{A}) dans $(\Theta, \mathcal{B}(\Theta))$ telle que pour tout $\omega \in \Omega$ (cf. Jennrich [22], Lemme 2)

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

Le théorème suivant établit des conditions suffisantes pour la convergence uniforme presque sûre 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 $\widehat{\varphi}_n(\omega; s)$. Le théorème reste valable si la convergence presque sûre est remplacée par la convergence en probabilité ou la convergence presque complète.

Théorème 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

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

$B_2)$ 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

$$\min_{\theta \in \Theta} L_\infty(s; \theta); \tag{1}$$

$B_3)$ si $\theta^* : C \rightarrow \Theta$ est une solution de (1) on a

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

Alors,

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

Remarque 1 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 (1).

3 Convergence uniforme 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 $\{Z_i, n \leq i \leq m\}$, nous considérons les coefficients de mélange d'ordre i de type ϕ et α définis par

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

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

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

Etant donné 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

$$L_n(s; \theta) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_n^d} K\left(\frac{s - S(Z_i)}{h_n}\right) \psi(Z_i; s; \theta), \quad (2)$$

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

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

et ψ est une application réelle vérifiant

$$\exists \eta > 0, c_\psi > 0 \text{ avec } E c_\psi(Z_0) < \infty : \forall s, s' \in C \forall \theta, \theta' \in \Theta :$$

$$|\psi(z; s; \theta) - \psi(z; s'; \theta')| \leq c_\psi(z) \|(s, \theta) - (s', \theta')\|^\eta,$$



et dont l'application partielle $\psi(\cdot; s; \theta) : z \rightarrow \psi(z; s; \theta)$ 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 2 Dans les conditions générales précédentes on suppose que:

H₁) il existe $\beta > 1$ tel que

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

H₂) il existe une suite entière (s_n) satisfaisant une des conditions suivantes:

$$\phi) \sum_{i=1}^{s_n} \phi(i)/s_n \xrightarrow{n \rightarrow +\infty} 0 \text{ et } \exists A > 0 : n\phi(s_n)/s_n \leq A, \text{ avec } s_n < n;$$

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

H₃) la suite (h_n) vérifie conjointement avec (s_n) et β l'hypothèse

$$\frac{n h_n^d}{M_n s_n \log n} \xrightarrow{n \rightarrow +\infty} +\infty,$$

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

H₄) il existe $\beta^* \in]1, \beta]$ tel que l'application

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

est bornée dans un ouvert contenant C ;

H₅) 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: $\forall \epsilon > 0 \exists \alpha > 0 : \sup_{s \in C} \sup_{\theta \in \Theta} \sup_{u: \|u-s\| \leq \alpha} |L(u; s; \theta) - L(s; s; \theta)| < \epsilon$;

H₆) les hypothèses B₂) et B₃) du Théorème 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 $\varphi_\infty(\cdot)$:

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

Remarque 2 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 $\exists \eta \in]0, 1 - \frac{1}{\beta}[: n^\eta h_n^d / \log n \xrightarrow{n \rightarrow +\infty} +\infty$ ($M_n = n^\gamma$ 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., $\exists C > 0, \omega > 0, \forall i \in \mathbb{N}^* : \phi(i) \leq Ci^{-\omega}$, et $\exists \eta \in]0, 1 - \frac{1}{\beta} - \frac{1}{1+\omega}[: n^\eta h_n^d / \log n \xrightarrow{n \rightarrow +\infty} +\infty$ 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 $\exists \eta \in]0, \frac{1}{2} - \frac{1}{\beta}[: n^\eta h_n^d / \log n \xrightarrow{n \rightarrow +\infty} +\infty$ où $\beta > 2$.

Remarque 3 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 $\frac{n h_n^d}{s_n \log n} \xrightarrow{n \rightarrow +\infty} +\infty$.

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 $n h_n^d / \log n \xrightarrow{n \rightarrow +\infty} +\infty$.

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

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

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

4 Application à l'estimation de la régression

Soit $(X_i, Y_i), i \in \mathbb{Z}$ un processus stochastique fortement stationnaire à valeurs dans $\mathbb{R}^d \times \mathbb{R}$. On suppose que X_0 admet une densité de probabilité f_X par rapport à la

mesure de Lebesgue sur \mathbb{R}^d . Soit de plus C un sous-ensemble compact C de \mathbb{R}^d , où la fonction de régression $r(\cdot) = E[Y_0|X_0 = \cdot]$ est supposée continue.

Considérons l'estimateur par polynômes locaux $r_n(\cdot)$ (cf. Tsybakov [30], Fan [12] et Ruppert et Wand [28]) défini pour $x \in C$, comme une des solutions pour α du problème

$$\min_{\substack{\alpha \in I \\ \beta \in R}} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) (Y_i - \alpha - \beta^T(x - X_i))^2,$$

où I et R sont des sous-ensembles compacts de \mathbb{R}^d avec $r(C) \subset I$ et, 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 les conditions introduites ci-avant. Remarquons qu'aucune condition de positivité est imposée sur K .

Comme application du Théorème 2, nous établissons maintenant la convergence uniforme presque sûre sur C , de $r_n(\cdot)$ vers $r(\cdot)$. Ce résultat est, à notre connaissance, nouveau. Par simplicité d'exposition on considère uniquement le cas où le processus des observations est géométriquement α -mélangeant.

Théorème 3 *Sous les conditions suivantes: $E \left| \sup_{\beta \in R} |\beta^T X_0| \right|^{4+\delta} < \infty$ et $E|Y_0|^{4+\delta} < \infty$ pour un $\delta > 0$; $f_X(\cdot)$ strictement positive dans C ; $r(\cdot)$, $E[Y_0^2|X_0 = \cdot]$ et $f_X(\cdot)$ continues et $E[|Y_0|^{2+\delta}|X_0 = \cdot]$ bornée, dans un ouvert contenant C ; alors s'il existe $\eta \in]0, \frac{1}{2} - \frac{2}{4+\delta}[$ tel que $n^\eta h_n^d / \log n \xrightarrow{n \rightarrow +\infty} +\infty$, on a*

$$\sup_{x \in C} |r_n(x) - r(x)| \xrightarrow[n \rightarrow +\infty]{p.s.} 0.$$

Lorsque que R se réduit à l'origine, $r_n(\cdot)$ est l'estimateur de Nadaraya-Watson (cf. Nadaraya [24] et Watson [31]). Voir Bosq [4], Györfi et al. [16] et Roussas [27] pour des résultats voisins sur la convergence uniforme presque sûre de cet estimateur.

5 Démonstrations

Le résultat suivant joue un rôle fondamental dans la démonstration du Théorème 1.

Lemme 4 *Sous les conditions générales du Théorème 1, admettons satisfaite la hypothèse B_2). Alors l'hypothèse B_3) est équivalente à*

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

$$|\varphi(s; \theta) - \varphi(s; \theta_\infty(s))| \geq \delta \Rightarrow L_\infty(s; \theta) - L_\infty(s; \theta_\infty(s)) > \epsilon. \quad (3)$$

Démonstration: Supposons que la condition (3) n'est pas satisfaite. Il existe alors $\delta > 0$ et deux suites $(s_n) \subset C$ et $(\theta_n) \subset \Theta$ tels que pour $n \in \mathbb{N}^*$:

$$|\varphi(s_n; \theta_n) - \varphi(s_n; \theta_\infty(s_n))| \geq \delta \text{ et } L_\infty(s_n; \theta_n) - L_\infty(s_n; \theta_\infty(s_n)) \leq 1/n. \quad (4)$$

Comme $C \times \Theta$ est une partie compacte de $\mathbb{R}^d \times \mathbb{R}^p$, la suite (s_n, θ_n) admet une sous-suite (s_{k_n}, θ_{k_n}) convergeant vers un point (s^*, θ^*) de $C \times \Theta$. Alors la condition B₃) n'est pas vérifiée puisque l'application $\theta^* : C \rightarrow \Theta$ définie par $\theta^*(s) = \theta_\infty(s)$ si $s \neq s^*$ et $\theta^*(s^*) = \theta^*$ est une solution de (1) distincte de θ_∞ . En effet, d'après (4) et tenant compte de la continuité de L_∞ et de θ_∞ , nous avons

$$L_\infty(s^*; \theta^*) = \lim_{n \rightarrow +\infty} L_\infty(s_{k_n}; \theta_{k_n}) = L_\infty(s^*; \theta_\infty(s^*)), \text{ et}$$

$$|\varphi(s^*; \theta^*(s^*)) - \varphi(s^*; \theta_\infty(s^*))| = \lim_{n \rightarrow +\infty} |\varphi(s_{k_n}; \theta_{k_n}) - \varphi(s_{k_n}; \theta_\infty(s_{k_n}))| \geq \delta.$$

Afin d'établir l'implication réciproque, admettons qu'il existe $\theta^* : C \rightarrow \Theta$ solution de (1) et $s^* \in C$ tels que $\varphi(s^*; \theta^*(s^*)) \neq \varphi(s^*; \theta_\infty(s^*))$. D'après la condition (3) et pour $\delta = |\varphi(s^*; \theta^*(s^*)) - \varphi(s^*; \theta_\infty(s^*))|$, il existe $\epsilon > 0$ tel que

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

ou encore

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

ce qui permet de conclure. \square

Démonstration du Théorème 1: D'après le lemme précédent, on conclut que pour tout $\delta > 0$ il existe $\epsilon > 0$ tel que

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

En effet, si $\omega \in \Omega$ est tel que $\sup_{s \in C} |\widehat{\varphi}_n(s) - \varphi_\infty(s)| \geq \delta$, il existe $s_0 \in C$ tel que $|\varphi(s_0; \widehat{\theta}_n(s_0)) - \varphi(s_0; \theta_\infty(s_0))| \geq \delta$. Ainsi, d'après le Lemme 4 on déduit qu'il existe $\epsilon > 0$ indépendant de s_0 tel que $L_\infty(s_0; \widehat{\theta}_n(s_0)) - L_\infty(s_0; \theta_\infty(s_0)) > \epsilon$.

De plus, si pour tout $s \in C$ et $\theta \in \Theta$, $|L_n(s; \theta) - L_\infty(s; \theta)| \leq \epsilon/2$, 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 2: En tenant compte du Théorème 1 et l'hypothèse H_6), il nous suffit d'établir la convergence

$$\sup_{s \in C} \sup_{\theta \in \Theta} |L_n(s; \theta) - L_\infty(s; \theta)| \xrightarrow[n \rightarrow +\infty]{p.s.} 0, \quad (5)$$

où $L_n(s; \theta)$ est donné par (2) et $L_\infty(s; \theta) = E[\psi(Z_0; s; \theta) | S(Z_0) = s]f_S(s)$. Pour cela, considérons pour $s \in C$ et $\theta \in \Theta$ la décomposition

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

Chaque un des termes précédents est étudié dans la suite.

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

Comme $h_n \rightarrow 0, n \rightarrow +\infty$, 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

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

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

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

où la suite (M_n) est donnée dans l'hypothèse H₃).

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 H₁) et H₄) il existe des constantes $a_0, b_0, c_0 > 0$ telles que pour $i = 1, \dots, n$, et n assez grand on a

$$EX_{n,i} = 0, |X_{n,i}| \leq \frac{a_0 M_n}{h_n^d}, E|X_{n,i}| \leq b_0 \text{ et } E(X_{n,i})^2 \leq \frac{c_0 M_n^{2-\beta^*}}{h_n^d}. \quad (6)$$

En conséquence de l'hypothèse H₂) nous établirons dans la suite que pour tout $\epsilon > 0$ il existe $a, b > 0$ tels que pour n assez grand

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

On le fera en considérant, dans un premier temps, valable la condition ϕ) de l'hypothèse H₂), 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 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 $i = 1, \dots, n$ et $n \geq n_0$ avec $n_0 \in \mathbb{N}^*$ fixe.

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

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

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

En tenant compte des majorations (6) et du lemme précédent en choisissant $k = s_n$, on obtient pour tout $\epsilon > 0$, n assez grand et $s \in C$ et $\theta \in \Theta$

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

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 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

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

pour $i = 1, \dots, n$ 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

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

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

Les majorations (6) 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 $\theta \in \Theta$

$$P\left(\left|L'_n(s; \theta) - EL'_n(s; \theta)\right| > \epsilon\right) \leq 2 \exp\left(-\frac{n h_n^d}{M_n s_n} \left[\frac{\epsilon}{4a_0 e} - \frac{c_0}{4a_0^2 M_n^{\beta-1} s_n e}\right] + \frac{2n \sum_{i=1}^{s_n} \alpha(i)}{e s_n^2} + 2\sqrt{\epsilon} n (\alpha(s_n))^{2s_n/3n} / s_n\right).$$

L'hypothèse α) permet maintenant de conclure.

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

D'après l'inégalité de Bienaymé-Chebyshev et l'hypothèse H_1), on a, pour $\epsilon > 0$,

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

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

$$\frac{\Psi(Z_n)}{M_n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0.$$

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

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

et donc

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

D'autre part,

$$EL_n''(s; \theta) = \frac{1}{h_n^d} \int_{\mathbb{R}^d} K\left(\frac{s-z}{h_n}\right) E[\psi(Z_0; s; \theta) \mathbb{I}_{|\psi(Z_0; s; \theta)| > M_n} | S(Z_0) = z] f_S(z) dz,$$

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

$$\begin{aligned} |EL_n''(s; \theta)| &\leq \frac{1}{h_n^d} \int_{\mathbb{R}^d} \left| K\left(\frac{s-u}{h_n}\right) \right| E^{1/\beta^*} [|\Psi(Z_0)|^{\beta^*} | S(Z_0) = z] \\ &\quad \times E^{1-1/\beta^*} [\mathbb{I}_{|\Psi(Z_0)| > M_n} | S(Z_0) = z] f_S(z) dz \\ &\leq \left(\frac{1}{M_n}\right)^{\beta^*-1} \int_{\mathbb{R}^d} |K(u)| E[|\Psi(Z_0)|^{\beta^*} | S(Z_0) = s - uh_n] f_S(s - uh_n) du. \end{aligned}$$

D'après l'hypothèse H_4) et du fait que K est à support borné on a alors

$$\sup_{s \in C} \sup_{\theta \in \Theta} |EL_n''(s; \theta)| = O\left(\frac{1}{M_n}\right)^{\beta^*-1},$$

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 $\theta, \theta' \in \Theta$,

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

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

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

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

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

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

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

Comme $nh_n^d \rightarrow +\infty$, $n \rightarrow +\infty$, 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

$$\begin{aligned}
 & P\left(\max_{\substack{\ell=1, \dots, N \\ \ell'=1, \dots, N'}} |L'_n(s_\ell; \theta_{\ell'}) - EL'_n(s_\ell; \theta_{\ell'})| > \epsilon\right) \\
 &\leq \sum_{\ell=1}^N \sum_{\ell'=1}^{N'} P\left(|L'_n(s_\ell; \theta_{\ell'}) - EL'_n(s_\ell; \theta_{\ell'})| > \epsilon\right) \\
 &\leq aNN' \exp\left(-b \frac{nh_n^d}{M_n s_n}\right)
 \end{aligned}$$

D'après l'hypothèse $\lim_{n \rightarrow +\infty} \frac{nh_n^d}{M_n s_n \log n} = +\infty$, 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

$$\sum_{n=1}^{\infty} P\left(\max_{\substack{\ell=1, \dots, N \\ \ell'=1, \dots, N'}} |L'_n(s_\ell; \theta_{\ell'}) - EL'_n(s_\ell; \theta_{\ell'})| > \epsilon\right) < +\infty,$$

ce qui entraîne la convergence presque sûre souhaitée.

La convergence (5) découle maintenant des points précédents. \square

Démonstration du Théorème 3: Conséquence immédiate du Théorème 2 en faisant pour $i \in \mathbb{Z}$, $Z_i = (X_i, Y_i)$, $S(Z_i) = X_i$, et en considérant $\theta = (\alpha, \beta)$, $\psi(Z_i; x; \theta) = (Y_i - \alpha - \beta^T(X_i - x))^2$ et $\varphi(x; \theta) = \alpha$. \square

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). Prédiction 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 Déviations 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 (à paraître dans *In Honour of P. Balestra*, Elsevier).
- [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] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.*, **33**, 73-101.
- [21] Ioffe, M.O. et Katkovnik, V.Ya. (1989). Pointwise and uniform convergence with probability 1 of nonparametric regression estimators. *Automat. Remote Control*, **50**, 1659-1667.
- [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] Ruppert, D. et Wand, M.P. (1994). Multivariate locally weighted least squares regression. *Ann. Statist.*, **22**, 1346-1370.
- [29] Stone, C.J. (1977). Consistent nonparametric regression. *Ann. Statist.*, **5**, 595-645.
- [30] Tsybakov, A.B. (1986). Robust reconstruction of functions by the local-approximation method. *Problems of Information Transmission*, **22**, 133-146.
- [31] Watson, G.S. (1964). Smooth regression analysis. *Sankhya A*, **26**, 359-372.
- [32] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, **50**, 1-25.

Departamento de Matemática, Universidade de Coimbra

Apartado 3008, 3000 Coimbra, Portugal

e-mail: tenreiro@mat.uc.pt