Estimation non paramétrique pour des modèles de diffusion et de régression

Jean-Yves Brua

Estimation non paramétrique pour des modèles de diffusion et de régression

J.-Y. BRUA[1]

[1] *E-mail address:* brua@math.u-strasbg.fr

Remerciements

Je souhaite en premier lieu remercier vivement mon directeur de thèse Leonid Galtchouk d'avoir accept d'encadrer mes travaux. A son contact, j'ai pu tirer parti de ses connaissances, de son expérience et de ses nombreuses idées pour développer ce mémoire.

Je tiens également à remercier les Professeurs Armelle Guillou, Dominique Fourdrinier et Serge Pergamenchtchikov, qui m'ont fait l'honneur d'être les rapporteurs de ma thèse, pour leur intérêt porté à mon travail, pour le temps consacré à l'examiner ainsi que pour leurs remarques pertinentes.

J'adresse aussi tous mes remerciements aux membres des équipes de Statistique et de Probabilités de l'Université Louis Pasteur pour leur accueil et leur disponibilité.

Mes sincères remerciements vont à tous ceux qui ont partagé ma vie durant ces trois années, que ce soit à l'université ou ailleurs. Un grand merci à Florian, Audrey, Jürgen et les autres doctorants ou étudiants pour leur amitié et leur dévouement.

Je remercie enfin tout particulièrement ma mère Edith qui m'a toujours grandement soutenu tout au long de mes études, et mes grands-parents Jeanne et Joseph qui ont toujours cru en moi. Ce travail de thèse leur est dédié.

Table des matières

Chapitre 1

Introduction

1.1 Problématique

1.1.1 Description générale

La théorie de l'estimation, comme d'autres sujets en statistique, se développe au travers de problèmes d'ordre pratique en mathématique financière, économétrie, recherche médicale, etc. Bien qu'un traitement non asymptotique de beaucoup de ces problèmes soit important, cela ne suffit pas à recouvrir la théorie mathématique générale. Ainsi se sont développées des procédures de construction d'estimateurs qui ne sont pas nécessairement optimaux pour une distribution donnée ou pour une taille d'échantillon donnée, mais qui deviennent optimaux lorsque certains paramètres du problème tendent vers des valeurs limites (taille de l'échantillon croissant vers l'infini, intensité du bruit diminuant vers 0, etc.). De tels estimateurs sont appelés asymptotiquement efficaces. De nombreux problèmes d'efficacité asymptotique ont été étudiés ces trente dernières années aussi bien dans un cadre paramétrique que non paramétrique (voir par exemple Ibragimov et Has'minskiĭ (1981)). Nous nous sommes attachés dans cette thèse à montrer l'efficacité asymptotique de certains estimateurs à noyau pour des problèmes de régression et de diffusion non paramétriques.

Le modèle de régression non paramétrique considéré est le suivant. On dispose de n observations y_k régies par :

$$y_k = S(x_k) + g(x_k, S)\xi_k, \qquad k \in \{1, \ldots, n\}, \tag{1.1}$$

où S est la fonction inconnue à estimer à partir des observations $(y_k)_{1\leqslant k\leqslant n}$ et $g : [0;1] \times C^1([0;1], \mathbb{R}) \longrightarrow \mathbb{R}_+^*$ une fonction inconnue représentant l'écart-type du bruit. Les variables aléatoires $(\xi_k)_{1\leqslant k\leqslant n}$ sont indépendantes, centrées et identiquement distribuées et de variance 1. Ce modèle est à pas fixe car nous supposerons que $x_k = k/n$ pour tout $k = 1, \ldots, n$.

Dans un deuxième temps, nous étudierons le modèle de diffusion non paramétrique suivant :

$$dX_t = S(X_t)dt + dB_t, \qquad 0 \leqslant t \leqslant T \tag{1.2}$$

où $(B_t)_{t\geqslant 0}$ est un mouvement brownien unidimensionnel et S est toujours la fonction inconnue à estimer à partir des observations $(X_t, 0 \leqslant t \leqslant T)$.

Pour chacun des modèles ci-dessus, on se propose d'estimer la fonction inconnue S en un point fixe en supposant qu'elle appartient à une classe Hölderienne. Pour définir le risque associé à l'emploi d'un estimateur et ainsi mesurer la qualité de celui-ci, on utilise la fonction de perte liée à l'erreur absolue. Enfin, dans l'optique de l'efficacité asymptotique, on suit l'approche minimax décrite ci-après.

1.1.2 Approche minimax

Avant de décrire cette approche, donnons la définition d'un estimateur pour les modèles considérés.

Définition 1.1.1 *Pour le modèle de régression (1.1), un estimateur de S au point z_0 est une variable aléatoire $\omega \mapsto \tilde{S}_n(z_0) = \tilde{S}_n(z_0, y_1, \ldots, y_n)$ mesurable par rapport à la tribu engendrée par $y_1, \ldots, y_n$.*
Pour le modèle de diffusion (1.2), un estimateur de S au point z_0 est une variable aléatoire $\omega \mapsto \tilde{S}_T(z_0) = \tilde{S}_T(z_0, \mathbf{x})$ mesurable par rapport à la tribu engendrée par $\mathbf{x} = (X_t, 0 \leqslant t \leqslant T)$.

Etant données $w : \mathbb{R}^+ \to \mathbb{R}^+$ une fonction de perte croissante, non identiquement nulle et nulle en zéro, et d une semi-distance, on définit le risque d'un estimateur $\tilde{S}$ d'une fonction S appartenant à la classe fonctionnelle $\mathcal{H}$ par $\mathbb{E}_S w(d(\tilde{S}, S))$, où $\mathbb{E}_S$ désigne l'espérance quand l'aléa est déterminé par le modèle (1.1) ou (1.2). Quelques exemples de fonction de perte sont :

$$w(u) = u^p,\ p > 0, \quad \text{ou} \quad w(u) = \mathbb{I}_{u \geqslant A},\ A > 0,$$

et de semi-distance :

$$d(f,g) = |f(x_0) - g(x_0)| \text{ pour un } x_0 \text{ fixé },\quad d(f,g) = \|f-g\|_2, \quad \text{ou} \quad d(f,g) = \|f-g\|_\infty.$$

Le risque maximal d'un estimateur $\tilde{S}$ sur une classe fonctionnelle $\mathcal{H}$ est donné par

$$\mathcal{R}(\tilde{S}) := \sup_{S \in \mathcal{H}} \mathbb{E}_S w(d(\tilde{S}, S)),$$

le risque minimax sur la classe $\mathcal{H}$ par

$$\inf_{\tilde{S}} \mathcal{R}(\tilde{S}) = \inf_{\tilde{S}} \sup_{S \in \mathcal{H}} \mathbb{E}_S w(d(\tilde{S}, S)),$$

l'infimum étant pris sur tous les estimateurs. Ce dernier sera noté $\mathcal{R}_n^*$ ou $\mathcal{R}_T^*$ selon les modèles considérés.

L'objectif premier de l'approche minimax est de trouver un estimateur $\hat{S}$ qui minimise le risque minimax. Un tel estimateur est dit minimax. Un estimateur $\hat{S}$ est dit asymptotiquement minimax si

$$\mathcal{R}(\hat{S}) \sim \inf_{\tilde{S}} \mathcal{R}(\tilde{S}),$$

lorsque la taille de l'échantillon n ou la durée d'observation T tend vers l'infini.

Plusieurs travaux ont conduit à l'existence d'estimateurs minimax ou à certains critères suffisant à un estimateur pour l'être (cf. Ibragimov et Has'minskiĭ (1981)) pour des modèles paramétriques. Par exemple il est connu que la moyenne des observations est minimax pour l'estimation de la moyenne non bornée d'une distribution normale réduite en dimension 1 et 2. De plus, pour le modèle de régression paramétrique

$$\begin{cases} y_{i1} = \theta + \varepsilon_{i1} \\ y_{i2} = \theta + r + \varepsilon_{i2}, \end{cases}$$

où $i = 1, \ldots, n$, $\theta \in \mathbb{R}$, $|r| \leqslant M$ et les ε_{ij} sont indépendantes normales centrées réduites, Sacks et Ylvisaker (1978) ont donné l'estimateur minimax parmi tous les estimateurs linéaires de θ. Mais d'après Sacks et Strawderman (1982), cet estimateur n'est pas minimax. Il est toutefois asymptotiquement minimax lorsque $M = \infty$.

Dans le cadre non paramétrique de l'estimation en un point fixe d'une fonction de régression lipschitzienne, Sacks et Ylvisaker (1978) ont fourni un estimateur minimax parmi tous les estimateurs linéaires, mais qui se révèle ni minimax, ni asymptotiquement minimax (cf. Sacks et Strawderman (1982)). Par ailleurs, pour l'estimation d'une densité quasi Höldérienne en un point fixe avec la perte quadratique, Sacks et Ylvisaker (1981) ont exhibé une suite d'estimateurs à noyau asymptotiquement minimax parmi les estimateurs à noyau. Puis Donoho et Liu (1991) ont montré que cet estimateur est asymptotiquement minimax parmi les estimateurs affines et le rapport du risque maximal de cet estimateur par le risque minimax est asymptotiquement majoré par 5/4.

On est donc amené à s'intéresser au comportement asymptotique du risque minimax. Considérons alors le cadre plus général pour lequel le risque maximal d'un estimateur $\tilde{S}_n$ (respectivement $\tilde{S}_T$) est défini par

$$\mathcal{R}(\tilde{S}_n) := \sup_{S \in \mathcal{H}} \mathbb{E}_S w(\varphi_n^{-1} d(\tilde{S}_n, S)) \quad \left(\textit{respectivement } \mathcal{R}(\tilde{S}_T) := \sup_{S \in \mathcal{H}} \mathbb{E}_S w(\varphi_T^{-1} d(\tilde{S}_T, S)) \right),$$

où la famille $(\varphi_n)_{n \in \mathbb{N}^*}$ (respectivement $(\varphi_T)_{T \in \mathbb{R}^*}$) est composée de réels strictement positifs. Le but de l'approche est ainsi de trouver de telles familles et des constantes $c > 0$ et $C < \infty$ telles que

$$\limsup_{n \to \infty} \mathcal{R}_n^* \leqslant C \quad \text{et} \quad \liminf_{n \to \infty} \mathcal{R}_n^* \geqslant c \tag{1.3}$$

$$(\text{respectivement } \limsup_{T \to \infty} \mathcal{R}_T^* \leqslant C \quad \text{et} \quad \liminf_{T \to \infty} \mathcal{R}_T^* \geqslant c). \tag{1.4}$$

Définition 1.1.2 *La famille $(\varphi_n)_{n \in \mathbb{N}^*}$ (respectivement $(\varphi_T)_{T \in \mathbb{R}^*}$) est dite vitesse de convergence minimax des estimateurs sur $\mathcal{H}$ si (1.3) (respectivement si (1.4)) est vérifiée.*

Définition 1.1.3 *Un estimateur S_n^* (respectivement S_T^*) vérifiant $c \leqslant \mathcal{R}(S_n^*) \leqslant C'$ (respectivement $c \leqslant \mathcal{R}(S_T^*) \leqslant C'$), où $(\varphi_n)_{n\in\mathbb{N}^*}$ (respectivement $(\varphi_T)_{T\in\mathbb{R}_+^*}$) est la vitesse de convergence minimax et $c > 0$ et $C' < \infty$ sont des constantes, est dit estimateur optimal en vitesse de convergence sur $\mathcal{H}$.*

Définition 1.1.4 *Un estimateur S_n^* (respectivement S_T^*) est dit asymptotiquement efficace sur $\mathcal{H}$ lorsque*

$$\lim_{n\to\infty} \frac{\mathcal{R}(S_n^*)}{\mathcal{R}_n^*} = 1 \quad (\textit{respectivement} \quad \lim_{n\to\infty} \frac{\mathcal{R}(S_T^*)}{\mathcal{R}_T^*} = 1).$$

Remarque 1.1.5 *Pour montrer qu'un estimateur est asymptotiquement efficace, il suffira d'obtenir une borne inférieure et une borne supérieure égales ($C' = c$ dans la Définition 1.1.3).*

1.1.3 Approche minimax adaptative

Cette approche est utilisée lorsqu'un des paramètres définissant la classe fonctionnelle $\mathcal{H}$ considérée est supposé inconnu, par exemple la régularité de la fonction de régression S dans le modèle (1.1). Nous ne traiterons dans cette thèse que le cas adaptatif pour ce modèle. Notons alors $\mathcal{H}(\beta)$ la classe fonctionnelle, où $\beta \in \mathcal{B}$, $\mathcal{B}$ étant un ensemble quelconque et

$$\mathcal{R}_\beta(\tilde{S}_n, \psi_n(\beta)) = \sup_{S\in\mathcal{H}(\beta)} \mathbb{E}_S w(\psi_n^{-1}(\beta) d(\tilde{S}_n, S)),$$

avec $\tilde{S}_n$ un estimateur et $(\psi_n(\beta))_{n\in\mathbb{N}^*}$ une suite de réels strictement positifs tendant vers 0.

Disposant de la vitesse de convergence minimax $\varphi_n(\beta)$ sur la classe $\mathcal{H}(\beta)$, on se demande premièrement s'il existe un estimateur optimal adaptatif en vitesse de convergence, c'est-à-dire un estimateur indépendant de $\beta \in \mathcal{B}$ qui converge à la vitesse $\varphi_n(\beta)$ sur chaque classe $\mathcal{H}(\beta)$. Plus précisément :

Définition 1.1.6 *Un estimateur S_n^*, indépendant de $\beta \in \mathcal{B}$, est dit optimal adaptatif en vitesse de convergence sur la famille $\{\mathcal{H}(\beta)\}_{\beta\in\mathcal{B}}$ pour la semi-distance d s'il existe une constante $C > 0$ telle que :*

$$\limsup_{n\to\infty} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(S_n^*, \varphi_n(\beta)) \leqslant C.$$

Puis, comme pour l'approche minimax précédente, on cherche un estimateur adaptatif (toujours indépendant de $\beta \in \mathcal{B}$) asymptotiquement exact, de même que la borne asymptotique exacte du risque minimax adaptatif

$$\inf_{\tilde{S}_n} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(\tilde{S}_n, \varphi_n(\beta)).$$

Définition 1.1.7 *Un estimateur S_n^* optimal adaptatif en vitesse de convergence est appelé adaptatif asymptotiquement exact sur la famille $\{\mathcal{H}(\beta)\}_{\beta\in\mathcal{B}}$ pour la semi-distance d s'il vérifie :*

$$\lim_{n\to\infty} \inf_{\tilde{S}_n} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(\tilde{S}_n, \varphi_n(\beta)) = \lim_{n\to\infty} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(S_n^*, \varphi_n(\beta)).$$

Cependant, des estimateurs optimaux adaptatifs en vitesse de convergence n'existent pas toujours. En effet, Lepskiĭ (1990) montre qu'il n'en existe pas pour l'estimation en un point fixe, dans un modèle de bruit blanc gaussien, d'une fonction Höldérienne appartenant à la classe $\Sigma(L,\beta)$, $\beta \in \mathcal{B} \subset \mathbb{R}_+^*$ définie en (1.5) et $\mathcal{B}$ contenant au moins deux éléments. Néanmoins, il se peut qu'on ait une relation du type

$$\limsup_{n\to\infty} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(S_n^*, \psi_n(\beta)) \leqslant C,$$

pour un certain estimateur S_n^*, alors que $\psi_n(\beta)$ n'est pas la vitesse de convergence minimax sur $\mathcal{H}(\beta)$.

Définition 1.1.8 *La famille $(\psi_n(\beta))_{n\in\mathbb{N}^*}$ est dite vitesse de convergence minimax adaptative des estimateurs sur la famille de classes $(\mathcal{H}(\beta))_{\beta\in\mathcal{B}}$ si*

- *pour un certain estimateur S_n^* et une constante $C > 0$, on a :*

$$\limsup_{n\to\infty} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(S_n^*, \psi_n(\beta)) \leqslant C;$$

- *il existe une constante $c > 0$ telle que :*

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \sup_{\beta\in\mathcal{B}} \mathcal{R}_\beta(\tilde{S}_n, \psi_n(\beta)) > c.$$

Un estimateur S_n^* vérifiant le premier point précédent, avec $\psi_n(\beta)$ la vitesse de convergence minimax adaptative est dit adaptatif en vitesse de convergence sur la famille $(\mathcal{H}(\beta))_{\beta\in\mathcal{B}}$.

Remarque 1.1.9 *Un problème plus délicat est la recherche d'un estimateur adaptatif en vitesse de convergence pour lequel les bornes inférieure et supérieure asymptotiques du risque coïncident.*

1.2 Etat de l'art

1.2.1 Cas non adaptatif

Durant les trente dernières années, l'approche minimax s'est développée de manière fructueuse. On peut distinguer quatre champs permettant de répertorier les résultats :

- le modèle d'observations (régression, densité, diffusion...),
- le type de risque (en un point fixe, global, local...),
- sa forme (L^1, L^2, L^∞...),

– la classe de fonctions (de Hölder, de Sobolev, de Lipschitz...).

C'est ainsi que la combinaison de ces champs a conduit à l'exploration de nombreux cas par l'approche minimax. En particulier les classes de Hölder sont largement étudiées.

Définition 1.2.1 *Soient $L > 0$ et $\beta > 0$. La classe de Hölder $\Sigma(L, \beta)$ est définie par*

$$\Sigma(L,\beta) = \left\{ S : \mathbb{R} \to \mathbb{R} : |S^{(m)}(x) - f^{(m)}(y)| \leqslant L|x-y|^{\beta - m}, \forall x, y \in \mathbb{R} \right\}, \tag{1.5}$$

où $m = \lfloor \beta \rfloor$ désigne le plus grand entier strictement plus petit que le réel β.

Les vitesses de convergence minimax ont été trouvées dans de nombreux cas. Pour l'estimation en norme L^p $(p > 1)$ d'une densité appartenant à un espace de type Sobolev, Bretagnolle et Huber (1979) ont donné un estimateur à noyau optimal en vitesse de convergence. De plus, dans le cadre de l'estimation d'une densité sur la classe de Hölder $\Sigma(L, \beta)$, la vitesse de convergence $n^{\beta/(2\beta+1)}$ est atteinte par un estimateur à noyau pour toute une famille de fonctions de perte (cf. Ibragimov et Has'minskiĭ (1981)). D'autres modèles de densité ont été étudiés, parmi lesquels on peut citer ceux traitant de densités appartenant à une classe de fonctions analytiques (cf. Ibragimov et Has'minskiĭ (1982b) et Golubev et Levit (1996)).

Remarquons toutefois que d'après Nemirovskiĭ et al. (1985) et Nemirovskiĭ (1986), aucun estimateur linéaire ne peut être optimal en vitesse de convergence pour l'estimation de fonctions dans certaines classes (fonctions décroissantes, espaces de Sobolev) et certaines fonctions de perte (par exemple en norme L^p, $p > 1$).

Pour le modèle de régression sur $\Sigma(L, \beta)$, les vitesses de convergence minimax $n^{\beta/(2\beta+1)}$ pour l'estimation en norme L^p $(p \in [1, \infty[)$ et $(n/\ln n)^{\beta/(2\beta+1)}$ pour l'estimation en norme L^∞ sont données par Ibragimov et Has'minskiĭ (1982a) et Stone (1982). Lorsque la fonction de régression est q fois continûment différentiable et pour le risque L^2, cette vitesse est $n^{q/(2q+1)}$ (cf. Stone (1982)). Pour des fonctions de régression appartenant à des espaces de Sobolev de régularité β, estimées avec une perte quadratique, Speckman (1985) a obtenu un estimateur optimal parmi tous les estimateurs linéaires avec la vitesse de convergence $n^{\beta/(2\beta+1)}$. Et par une procédure basée sur la sélection de modèles, Fourdrinier et Pergamenshchikov (2007) donnent des estimateurs optimaux en vitesse de convergence pour un problème semblable, en supposant les erreurs dépendantes, de loi radiale.

Les vitesses de convergence sont également calculées dans des modèles de bruit blanc gaussien pour l'estimation d'un signal Höldérien ou analytique en norme L^p $(p \in [1, \infty[)$ et en norme L^∞ par Ibragimov et Has'minskiĭ (1980, 1981, 1984).

Dans le modèle de diffusion (1.2), la vitesse de convergence minimax a été obtenue dans Galtchouk et Pergamenshchikov (2004) pour l'estimation du coefficient de dérive en norme L^2 et dans Galtchouk et Pergamenshchikov (2005b) pour l'estimation en un point fixe.

En ce qui concerne la recherche de la borne asymptotique inférieure exacte du risque minimax et d'estimateurs asymptotiquement efficaces, si Ibragimov et Has'minskiĭ (1981) en donnent pour des modèles paramétriques, Pinsker (1980) est le premier à fournir la constante exacte et un estimateur linéaire asymptotiquement efficace pour l'estimation en norme L^2 d'un signal appartenant à un ellipsoïde de Sobolev de $L^2([0;1])$ dans un modèle de bruit blanc gaussien. Efroĭmovich et Pinsker (1981, 1982) ont ensuite prolongé ces résultats pour l'estimation de densités et de densités spectrales quand le nombre d'observations tend vers l'infini.

Pour un premier problème de régression non paramétrique, Golubev (1984) a obtenu un estimateur asymptotiquement efficace pour le risque L^2 lorsque l'intervalle sur lequel les observations sont effectuées grandit. Nussbaum (1985) s'est quant à lui inspiré de la méthode de Pinsker (1980) pour trouver un estimateur linéaire asymptotiquement efficace et la borne asymptotique exacte du risque minimax pour l'estimation d'une fonction de régression appartenant à un espace de Sobolev (voir aussi Golubev et Nussbaum (1990a,b); Efroĭmovich (1996)).

Le deuxième cas pour lequel le comportement asymptotique exact du risque minimax a été découvert est l'estimation de fonctions Höldériennes avec le risque L^∞. En effet, Korostelev (1993) fournit la borne asymptotique exacte du risque minimax ainsi qu'un estimateur asymptotiquement efficace d'une fonction de régression appartenant à $\Sigma(L,\beta)$, $\beta \in]0;1]$. Par la suite, toujours pour l'estimation d'une fonction de $\Sigma(L,\beta)$, $\beta > 0$, ou de ses dérivées en norme L^∞, Donoho (1994a) dans un modèle de bruit blanc gaussien et Korostelev et Nussbaum (1999) dans un modèle de densité obtiennent des résultats similaires. En s'intéressant à l'estimation d'une fonction de régression Höldérienne de régularité $\beta \in]1;2[$ avec le risque lié à la fonction de perte absolue, Galtchouk et Pergamenshchikov (2006a) ont établi l'efficacité asymptotique d'un estimateur à noyau et la borne asymptotique exacte du risque minimax sur une classe Höldérienne plus faible, la vitesse de convergence optimale étant $n^{\beta/(2\beta+1)}$.

Enfin, le troisième exemple de comportement asymptotique exact du risque minimax provient de l'estimation de fonctions de régression analytiques (cf. Golubev et al. (1996)) ou d'une densité analytique (cf. Golubev et Levit (1996)) avec le risque L^∞. Ces résultats ont été étendus par Guerre et Tsybakov (1998) au modèle de bruit blanc gaussien avec le risque L^p, $p \in [1;\infty[$.

Des résultats concernant l'efficacité asymptotique d'estimateurs pour un modèle de diffusion sont plus récents. Lorsque la dérive est k fois dérivable, la densité ergodique est estimée avec la vitesse $\sqrt{T}$ (cf. Kutoyants (1998)) tandis que sa dérivée l'est avec la vitesse $T^{k/(2k+1)}$ (cf. Dalayan et Kutoyants (2002); Dalalyan et Kutoyants (2003)). Plusieurs estimateurs asymptotiquement efficaces sont construits dans les trois dernières références citées. Par ailleurs, puisque la dérive est le rapport de la dérivée de la densité ergodique sur deux fois celle-ci, ces résultats en amènent à d'autres sur l'estimation de la dérive. Plus précisément, Dalayan et Kutoyants (2002) donnent une borne inférieure asymptotique d'un risque minimax local de type L^2 pondéré (la fonction de pondération étant la densité ergodique inconnue) et un estimateur asymptotiquement efficace avec la vitesse $T^{k/(2k+1)}$. En utilisant des inégalités de concentration pour la densité ergodique, sans toutefois donner d'estimateur asymptotiquement efficace de celle-ci ou de sa dérivée, on peut construire une procédure séquentielle aboutissant aussi à l'efficacité asymptotique d'estimateurs de la dérive lorsque le risque est pris localement sur des voisinages centrés sur une fonction Höldérienne (cf. Galtchouk et Pergamenshchikov (2005b)) ou lipschitzienne (cf. Galtchouk et Pergamenshchikov (2006b)).

1.2.2 Cas adaptatif

De nombreux travaux ont été consacrés à la recherche de la vitesse optimale de convergence ou d'un estimateur asymptotiquement efficace lorsqu'un ou plusieurs paramètres du modèle sont supposés inconnus, en particulier la régularité de la fonction à estimer. Ce cas, dit adaptatif, a engendré des premiers résultats sur la vitesse de convergence minimax adaptative comme dans Efroĭmovich et Pinsker (1984) pour un modèle de bruit blanc gaussien, Härdle et Marron (1985) pour un modèle de régression et Efroĭmovich (1985) pour l'estimation d'une densité.

Un estimateur optimal adaptatif en vitesse de convergence n'existe pas nécessairement. Lepskiĭ (1990) et Brown et Low (1996) ont prouvé qu'il n'en existe pas pour des classes Höldériennes quand le signal dans un modèle de bruit blanc gaussien est estimé en un point fixe. De plus, pour des classes de Sobolev, Tsybakov (1998) a obtenu ce résultat dans le même contexte et Butucea (2000a,b) dans le cadre de l'estimation d'une densité en un point fixe.

Par contre, pour certaines classes de Hölder, de Sobolev ou de Besov avec des risques de type L^p, on peut construire des estimateurs optimaux adaptatif en vitesse de convergence par différentes méthodes : à partir de noyaux, de splines, de polynômes par morceaux ou d'ondelettes (cf. Lepskiĭ (1991, 1992a); Donoho et al. (1995) pour des modèles de bruit blanc gaussien, et Lepskiĭ et al. (1997); Goldenshluger et Nemirovski (1997); Juditsky (1997) pour des modèles de régression). Certains travaux ont même conduit à la découverte d'estimateurs adaptatifs asymptotiquement exacts : par exemple pour l'estimation de fonctions appartenant à des classes de Sobolev, on peut citer Efroĭmovich et Pinsker (1984) (norme L^2) et Tsybakov (1998) (norme L^∞) pour des modèles de bruit blanc gaussien, Golubev et Nussbaum (1992) (norme L^2) pour l'estimation d'une fonction de régression, ou encore Dalalyan (2005) pour l'estimation de la dérive d'une diffusion ergodique avec un risque local.

A l'instar du résultat de Tsybakov (1998) pour l'estimation en un point fixe, on constate parfois que la vitesse de convergence minimax n'est pas optimale (au sens adaptatif). On cherche alors la vitesse de convergence minimax adaptative ainsi qu'un estimateur adaptatif en vitesse de convergence. Pour l'estimation en un point fixe d'une densité appartenant à une classe de Sobolev, le problème est traité dans Butucea (2000a,b). Tsybakov (1998) parvient même à obtenir la limite exacte du risque minimax adaptatif normalisé par la vitesse de convergence minimax adaptative (les bornes inférieure et supérieure de ce risque se révélant identiques). Il en est de même pour Lepskiĭ et Spokoiny (1997) dans un modèle de bruit blanc gaussien lorsque le signal à estimer (en un point) appartient à $\Sigma(\beta, L)$, les paramètres β et L étant supposés tous deux inconnus. Enfin, Galtchouk et Pergamenshchikov (2001) obtiennent la vitesse de convergence minimax adaptative ainsi qu'un estimateur adaptatif en vitesse de convergence de la dérive d'une diffusion, appartenant à une classe Höldérienne.

1.3 Description des résultats obtenus

1.3.1 Modèles de régression

Cas non adaptatif

On considère le modèle de régression non paramétrique (1.1), la fonction de régression étant à estimer en un point fixe x_0. Les résultats du Chapitre 2 sont dans le prolongement de ceux obtenus par Galtchouk et Pergamenshchikov (2006a) et font l'objet d'un article (Brua (2008a)).

Le problème de l'estimation asymptotique efficace, avec la perte liée à l'erreur absolue, pour lequel la fonction de régression S appartient à la classe

$$\mathcal{H}(\beta) = \bigcup_{M>0,K>0} \mathcal{H}(M, K, \beta),$$

où $\mathcal{H}(M, K, \beta)$ est la classe de Hölder de régularité $\beta = 1 + \alpha, \quad \alpha \in]0; 1[$ définie par

$$\mathcal{H}(M, K, \beta) = \left\{ S \in C^1([0; 1]) \ : \ \|S'\|_\infty \leqslant M, \sup_{x\in[0;1]} \frac{|S'(x) - S'(x_0)|}{|x - x_0|^\alpha} \leqslant K \right\},$$

reste ouvert. Pour notre part, on étudie le risque minimax pris sur une classe plus large qu'une classe de Hölder, appelée classe Höldérienne faible au point x_0 et définie, pour $\delta \in]0; 1[$ par

$$\mathcal{U}_{x_0,\delta,n} = \left\{ S \in C^1([0; 1]) \ : \ \|S'\|_\infty \leqslant \delta^{-1}; \left| \int_{-1}^{1} (S(x_0 + h_n u) - S(x_0))\, du \right| \leqslant \delta h_n^\beta \right\},$$

où le paramètre β est supposé connu et $h_n = n^{-1/(2\beta+1)}$, n étant le nombre d'observations dans notre modèle de régression (1.1).
Remarquons que si $S \in \mathcal{H}(M, K, \beta)$, avec $M < \delta^{-1}$ et $2K/(\beta(\beta+1)) < \delta$, alors $S \in \mathcal{U}_{x_0,\delta,n}$. Mais une classe Höldérienne faible contient aussi des fonctions qui n'appartiennent pas à une classe de Hölder.

Le risque d'un estimateur $\tilde{S}_n$ de $S(x_0)$ est défini par

$$\mathcal{R}_{x_0,\delta,n}(\tilde{S}_n) = \sup_{S\in\mathcal{U}_{x_0,\delta,n}} \mathbb{E}_S \varphi_n \frac{|\tilde{S}_n - S(x_0)|}{g(x_0, S)},$$

où $\mathbb{E}_S$ est l'espérance calculée par rapport à la loi $\mathbb{P}_S$ quand S est la fonction de régression dans le modèle (1.1).

Dans le cas d'un bruit gaussien homogène, c'est-à-dire quand $\xi_k \sim \mathcal{N}(0,1)$ et $g(x,S) = \sigma > 0$ pour tous $x \in [0;1]$ et $S \in C^1([0;1])$, Galtchouk et Pergamenshchikov (2006a) ont obtenu la borne asymptotique inférieure exacte du risque minimax ainsi qu'un estimateur asymptotiquement efficace.

Théorème 1.3.1 *(Galtchouk et Pergamenshchikov, 2006a)*
Pour tout $\delta \in]0;1[$, on a :

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \mathcal{R}_{x_0,\delta,n}(\tilde{S}_n) \geqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0,1/2).$$

Théorème 1.3.2 *(Galtchouk et Pergamenshchikov, 2006a)*
Soit $Q = \mathbb{I}_{[-1;1]}$. Alors l'estimateur à noyau Q défini par

$$\hat{S}_n = \left(\sum_{k=1}^n Q\left(\frac{x_k - x_0}{h_n}\right)\right)^{-1} \sum_{k=1}^n Q\left(\frac{x_k - x_0}{h_n}\right) y_k \tag{1.6}$$

est asymptotiquement efficace car il vérifie la relation

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \mathcal{R}_{x_0,\delta,n}(\hat{S}_n) \leqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0,1/2).$$

Précisons que la borne inférieure a été obtenue en considérant la famille $\tilde{\Sigma}_n$ composée des fonctions

$$S_\nu(x) = \varphi_n^{-1} V_\nu\left(\frac{x - x_0}{h_n}\right),$$

où

$$\begin{aligned} V_\nu(x) &= \nu^{-1}\int_{-\infty}^{\infty} \tilde{Q}_\nu(u)\, l\left(\frac{u-x}{\nu}\right) du, \\ \tilde{Q}_\nu(u) &= \mathbb{I}_{\{|u|\leqslant 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\leqslant|u|\leqslant 1-\nu\}}, \\ l(z) &= c\exp\left(-(1-z^2)^{-1}\right)\mathbb{I}_{\{|z|\leqslant 1\}}, \end{aligned}$$

avec $0 < \nu < 1/4$ et c la constante de normalisation telle que $\displaystyle\int_{-1}^{1} l(z)dz = 1$.
La constante de Hölder des fonctions V_ν est de l'ordre de ν^{-2} quand le paramètre ν tend vers 0. Contrairement à précédemment, il n'existe donc pas de classe de Hölder $\mathcal{H}(M,K,\beta)$ contenant toute la famille $\tilde{\Sigma}_n$.

Remarquons que l'idée principale pour obtenir la borne inférieure du risque minimax est de remplacer le supremum sur toute la classe fonctionnelle par celui pris sur une sous-classe paramétrique (impliquant ici la famille $\tilde{\Sigma}_n$) puis d'appliquer l'inégalité de Hajek-Le Cam (cf. Ibragimov et Has'minskiĭ (1981)). Cette démarche nécessite dans Galtchouk et Pergamenshchikov (2006a) de faire tendre la constante de Hölder des fonctions considérées vers l'infini (quand $\nu \to 0$).

La constante δ majorant l'expression $\varphi_n \int_{-1}^{1} (S(x_0 + uh_n) - S(x_0))\, du$ dans la définition de $\mathcal{U}_{x_0,\delta,n}$, appelée constante Höldérienne faible, est amenée à tendre vers 0. Cette propriété nous permet d'atteindre la borne supérieure exacte avec un estimateur à noyau. Notons enfin que, tout comme dans le cas paramétrique, la constante exacte obtenue est égale à l'espérance de la fonction de perte prise en une variable aléatoire gaussienne.

Par ailleurs la procédure décrite ci-dessus est robuste par rapport au bruit. En effet, notons $\mathcal{P}_{\epsilon,L}$ l'ensemble des lois de probabilité de moyenne nulle et de variance 1 et telles que $\mathbb{E}|\xi|^{2+\epsilon} \leqslant L$ si ξ suit cette loi (avec L suffisamment grand pour que la loi normale standard y figure). On suppose que les variables aléatoires i.i.d. (ξ_k) du modèle (1.1) suivent une loi appartenant à $\mathcal{P}_{\epsilon,L}$. On définit alors le risque d'un estimateur $\tilde{S}_n$ par

$$\tilde{\mathcal{R}}_{x_0,\delta,n}(\tilde{S}_n) = \sup_{\mathcal{P}_{\epsilon,L}} \sup_{S \in \mathcal{U}_{x_0,\delta,n}} \mathbb{E}_S \varphi_n \frac{|\tilde{S}_n - S(x_0)|}{g(z_0, S)}.$$

Toujours dans le cas d'un bruit homogène ($g \equiv \sigma$), la borne inférieure du risque minimax correspondant se déduit immédiatement du Théorème 1.3.1. Pour tout $\delta \in]0;1[$, on a :

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \tilde{\mathcal{R}}_{x_0,\delta,n}(\tilde{S}_n) \geqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Finalement, Galtchouk et Pergamenshchikov (2006a) montrent que la borne supérieure du risque maximal de l'estimateur (1.6) est identique :

Théorème 1.3.3 *(Galtchouk et Pergamenshchikov, 2006a)*

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \tilde{\mathcal{R}}_{x_0,\delta,n}(\hat{S}_n) \leqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Détaillons maintenant les résultats obtenus pour un bruit hétéroscedastique. Nous avons tout d'abord besoin de quelques hypothèses. On suppose que la fonction g est uniformément bornée, plus précisément qu'il existe deux constantes $g_* > 0$ et $g^* < \infty$ telles que :

$$g_* \leqslant \inf_{x\in[0;1]} \inf_{S\in C^1([0;1])} g(x,S) \leqslant \sup_{x\in[0;1]} \sup_{S\in C^1([0;1])} g(x,S) \leqslant g^*. \tag{1.7}$$

On suppose de plus que la fonction g est différentiable au sens de Fréchet par rapport à la variable S, uniformément sur $[0;1]$, c'est-à-dire pour tous $S, S_0 \in C^1([0;1])$ et $x \in [0;1]$,

$$g(x,S) = g(x,S_0) + L_{x,S_0}(S - S_0) + \Gamma_{x,S_0}(S - S_0), \tag{1.8}$$

où l'application linéaire L_{x,S_0} est bornée sur $C^1([0;1])$ uniformément sur $[0;1]$, i.e. il existe une constante strictement positive C_{S_0} telle que

$$\sup_{x\in[0;1]} \sup_{S\in C^1([0;1]), \|S\|\neq 0} |L_{x,S_0}(S)|/\|S\| \leqslant C_{S_0} \tag{1.9}$$

et le terme résiduel $\Gamma_{x,S_0}(S)$ vérifie

$$\lim_{\|S\|\to 0} \sup_{x\in[0;1]} \Gamma_{x,S_0}(S)/\|S\| = 0. \tag{1.10}$$

Enfin on supposera que

$$\lim_{n\to\infty} \sup_{S\in\mathcal{U}_{x_0,\delta,n}} \left| \left(\frac{1}{q_n} \sum_{k=1}^{n} Q\left(\frac{x_k - x_0}{h_n}\right) g^2(x_k, S) \right)^{1/2} - g(x_0, S) \right| = 0, \tag{1.11}$$

où $q_n = \sum_{k=1}^{n} Q\left(\frac{x_k - x_0}{h_n}\right)$.

Remarque 1.3.4 *1. L'hypothèse (1.11) est vérifiée lorsque pour tout $\varepsilon > 0$, il existe $\iota > 0$ tel que si $|x - x_0| \leqslant \iota$, alors $\sup_{S\in C^1([0;1])} |g(x,S) - g(x_0,S)| \leqslant \varepsilon$.*

2. En particulier, la fonction g satisfait à cette propriété si elle est uniformément continue par rapport à ses deux variables.

3. Dans ce cas, les résultats suivants restent valables sans avoir recours aux hypothèses (1.7–1.10).

Les théorèmes suivants seront démontrés au Chapitre 2.

Théorème 1.3.5 *Si les variables aléatoires (ξ_k) sont gaussiennes standard, alors pour tout $\delta \in]0;1[$, on a :*

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \mathcal{R}_{x_0,\delta,n}(\tilde{S}_n) \geqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Théorème 1.3.6 *Si les variables aléatoires (ξ_k) sont gaussiennes standard, alors l'estimateur (1.6) est asymptotiquement efficace puisqu'il vérifie*

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \mathcal{R}_{x_0,\delta,n}(\hat{S}_n) \leqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Toujours sous les hypothèses (1.7–1.11) et pour un bruit hétéroscedastique, on peut également supposer que les lois des variables aléatoires (ξ_k) appartiennent à $\mathcal{P}_{\epsilon,L}$. Dans ce cas, la borne inférieure asymptotique du risque minimax $\inf_{\tilde{S}_n} \tilde{\mathcal{R}}_{x_0,\delta,n}(\tilde{S}_n)$ découle du Théorème (1.3.5). Pour tout $\delta \in]0;1[$, on a :

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \tilde{\mathcal{R}}_{x_0,\delta,n}(\tilde{S}_n) \geqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Finalement, on montre aussi que l'estimateur (1.6) est asymptotiquement efficace.

Théorème 1.3.7 *On a :*

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \tilde{\mathcal{R}}_{x_0,\delta,n}(\hat{S}_n) \leqslant \mathbb{E}|\eta|, \qquad \eta \sim \mathcal{N}(0, 1/2).$$

Cas adaptatif

Dans cette partie, on suppose inconnue la régularité de la fonction de régression S, et on se place sur la classe Höldérienne forte $\mathcal{H}(M, K, \beta)$ correspondant à la vraie valeur du paramètre β qu'on situe dans un segment $[\beta_*; \beta^*]$ connu. On conserve l'hypothèse (1.7) sur les bornes de g. Les résultats présentés ici seront démontrés au Chapitre 3 et sont relatés dans Brua (2008b).

Le risque d'un estimateur $\tilde{S}_n$ de $S(x_0)$ est défini par

$$\mathcal{RA}_{x_0,n}(\tilde{S}_n) = \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(x_0, S)} \mathbb{E}_S|\tilde{S}_n - S(x_0)|,$$

où $N(\beta) = \left(\dfrac{n}{\ln n}\right)^{\beta/(2\beta+1)}$.

Pour montrer que $N(\beta)$ est la vitesse de convergence minimax adaptative des estimateurs sur la famille de classe $(\mathcal{H}(M, K, \beta))_{\beta\in[\beta_*;\beta^*]}$, on donne une borne inférieure du risque minimax adaptatif sous les hypothèses suivantes. On suppose que la fonction g est différentiable au sens de Fréchet en un point $S_0 \in C^2([0; 1])$ par rapport à $S \in C^1([0; 1])$ uniformément en $x \in [0; 1]$, i.e. pour tous $S \in C^1([0; 1])$

$$g(x, S) = g(x, S_0) + L_{x,S_0}(S - S_0) + \Gamma_{x,S_0}(S - S_0),$$

où l'application linéaire L_{x,S_0} est nulle sur $C^1([0; 1])$ pour tout $x \in [0; 1]$, et le terme résiduel $\Gamma_{x,S_0}(S)$ satisfait à la propriété

$$\lim_{\|S\|\to 0} \sup_{x\in[0;1]} \Gamma_{x,S_0}(S)/ \parallel S \parallel = 0.$$

On suppose aussi qu'il existe une constante positive G^* telle que pour tous $S \in C^1([0; 1])$ et $x \in [0; 1]$, on a :

$$\left|\tilde{\Gamma}_{x,S_0}(S)\right| \leqslant G^*|S(x)|^2.$$

Théorème 1.3.8 *Si les variables aléatoires* (ξ_k) *sont gaussiennes standard, alors pour des constantes* M *et* K *suffisamment grandes, on a :*

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \mathcal{RA}_{x_0,n}(\tilde{S}_n) > 0.$$

Afin de construire un estimateur adaptatif en vitesse de convergence, nous ne pouvons plus considérer l'estimateur à noyau (2.12)

$$S_h^* = \frac{1}{q_n(h)} \sum_{k=1}^{n} Q\left(\frac{x_k - x_0}{h}\right) y_k,$$

car sa fenêtre $h = h_n$ dépend de la régularité β désormais inconnue. C'est pourquoi on procède suivant la méthode de Lepskiĭ (1990). On partitionne l'intervalle $[\beta_*;\beta^*]$ de la manière suivante :

$$\beta_l = \beta_* + l\frac{\beta^* - \beta_*}{[\ln n]}, \quad l = 0,\dots,[\ln n],$$

où $[a]$ désigne la partie entière du réel a.
A ces valeurs, on associe les fenêtres correspondantes $h_l = \left(\frac{n}{\ln n}\right)^{-1/(2\beta_l+1)}$ et les vitesses $N_l = \left(\frac{n}{\ln n}\right)^{\beta_l/(2\beta_l+1)}$. Finalement, pour $\lambda = 2 + 2\sqrt{2}g^*\left(\frac{\beta^*}{2\beta^*+1} - \frac{\beta_*}{2\beta_*+1}\right)^{1/2}$, on pose

$$\hat{l} = \max\left\{0 \leqslant l \leqslant [\ln n] : \max_{0\leqslant j \leqslant l}\left(\left|S^*_{h_l} - S^*_{h_j}\right| - \frac{\lambda}{N_j}\right) \leqslant 0\right\}.$$

L'estimateur utilisé sera alors $\hat{S}^*_n = S^*_{h_{\hat{l}}}$.

Théorème 1.3.9 *On a :*

$$\limsup_{n\to\infty} \mathcal{RA}_{x_0,n}(\hat{S}^*_n) < \infty,$$

ce qui fait que l'estimateur $\hat{S}^*_n$ est adaptatif en vitesse de convergence.

1.3.2 Modèle de diffusion

On considère dans ce paragraphe le modèle de diffusion (1.2) pour lequel la dérive S est à estimer en un point fixe x_0 à partir des observations $\{X_t, 0 \leqslant t \leqslant T\}$. Les résultats obtenus, démontrés au Chapitre 4, sont dans le prolongement de ceux de Galtchouk et Pergamenshchikov (2006b) et font l'objet d'un article (cf. Brua (2008c)).

On s'intéresse au cas où la dérive S appartient à la classe lipschitzienne

$$\Sigma_{L,M} = \left\{S:\ |S(0)| \leqslant L,\ -L \leqslant \frac{S(x)-S(y)}{x-y} \leqslant -M,\ \forall x,y \in \mathbb{R}\right\},$$

avec $0 < M < L$.
Trouver un estimateur asymptotiquement efficace pour l'ensemble de la classe $\Sigma_{L,M}$ lorsque $T \to \infty$ reste un problème ouvert. Galtchouk et Pergamenshchikov (2006b) utilisent un risque local pris sur un certain voisinage d'une fonction $S_0 \in \Sigma_{L,M}$. Plus précisément, pour $\delta,\varepsilon \in]0;1[$, ils définissent

$$\mathcal{U}_{\delta,\varepsilon,T}(S_0) = \left\{S:\ S = S_0 + D,\ D \in \mathcal{H}^w_{x_0,h}(\delta,\beta),\ \sup_{x\in\mathbb{R}}(|D(x)| + |D'(x)|) \leqslant \varepsilon\right\},$$

où $h = h_T = T^{-1/(2\beta+1)}$ et

$$\mathcal{H}^w_{x_0,h}(\delta,\beta) = \left\{S \in C^1(\mathbb{R}) :\ \left|\int_{-1}^1 z \int_0^1 (S'(x_0 + uh_T z) - S'(x_0))\, du dz\right| \leqslant \delta h_T^\alpha\right\},$$

avec $\beta = 1 + \alpha$, $\alpha \in]0;1[$.
On peut qualifier la classe $\mathcal{H}^{w}_{x_0,h}(\delta,\beta)$ de Höldérienne faible, puisque comme pour le modèle de régression précédent, la condition en jeu implique

$$\left| \int_{-1}^{1} \left(S(x_0 + h_T z) - S(x_0) \right) dz \right| \leqslant \delta h_T^{\beta}.$$

Le paramètre ε est le rayon du voisinage pour la norme usuelle de $C^1(\mathbb{R})$ et le paramètre δ est la borne supérieure de la constante faible de Hölder. Galtchouk et Pergamenshchikov (2006b) font tendre ces deux constantes vers 0 ; cela veut dire en particulier que le voisinage se réduit à la limite au point $\{S_0\}$.

Dans ces conditions ($S \in \mathcal{U}_{\delta,\varepsilon,T}(S_0)$, $S_0 \in \Sigma_{L,M}$), le processus observé $\{X_t, 0 \leqslant t \leqslant T\}$ est ergodique. Il retourne ainsi une infinité de fois dans n'importe quel voisinage du point x_0, ce qui nous permet d'avoir une bonne estimation du coefficient de dérive $S(x_0)$. De plus, il existe une densité ergodique donnée par (cf. Gihman et Skorohod (1968))

$$q_S(x) = \frac{\exp\left(2\int_0^x S(z)dz\right)}{\int_{-\infty}^{\infty} \exp\left(2\int_0^y S(z)dz\right) dy}. \tag{1.12}$$

Le risque maximal local d'un estimateur $\tilde{S}_T$ de $S(x_0)$ est défini par

$$\mathcal{R}_{\delta,\varepsilon,T}(\tilde{S}_T, S_0) = \sup_{S \in \mathcal{U}_{\delta,\varepsilon,T}(S_0)} \mathbb{E}_S \varphi_T |\tilde{S}_T - S(x_0)|, \qquad \varphi_T = T^{\beta/(2\beta+1)}.$$

La borne inférieure du risque minimax est alors donnée par le théorème suivant.

Théorème 1.3.10 *(Galtchouk et Pergamenshchikov, 2006b)*
Si $S_0 \in \Sigma_{L,M}$, alors pour tous $\delta, \varepsilon \in]0;1[$, on a :

$$\liminf_{T\to\infty} \inf_{\tilde{S}_T} \mathcal{R}_{\delta,\varepsilon,T}(\tilde{S}_T, S_0) \geqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}\left(0, 1/(2q_{S_0}(x_0))\right).$$

Soulignons le fait que pour obtenir cette borne inférieure, la classe de fonctions $\tilde{\Sigma}_T$ utilisée pour minorer le risque minimax est une version similaire en temps continu de la classe $\tilde{\Sigma}_n$. Plus précisément, $\tilde{\Sigma}_T$ est la famille contenant les fonctions

$$S_\nu(x) = \varphi_T^{-1} V_\nu\left(\frac{x - x_0}{h_T}\right),$$

avec

$$\begin{aligned} V_\nu(x) &= \nu^{-1} \int_{-\infty}^{\infty} \tilde{Q}_\nu(u) l\left(\frac{u-x}{\nu}\right) du, \\ \tilde{Q}_\nu(u) &= \mathbb{I}_{\{|u|\leqslant 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\leqslant|u|\leqslant 1-\nu\}}, \\ l(z) &= c\exp\left(-(1-z^2)^{-1}\right)\mathbb{I}_{\{|z|\leqslant 1\}}, \end{aligned}$$

pour $0 < \nu < 1/4$ et c la constante de normalisation telle que $\int_{-1}^{1} l(z)dz = 1$.

Afin de construire un estimateur asymptotiquement efficace de $S(x_0)$, Galtchouk et Pergamenshchikov (2006b) estiment d'abord la densité ergodique q_S au point x_0 à l'aide des observations $\{X_t, t \leqslant t_0\}$, avec $t_0 = T^{2\gamma}$, $\gamma_* < \gamma < 1/2$ et $\gamma_* = \alpha/(2\beta+1)$. Cette estimation est réalisée par la quantité

$$\hat{q}_T(x_0) = \frac{1}{2t_0 l_T} \int_0^{t_0} Q\left(\frac{X_t - x_0}{l_T}\right) dt,$$

où $Q(x) = \mathbb{I}_{|x|\leqslant 1}$, $l_T = o(T^{-1/2})$ quand $T \to \infty$.
La procédure séquentielle aboutissant à l'estimation de $S(x_0)$ est la suivante. On se donne premièrement un niveau $H > 0$. Puis, pour une fenêtre $h > 0$, on définit le temps d'atteinte de ce niveau par

$$\tau_H = \inf\left\{t \geqslant t_0 : \int_{t_0}^{t} Q\left(\frac{X_t - x_0}{h}\right) dt \geqslant H\right\},$$

et l'estimateur de $S(x_0)$ par

$$S_T^*(x_0) = \frac{1}{H} \int_{t_0}^{\tau_H} Q\left(\frac{X_t - x_0}{h}\right) dX_t \mathbb{I}_{\{\tau_H \leqslant T\}}. \tag{1.13}$$

Finalement, les bons choix de fenêtre h et de niveau H se révèlent être

$$h = h_T = T^{-1/(2\beta+1)}, \quad H = H_T = (T - t_0)(2\tilde{q}_T(x_0) - \epsilon_T)h_T, \tag{1.14}$$

$$\tilde{q}_T(x_0) = \max\left(\hat{q}_T(x_0), \nu_T^{-1/2}\right), \quad \epsilon_T = 1/(\nu_T T^{\gamma_*}), \quad \nu_T = \ln T. \tag{1.15}$$

Une hypothèse supplémentaire sur le centre du voisinage s'avère utile.

Définition 1.3.11 *On dit qu'une fonction f satisfait à la condition "zéro-constante" de Hölder d'exposant $\iota > 0$ au point x_0 si*

$$\lim_{y \to x_0} \frac{f(y) - f(x_0)}{|y - x_0|^\iota} = 0.$$

Cette condition, appliquée à la dérivée d'une fonction, signifie en particulier que celle-ci appartient à une classe de Hölder d'exposant $1 + \iota$ au point x_0.

Théorème 1.3.12 *(Galtchouk et Pergamenshchikov, 2006b)*
Soient $S_0 \in \Sigma_{L,M}$ et $\beta = 1 + \alpha$, $\alpha \in]0;1[$. Supposons que S_0' satisfait à la condition "zéro-constante" de Hölder d'exposant α au point x_0. Alors l'estimateur (1.13) vérifie

$$\limsup_{\varepsilon \to 0} \limsup_{\delta \to 0} \limsup_{T \to \infty} \mathcal{R}_{\delta,\varepsilon,T}(S_T^*, S_0) \leqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}\left(0, 1/(2q_{S_0}(x_0))\right).$$

Une évolution de ces résultats est de considérer des ensembles de fonctions qui ne se réduisent pas à la limite au point $\{S_0\}$. On s'intéresse toujours à des fonctions S "proches" de $S_0 \in \Sigma_{L,M}$. Plus précisément, l'ensemble sur lequel sera défini le risque maximal d'un estimateur est

$$\mathcal{U}_{\delta,\beta,T}(S_0) = \left\{S : \; S = S_0 + D, D \in \mathcal{H}^w_{x_0}(\delta,\beta)\right\},$$

avec

$$\begin{aligned}\mathcal{H}^w_{x_0}(\delta,\beta) \quad = \quad & \Bigg\{ D \text{ dérivable : } \sup_{x\in\mathbb{R}} (|D(x)| + |D'(x)|) \leqslant B; \\ & \left|\int_{-1}^{1} (D(x_0 + h_T z) - D(x_0))dz\right| \leqslant \delta h_T^{\beta} \Bigg\},\end{aligned}$$

où $\beta \in]1;2[$, $\delta \in]0;1[$ et $B \in]0;M[$ est une constante fixe.

Si la dérive régissant la diffusion du modèle (1.2) appartient à $\mathcal{U}_{\delta,\beta,T}(S_0)$, alors le processus $\{X_t, 0 \leqslant t \leqslant T\}$ est ergodique et la densité ergodique q_S est donnée par (1.12). Le risque maximal d'un estimateur $\tilde{S}_T$ de $S(x_0)$ est défini par

$$\mathcal{R}_{\delta,\beta,T}(\tilde{S}_T, S_0) = \sup_{S\in\mathcal{U}_{\delta,\beta,T}(S_0)} \sqrt{2q_S(x_0)}\varphi_T \mathbb{E}_S|\tilde{S}_T - S(x_0)|, \qquad \varphi_T = T^{\beta/(2\beta+1)}.$$

Le terme $\sqrt{2q_S(x_0)}$ a été introduit dans la définition du risque afin d'uniformiser la constante exacte sur tout l'ensemble $\mathcal{U}_{\delta,\beta,T}(S_0)$, car les résultats prouvés par Galtchouk et Pergamenshchikov (2006b) nous montrent que cette constante change selon le centre du voisinage à travers l'écart-type de la "loi limite". Puisqu'on a l'inclusion $\mathcal{U}_{\delta,\beta,T}(S_0) \subset \Sigma_{L+B,M-B}$, en multipliant le risque par l'inverse de cet écart-type pris en S, on s'attend donc à obtenir une constante uniforme.

La borne inférieure du risque minimax s'obtient, modulo quelques attentions, de la même manière que précédemment.

Théorème 1.3.13 *Si $S_0 \in \Sigma_{L,M}$, alors pour tout $\delta \in]0;1[$, on a :*

$$\liminf_{T\to\infty} \inf_{\tilde{S}_T} \mathcal{R}_{\delta,\beta,T}(\tilde{S}_T, S_0) \geqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}(0,1).$$

Finalement, on montre en fait que le même estimateur $S^*_T(x_0)$ donné par (1.13) est asymptotiquement efficace.

Théorème 1.3.14 *Soient $S_0 \in \Sigma_{L,M}$ et $\beta = 1 + \alpha$, $\alpha \in]0;1[$. Si S'_0 vérifie la condition de Hölder "zéro-constante" d'exposant α au point x_0, alors*

$$\limsup_{\delta\to 0} \limsup_{T\to\infty} \mathcal{R}_{\delta,\beta,T}(S^*_T(x_0), S_0) \leqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}(0,1).$$

Chapitre 2

Modèles de régression : cas non adaptatif

2.1 Introduction

On considère le problème de l'estimation d'une fonction de régression S en un point fixe $z_0 \in]0; 1[$, où l'on dispose des observations

$$y_k = S(x_k) + g(x_k, S)\xi_k, \qquad k \in \{1, \ldots, n\}, \tag{2.1}$$

les régresseurs $x_k = k/n$ étant déterministes, les variables aléatoires ξ_k indépendantes, identiquement distribuées, centrées et de variance 1. On supposera dans un premier temps qu'elles sont gaussiennes standard puis qu'on ne connait pas leur loi. Notons que la variance g^2 des bruits est inconnue et dépend de la fonction de régression inconnue S ainsi que des régresseurs x_k.

De tels modèles hétéroscédastiques de regression se rencontrent dans des études de budgets dévolus à la consommation basées sur des observations prises sur des individus ayant différents revenus, dans des analyses des investissements d'entreprises de différentes tailles et plus récemment en recherche médicale. Par exemple, Goldfeld et Quandt (1972) considèrent des modèles de régression polynomiaux tels que $y_k = \alpha + \beta x_k + u_k$, $\mathbb{E}(u_k^2) = a + bx_k + cx_k^2$, qui est un cas particulier de notre modèle (2.1) en supposant que la fonction de régression inconnue s'écrit $S(x) = \alpha + \beta x$ et $g^2(x, S) = (a - \frac{\alpha c}{\beta^2}) + \left(b - 2\frac{\alpha c}{\beta}\right) x + \frac{c}{\beta^2} S^2(x)$. D'autres modèles de régression hétéroscédastiques sont étudiés par exemple dans Efroĭmovich et Pinsker (1996), Galtchouk et Pergamenshchikov (2005a), Belomestny et Reiß (2006) et Efroĭmovich (2007).

Le problème de l'estimation d'une fonction de régression Höldérienne a été étudié par de nombreux auteurs. Dans le cas d'une fonction de régression appartenant à une classe quasi Höldérienne et estimée en un point fixe avec une perte quadratique, Sacks et Ylvisaker (1981) ont montré que l'estimateur linéaire minimax est un estimateur à noyau. Donoho et Liu (1991) ont ensuite prouvé que le risque de cet estimateur diffère de moins de 17 pourcent de celui de l'estimateur asymptotiquement minimax pour toutes les procédures et ont obtenu des noyaux optimaux pour des classes Höldériennes. En ce qui concerne l'estimation de la fonction ou de ses k-ièmes dérivées avec la perte globale associée à la norme sup, Korostelev (1993) et Donoho (1994a) ont montré qu'un certain estimateur à noyau est asymptotiquement efficace.

On traite ici de l'estimation non paramétrique d'une fonction de régression appartenant à une classe Höldérienne faible. Le risque d'un estimateur est basé sur la perte associée à l'erreur absolue. L'objectif est de trouver un estimateur asymptotiquement efficace. Dans ce but, on utilise la méthode développée par Galtchouk et Pergamenshchikov (2006a) qui ont introduit les classes Höldériennes faibles pour définir le risque d'un estimateur. On travaille donc sur les classes $\mathcal{U}_{z_0,\delta,n}$ qui autorisent les fonctions à posséder une dérivée arbitrairement grande mais qui contraignent ces fonctions à une condition Höldérienne basée sur une constante Höldérienne faible tendant vers zéro (voir (2.2)). Puis on définit le risque $\mathcal{R}_{z_0,\delta,n}(\tilde{S})$ d'un estimateur $\tilde{S}$ de $S(z_0)$ et le risque minimax $\inf_{\tilde{S}} \mathcal{R}_{z_0,\delta,n}(\tilde{S})$ (voir (2.11)).

La prochaine section présente le problème dans le cas de bruits gaussiens, les hypothèses requises et tous les objets mathématiques nécessaires. La borne inférieure asymptotique du risque minimax et un estimateur asymptotiquement efficace sont obtenus à la Section 3 pour des bruits gaussiens. La Section 4 décrit des résultats similaires dans le cas de bruits inconnus. Enfin l'Annexe A contient des résultats techniques utiles dans les démonstrations.

2.2 Description du problème

Considérons le modèle (2.1) où $g : [0;1] \times C^1([0;1]) \longrightarrow \mathbb{R}_+^*$ et S sont des fonctions inconnues. On suppose que les bruits $(\xi_k)_{1\leqslant k\leqslant n}$ sont indépendants identiquement distribués de loi $\mathcal{N}(0,1)$.

Dans le cas où S appartient à un espace Höldérien

$$\mathcal{H}(M,K,\beta) = \left\{ S \in C^1([0;1],\mathbb{R}) \ : \| S' \| \leqslant M, \ \sup_{y\in[0;1]} \frac{|S'(y) - S'(z_0)|}{|y-z_0|^\alpha} \leqslant K \right\},$$

avec $\beta = 1+\alpha$, $\alpha \in]0;1]$ et $\| f \| = \sup_{x\in[0;1]} |f(x)|$, le problème de l'estimation asymptotiquement efficace avec la perte liée à l'erreur absolue reste ouvert.

En conséquence on travaille avec un risque minimax pris sur une classe plus large, appelée classe Höldérienne faible au point z_0. Elle est définie, pour $0 < \delta < 1$, par

$$\mathcal{U}_{z_0,\delta,n} = \left\{ S \in C^1([0;1]) : \|S'\| \leqslant \delta^{-1}; \left| \int_{-1}^{1} \big(S(z_0 + h_n u) - S(z_0)\big) du \right| \leqslant \delta h_n^\beta \right\}, \quad (2.2)$$

où $h_n = n^{-1/(2\beta+1)}$.
Précisons que la régularité β de la fonction S est supposée connue dans tout ce chapitre.

En remarquant que

$$\int_{-1}^{1}\big(S(z_0+h_n u)-S(z_0)\big)du=\int_{-1}^{1}\left(\int_{z_0}^{z_0+uh_n}(S'(t)-S'(z_0))dt\right)du, \tag{2.3}$$

on voit que si S est Höldérienne, $S\in\mathcal{H}(M,K,\beta)$ avec $M<\delta^{-1}$ et $2K/(\beta(\beta+1))<\delta$ alors $S\in\mathcal{U}_{z_0,\delta,n}$. Mais $\mathcal{U}_{z_0,\delta,n}$ contient aussi des fonctions qui ne sont pas Höldériennes. C'est la raison pour laquelle la classe $\mathcal{U}_{z_0,\delta,n}$ est appelée classe Höldérienne faible.

Donnons maintenant les hypothèses nécessaires. Tout d'abord on suppose que

$$\lim_{n\to\infty}\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\left|\left(\frac{1}{q_n}\sum_{k=1}^{n}Q\big(\frac{x_k-z_0}{h_n}\big)g^2(x_k,S)\right)^{\frac{1}{2}}-g(z_0,S)\right|=0, \tag{2.4}$$

avec

$$q_n=\sum_{k=1}^{n}Q\big(\frac{x_k-z_0}{h_n}\big),\quad \text{et } Q=\mathbb{I}_{[-1;1]}.$$

De plus, on suppose qu'il existe $g_*>0$ et $g^*<\infty$ tels que

$$g_*\leqslant\inf_{0\leqslant x\leqslant 1}\inf_{S\in C^1([0;1])}g(x,S)\leqslant\sup_{0\leqslant x\leqslant 1}\sup_{S\in C^1([0;1])}g(x,S)\leqslant g^*. \tag{2.5}$$

Soit S_0 la fonction identiquement nulle. On suppose la fonction g est différentiable au sens de Fréchet par rapport à $S\in C^1([0;1])$ en S_0 uniformément en $x\in[0;1]$, i.e. pour tous $S\in C^1([0;1])$

$$g(x,S)=g(x,S_0)+L_{x,S_0}(S)+\Gamma_{x,S_0}(S), \tag{2.6}$$

où l'application linéaire L_{x,S_0} est bornée sur $C^1([0;1])$ uniformément en $x\in[0;1]$, i.e. il existe une constante strictement positive C_{S_0} telle que

$$\sup_{x\in[0;1]}\sup_{S\in C^1([0;1]),\ \|S\|\neq 0}|L_{x,S_0}(S)|/\parallel S\parallel\leqslant C_{S_0} \tag{2.7}$$

et le terme résiduel $\Gamma_{x,S_0}(S)$ satisfait à la propriété

$$\lim_{\|S\|\to 0}\sup_{x\in[0;1]}\Gamma_{x,S_0}(S)/\parallel S\parallel=0. \tag{2.8}$$

Remarque 2.2.1
Notons que l'hypothèse (2.4) est vérifiée lorsque pour tout $\varepsilon>0$, il existe $\eta>0$ tel que si $|x-z_0|\leqslant\eta$, alors $\sup_{S\in C^1([0;1],\mathbb{R})}|g(x,S)-g(z_0,S)|\leqslant\varepsilon$.
En particulier une fonction g satisfait à cette propriété si elle est uniformément continue par rapport à ses deux variables. Par ailleurs, les résultats de ce chapitre restent valables lorsque la fonction g est uniformément continue par rapport à ses deux variables et bornée (2.5) sans avoir recours aux hypothèses (2.6-2.8).

Remarque 2.2.2
Donnons un exemple général d'une fonction g satisfaisant aux hypothèses (2.4)–(2.8) ci-dessus. Soient $V : \mathbb{R} \longrightarrow \mathbb{R}_+$ et $G : [0;1] \times \mathbb{R} \longrightarrow \mathbb{R}_+$ deux fonctions différentiables telles que

$$\|V'\|_\infty < \infty, \;\; G_* = \inf_{x\in[0,1],\;y\in\mathbb{R}} G(x,y) > 0, \;\; G'_* = \sup_{x\in[0;1],\;y\in\mathbb{R}} \left|\frac{\partial G}{\partial y}(x,y)\right| < \infty.$$

Posons

$$g^2(x,S) = G(x,S(x)) + \int_0^1 V(S(t))dt. \tag{2.9}$$

La dérivée au sens de Fréchet de g s'écrit

$$L_{x,S}(f) = \frac{1}{2g(x,S)}\frac{\partial G}{\partial y}(x,S(x))f(x) + \frac{1}{2g(x,S)}\int_0^1 V'(S(t))f(t)dt,$$

donc on a

$$\sup_{x\in[0;1]}\;\sup_{S\in C^1([0;1]),\;\|S\|\neq 0} \frac{|L_{x,S}(f)|}{\|f\|_\infty} \leqslant \frac{G'_* + \|V'\|_\infty}{2\sqrt{G_*}}.$$

En écrivant les développements de Taylor des fonctions $y \mapsto G(x,y)$ au point $(x,S(x))$ et V au point $S(t)$ au premier ordre :

$$\begin{aligned} G(x,S(x)+f(x)) &= G(x,S(x)) + \frac{\partial G}{\partial y}(x,S(x))f(x) + f(x)\varepsilon_{x,S}(f(x)), \\ V(S(t)+f(t)) &= V(S(t)) + V'(S(t))f(t) + f(t)\tilde{\varepsilon}_{t,S}(f(t)), \end{aligned}$$

on peut aisément montrer que

$$\begin{aligned} \frac{|\Gamma_{x,S}(f)|}{\|f\|_\infty} &\leqslant \frac{G'_* + \|V'\|_\infty}{8G_*^{3/2}} \left|g^2(x,S+f) - g^2(x,S)\right| \\ &\quad + \frac{1}{2\sqrt{G_*}}\left(|\varepsilon_{x,S}(f(x))| + \int_0^1 |\tilde{\varepsilon}_{t,S}(f(t))|dt\right). \end{aligned} \tag{2.10}$$

Par conséquent, si nous prenons $G(x,y) = \alpha_0 + \alpha_1 x + \alpha_2 \sin^2 y$ et $V(y) = \alpha_3 \sin^2 y$ pour tous $(x,y) \in [0,1] \times \mathbb{R}$, avec $\alpha_0 > 0$ et $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}_+$, alors la fonction g définie par (2.9) est uniformément continue, bornée par $\sqrt{\alpha_0}$ et $\sqrt{\alpha_0+\alpha_1+\alpha_2+\alpha_3}$. De plus, en développant explicitement les fonctions $\varepsilon_{x,S}$ et $\tilde{\varepsilon}_{x,S}$ dans ce cas, on peut prouver par (2.10) que g satisfait à l'hypothèse (2.8). Ainsi nous avons exhibé une fonction g vérifiant toutes les hypothèses voulues.

Pour tout estimateur $\tilde{S}_n(z_0)$ de $S(z_0)$ on définit le risque

$$\mathcal{R}_{z_0,\delta,n}(\tilde{S}_n) = \sup_{S\in\mathcal{U}_{z_0,\delta,n}} \mathbb{E}_S \varphi_n \frac{|\tilde{S}_n(z_0) - S(z_0)|}{g(z_0,S)}, \tag{2.11}$$

où $\mathbb{E}_S$ désigne l'espérance correspondant à la loi $\mathbb{P}_S$ dans (2.1) et $\varphi_n = n^{\frac{\beta}{2\beta+1}}$.
L'objectif est d'atteindre la constante asymptotique exacte avec cette vitesse φ_n. On suppose seulement que $\beta \in]1;2]$ car si $\beta > 2$ on devrait utiliser un noyau Q d'ordre $\lfloor\beta\rfloor$ i.e. tel que $\int u^j Q(u)du = 0$ pour $j = 1, 2, \ldots, \lfloor\beta\rfloor$ et $\int Q(u)du < \infty$, où $\lfloor a\rfloor$ désigne le plus petit entier strictement plus petit que a.

2.3 Bornes asymptotiques pour des bruits gaussiens

On donne dans ce paragraphe la borne inférieure du risque minimax et on montre que l'estimateur à noyau $\hat{S}_n(z_0)$, defini par

$$\hat{S}_n(z_0) = \frac{1}{q_n}\sum_{k=1}^{n} Q\big(\frac{x_k - z_0}{h_n}\big)y_k, \quad h_n = n^{-1/(2\beta+1)}, \tag{2.12}$$

est asymptotiquement efficace.

Théorème 2.3.1 *Si les variables aléatoires (ξ_k) sont gaussiennes $\mathcal{N}(0,1)$, alors pour tout $\delta \in]0;1[$, on a*

$$\liminf_{n\to\infty} \inf_{\tilde{S}} \mathcal{R}_{z_0,\delta,n}(\tilde{S}) \geqslant \frac{\mathbb{E}|\xi|}{\sqrt{2}}, \qquad \xi \sim \mathcal{N}(0,1),$$

où l'infimum est pris sur tous les estimateurs $\tilde{S}$ de $S(z_0)$.

Démonstration: Pour tout $\nu \in \,]0;1/4[$, on pose $S_\nu(x) = \varphi_n^{-1} V_\nu\left(\frac{x - z_0}{h_n}\right)$, où la fonction V_ν est définie par

$$\begin{aligned} V_\nu(x) &= \nu^{-1}\int_{-\infty}^{\infty} \tilde{Q}_\nu(u) l\left(\frac{u-x}{\nu}\right)du, \\ \tilde{Q}_\nu(u) &= \mathbb{I}_{\{|u|\leqslant 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\leqslant|u|\leqslant 1-\nu\}}, \\ l(z) &= c\exp\left(-(1-z^2)^{-1}\right)\mathbb{I}_{\{|z|\leqslant 1\}}, \end{aligned}$$

avec c la constante de normalisation telle que $\displaystyle\int_{-1}^{1} l(z)dz = 1$.
V_ν est une fonction de classe C^∞ à support compact $[-1;1]$.

Soient $b > 0$ et $\delta \in]0;1[$. Notons, pour $x \in \mathbb{R}$ et $u \in [-b;b]$, $S_{\nu,u}(x) = uS_\nu(x)$. D'après le Lemme 2.5.1, il existe un entier $n_{b,\delta} > 0$ tel que $S_{\nu,u} \in \mathcal{U}_{z_0,\delta,n}$ pour tous $n \geqslant n_{b,\delta}$ et $u \in [-b;b]$. Par conséquent, si $\tilde{S}$ est un estimateur de $S(z_0)$, on a pour $n \geqslant n_{b,\delta}$,

$$\begin{aligned} \mathcal{R}_{z_0,\delta,n}(\tilde{S}) &\geqslant \sup_{|u|\leqslant b} \frac{1}{g(z_0, S_{\nu,u})}\mathbb{E}_{S_{\nu,u}}\varphi_n|\tilde{S} - S_{\nu,u}(z_0)| \\ &\geqslant \frac{1}{2b}\int_{-b}^{b} \frac{1}{g(z_0, S_{\nu,u})}\mathbb{E}_{S_{\nu,u}} v_a\left(\varphi_n(\tilde{S} - S_{\nu,u}(z_0))\right)du =: I_n(a,b), \end{aligned}$$

où $v_a(x) = |x| \wedge a, a > 0$.

Notons $\mathbb{P}_{S_{\nu,u}}$ la loi de $(y_k^{(1)})_{k=1,\dots,n}$, où $y_k^{(1)} = S_{\nu,u}(x_k) + g(x_k, S_{\nu,u})\xi_k$, et $\mathbb{P}$ la loi de $(y_k^{(0)})_{k=1,\dots,n}$, où $y_k^{(0)} = g(x_k, S_{\nu,u})\xi_k$. Ces deux mesures sont équivalentes et la dérivée de Radon-Nikodym au point $(y_1, \dots, y_n)$ correspondante s'écrit

$$\begin{aligned}
\rho_n(u) &= \frac{d\mathbb{P}_{S_{\nu,u}}}{d\mathbb{P}}(y_1, \dots, y_n) \\
&= \exp\left\{ -\frac{1}{2} \sum_{k=1}^n \left(\left(\frac{y_k - S_{\nu,u}(x_k)}{g(x_k, S_{\nu,u})} \right)^2 - \left(\frac{y_k}{g(x_k, S_{\nu,u})} \right)^2 \right) \right\} \\
&= \exp\left(u\varsigma_n \eta_n - \frac{u^2}{2}\varsigma_n^2 \right),
\end{aligned}$$

où $\varsigma_n^2 = \dfrac{1}{\varphi_n^2} \displaystyle\sum_{k=1}^n \frac{V_\nu^2\left(\frac{x_k - z_0}{h_n}\right)}{g^2(x_k, S_{\nu,u})}$ et $\eta_n = \dfrac{1}{\varsigma_n \varphi_n} \displaystyle\sum_{k=1}^n \frac{V_\nu\left(\frac{x_k - z_0}{h_n}\right)}{g^2(x_k, S_{\nu,u})} y_k$.
Sous la loi $\mathbb{P}$, η_n est une variable aléatoire gaussienne standard.

On montre dans le Lemme 2.5.2 que

$$\varsigma_n^2 \xrightarrow[n\to\infty]{} \frac{1}{g^2(z_0, 0)} \int_{-1}^1 V_\nu^2(z)dz =: \sigma_\nu^2. \tag{2.13}$$

Ainsi on réécrit $\rho_n(u) = \exp\left(u\sigma_\nu\eta_n - \frac{u^2\sigma_\nu^2}{2} + r_n\right)$, où r_n converge en $\mathbb{P}$-probabilité vers zéro.

En notant $\psi_{a,n}(\tilde{S}, S_{\nu,u}) = v_a(\varphi_n(\tilde{S}_n(z_0) - S_{\nu,u}(z_0)))$ et $\mathbb{E}$ l'espérance correspondant à la mesure de probabilité $\mathbb{P}$, on a

$$I_n(a,b) \geqslant \frac{1}{2b} \int_{-b}^b \mathbb{E}\mathbb{I}_{B_d} \frac{\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})} \varrho_n(u)du + \delta_n(a,b) =: J_n(a,b) + \delta_n(a,b), \tag{2.14}$$

où

$$\begin{aligned}
B_d &= \{|\eta_n| \leqslant d\} \quad \text{et} \quad d = \sigma_\nu(b - \sqrt{b}), b > 1, \\
\varrho_n(u) &= \exp\left(u\sigma_\nu\eta_n - \frac{u^2\sigma_\nu^2}{2} \right), \\
\delta_n(a,b) &= \frac{1}{2b} \int_{-b}^b \mathbb{E}\mathbb{I}_{B_d} \frac{\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})} \theta_n(u)du, \\
\theta_n(u) &= \rho_n(u) - \varrho_n(u).
\end{aligned}$$

Remarquons que $\rho_n(u) \xrightarrow[n\to\infty]{\mathcal{L}} \rho_\infty(u) = \exp\left(u\sigma_\nu\eta - \frac{u^2\sigma_\nu^2}{2}\right)$. On montre aisément que $\mathbb{E}\rho_\infty(u) = 1$ et on a aussi $\mathbb{E}\rho_n(u) = 1$ car $\rho_n(u)$ est une densité. Donc, en utilisant le Lemme 2.5.3, $\{\rho_n(u), n \geqslant 1\}$ est uniformément intégrable. Comme $\varrho_n(u)$ est borné sur

B_d, on obtient l'intégrabilité uniforme de la famille $\{\mathbb{I}_{B_d}\psi_{a,n}(\tilde{S}, S_{\nu,u})\theta_n(u), n \geqslant 1\}$.

Ecrivons désormais $\theta_n(u) = \exp\left(u\sigma_\nu\eta_n - \frac{u^2\sigma_\nu^2}{2}\right)(e^{r_n}-1)$ et notons que $\exp\left(u\sigma_\nu\eta_n - \frac{u^2\sigma_\nu^2}{2}\right)$ est bornée sur B_d et que $e^{r_n} - 1 \xrightarrow[n\to\infty]{\mathbb{P}} 0$. En conséquence on a

$$\frac{\mathbb{I}_{B_d}\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})}\theta_n(u) \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Puis il s'en suit : $\frac{\mathbb{I}_{B_d}\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})}\theta_n(u) \xrightarrow[n\to\infty]{\mathbb{L}^1} 0$ et $\mathbb{E}\frac{\mathbb{I}_{B_d}\psi_{a,n}(\tilde{S}, S_{\nu,u})}{g(z_0, S_{\nu,u})}\theta_n(u) \xrightarrow[n\to\infty]{} 0$. Finalement, par convergence dominée, il vient $\sup_{\tilde{S}} |\delta_n(a,b)| \xrightarrow[n\to\infty]{} 0$ dans (2.14).

Intéressons-nous maintenant au terme $J_n(a,b)$ dans (2.14). Réécrivons d'abord $\varrho_n(u) = \zeta_n e^{-\sigma_\nu^2(u-\tilde{\eta}_n)^2/2}$ avec $\zeta_n = e^{\eta_n^2/2}$ et $\tilde{\eta}_n = \frac{\eta_n}{\sigma_\nu}$. Ensuite si $\xi \sim \mathcal{N}(0,1)$ on note $\tilde{\xi} = \frac{\xi}{\sigma_\nu}$, $\zeta = e^{\xi^2/2}$, $\tilde{B}_d = \{|\xi| \leqslant d\}$ et $\tilde{\mathbb{E}}$ l'espérance correspondant à la loi de probabilité de ξ. En posant $t_n = \varphi_n\tilde{S}_n(z_0)$, on obtient successivement

$$\begin{aligned}
J_n(a,b) &= \frac{1}{2b}\int_{-b}^{b} \mathbb{E}\mathbb{I}_{B_d}\zeta_n \frac{v_a(u-t_n)}{g(z_0, S_{\nu,u})}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\eta}_n)^2\right)du \\
&= \frac{1}{2b}\int_{-b}^{b} \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta \frac{v_a(u-t_n)}{g(z_0, S_{\nu,u})}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)du \\
&= \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} \frac{v_a(u-t_n)}{g(z_0, S_{\nu,u})}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)du.
\end{aligned}$$

On a la convergence suivante :

$$\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} v_a(u-t_n)\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)\left(\frac{1}{g(z_0, S_{\nu,u})} - \frac{1}{g(z_0,0)}\right)du \xrightarrow[n\to\infty]{} 0. \quad (2.15)$$

En effet, en utilisant les hypothèses (2.6) et (2.7) on obtient

$$\begin{aligned}
&\left|\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} v_a(u-t_n)\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)\left(\frac{1}{g(z_0, S_{\nu,u})} - \frac{1}{g(z_0,0)}\right)du\right| \\
\leqslant\ & \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} v_a(u-t_n)\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)\left|\frac{\Gamma_{z_0,0}(S_{\nu,u}) - L_{z_0,0}(S_{\nu,u})}{g_*^2}\right|du \\
\leqslant\ & \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} a\frac{C_0 \parallel S_{\nu,u} \parallel + |\Gamma_{z_0,0}(S_{\nu,u})|}{g_*^2}du.
\end{aligned}$$

Puisque $\parallel S_{\nu,u} \parallel$ converge vers zéro quand n tend vers l'infini, les hypothèses (2.8) et (2.15) nous permettent d'avoir

$$\begin{aligned}
\liminf_{n\to\infty} J_n(a,b) &= \liminf_{n\to\infty} \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta\frac{1}{2b}\int_{-b}^{b} \frac{v_a(u-t_n)}{g(z_0,0)}\exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right)du \\
&=: \liminf_{n\to\infty} H_n(a,b).
\end{aligned}$$

En réalité, on a même montré que

$$\sup_{\tilde{S}} |J_n(a,b) - H_n(a,b)| \xrightarrow[n\to\infty]{} 0.$$

Nous avons les inégalités

$$\begin{aligned}
& \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta \frac{1}{2b}\int_{-b}^{b} \frac{v_a(u-t_n)}{g(z_0,0)} \exp\left(-\frac{\sigma_\nu^2}{2}(u-\tilde{\xi})^2\right) du \\
\geqslant\ & \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta \frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} \frac{v_a(t-t_n+\tilde{\xi})}{g(z_0,0)} \exp\left(-\frac{\sigma_\nu^2}{2}t^2\right) dt \\
\geqslant\ & \tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta \frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} \frac{v_a(t)}{g(z_0,0)} \exp\left(-\frac{\sigma_\nu^2}{2}t^2\right) dt,
\end{aligned}$$

la dernière provenant du Lemme 2.5.4.
Ainsi, en utilisant le fait que $\tilde{\mathbb{E}}\mathbb{I}_{\tilde{B}_d}\zeta = \dfrac{2\sigma_\nu(b-\sqrt{b})}{\sqrt{2\pi}}$, il vient

$$\liminf_{a\to\infty}\liminf_{n\to\infty}\inf_{\tilde{S}} H_n(a,b) \geqslant \frac{\sigma_\nu}{\sqrt{2\pi}}\frac{b-\sqrt{b}}{b}\int_{-\sqrt{b}}^{\sqrt{b}} \frac{|t|}{g(z_0,0)} \exp\left(-\frac{\sigma_\nu^2}{2}t^2\right) dt.$$

En passant à la limite $b \to \infty$ puis $\nu \to 0$, en se rappellant que $\sigma_\nu^2 \xrightarrow[\nu\to 0]{} \dfrac{2}{g^2(z_0,0)}$, et en effectuant un changement de variable trivial dans l'intégrale, on obtient

$$\liminf_{\nu\to 0}\liminf_{b\to\infty}\liminf_{a\to\infty}\liminf_{n\to\infty}\inf_{\tilde{S}} H_n(a,b) \geqslant \mathbb{E}|\xi|.$$

Finalement, puisque pour $n \geqslant n_{b,\delta}$ on a

$$\inf_{\tilde{S}} \mathcal{R}_{z_0,\delta,n}(\tilde{S}) \geqslant -\sup_{\tilde{S}} |J_n(a,b) - H_n(a,b)| - \sup_{\tilde{S}} |\delta_n(a,b)| + \inf_{\tilde{S}} H_n(a,b),$$

on en déduit le théorème. □

Théorème 2.3.2 *Si les variables aléatoires* (ξ_k) *sont gaussiennes* $\mathcal{N}(0,1)$*, alors l'estimateur* $\hat{S}_n(z_0)$ *donné par (2.12) vérifie l'inégalité suivante :*

$$\limsup_{\delta\to 0}\limsup_{n\to\infty} \mathcal{R}_{z_0,\delta,n}(\hat{S}_n(z_0)) \leqslant \frac{\mathbb{E}|\xi|}{\sqrt{2}}, \qquad \xi \sim \mathcal{N}(0,1).$$

Démonstration: En premier lieu, nous réécrivons l'erreur de l'estimation sous la forme $\hat{S}_n(z_0) - S(z_0) = B_n + \dfrac{1}{\sqrt{q_n}}\zeta_n$ avec

$$B_n = \frac{1}{q_n}\sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_n}\right)(S(x_k) - S(z_0)), \tag{2.16}$$

$$\zeta_n = \frac{1}{\sqrt{q_n}}\sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_n}\right)g(x_k,S)\xi_k. \tag{2.17}$$

Occupons-nous premièrement du terme $\dfrac{\zeta_n}{\sqrt{q_n}}$. Il est clair d'après (2.17) que ζ_n est une variable aléatoire gaussienne $\mathcal{N}\big(0,\sigma_n^2(S)\big)$ où $\sigma_n^2(S) = \dfrac{1}{q_n}\sum_{k=1}^n Q\big(\dfrac{x_k-z_0}{h_n}\big)g^2(x_k,S)$. On montre dans le Lemme 2.5.5 que la variance $\sigma_n^2(S)$ vérifie $\sigma_n^2(S) \xrightarrow[n\to\infty]{} g^2(z_0,S)$. Si $\xi\sim\mathcal{N}(0,1)$, on a

$$\begin{aligned}\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\frac{1}{g(z_0,S)}\mathbb{E}_S\Big|\frac{\varphi_n}{\sqrt{q_n}}\zeta_n\Big| &= \frac{\varphi_n}{\sqrt{q_n}}\mathbb{E}|\xi|\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\frac{\sigma_n(S)}{g(z_0,S)}\\ &\leqslant \frac{\varphi_n}{\sqrt{q_n}}\frac{\mathbb{E}|\xi|}{g_*}\left(\sup_{S\in\mathcal{U}_{z_0,\delta,n}}|\sigma_n(S)-g(z_0,S)|+g_*\right).\end{aligned}$$

Utilisant l'hypothèse (2.4) et le fait que $\dfrac{q_n}{\varphi_n^2}=\dfrac{q_n}{nh_n}\xrightarrow[n\to\infty]{} 2$, il vient

$$\limsup_{n\to\infty}\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\mathbb{E}_S\frac{\varphi_n}{g(z_0,S)}\frac{|\zeta_n|}{\sqrt{q_n}}\leqslant\frac{\mathbb{E}|\xi|}{\sqrt{2}}. \tag{2.18}$$

Notons désormais $u_k=\dfrac{x_k-z_0}{h_n}$, $\Delta u_k=\dfrac{1}{nh_n}$ et réécrivons (2.16) comme suit :

$$\begin{aligned}B_n &= \frac{\varphi_n^2}{q_n}\sum_{k=1}^n Q(u_k)\big(S(z_0+h_nu_k)-S(z_0)\big)\Delta u_k &(2.19)\\ &= \frac{\varphi_n^2}{q_n}\int_{-1}^1\big(S(z_0+h_nu)-S(z_0)\big)du+\frac{\varphi_n^2}{q_n}R_n, &(2.20)\end{aligned}$$

avec

$$\begin{aligned}R_n &= \sum_{k=1}^n Q(u_k)\big(S(z_0+h_nu_k)-S(z_0)\big)\Delta u_k-\int_{-1}^1\big(S(z_0+h_nu)-S(z_0)\big)du\\ &= \sum_{k=k_*}^{k^*}\int_{u_{k-1}}^{u_k}\big(S(z_0+h_nu_k)-S(z_0+h_nu)\big)du\\ &- \int_{u_{k^*}}^1\big(S(z_0+h_nu)-S(z_0)\big)du+\int_{u_{k_*-1}}^{-1}\big(S(z_0+h_nu)-S(z_0)\big)du,\end{aligned}$$

où $k^*=[n(z_0+h_n)]$ et $k_*=[n(z_0-h_n)]+1$.

On peut borner R_n de la manière suivante :

$$\begin{aligned}|R_n| &\leqslant \sum_{k=k_*}^{k^*}\int_{u_{k-1}}^{u_k}h_n(u_k-u)\delta^{-1}du+\int_{u_{k^*}}^1 h_n\delta^{-1}udu+\int_{u_{k_*-1}}^{-1}h_n\delta^{-1}|u|du\\ &\leqslant h_n\delta^{-1}\Big(\sum_{k=k_*}^{k^*}(u_k-u_{k-1})\frac{1}{nh_n}+(1-u_{k^*})+2(-1-u_{k_*-1})\Big)\leqslant\frac{6\delta^{-1}}{n}.\end{aligned}$$

D'où

$$\limsup_{n\to\infty} \sup_{S\in\mathcal{U}_{z_0,\delta,n}} \mathbb{E}_S \varphi_n \Big| \frac{\varphi_n^2}{q_n} R_n \Big| = 0. \tag{2.21}$$

En considérant le terme $\dfrac{\varphi_n^2}{q_n} \displaystyle\int_{-1}^{1} \big(S(z_0+h_n u) - S(z_0)\big) du$ dans (2.20), on a

$$\Big| \frac{\varphi_n^2}{q_n} \int_{-1}^{1} \big(S(z_0+h_n u) - S(z_0)\big) du \Big| \leqslant \frac{\varphi_n^2}{q_n} \delta n^{\frac{-\beta}{2\beta+1}} = \delta \frac{\varphi_n}{q_n}.$$

Puis d'après la définition de $\mathcal{U}_{z_0,\delta,n}$ on obtient

$$\limsup_{n\to\infty} \sup_{S\in\mathcal{U}_{z_0,\delta,n}} \mathbb{E}_S \varphi_n \Big| \frac{\varphi_n^2}{q_n} \int_{-1}^{1} \big(S(z_0+h_n u) - S(z_0)\big) du \Big| \leqslant \frac{\delta}{2}. \tag{2.22}$$

Finalement (2.18), (2.21) et un passage à la limite $\delta \to 0$ dans (2.22) induisent

$$\limsup_{\delta\to 0} \limsup_{n\to\infty} \mathcal{R}_{z_0,\delta,n}(\hat{S}_n(z_0)) \leqslant \frac{\mathbb{E}|\xi|}{\sqrt{2}}.$$

□

2.4 Bornes asymptotiques pour des bruits de loi inconnue

Dans ce paragraphe, on suppose que les variables aléatoires (ξ_k) dans le modèle (2.1) sont indépendantes et identiquement distribuées selon une loi inconnue d'espérance nulle, de variance 1 et telles que $\mathbb{E}|\xi_1|^{2+\epsilon} \leqslant L$ pour certaines constantes strictement positives ϵ et L. On note $\mathcal{P}_{\epsilon,L}$ l'ensemble de toutes les lois vérifiant ces conditions avec $L > 0$ suffisamment grand pour que la loi gaussienne standard y figure. On définit le risque robuste d'un estimateur $\tilde{S}_n$ de $S(z_0)$ correspondant à ce cas par

$$\tilde{\mathcal{R}}_{z_0,\delta,n}(\tilde{S}_n) = \sup_{\mathcal{P}_{\epsilon,L}} \sup_{S\in\mathcal{U}_{z_0,\delta,n}} \mathbb{E}_S \varphi_n \frac{|\tilde{S}_n - S(z_0)|}{g(z_0,S)}.$$

Les théorèmes suivants établissent la constante asymptotique exacte du risque minimax pris sur tous les estimateurs ainsi que l'efficacité asymptotique de l'estimateur à noyau $\hat{S}_n(z_0)$ de $S(z_0)$ défini par (2.12).

Théorème 2.4.1 *Pour tout $\delta \in]0;1[$, on a*

$$\liminf_{n\to\infty} \inf_{\tilde{S}} \tilde{\mathcal{R}}_{z_0,\delta,n}(\tilde{S}) \geqslant \frac{\mathbb{E}|\eta|}{\sqrt{2}}, \qquad \eta \sim \mathcal{N}(0,1),$$

où l'infimum est pris sur tous les estimateurs $\tilde{S}$ de $S(z_0)$.

Démonstration: C'est une conséquence directe du Théorème 2.3.1 qui donne cette même borne asymptotique inférieure dans le cas de bruits gaussiens d'espérance nulle et de variance dépendant des régresseurs et de la fonction de régression. Le risque correspondant $\mathcal{R}_{z_0,\delta,n}$ est plus petit que le risque $\tilde{\mathcal{R}}_{z_0,\delta,n}$ car la loi gaussienne standard appartient à $\mathcal{P}_{\varepsilon,L}$. Le Théorème 2.4.1 en découle. □

Théorème 2.4.2 *L'estimateur à noyau (2.12) est asymptotiquement efficace. En effet, il vérifie l'inégalité :*

$$\limsup_{\delta\to 0}\limsup_{n\to\infty}\tilde{\mathcal{R}}_{z_0,\delta,n}(\hat{S}_n(z_0))\leqslant\frac{\mathbb{E}|\eta|}{\sqrt{2}},\qquad \eta\sim\mathcal{N}(0,1).$$

Démonstration: En écrivant à nouveau $\hat{S}_n(z_0)-S(z_0)=B_n+\zeta_n/\sqrt{q_n}$, avec B_n et ζ_n définis en (2.16) et (2.17), on remarque que B_n ne dépend pas des distributions des variables aléatoires ξ_k. C'est pourquoi (2.21) et (2.22) restent valables et impliquent pour tout $\delta\in]0;1[$:

$$\limsup_{n\to\infty}\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\varphi_n|B_n|\leqslant\delta/2.$$

Il suffit donc de montrer que

$$\lim_{n\to\infty}\sup_{p\in\mathcal{P}_{\epsilon,L}}\sup_{S\in\mathcal{U}_{z_0,\delta,n}}\left|\frac{\mathbb{E}_S|\zeta_n|}{g(z_0,S)}-\mathbb{E}|\eta|\right|=0,\tag{2.23}$$

avec $\eta\sim\mathcal{N}(0,1)$.

Notons $\tilde{\zeta}_n=\zeta_n/g(z_0,S)=\sum_{k=1}^n u_k$, où $u_k=\frac{1}{\sqrt{q_n}}Q\left(\frac{x_k-z_0}{h_n}\right)\frac{g(x_k,S)}{g(z_0,S)}\xi_k$, et réécrivons $\frac{g(x_k,S)}{g(z_0,S)}\xi_k=\xi'_k+\xi''_k$, où

$$\begin{aligned}\xi'_k &= \frac{g(x_k,S)}{g(z_0,S)}\xi_k\mathbb{I}_{|\xi_k|\leqslant q_n^{1/4}}-\frac{g(x_k,S)}{g(z_0,S)}\mathbb{E}\left(\xi_1\mathbb{I}_{|\xi_1|\leqslant q_n^{1/4}}\right),\\ \xi''_k &= \frac{g(x_k,S)}{g(z_0,S)}\xi_k\mathbb{I}_{|\xi_k|> q_n^{1/4}}-\frac{g(x_k,S)}{g(z_0,S)}\mathbb{E}\left(\xi_1\mathbb{I}_{|\xi_1|> q_n^{1/4}}\right).\end{aligned}$$

Soient $u'_k=\frac{1}{\sqrt{q_n}}Q\left(\frac{x_k-z_0}{h_n}\right)\xi'_k$ et $u''_k=\frac{1}{\sqrt{q_n}}Q\left(\frac{x_k-z_0}{h_n}\right)\xi''_k$, alors nous avons $\tilde{\zeta}_n=\tilde{\zeta}'_n+\tilde{\zeta}''_n=\sum_{k=1}^n u'_k+\sum_{k=1}^n u''_k$. De plus, $(u'_k)_{k\geqslant 1}$ est une "martingale difference" et pour tout $k\geqslant 1$, on a $|u'_k|\leqslant 2\frac{g^*}{g_*}q_n^{-1/4}$ et

$$\mathbb{E}_S\left((u'_k)^2|\mathcal{F}_{k-1}\right)=\frac{1}{q_n}Q\left(\frac{x_k-z_0}{h_n}\right)\frac{g^2(x_k,S)}{g^2(z_0,S)}Var\left(\xi_1\mathbb{I}_{|\xi_1|\leqslant q_n^{1/4}}\right),$$

où $\mathcal{F}_k = \sigma(\xi_i, 1 \leqslant i \leqslant k)$.
Ecrivons

$$\sum_{i=1}^{n} \mathbb{E}_S \left((u_i')^2 | \mathcal{F}_{i-1}\right) = \frac{Var\left(\xi_1 \mathbb{I}_{|\xi_1| \leqslant q_n^{1/4}}\right)}{q_n} \sum_{i=1}^{n} Q\left(\frac{x_i - z_0}{h_n}\right) \frac{g^2(x_i, S)}{g^2(z_0, S)} = \frac{G_n(S)}{q_n} a_n,$$

où $G_n(S) = \sum_{i=1}^{n} Q\left(\frac{x_i - z_0}{h_n}\right) \frac{g^2(x_i, S)}{g^2(z_0, S)}$ et $a_n = Var\left(\xi_1 \mathbb{I}_{|\xi_1| \leqslant q_n^{1/4}}\right)$.

En notant $r_n(S) = \frac{G_n(S)}{q_n} a_n$ et $\tau_n = \inf\left\{k : \sum_{i=1}^{k} \mathbb{E}_S \left(u_i'^2 | \mathcal{F}_{i-1}\right) \geqslant r_n(S)\right\}$, on obtient $\tau_n = \inf\left\{k : \sum_{i=1}^{k} Q\left(\frac{x_i - z_0}{h_n}\right) \geqslant q_n\right\}$ et $\tilde{\zeta}_n' = \sum_{k=1}^{\tau_n} u_k'$.

Montrons maintenant que a_n ainsi que $r_n(S)$ convergent vers 1 uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$. On a premièrement

$$|a_n - 1| = |\mathbb{E}\left(\xi_1^2 \mathbb{I}_{|\xi_1| \leqslant q_n^{1/4}}\right) - \left(\mathbb{E}\xi_1 \mathbb{I}_{|\xi_1| \leqslant q_n^{1/4}}\right)^2 - 1| \leqslant 2\mathbb{E}\left(\xi_1^2 \mathbb{I}_{|\xi_1| > q_n^{1/4}}\right). \tag{2.24}$$

Puis d'après la définition de l'ensemble $\mathcal{P}_{\epsilon,L}$, il vient

$$\sup_{\mathcal{P}_{\epsilon,L}} \mathbb{E}\left(\xi_1^2 \mathbb{I}_{|\xi_1| > q_n^{1/4}}\right) \xrightarrow[n \to \infty]{} 0. \tag{2.25}$$

Ainsi le terme de gauche de l'inégalité (2.24) tend vers zéro quand n tend vers l'infini uniformément sur $\mathcal{P}_{\epsilon,L}$.
En tenant compte de l'hypothèse (2.4) et de l'inégalité

$$|r_n(S) - 1| \leqslant \left|\frac{G_n(S)}{q_n} - 1\right| + \frac{G_n(S)}{q_n}|a_n - 1|,$$

on prouve la convergence de $r_n(S)$ vers 1 uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$.

En appliquant le Lemme 2.5.6, cela montre d'une part la convergence en loi de ζ_n' vers $\eta \sim \mathcal{N}(0,1)$ uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$ car la fonction ρ du Lemme 2.5.6 ne dépend pas de la loi de la "martingale difference" en question. En fait, si Φ désigne la fonction de répartition de la loi gaussienne standard, on a

$$\begin{aligned} & \left|\mathbb{P}\left(\sum_{k=1}^{\tau_n} u_k' \leqslant x\right) - \Phi(x)\right| \\ \leqslant & \left|\mathbb{P}\left(\sum_{k=1}^{\tau_n} u_k' \leqslant x\right) - \Phi(x/\sqrt{r_n(S)})\right| + \left|\Phi(x) - \Phi(x/\sqrt{r_n(S)})\right|. \end{aligned}$$

Le second terme de droite de cette inégalité converge vers zéro uniformément sur $\mathcal{P}_{\epsilon,L}$, sur $\mathcal{U}_{z_0,\delta,n}$ et en x puisque $r_n(S) \to 1$ uniformément sur $\mathcal{P}_{\epsilon,L}$, sur $\mathcal{U}_{z_0,\delta,n}$ et car Φ est uniformément continue sur $\mathbb{R}$.

D'autre part, on a $\mathbb{E}|\tilde{\zeta}_n''| \to 0$ uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$. En effet, l'égalité $\mathbb{E}(\tilde{\zeta}_n''^2) = \dfrac{G_n(S)}{q_n}\mathbb{E}\left(\xi_1^2 \mathbb{I}_{|\xi_1|>q_n^{1/4}}\right)$ est immédiate. Ainsi (2.25) et l'inégalité de Cauchy-Schwarz impliquent

$$\sup_{\mathcal{P}_{\epsilon,L}} \sup_{S\in\mathcal{U}_{z_0,\delta,n}} \mathbb{E}_S|\tilde{\zeta}_n''| \to 0.$$

En utilisant l'inégalité de Markov, on montre que $(\tilde{\zeta}_n'')$ converge vers 0 en probabilité uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$.

Par conséquent $\tilde{\zeta}_n = \tilde{\zeta}_n' + \tilde{\zeta}_n''$ converge en loi vers $\eta \sim \mathcal{N}(0,1)$ uniformément sur $\mathcal{P}_{\epsilon,L}$ et sur $\mathcal{U}_{z_0,\delta,n}$. D'où l'assertion (2.23) et le théorème. □

2.5 Annexe A

Lemme 2.5.1 *Soient $\delta \in]0;1[$ et $\nu \in]0;1/4[$. Alors il existe un entier $n_{\delta,\nu} > 0$ tel que si $n \geqslant n_{\delta,\nu}$, on a $S_\nu \in \mathcal{U}_{z_0,\delta,n}$.*

Démonstration: Puisque $V_\nu(0) = 1$ et $\displaystyle\int_{-1}^{1} V_\nu(z)dz = 2$, il est clair que

$$\int_{-1}^{1} \left(S_\nu(z_0 + uh_n) - S_\nu(z_0)\right) du = 0.$$

Par ailleurs, on a de suite

$$|S_\nu'(x)| = \varphi_n^{-1} h_n^{-1} \left| V_\nu'\left(\frac{x - z_0}{h_n}\right)\right| \leqslant n^{-\alpha/(2\beta+1)} 2\nu^{-1} \int_{-1}^{1} |l'(z)|dz.$$

Donc $S_\nu \in \mathcal{U}_{z_0,\delta,n}$ pour

$$n \geqslant \left(\frac{2\delta}{\nu}\int_{-1}^{1}|l'(z)|dz\right)^{(2\beta+1)/\alpha}.$$

D'où le lemme. □

Lemme 2.5.2 *Pour tout $\nu \in]0;1/4[$, on a la convergence suivante :*

$$\varsigma_n^2 \xrightarrow[n\to\infty]{} \frac{1}{g^2(z_0,0)} \int_{-1}^{1} V_\nu^2(z)dz.$$

Démonstration: Fixons $\nu \in]0;1/4[$. Pour un entier n suffisamment grand pour que l'intervalle $[z_0 - h_n; z_0 + h_n]$ soit inclus dans $[0;1]$, on a

$$\varsigma_n^2 = \frac{1}{nh_n}\sum_{k=1}^{n} \frac{V_\nu^2\left(\frac{x_k - z_0}{h_n}\right)}{g^2(x_k, S_{\nu,u})} = \frac{1}{h_n}\int_{z_0-h_n}^{z_0+h_n} \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(x, S_{\nu,u})}\mu_n(dx) = \int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(x, S_{\nu,u})}\nu_n(dx)$$

avec $\mu_n = \dfrac{1}{n}\sum_{k=1}^{n}\delta_{k/n} =$ et $\nu_n = \dfrac{\mathbb{I}_{[z_0-h_n;z_0+h_n]}}{h_n}\mu_n$.
En utilisant les hypothèses (2.6) et (2.7) sur la fonction g, nous pouvons écrire pour tout $x \in [0;1]$:

$$\begin{aligned}
\left|\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(x,0)}\right| &\leqslant \frac{1}{g_*^4}\left|2g(x,0)L_{x,0}(S_{\nu,u}) + L^2_{x,0}(S_{\nu,u}) + \Gamma^2_{x,0}(S_{\nu,u})\right.\\
&+ \left. 2g(x,0)\Gamma_{x,0}(S_{\nu,u}) + 2L_{x,0}(S_{\nu,u})\Gamma_{x,0}(S_{\nu,u})\right|\\
&\leqslant \frac{1}{g_*^4}\left(2g^*C_0\|S_{\nu,u}\| + C_0^2\|S_{\nu,u}\|^2 + |\Gamma_{x,0}(S_{\nu,u})|^2\right.\\
&+ \left. 2g^*|\Gamma_{x,0}(S_{\nu,u})| + 2C_0\|S_{\nu,u}\||\Gamma_{x,0}(S_{\nu,u})|\right).
\end{aligned}$$

D'où

$$\begin{aligned}
&\left|\int_0^1\left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(x,0)}\right)\nu_n(dx)\right|\\
\leqslant\ & \frac{\|S_{\nu,u}\|}{g_*^4}\int_0^1\nu_n(dx)\left(2g^*C_0 + C_0^2\|S_{\nu,u}\| + \left(\sup_{x\in[0;1]}\frac{|\Gamma_{x,0}(S_{\nu,u})|}{\|S_{\nu,u}\|}\right)^2\|S_{\nu,u}\|\right.\\
+\ & \left. 2g^*\left(\sup_{x\in[0;1]}\frac{|\Gamma_{x,0}(S_{\nu,u})|}{\|S_{\nu,u}\|}\right) + 2C_0\left(\sup_{x\in[0;1]}\frac{|\Gamma_{x,0}(S_{\nu,u})|}{\|S_{\nu,u}\|}\right)\|S_{\nu,u}\|\right).
\end{aligned}$$

Comme (ν_n) converge étroitement vers la mesure de Dirac $2\delta_{z_0}$ quand $n\to\infty$, on a

$$\lim_{n\to\infty}\int_0^1\nu_n(dx) = 2 \quad\text{et}\quad \lim_{n\to\infty}\int_0^1\left(\frac{1}{g^2(x,0)} - \frac{1}{g^2(z_0,0)}\right)\nu_n(dx) = 0.$$

Ainsi, en tenant compte de l'hypothèse (2.8) et du fait que $\| S_{\nu,u} \|$ tende vers 0 quand $n\to\infty$, on obtient d'une part

$$\int_0^1\left(\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(z_0,0)}\right)\nu_n(dx) \xrightarrow[n\to\infty]{} 0.$$

D'autre part

$$\int_0^1\frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(z_0,0)}\nu_n(dx) = \frac{1}{nh_n}\sum_{k=1}^{n}\frac{V_\nu^2\left(\frac{x_k-z_0}{h_n}\right)}{g^2(z_0,0)} \xrightarrow[n\to\infty]{} \frac{1}{g^2(z_0,0)}\int_{-1}^1 V_\nu^2(y)dy.$$

A partir de là, si V_ν^* désigne le maximum de V_ν^2 sur $\mathbb{R}$, on a

$$\begin{aligned}
\left|\varsigma_n^2 - \frac{2}{g^2(z_0,0)}\right| &\leqslant \left|\int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(x,S_{\nu,u})}\nu_n(dx) - \int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(z_0,0)}\nu_n(dx)\right| \\
&+ \left|\int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(z_0,0)}\nu_n(dx) - \int_{-1}^1 \frac{V_\nu^2(z)}{g^2(z_0,0)}dz\right| \\
&\leqslant V_\nu^* \int_0^1 \left|\frac{1}{g^2(x,S_{\nu,u})} - \frac{1}{g^2(z_0,0)}\right| \nu_n(dx) \\
&+ \left|\int_0^1 \frac{V_\nu^2\left(\frac{x-z_0}{h_n}\right)}{g^2(z_0,0)}\nu_n(dx) - \int_{-1}^1 \frac{V_\nu^2(z)}{g^2(z_0,0)}dz\right|.
\end{aligned}$$

En faisant tendre n vers l'infini, ceci termine la démonstration du lemme. □

Lemme 2.5.3 *(Billingsley, 1999, Théorème 3.6, p. 32)*
Si X et X_n, $n \in \mathbb{N}$, sont des variables aléatoires positives et intégrables telles que

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}} X \quad et \quad \mathbb{E}(X_n) \xrightarrow[n\to\infty]{} \mathbb{E}(X),$$

alors la famille $(X_n)_{n\in\mathbb{N}}$ est uniformément intégrable.

Lemme 2.5.4 *(Ibragimov et Has'minskiĭ, 1981, Lemme 10.2, p. 157)*
Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$ de loi symétrique possédant une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Soit l une fonction définie sur $\mathbb{R}^d$, positive satisfaisant aux conditions

$$l(0) = 0 \quad et \quad l(x) = l(-x) \quad pour\ tout\ x \in \mathbb{R}^d,$$

et telle que pour tout $c > 0$, l'ensemble $\{x \in \mathbb{R}^d : l(x) < c\}$ soit convexe et $\mathbb{E}l(X+y) < \infty$ pour tout $y \in \mathbb{R}^d$.
Alors pour tout $y \in \mathbb{R}^d$, on a

$$\mathbb{E}l(X+y) \geqslant \mathbb{E}l(X).$$

Lemme 2.5.5 *La variance $\sigma_n^2(S)$ de ζ_n vérifie*

$$\sigma_n^2(S) \xrightarrow[n\to\infty]{} g^2(z_0,S).$$

Démonstration: Ecrivons

$$\sum_{k=1}^{n} Q\big(\frac{x_k - z_0}{h_n}\big) g^2(x_k, S) = n \int_{z_0-h_n}^{z_0+h_n} g^2(x, S)\mu_n(dx)$$

avec la mesure $\mu_n = \dfrac{1}{n}\sum_{k=1}^{n} \delta_{k/n}$.
On sait que $(\mu_n)_{n\geqslant 1}$ converge étroitement vers la mesure uniforme sur $[0;1]$.
De plus, pour n suffisamment grand,

$$\frac{1}{nh_n}\sum_{k=1}^{n} Q\big(\frac{x_k - z_0}{h_n}\big) g^2(x_k, S) = \int_{z_0-h_n}^{z_0+h_n} g^2(x, S)\nu_n(dx)$$

avec $\nu_n = \frac{\mu_n \mathbb{I}_{[z_0-h_n;z_0+h_n]}}{h}$.
Ainsi $(\nu_n)_{n\geqslant 1}$ converge étroitement vers la mesure de Dirac $2\delta_{z_0}$ quand $n \to \infty$.
On peut donc conclure en rappelant que $\dfrac{q_n}{\varphi_n^2} \xrightarrow[n\to\infty]{} 2$ et que $nh_n = \varphi_n^2$. □

Lemme 2.5.6 *(Freedman, 1971, pp. 90-91)*
Soient $\delta \in]0;1[$ et $r > 0$. Supposons que $(u_k)_{k\geqslant 0}$ est une "martingale difference" par rapport à la filtration $(\mathcal{F}_k)_{k\geqslant 0}$ telle que $|u_k| \leqslant \delta$ pour tout k et $\sum_{k=1}^{\infty} \mathbb{E}(u_k^2|\mathcal{F}_{k-1}) \geqslant r$.

Soit $\tau = \inf\left\{n : \sum_{k=1}^{n} \mathbb{E}(u_k^2|\mathcal{F}_{k-1}) \geqslant r\right\}$.
Alors il existe une fonction $\rho :]0;+\infty[\to [0;2]$, qui ne dépend pas de la distribution de la "martingale difference", telle que $\lim_{x\to 0} \rho(x) = 0$ et

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}\left(\sum_{k=1}^{\tau} u_k \leqslant x\right) - \Phi(x/\sqrt{r}) \right| \leqslant \rho(\delta/\sqrt{r}),$$

où Φ est la fonction de répartition de la loi gaussienne standard.

Chapitre 3

Modèle de régression : cas adaptatif

3.1 Introduction

Notre problème de régression est maintenant le suivant. Supposons qu'on observe des données à partir du modèle :

$$y_k = S(x_k) + g(x_k, S)\xi_k, \qquad k \in \{1, \ldots, n\}, \tag{3.1}$$

où $x_k = k/n$, $(\xi_k)_{k\in\{1,\ldots,n\}}$ sont des variables aléatoires indépendantes et identiquement distribuées selon la loi gaussienne standard. On s'intéresse à l'estimation de la fonction de régression S en un point fixe $z_0 \in]0;1[$. Soulignons encore que dans ce modèle de régression hétéroscédastique, la variance des bruits dépend de la fonction inconnue S et des régresseurs x_k. Ce type de modèle est utilisé en analyse financière ou encore en recherche médicale (voir par exemple Goldfeld et Quandt (1972)). Plus récemment, on retrouve ce genre de modèle dans Cai et Wang (2008) où les auteurs réduisent leur modèle de régression classique avec V pour variance dépendant uniquement des régresseurs à un modèle spécifique de régression où la fonction de régression est à peu de chose près V, tandis que les bruits dépendent de V.

On suppose que la fonction de régression appartient à une classe Höldérienne forte mais sa régularité β est inconnue. On travaille aussi avec la perte liée à l'erreur absolue et le risque correspondant. L'objectif est de trouver une vitesse de convergence adaptative et pour celle-ci de construire un estimateur adaptatif. Parce que β est inconnu, cette vitesse diffèrera ici de la vitesse de convergence obtenue dans le cas contraire. De nombreux travaux traitent de tels problèmes adaptatifs, on peut se référer par exemple à Barron, Birgé et Massart (1999), Galtchouk et Pergamenshchikov (2005a) et Lepskiĭ (1990, 1991, 1992a). Notre construction est basée sur celle que l'on peut trouver dans Lepskiĭ (1990) et Galtchouk et Pergamenshchikov (2001) pour l'estimation adaptative du coefficient de dérive d'un processus de diffusion. Comme dans le chapitre précédent, pour définir le risque d'un estimateur, on suit la méthode proposée par Galtchouk et Pergamenshchikov (2006a) dans le cas homoscédastique et non adaptatif.

Nous décrivons en détail le problème et les hypothèses formulées dans le prochain paragraphe. Nous donnons ensuite une borne inférieure asymptotique du risque minimax adaptatif. Puis nous obtenons une borne supérieure asymptotique du risque d'un estimateur à noyau.

3.2 Description du problème

On considère le modèle (3.1) où $S \in C^1([0;1])$ et $g : [0;1] \times C^1([0;1]) \longrightarrow \mathbb{R}_+^*$ sont des fonctions inconnues. On souhaite estimer la fonction de régression S en un point fixe $z_0 \in]0;1[$. On suppose que

$$S \in \bigcup_{M,K>0} \mathcal{H}(M,K,\beta), \text{où } \beta \in [\beta_*;\beta^*] \subset]1;2],$$

$$\mathcal{H}(M,K,\beta) = \left\{ S \in C^1[0;1] : ||S'|| \leqslant M, \sup_{y\in[0;1]} \frac{|S'(y)-S'(z_0)|}{|y-z_0|^\alpha} \leqslant K \right\},$$

avec $\beta = 1+\alpha$ et $||f|| = \sup_{x\in[0;1]} |f(x)|$.

Le paramètre de régularité β est supposé inconnu tandis que l'intervalle $[\beta_*;\beta^*]$ le contenant est considéré connu.

Le risque d'un estimateur $\hat{S}$ de $S(z_0)$ est défini pour une famille de classes $\mathcal{H}(M,K,\beta)$, $\beta \in [\beta_*;\beta^*]$, par

$$\mathcal{R}_{z_0,n}(\hat{S}) = \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S|\hat{S}-S(z_0)|,$$

où $N(\beta) = N_n(\beta) = \left(\frac{n}{\ln n}\right)^{\beta/(2\beta+1)}$.

Nous supposons que la fonction g est différentiable au sens de Fréchet en un point $\tilde{S}_0 \in C^2([0;1])$ par rapport à $S \in C^1([0;1])$ uniformément en $x \in [0;1]$, i.e. pour tous $S \in C^1([0;1])$

$$g(x,S) = g(x,\tilde{S}_0) + L_{x,\tilde{S}_0}(S-\tilde{S}_0) + \Gamma_{x,\tilde{S}_0}(S-\tilde{S}_0),$$

où l'application linéaire $L_{x,\tilde{S}_0}$ est nulle sur $C^1([0;1])$ pour tout $x \in [0;1]$, et le terme résiduel $\Gamma_{x,\tilde{S}_0}(S)$ satisfait à la propriété

$$\lim_{\|S\|\to 0} \sup_{x\in[0;1]} \Gamma_{x,\tilde{S}_0}(S)/\parallel S \parallel = 0.$$

Ainsi pour tous $S \in C^1([0;1])$ et $x \in [0;1]$, on a :

$$g^2(x,S) = g^2(x,\tilde{S}_0) + \tilde{\Gamma}_{x,\tilde{S}_0}(S-\tilde{S}_0),$$

où

$$\lim_{\|S\|\to 0} \sup_{x\in[0;1]} \tilde{\Gamma}_{x,\tilde{S}_0}(S)/ \parallel S \parallel = 0.$$

On suppose aussi qu'il existe une constante positive G^* telle que pour tous $S \in C^1([0;1])$ et $x \in [0;1]$, on a :

$$\left|\tilde{\Gamma}_{x,S_0}(S)\right| \leqslant G^* |S(x)|^2. \tag{3.2}$$

Enfin nous supposons connues deux constantes $g_* > 0$ et $g^* < \infty$ telles que

$$g_* \leqslant \inf_{0\leqslant x\leqslant 1} \inf_{S\in C^1([0;1])} g(x,S) \leqslant \sup_{0\leqslant x\leqslant 1} \sup_{S\in C^1([0;1])} g(x,S) \leqslant g^*.$$

3.3 Borne inférieure

Dans ce paragraphe, on notera $d_n = c\dfrac{n}{\ln n}$, où $c > 0$ sera précisée ultérieurement, et $\tilde{N}(\beta) = d_n^{\beta/(2\beta+1)}$, $\tilde{h}(\beta) = d_n^{-1/(2\beta+1)}$. On montre qu'avec cette vitesse $\tilde{N}(\beta)$, la borne inférieure du risque minimax est strictement positive, ce qui implique qu'elle l'est aussi pour la vitesse $N(\beta) = \left(\dfrac{n}{\ln n}\right)^{\beta/(2\beta+1)}$.

Théorème 3.3.1 *Pour des constantes M et K suffisamment grandes, on a :*

$$\liminf_{n\to\infty} \inf_{\tilde{S}_n} \mathcal{R}_{z_0,n}(\tilde{S}_n) > 0.$$

Démonstration: Pour alléger les notations, on notera désormais $\tilde{N}(\beta_*) = N_*$, $\tilde{N}(\beta^*) = N^*$ et $\tilde{h}(\beta_*) = h_*$.
Remarquons tout d'abord que, pour tout estimateur $\tilde{S}_n$ de $S(z_0)$,

$$\mathcal{R}_{z_0,n}(\tilde{S}_n) \geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{g^*} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \mathbb{E}_S \tilde{N}(\beta) |\tilde{S}_n - S(z_0)|.$$

Notons pour $\nu \in]0;1/4[$,

$$S_\nu(y) = \frac{1}{N^*}\tilde{S}_0\left(\frac{y-z_0}{h^*}\right) + \frac{1}{N_*} V_\nu\left(\frac{y-z_0}{h_*}\right),$$

avec

$$\begin{aligned} V_\nu(x) &= \nu^{-1}\int_{-\infty}^{\infty} \tilde{Q}_\nu(u) l\left(\frac{u-x}{\nu}\right) du, \\ \tilde{Q}_\nu(u) &= \mathbb{I}_{\{|u|\leqslant 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\leqslant|u|\leqslant 1-\nu\}}, \\ l(z) &= L\exp\left(-(1-|z|^2)^{-1}\right)\mathbb{I}_{\{|z|\leqslant 1\}}, \end{aligned}$$

et L la constante de normalisation telle que $\int_{-1}^{1} l(z)dz = 1$.
On note aussi

$$S_0(y) = \frac{1}{N^*}\tilde{S}_0\left(\frac{y-z_0}{h^*}\right).$$

En se servant en particulier du fait que $\tilde{S}_0$ est de classe C^2, on montre aisément que pour certains M et K suffisamment grand, $S_\nu \in \mathcal{H}(M,K,\beta_*)$ et $S_0 \in \mathcal{H}(M,K,\beta^*)$.

Soit $\mathbb{P}_0$ (respectivement $\mathbb{P}_\nu$) la loi du vecteur $(y_k)_{k=1,\dots,n}$ lorsque $y_k = S_0(x_k)+g(x_k,S_0)\xi_k$ (respectivement $y_k = S_\nu(x_k)+g(x_k,S_\nu)\xi_k$). On note $\frac{d\mathbb{P}_0}{d\mathbb{P}_\nu}(y)$ la dérivée de Radon-Nikodym de $\mathbb{P}_0$ par rapport à $\mathbb{P}_\nu$ au point $y=(y_k)_{k=1,\dots,n}$. On écrit alors avec $\theta_n = N_*(\tilde{S}_n - S_0(z_0))$:

$$\begin{aligned}
\mathcal{R}_{z_0,n}(\tilde{S}_n) &\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{g^*}\max\left(\mathbb{E}_{S_0}N^*|\tilde{S}_n - S_0(z_0)|, \mathbb{E}_{S_\nu}N_*|\tilde{S}_n - S_\nu(z_0)|\right)\\
&= \frac{c^{-\beta^*/(2\beta^*+1)}}{g^*}\max\left(\mathbb{E}_{S_0}\frac{N^*}{N_*}|\theta_n|, \mathbb{E}_{S_\nu}|1-\theta_n|\right),\\
&\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{2g^*}\left(\frac{N^*}{N_*}\mathbb{E}_{S_\nu}|\theta_n|\frac{d\mathbb{P}_0}{d\mathbb{P}_\nu}(y) + \mathbb{E}_{S_\nu}|1-\theta_n|\right).
\end{aligned}$$

Ecrivons

$$\frac{d\mathbb{P}_0}{d\mathbb{P}_\nu}(y) = \prod_{k=1}^{n}\frac{g(x_k,S_\nu)}{g(x_k,S_0)}\exp\left(-\frac{1}{2}\sum_{k=1}^{n}\left(\frac{y_k - S_0(x_k)}{g(x_k,S_0)}\right)^2 - \left(\frac{y_k - S_\nu(x_k)}{g(x_k,S_\nu)}\right)^2\right),$$

puis sous la loi $\mathbb{P}_\nu$, en notant $\zeta_k = \frac{y_k - S_\nu(x_k)}{g(x_k,S_\nu)}$, il vient :

$$\begin{aligned}
\ln\frac{d\mathbb{P}_0}{d\mathbb{P}_\nu}(y) &= \sum_{k=1}^{n}\ln\frac{g(x_k,S_\nu)}{g(x_k,S_0)} - \frac{1}{2}\sum_{k=1}^{n}\left(\left(\frac{y_k - S_0(x_k)}{g(x_k,S_0)}\right)^2 - \left(\frac{y_k - S_\nu(x_k)}{g(x_k,S_\nu)}\right)^2\right)\\
&= \frac{1}{2}\sum_{k=1}^{n}\ln\left(1+\frac{g^2(x_k,S_\nu)}{g^2(x_k,S_0)}-1\right) - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{g^2(x_k,S_\nu)}{g^2(x_k,S_0)}-1\right)\zeta_k^2\\
&- \sum_{k=1}^{n}\frac{g(x_k,S_\nu)}{g^2(x_k,S_0)}(S_\nu(x_k)-S_0(x_k))\zeta_k - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{S_\nu(x_k)-S_0(x_k)}{g(x_k,S_0)}\right)^2\\
&= \frac{1}{2}\sum_{k=1}^{n}\ln\left(1+\frac{\tilde{\Gamma}_{x_k,S_0}(S_\nu - S_0)}{g(x_k,S_0)}\right) - \frac{1}{2}\sum_{k=1}^{n}\frac{\tilde{\Gamma}_{x_k,S_\nu}(S_\nu - S_0)}{g(x_k,S_0)}\zeta_k^2\\
&- \sum_{k=1}^{n}\frac{g(x_k,S_\nu)}{g^2(x_k,S_0)}(S_\nu(x_k)-S_0(x_k))\zeta_k - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{S_\nu(x_k)-S_0(x_k)}{g(x_k,S_0)}\right)^2\\
&= \frac{1}{2}A_1 - \frac{1}{2}A_2 + A_3 - \frac{1}{2}A_4,
\end{aligned}$$

où

$$A_1 = \sum_{k=1}^{n} \ln\left(1 + \frac{\tilde{\Gamma}_{x_k,S_0}(S_\nu - S_0)}{g(x_k,S_0)}\right) \quad , \qquad A_2 = \sum_{k=1}^{n} \frac{\tilde{\Gamma}_{x_k,S_\nu}(S_\nu - S_0)}{g(x_k,S_0)}\zeta_k^2,$$

$$A_3 = -\sum_{k=1}^{n} \frac{g(x_k,S_\nu)}{g^2(x_k,S_0)}(S_\nu(x_k) - S_0(x_k))\zeta_k \quad \text{et} \quad A_4 = \sum_{k=1}^{n}\left(\frac{S_\nu(x_k) - S_0(x_k)}{g(x_k,S_0)}\right)^2.$$

Soit C_n l'événement $\left\{\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right)\zeta_k^2 < 2\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right)\right\}$.

Posons $a_\nu := \int_{-1}^{1} V_\nu^2(z)dz$. Sur l'événement C_n, on a :

$$|A_2| \leqslant \frac{G^*}{g_* N_*^2}\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right)\zeta_k^2 \leqslant \frac{2G^*}{g_* N_*^2}\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right), \tag{3.3}$$

ce dernier majorant étant équivalent à $\frac{2a_\nu G^*}{g_*}\frac{nh_*}{N_*^2}$ lorsque $n \to \infty$.
D'autre part d'après l'hypothèse (3.2),

$$\begin{aligned} |A_1| &\leqslant \sum_{k=1}^{n}\left|\frac{\tilde{\Gamma}_{x_k,S_0}(S_\nu - S_0)}{g(x_k,S_0)}\right| \\ &\leqslant \frac{G^*}{g_*}\sum_{k=1}^{n}|S_\nu(x_k) - S_0(x_k)|^2 \leqslant \frac{G^*}{g_* N_*^2}\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right), \end{aligned} \tag{3.4}$$

ce dernier majorant étant équivalent à $\frac{a_\nu G^*}{g_*}\frac{nh_*}{N_*^2}$ lorsque $n \to \infty$.
Puis

$$|A_4| \leqslant \frac{1}{g_*^2}\sum_{k=1}^{n}(S_\nu(x_k) - S_0(x_k))^2 \leqslant \frac{1}{g_*^2 N_*^2}\sum_{k=1}^{n} V_\nu^2\left(\frac{x_k - z_0}{h_*}\right), \tag{3.5}$$

ce dernier majorant étant équivalent à $\frac{a_\nu}{g_*^2}\frac{nh_*}{N_*^2}$ lorsque $n \to \infty$.
Enfin sous $\mathbb{P}_\nu$, A_3 est une variable aléatoire normale centrée de variance

$$v_n^2 := \sum_{k=1}^{n} \frac{g^2(x_k,S_\nu)}{g^4(x_k,S_0)}(S_\nu(x_k) - S_0(x_k))^2,$$

donc on peut écrire $A_3 = v_n\eta_n$, avec $\eta_n \sim \mathcal{N}(0,1)$ sous la probabilité $\mathbb{P}_\nu$.

Ainsi sur l'événement C_n, en utilisant (3.3),(3.4) et (3.5) on obtient :

$$\begin{aligned}
\ln \frac{d\mathbb{P}_0}{d\mathbb{P}_\nu}(y) &\geqslant v_n \eta_n - \frac{1}{2}(|A_1| + |A_2| + |A_4|) \\
&\geqslant v_n \eta_n - \frac{a_\nu + o_n(1)}{2} \left(\frac{3G^*}{g_*} + \frac{1}{g_*^2} \right) \frac{n h_*}{N_*^2} \\
&= v_n \eta_n - \frac{a_\nu + o_n(1)}{2} \left(\frac{3G^*}{g_*} + \frac{1}{g_*^2} \right) \frac{\ln n}{c}.
\end{aligned}$$

Introduisons maintenant l'événement $B_n = \{\eta_n > 0\}$, la notation $\overline{\beta} = \dfrac{\beta^*}{2\beta^* + 1} - \dfrac{\beta_*}{2\beta_* + 1}$ et remarquons que $a_\nu \in [0; 8]$. Choisissons $c > \left(\dfrac{3G^*}{g_*} + \dfrac{1}{g_*^2} \right) \dfrac{4}{\overline{\beta}}$. Alors pour n suffisamment grand, il vient :

$$\begin{aligned}
\mathcal{R}_{z_0,n}(\tilde{S}_n) &\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{2g_*} \left(\mathbb{E}_{S_\nu} d_n^{\overline{\beta}} \frac{d\mathbb{P}_0}{d\mathbb{P}_{S_\nu}}(y) \mathbb{I}_{B_n \cap C_n} |\theta_n| + \mathbb{E}_{S_\nu} |\theta_n - 1| \right) \\
&\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{2g_*} \left(e^{\left(\overline{\beta} \ln d_n - \frac{a_\nu + o_n(1)}{2c} \left(\frac{3G^*}{g_*} + \frac{1}{g_*^2} \right) \ln n \right)} \mathbb{E}_{S_\nu} |\theta_n| \mathbb{I}_{B_n \cap C_n} + \mathbb{E}_{S_\nu} |\theta_n - 1| \right) \\
&\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{2g_*} \left(\mathbb{E}_{S_\nu} |\theta_n| \mathbb{I}_{B_n \cap C_n} + \mathbb{E}_{S_\nu} |\theta_n - 1| \mathbb{I}_{B_n \cap C_n} \right) \\
&\geqslant \frac{c^{-\beta^*/(2\beta^*+1)}}{2g_*} \left(\mathbb{E}_{S_\nu} \mathbb{I}_{B_n \cap C_n} (|\theta_n| + 1 - |\theta_n|) \right) \\
&= \frac{c^{-\beta^*/(2\beta^*+1)}}{2g_*} \mathbb{P}_{S_\nu}(B_n \cap C_n). \qquad (3.6)
\end{aligned}$$

Or

$$\begin{aligned}
\mathbb{P}_\nu(C_n^c) &= \mathbb{P}_\nu \left(\sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) \zeta_k^2 \geqslant 2 \sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) \right) \\
&= \mathbb{P}_\nu \left(\sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) (\zeta_k^2 - 1) \geqslant \sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) \right) \\
&= \mathbb{P}_\nu \left(\sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) (\zeta_k^2 - 1) \geqslant n h_* (a_\nu + o_n(1)) \right) \\
&\leqslant \frac{1}{(n h_*)^2 (a_\nu + o_n(1))^2} \mathbb{E}_\nu \left(\sum_{k=1}^n V_\nu^2 \left(\frac{x_k - z_0}{h_*} \right) (\zeta_k^2 - 1) \right)^2 \\
&= \frac{3}{n h_* (a_\nu + o_n(1))^2} \frac{1}{n h_*} \sum_{k=1}^n V_\nu^4 \left(\frac{x_k - z_0}{h_*} \right).
\end{aligned}$$

Comme $\frac{1}{nh_*}\sum_{k=1}^{n} V_\nu^4\left(\frac{x_k - z_0}{h_*}\right) \xrightarrow[n\to\infty]{} \int_{-1}^{1} V_\nu^4(x)dx$, il vient $\mathbb{P}_\nu(C_n^c) \xrightarrow[n\to\infty]{} 0$.
Puisque $\mathbb{P}_\nu(B_n) = 1/2$, on en déduit que $\mathbb{P}_\nu(B_n \cap C_n) \xrightarrow[n\to\infty]{} 1/2$.
En passant à la limite quand $n \to \infty$ dans (3.6), on obtient le théorème. □

3.4 Estimation adaptative

Les travaux de Galtchouk et Pergamenshchikov (2006a) traitent du cas homoscédastique et non adaptatif, en étudiant l'estimateur à noyau

$$S_h^*(z_0) = \frac{1}{q_n(h)}\sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h}\right) y_k,$$

où $Q = \mathbb{I}_{[-1,1]}$, $h = n^{-1/(2\beta+1)}$ et $q_n(h) = \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h}\right)$. En prenant en compte le fait que β est inconnu, on ne peut pas utiliser un tel estimateur à noyau car sa fenêtre h dépend de β. C'est la raison pour laquelle nous formons la partition de l'intervalle $[\beta_*; \beta^*]$ de la manière suivante :

$$\beta_l = \beta_* + l\frac{\beta^* - \beta_*}{[\ln n]}, \quad l = 0, \ldots, [\ln n],$$

où $[a]$ désigne la partie entière du réel a, et on définit les fenêtres correspondantes $h_l = h(\beta_l) = \left(\frac{n}{\ln n}\right)^{-1/(2\beta_l+1)}$. Posons ensuite

$$\hat{l} = \max\left\{0 \leqslant l \leqslant [\ln n] : \max_{0\leqslant j\leqslant l}\left(\left|S_{h_l}^*(z_0) - S_{h_j}^*(z_0)\right| - \frac{\lambda}{N_j}\right) \leqslant 0\right\},$$

où $N_j = N(\beta_j)$ et $\lambda > \frac{4K}{\beta_*(\beta_* + 1)} + 2\sqrt{2}g^*\left(\frac{\beta^*}{2\beta^* + 1} - \frac{\beta_*}{2\beta_* + 1}\right)^{1/2}$.
Remarquons que $\hat{l}$ existe bien puisque l'ensemble ci-dessus contient l'indice 0.
L'estimateur adaptatif est alors défini par $\hat{S}_n = S_{h_{\hat{l}}}^*(z_0)$.

Par ailleurs, on associe au paramètre inconnu β l'unique entier $l(\beta) \in \{0, \ldots, [\ln n]-1\}$ tel que $\beta_{l(\beta)} \leqslant \beta < \beta_{l(\beta)+1}$.

Le théorème suivant donne une borne supérieure du risque de l'estimateur construit plus haut.

Théorème 3.4.1 *On a*

$$\limsup_{n\to\infty} \mathcal{R}_{z_0,n}(\hat{S}_n) \leqslant \lambda\frac{e^{(\beta^*-\beta_*)/(2\beta_*+1)^2}}{g_*} + \frac{K}{g_*\beta_*(\beta_* + 1)},$$

donc l'estimateur $\hat{S}_n$ est adaptatif en vitesse de convergence.

DÉMONSTRATION: Ecrivons l'erreur de l'estimateur $\hat{S}_n$:

$$|\hat{S}_n - S(z_0)| = |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}\geqslant l(\beta)\}} + |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}<l(\beta)\}} =: I_1 + I_2,$$

où $I_1 = \sum_{j=l(\beta)}^{[\ln n]} |\hat{S}_n - S(z_0)|\mathbb{I}_{\{\hat{l}=j\}} = \sum_{j=l(\beta)}^{[\ln n]} |S^*_{h_j}(z_0) - S(z_0)|\mathbb{I}_{\{\hat{l}=j\}}$.

Pour tout $j = 0, \dots, [\ln n]$, on note $S^*_{h_j}(z_0) - S(z_0) = \dfrac{\zeta_n(\beta_j)}{\sqrt{q_n(\beta_j)}} + B_n(\beta_j)$, avec

$$\begin{aligned} q_n(\beta_j) &= q_n(h(\beta_j)), \\ \zeta_n(\beta_j) &= \frac{1}{\sqrt{q_n(\beta_j)}} \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_j}\right) g(x_k, S)\xi_k, \\ B_n(\beta_j) &= \frac{1}{q_n(\beta_j)} \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_j}\right) (S(x_k) - S(z_0)). \end{aligned}$$

La variance de $\zeta_n(\beta_j)$ est alors $\sigma_n^2(\beta_j, S) := \dfrac{1}{q_n(\beta_j)} \sum_{k=1}^{n} Q\left(\dfrac{x_k - z_0}{h_j}\right) g^2(x_k, S)$.

Nous avons $\left|\dfrac{\zeta_n(\beta_j)}{\sqrt{q_n(\beta_j)}}\right| = |Z_n(\beta_j)|\dfrac{\sigma_n(\beta_j, S)}{\sqrt{q_n(\beta_j)}}$, avec $Z_n(\beta_j) \sim \mathcal{N}(0,1)$.
De plus, en reprenant la décomposition établie dans la démonstration du Théorème 2.3.2, on a

$$B_n(\beta_j) = \frac{nh_j}{q_n(\beta_j)} \int_{-1}^{1} (S(z_0 + h_j u) - S(z_0))\, du + \frac{nh_j}{q_n(\beta_j)} R_n(\beta_j),$$

où

$$|R_n(\beta_j)| \leqslant \frac{6M}{n}.$$

Rappelons également qu'en utilisant la condition Höldérienne dans l'expression

$$\int_{-1}^{1} \big(S(z_0 + h_j u) - S(z_0)\big) du = \int_{-1}^{1} \left(\int_{z_0}^{z_0 + uh_j} (S'(t) - S'(z_0)) dt \right) du, \tag{3.7}$$

on obtient,

$$\left| \int_{-1}^{1} \big(S(z_0 + h_j u) - S(z_0)\big) du \right| \leqslant \frac{2K}{\beta(\beta+1)} h_j^{\beta}.$$

Il vient alors $|B_n(\beta_j)| \leqslant \dfrac{2K}{\beta(\beta+1)} \dfrac{nh_j^{1+\beta}}{q_n(\beta_j)} + 6M \dfrac{h_j}{q_n(\beta_j)}$, puis

$$|S^*_{h_j}(z_0) - S(z_0)| \leqslant \frac{2K}{\beta(\beta+1)} \frac{nh_j^{1+\beta}}{q_n(\beta_j)} + 6M \frac{h_j}{q_n(\beta_j)} + |Z_n(\beta_j)| \frac{\sigma_n(\beta_j, S)}{\sqrt{q_n(\beta_j)}}. \tag{3.8}$$

D'où

$$\begin{aligned}
I_1 &\leqslant \sum_{j=l(\beta)}^{[\ln n]} \Big(|S^*_{h_j}(z_0) - S^*_{h_{l(\beta)}}(z_0)| \mathbb{I}_{\{\hat{l}=j\}} \Big) + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \mathbb{I}_{\{\hat{l} \geqslant l(\beta)\}} \\
&\leqslant \sum_{j=l(\beta)}^{[\ln n]} \frac{\lambda}{N_j} \mathbb{I}_{\{\hat{l}=j\}} + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \\
&\leqslant \frac{\lambda}{N_{l(\beta)}} + |S^*_{h_{l(\beta)}}(z_0) - S(z_0)| \\
&\leqslant \frac{\lambda}{N(\beta)} e^{(\beta^* - \beta_*)/(2\beta_*+1)^2} + \frac{2Knh(\beta_{l(\beta)})^{1+\beta}}{\beta(\beta+1) q_n(\beta_{l(\beta)})} + \frac{6Mh(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} + |Z_n(\beta_{l_(\beta)})| \frac{\sigma_n(\beta_{l(\beta)}, S)}{\sqrt{q_n(\beta_{l(\beta)})}}.
\end{aligned}$$

Montrons que d'une part

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S I_1 \leqslant \lambda \frac{e^{(\beta^*-\beta_*)/(2\beta_*+1)^2}}{g_*} + \frac{K}{g_*\beta_*(\beta_*+1)}. \tag{3.9}$$

Comme $N(\beta)h(\beta_{l(\beta)})^\beta \leqslant 1$ et $q_n(\beta_{l(\beta)}) \sim 2nh(\beta_{l(\beta)})$, on a la limite suivante :

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{2K}{\beta(\beta+1)} \frac{N(\beta)}{g(z_0,S)} \frac{nh(\beta_{l(\beta)})^{1+\beta}}{q_n(\beta_{l(\beta)})} \leqslant \frac{K}{g_*\beta_*(\beta_*+1)}. \tag{3.10}$$

De plus, il n'est pas difficile de voir que

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(z_0,S)} 6M \frac{h(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} = 0. \tag{3.11}$$

Puisque $\sigma_n(\beta_{l(\beta)}, S)$ est borné par g^* et comme

$$\frac{N(\beta)^2}{q_n(\beta_{l(\beta)})} = \frac{nh(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} \left(\frac{n}{\ln n}\right)^{2(\beta-\beta_{l(\beta)})/(2\beta+1)(2\beta_{l(\beta)}+1)} \frac{1}{\ln n} \leqslant \frac{nh(\beta_{l(\beta)})}{q_n(\beta_{l(\beta)})} e^{2(\beta^*-\beta_*)} \frac{1}{\ln n},$$

il vient

$$\limsup_{n\to\infty} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S |Z_n(\beta_{l(\beta)})| \frac{\sigma_n(\beta_{l(\beta)}, S)}{\sqrt{q_n(\beta_{l(\beta)})}} = 0. \tag{3.12}$$

Finalement, (3.9) découle de (3.10) – (3.12).

Ici rappelons que $I_2 = |\hat{S}_n - S(z_0)| \mathbb{I}_{\{\hat{l} < l(\beta)\}}$ et montrons d'autre part que

$$\lim_{n\to\infty} \sup_{\beta\in[\beta_*;\beta^*]} \sup_{S\in\mathcal{H}(M,K,\beta)} \frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S I_2 = 0. \tag{3.13}$$

On a simplement $\{\hat{l} < l(\beta)\} = \bigcup_{j=0}^{l(\beta)-1} \{\hat{l} = j\}$ et par définition de $\hat{l}$,

$$\{\hat{l} = j\} \subset \bigcup_{i=0}^{j+1} \left(\left| S^*_{h_i}(z_0) - S^*_{h_{j+1}}(z_0) \right| > \frac{\lambda}{N_i} \right).$$

Il est utile de remarquer que pour $0 \leqslant i \leqslant l \leqslant l(\beta)$, on a $h_i \leqslant h_l \leqslant h(\beta)$ et $q_n(\beta_i) \leqslant q_n(\beta_l)$. En notant alors $Z^*_n = \max_l |Z_n(\beta_l)|$ et utilisant (3.8) ainsi que le Lemme 3.4.2, on obtient pour $0 \leqslant i \leqslant l \leqslant l(\beta)$:

$$\begin{aligned}
&\left\{ \left| S^*_{h_l}(z_0) - S^*_{h_i}(z_0) \right| > \frac{\lambda}{N_i} \right\} \subset \left\{ \left| S^*_{h_l}(z_0) - S(z_0) \right| + \left| S(z_0) - S^*_{h_i}(z_0) \right| > \frac{\lambda}{N_i} \right\} \\
&\subset \left\{ \frac{4K}{\beta(\beta+1)} h(\beta)^\beta + \frac{12M}{nh_i} h(\beta) + \frac{2g^*}{\sqrt{q_n(\beta_i)}} Z^*_n > \frac{\lambda}{N_i} \right\} \\
&\subset \left\{ \frac{2g^*}{\sqrt{q_n(\beta_i)}} Z^*_n > \frac{\lambda - \frac{4K}{\beta_*(\beta_*+1)} - 12Mh(\beta^*)}{N_i} \right\},
\end{aligned}$$

car $nh_i > N_i$ et $N_i < h(\beta)^{-\beta}$.
Signalons que les inclusions précédentes sont vraies pour n suffisamment grand et nous considérerons que tel est le cas jusqu'à la fin de la démonstration.
Comme $\frac{q_n(\beta_i)}{N_i^2} = \frac{q_n(\beta_i)}{nh_i} \ln n$, le Lemme 3.4.2 nous donne $\sqrt{\ln n} \leqslant \frac{\sqrt{q_n(\beta_i)}}{N_i} \leqslant \sqrt{3 \ln n}$.
En posant $\Lambda^*_n := \lambda - \frac{4K}{\beta_*(\beta_*+1)} - 12Mh(\beta^*)$, on peut écrire

$$(\hat{l} = j) \ \subset \ \bigcup_{i=0}^{j+1} \left(\left| S^*_{h_i}(z_0) - S^*_{h_{j+1}}(z_0) \right| > \frac{\lambda}{N_i} \right) \subset \left\{ Z^*_n > \frac{\Lambda^*_n}{2g^*} \sqrt{\ln n} \right\} =: A^*_n$$

et

$$I_2 \leqslant \sum_{j=0}^{l(\beta)-1} |S^*_{h_j}(z_0) - S(z_0)| \mathbb{I}_{\{A^*_n\}}.$$

Pour obtenir (3.13), on écrit les inégalités

$$\begin{aligned}
&\frac{N(\beta)}{g(z_0,S)} \mathbb{E}_S I_2 \leqslant \frac{N(\beta)}{g(z_0,S)} \sum_{j=0}^{l(\beta)-1} \mathbb{E}_S \left(\frac{2Knh_j^{1+\beta}}{\beta(\beta+1) q_n(\beta_j)} + \frac{6Mh_j}{q_n(\beta_j)} + |Z_n(\beta_j)| \frac{\sigma(\beta_j,S)}{\sqrt{q_n(\beta_j)}} \right) \mathbb{I}_{\{A^*_n\}} \\
&\leqslant \ \sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{g(z_0,S)} \frac{2K}{\beta(\beta+1)} \frac{nh_j^{1+\beta}}{q_n(\beta_j)} \mathbb{P}_S\left(A^*_n\right) + \sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{g(z_0,S)} \frac{6Mh_j}{q_n(\beta_j)} \mathbb{P}_S(A^*_n) \\
&+ \ \sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{g(z_0,S)} \frac{\sigma_n(\beta_j,S)}{\sqrt{q_n(\beta_j)}} \mathbb{E}_S\left(|Z_n(\beta_j)| \mathbb{I}_{(A^*_n)}\right)
\end{aligned}$$

et on étudie le comportement asymptotique de chaque terme du membre de droite. Soient $Z \sim \mathcal{N}(0,1)$ et $A_n := \left\{|Z| > \frac{\Lambda_n^*}{2g^*}\sqrt{\ln n}\right\}$. Pour le premier terme, on a

$$\begin{aligned}
&\sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{g(z_0,S)} \frac{2K}{\beta(\beta+1)} \frac{nh_j^{1+\beta}}{q_n(\beta_j)} \mathbb{P}_S\left(A_n^*\right) \leqslant \frac{2K}{\beta(\beta+1)g_*} \sum_{j=0}^{l(\beta)-1} N(\beta) h_j^{\beta} \mathbb{P}_S\left(A_n^*\right) \\
\leqslant\ & \frac{2K}{\beta(\beta+1)g_*} \sum_{j=0}^{l(\beta)-1} \left(\frac{n}{\ln n}\right)^{2\beta(\beta_j-\beta)/(2\beta_j+1)(2\beta+1)} \mathbb{P}_S\left(A_n^*\right) \\
\leqslant\ & [\ln n]^2 \frac{2K}{\beta(\beta+1)g_*} \left(\frac{n}{\ln n}\right)^{2\beta(\beta_{l(\beta)-1}-\beta)/(2\beta_{l(\beta)-1}+1)(2\beta+1)} \mathbb{P}_S\left(A_n\right) \\
\leqslant\ & [\ln n]^2 \frac{2K}{\beta(\beta+1)g_*} \left(\frac{n}{\ln n}\right)^{2\beta(\beta_{l(\beta)-1}-\beta)/(2\beta+1)^2} 2\frac{2g^*}{\Lambda_n^*\sqrt{\ln n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\Lambda_n^*)^2}{4g^{*2}}\ln n} \\
\leqslant\ & \frac{8Kg^*}{\beta_*(\beta_*+1)\sqrt{2\pi}g_*\Lambda_n^*} e^{-2\beta_*/(2\beta^*+1)^2} e^{2\beta^*/(2\beta_*+1)^2} \frac{[\ln n]^2}{\sqrt{\ln n}} n^{-\frac{(\Lambda_n^*)^2}{4g^{*2}}}.
\end{aligned}$$

Puisque $\lambda > \dfrac{4K}{\beta_*(\beta_*+1)}$, cette dernière expression tend vers zéro quand n tend vers l'infini.
On procède de même avec le second terme.
En utilisant l'inégalité de Cauchy-Schwarz pour le troisième, il vient

$$\begin{aligned}
&\frac{N(\beta)}{g(z_0,S)} \sum_{j=0}^{l(\beta)-1} \mathbb{E}_S\left(|Z_n(\beta_j)| \frac{\sigma_n(\beta_j,S)}{\sqrt{q_n(\beta_j)}} \mathbb{I}_{\{A_n^*\}}\right) \leqslant \frac{g^*}{g_*} \sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{\sqrt{q_n(\beta_j)}} \left(\mathbb{P}_S\left(A_n^*\right)\right)^{1/2} \\
\leqslant\ & \frac{g^*}{g_*} \sum_{j=0}^{l(\beta)-1} \frac{N(\beta)}{\sqrt{q_n(\beta_j)}} [\ln n]^{1/2} \left(\mathbb{P}_S\left(A_n\right)\right)^{1/2} \\
\leqslant\ & \frac{2(g^*)^{3/2}}{g_*\sqrt{\Lambda_n^*\sqrt{2\pi}}} \frac{[\ln n]^{3/2}}{\sqrt{\sqrt{\ln n}}\sqrt{q_n(\beta_*)}} N(\beta^*) n^{-(\Lambda_n^*)^2/8g^{*2}}.
\end{aligned}$$

D'après l'expression de λ, cette dernière expression converge vers zéro quand n tend vers l'infini. Finalement, nous avons prouvé (3.13) qui, mis en relation avec (3.9), complète la démonstration. □

Lemme 3.4.2 *Il existe un entier N_* tel que si $n \geqslant N_*$, alors*

$$1 \leqslant \frac{q_n(\beta_i)}{nh_i} \leqslant 3,$$

pour tout $i = 0, \ldots, [\ln n]$.

Démonstration: En écrivant

$$\frac{q_n(\beta_i)}{nh_i} = \frac{1}{nh_i}\sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_i}\right) = \frac{[n(z_0+h_i)] - [n(z_0-h_i)]}{nh_i},$$

les inégalités suivantes

$$\begin{array}{rcl} [n(z_0+h_i)] \leqslant & n(z_0+h_i) & < [n(z_0+h_i)] + 1, \\ -1 - [n(z_0-h_i)] < & -n(z_0-h_i) & \leqslant -[n(z_0-h_i)], \end{array}$$

montrent que $2 - \dfrac{1}{nh_i} \leqslant \dfrac{q_n(\beta_i)}{nh_i} \leqslant 2 + \dfrac{1}{nh_i}$.
Or il existe un entier N_* tel que si $n \geqslant N_*$, alors $nh(\beta_*) \geqslant 1$. Comme h est une fonction croissante du paramètre β, on obtient pour $n \geqslant N^*$:

$$2 - \frac{1}{nh(\beta_*)} \leqslant \frac{q_n(\beta_i)}{nh_i} \leqslant 2 + \frac{1}{nh(\beta_*)},$$

et ainsi le résultat attendu. □

Chapitre 4

Modèle de diffusion

4.1 Introduction

Ce chapitre est consacré au problème de l'estimation du coefficient de dérive d'une diffusion ergodique qui est solution de l'équation différentielle stochastique

$$dX_t = S(X_t)dt + dB_t, \qquad 0 \leqslant t \leqslant T, \tag{4.1}$$

où $(B_t)_{t\geqslant 0}$ est un mouvement brownien standard réel. En supposant que l'on observe continûment les données $(X_t, 0 \leqslant t \leqslant T)$ et que l'on connait la régularité de S, on souhaite estimer la fonction S en un point fixe $x_0 \in \mathbb{R}$.

Les domaines d'application de notre modèle (4.1) sont nombreux, en particulier les mathématiques financières, l'économétrie, la théorie du contrôle stochastique et du filtrage, etc. (voir par exemple Liptser et Shiryaev (1978), Jiang et Knight (1997) et Aït-Sahalia (2002)). Beaucoup de travaux sont dédiés à l'étude du problème de l'estimation non paramétrique du coefficient de dérive d'une diffusion. Pour notre modèle (4.1), Banon (1978) a prouvé la consistance tandis que Pham (1981) a considéré la vitesse de convergence d'estimateurs à noyaux. On renvoie par exemple aux articles de Pinsker (1980), Efroĭmovich et Pinsker (1984), Tsybakov (1997) et Spokoiny (2000) dans lesquels d'autres problèmes d'estimation d'un signal en temps continu sont étudiés.

Notre objectif ici est d'estimer la fonction de dérive en un point fixe en utilisant la perte liée à l'erreur absolue pour mesurer la performance d'un estimateur à travers son risque maximal correspondant. Nous nous efforçons d'abord d'obtenir le comportement asymptotique exact du risque minimax, plus particulièrement de trouver sa meilleure borne asymptotique inférieure. Dans un deuxième temps, on construit un estimateur du coefficient de dérive pour lequel le risque maximal est asymptotiquement majoré par la même constante que celle trouvée pour le risque minimax. Rappelons qu'un tel estimateur est dit asymptotiquement efficace.

Le problème de l'estimation asymptotique exacte du coefficient de dérive dans des modèles de diffusion a été étudié pour des classes de Sobolev : Dalayan et Kutoyants (2002) en supposant la régularité de la dérive connue, puis dans le cas contraire Dalalyan

(2005) ont proposé des estimateurs asymptotiquement efficaces du coefficient de dérive dans le modèle (1.2) pour un risque de type L^2.

L'estimation du coefficient de dérive appartenant à une classe Höldérienne a été aussi largement traitée. Galtchouk et Pergamenshchikov (2004) ont obtenu la vitesse de convergence optimale du risque minimax pour une perte de type $L^2([a,b],dx)$ lorsque la régularité de la dérive est connue ou également lorsqu'elle demeure inconnue. La vitesse optimale de convergence du risque minimax est aussi fournie par Galtchouk et Pergamenshchikov (2001) dans le cadre de l'estimation ponctuelle, avec la perte liée à l'erreur absolue, du coefficient de dérive quand sa régularité est supposée inconnue. Pour toute puissance strictement positive de la perte liée à l'erreur absolue et les mêmes classes de Hölder, la constante asymptotique exacte pour le risque minimax est donnée par Galtchouk et Pergamenshchikov (2005b) tout comme un estimateur asymptotiquement efficace du coefficient de dérive de régularité connue.

On considère ici l'estimation ponctuelle du coefficient de dérive lorsque celle-ci appartient à une classe Höldérienne de régularité connue en utilisant la perte liée à l'erreur absolue. Pour ce problème, Galtchouk et Pergamenshchikov (2006b) ont donné la constante asymptotique exacte du risque minimax local et un estimateur à noyau asymptotiquement efficace. Plus précisément, il y est supposé que la dérive appartient à un voisinage centré en une certaine fonction lipschitzienne. Ce voisinage est composé de toutes les fonctions s'écrivant comme la somme du centre et d'une autre fonction satisfaisant à une condition Höldérienne faible (mettant en jeu une constante Höldérienne faible) et ayant une petite norme (usuelle sur $C^1(\mathbb{R})$). Les résultats asymptotiques y ont été fournis en faisant tendre le temps d'observation vers l'infini, la constante Höldérienne faible et le diamètre du voisinage vers zéro. Nous nous proposons de trouver la borne asymptotique inférieure exacte du risque minimax ainsi qu'un estimateur asymptotiquement efficace sans avoir à faire tendre le voisinage vers son centre, mais en conservant son diamètre constant.

Dans la prochaine section, nous décrivons le problème et définissons explicitement la classe lipschitzienne de fonctions, les voisinages et le risque maximal d'un estimateur considérés. La borne asymptotique inférieure du risque minimax est donnée à la Section 3. Enfin la procédure séquentielle aboutissant à l'estimateur asymptotiquement efficace est développée le long de la Section 4.

4.2 Description du problème

Dans le modèle (4.1), nous nous intéressons à l'estimation de la fonction inconnue S en un point fixe $x_0 \in \mathbb{R}$ en supposant que S appartient à un voisinage d'une fonction S_0 elle-même comprise dans l'ensemble

$$\Sigma_{L,M} := \left\{ f : |f(0)| \leqslant L, -L \leqslant \frac{f(x)-f(y)}{x-y} \leqslant -M, \forall x,y \in \mathbb{R} \right\},$$

avec $0 < M < L$.
Comme mentionné dans l'introduction, nous avons besoin de construire des voisinages de

la fonction S_0 avec la perspective de définir le risque maximal d'un estimateur de $S(x_0)$ pris lorsque S décrit tout ce voisinage. Soit $S_0 \in \Sigma_{L,M}$, posons

$$\mathcal{U}_{\delta,\beta,T}(S_0) = \left\{S : S = S_0 + D, D \in \mathcal{H}^w_{x_0}(\delta,\beta)\right\}$$

où

$$\mathcal{H}^w_{x_0}(\delta,\beta) = \left\{D \text{ dérivable} : \sup_{x\in\mathbb{R}} \Big(|D(x)| + |\dot{D}(x)|\Big) \leqslant B; \right.$$
$$\left. \left|\int_{-1}^{1} (D(x_0 + zh_T) - D(x_0))dz\right| \leqslant \delta h_T^\beta \right\},$$

avec $\beta \in]1; 2[$, $h_T = T^{-1/(2\beta+1)}$, $0 < \delta < 1$ et $0 < B < M$.

Remarque 4.2.1 *Si $S_0 \in \Sigma_{L,M}$ on a $\mathcal{U}_{\delta,\beta,T}(S_0) \subset \Sigma_{L+B,M-B}$, de manière à ce que pour $S \in \mathcal{U}_{\delta,\beta,T}(S_0)$, le processus $(X_t)_{t\geqslant 0}$ est ergodique et qu'il existe une densité ergodique (cf. Gihman et Skorohod (1968)) donnée par*

$$q_S(x) = \frac{\exp\left(2\int_0^x S(z)dz\right)}{\displaystyle\int_{-\infty}^{+\infty} \exp\left(2\int_0^y S(z)dz\right) dy}.$$

Rappelons que l'ergodicité d'un processus garantit le fait qu'il retourne dans n'importe quel voisinage de x_0 une infinité de fois.

Le risque d'un estimateur $\tilde{S}_T(x_0)$ de $S(x_0)$ est défini par

$$\mathcal{R}_{\delta,\beta,T}(\tilde{S}_T, S_0) = \sup_{S\in\mathcal{U}_{\delta,\beta,T}(S_0)} \sqrt{2q_S(x_0)}\varphi_T \mathbb{E}_S|\tilde{S}_T(x_0) - S(x_0)|, \qquad \varphi_T = T^{\beta/(2\beta+1)}.$$

4.3 Borne inférieure du risque

Le théorème suivant fournit la borne inférieure asymptotique du risque minimax.

Théorème 4.3.1 *Si $S_0 \in \Sigma_{L,M}$, on a pour tout $\delta \in]0; 1[$*

$$\liminf_{T\to\infty} \inf_{\tilde{S}} \mathcal{R}_{\delta,\beta,T}(\tilde{S}_T, S_0) \geqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}(0,1),$$

où l'infimum est pris sur tous les estimateurs de $S(x_0)$.

DÉMONSTRATION: Pour tout $\nu \in]0; 1/4[$, on pose $D_\nu(x) = \varphi_T^{-1} V_\nu \left(\frac{x - z_0}{h_T}\right)$, où la fonction V_ν est définie sur $\mathbb{R}$ par

$$\begin{aligned}
V_\nu(x) &= \nu^{-1} \int_{-\infty}^{\infty} \tilde{Q}_\nu(u) g\left(\frac{u-x}{\nu}\right) du \\
\tilde{Q}_\nu(u) &= \mathbb{I}_{\{|u|\leqslant 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu\leqslant|u|\leqslant 1-\nu\}} \\
g(z) &= c\exp\left(-(1-z^2)^{-1}\right) \mathbb{I}_{\{|z|\leqslant 1\}},
\end{aligned}$$

avec c la constante de normalisation telle que $\int_{-1}^{1} g(z)dz = 1$.
V_ν est une fonction de classe C^∞ à support compact $[1;1]$.
Enfin pour $u > 0$ et $\nu \in]0;1/4[$, notons $S_{\nu,u}(x) = S_0(x) + uD_\nu(x)$.

Fixons maintenant $b > 0$ et $\delta > 0$. D'après le Lemme 2.5.1 adapté à une version en temps continu, il existe $T_{b,\delta} > 0$ tel que pour tous $|u| \leqslant b$ et $T \geqslant T_{b,\delta}$ on a $S_{\nu,u} \in \mathcal{U}_{\delta,\beta,T}(S_0)$.

Par conséquent, on peut écrire pour tout $T \geqslant T_{b,\delta}$:

$$\begin{aligned}
\mathcal{R}_{\delta,\beta,T}(\tilde{S}_T, S_0) &\geqslant \sup_{|u|\leqslant b} \varphi_T \sqrt{2q_{S_{\nu,u}}(x_0)} \mathbb{E}_{S_{\nu,u}} |\tilde{S}_T(x_0) - S_{\nu,u}(x_0)| \\
&\geqslant \frac{1}{2b} \int_{-b}^{b} \sqrt{2q_{S_{\nu,u}}(x_0)} \mathbb{E}_{S_{\nu,u}} \Psi_{a,T}(\tilde{S}_T, S_{\nu,u}) du \\
&=: I_T(a,b),
\end{aligned}$$

avec $\Psi_{a,T}(\tilde{S}_T, S) = v_a \left(\varphi_T(\tilde{S}_T(x_0) - S(x_0)) \right)$ et $v_a(x) = a \wedge |x|$, $a > 0$.

En notant $\mathbb{P}_S$ la distribution du processus (X_t) dans $C([0;T])$ lorsque la dérive est S, le Lemme 4.3.2 donné à la fin de cette section nous dit que

$$\rho_T(u) := \frac{d\mathbb{P}_{S_{\nu,u}}}{d\mathbb{P}_{S_0}} = \exp\left(u\Delta_T - \frac{1}{2}u^2\sigma_\nu^2 + r_T(u) \right), \quad \forall u > 0,$$

où $\Delta_T = \int_0^T D_\nu(X_t) dB_t$ et $\sigma_\nu^2 = q_{S_0}(x_0) \int_{-1}^{1} V_\nu^2(z)dz$.
De plus,

$$\Delta_T \xrightarrow[T\to\infty]{\mathcal{L}_{\mathbb{P}_{S_0}}} \xi_\nu, \quad \xi_\nu \sim \mathcal{N}(0, \sigma_\nu^2) \quad \text{et} \quad r_T(u) \xrightarrow[T\to\infty]{\mathbb{P}_{S_0}} 0. \tag{4.2}$$

Avec ces quelques notations, on décompose

$$\begin{aligned}
I_T(a,b) &= \frac{1}{2b} \int_{-b}^{b} \sqrt{2q_{S_{\nu,u}}(x_0)} \mathbb{E}_{S_0} \Psi_{a,T}(\tilde{S}_T, S_{\nu,u}) \rho_T(u) du \\
&\geqslant \frac{1}{2b} \int_{-b}^{b} \sqrt{2q_{S_{\nu,u}}(x_0)} \mathbb{E}_{S_0} \mathbb{I}_{B_d} \Psi_{a,T}(\tilde{S}_T, S_{\nu,u}) \rho_T^0(u) du + \delta_T^0(a,b) \\
&=: J_T(a,b) + \delta_T^0(a,b),
\end{aligned}$$

où $B_d = \{|\Delta_T| \leqslant d\}$, $d = \sigma_\nu^2(b - \sqrt{b})$, $\rho_T^0(u) = \exp(u\Delta_T - u^2\sigma_\nu^2/2)$ et

$$\delta_T^0(a,b) = \frac{1}{2b} \int_{-b}^{b} \sqrt{2q_{S_{\nu,u}}(x_0)} \mathbb{E}_{S_0} \mathbb{I}_{B_d} \Psi_{a,T}((\tilde{S}_T, S_{\nu,u})(\rho_T(u) - \rho_T^0(u)) du.$$

La famille de variables aléatoires $\{\rho_T(u), T > 0\}$ est uniformément intégrable. En effet, pour tout $T > 0$, nous avons immédiatement $\mathbb{E}_{S_0}\rho_T(u) = 1$ et en utilisant (4.2) nous pouvons montrer que

$$\rho_T(u) \xrightarrow[T\to\infty]{\mathcal{L}} \eta(u) = \exp(u\xi_\nu - u^2\sigma_\nu^2/2), \quad \xi_\nu \sim \mathcal{N}(0, \sigma_\nu^2),$$

avec $\mathbb{E}\eta(u) = 1$.
Le critère d'intégrabilité uniforme du Lemme 2.5.3 nous permet alors de conclure.

Puisque $\rho_T^0(u)$ est borné sur B_d et $\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})$ est majoré par a, la famille de variables aléatoires $\{\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})(\rho_T(u) - \rho_T^0(u)), T > 0\}$ est aussi uniformément intégrable. Remarquons que $\rho_T(u) - \rho_T^0(u) = \rho_T^0(u)\,(\exp(r_T(u)) - 1)$. On peut prouver, via (4.2), que d'une part $\exp(r_T(u)) \xrightarrow[T\to\infty]{\mathbb{P}_{S_0}} 1$. D'autre part, puisque $\rho_T^0(u)$ est borné sur B_d et $\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})$ est majoré par a, il vient

$$\left(\rho_T(u) - \rho_T^0(u)\right) \Psi_{a,T}(\tilde{S}_T, S_{\nu,u})\mathbb{I}_{B_d} \xrightarrow[T\to\infty]{\mathbb{P}_{S_0}} 0.$$

En prenant alors en compte l'intégrabilité uniforme de la famille de variables aléatoires correspondante, on obtient pour tout $|u| \leqslant b$,

$$\mathbb{E}_{S_0}\left|\left(\rho_T(u) - \rho_T^0(u)\right) \Psi_{a,T}(\tilde{S}_T, S_{\nu,u})\mathbb{I}_{B_d}\right| \xrightarrow[T\to\infty]{} 0.$$

Sur l'événement B_d, on a $\rho_T(u) \leqslant \exp(|u|d)$ et $\rho_T^0(u) \leqslant \exp(|u|d - u^2\sigma_\nu^2/2)$. D'où $\mathbb{E}_{S_0}\left(\rho_T(u) - \rho_T^0(u)\right)\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})\mathbb{I}_{B_d} \leqslant 2a\exp(|u|d)$, qui est une fonction intégrable de la variable u sur l'intervalle $[-b; b]$. D'après le théorème de convergence dominée de Lebesgue, on en déduit que

$$\sup_{\tilde{S}_T}\left|\delta_T^0(a,b)\right| \xrightarrow[T\to\infty]{} 0. \tag{4.3}$$

Considérons maintenant la quantité

$$\frac{1}{2b}\int_{-b}^{b} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})\left(\sqrt{2q_{S_{\nu,u}}(x_0)} - \sqrt{2q_{S_0}(x_0)}\right)\rho_T^0(u)du =: J_T(a,b) - K_T(a,b),$$

où $K_T(a,b) = \dfrac{1}{2b}\displaystyle\int_{-b}^{b} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\Psi_{a,T}(\tilde{S}_T, S_{\nu,u})\sqrt{2q_{S_0}(x_0)}\rho_T^0(u)\mathrm{d}u$.
Rappelons que $S_{\nu,u}(z) = S_0(z) + uD_\nu(z)$, $S_{\nu,u}(x_0) = S_0(x_0) + u\varphi_T^{-1}$ et précisons que $\|D_\nu\|_\infty \leqslant \varphi_T^{-1}\|V_\nu\|_\infty$. Posons

$$c_T(y) := \exp\left(2\int_0^y uD_\nu(z)dz\right),$$

et réécrivons

$$q_{S_{\nu,u}}(x_0) = \frac{\exp\left(2\int_0^{x_0} S_0(z)dz\right)c_{T,u}(x_0)}{\int_{-\infty}^{\infty}\exp\left(2\int_0^y S_0(z)dz\right)c_{T,u}(y)dy}.$$

Par convergence dominée, on a $c_{T,u}(y) \xrightarrow[T\to\infty]{} 1$, pour tout $y \in \mathbb{R}$.
Il n'est pas difficile de montrer que si $y > 0$,

$$\exp(-Ly^2 + 2Ly) \leqslant \exp\left(2\int_0^y S_0(z)dz\right) \leqslant \exp(-My^2 - 2Ly);$$

et si $y < 0$,

$$\exp(-Ly^2 - 2Ly) \leqslant \exp\left(2\int_0^y S_0(z)dz\right) \leqslant \exp(-My^2 + 2Ly).$$

Ajoutons que pour tout $y \in \mathbb{R}$, on a $|c_{T,u}(y)| \leqslant \exp(6|u||y|)$, de manière à ce que la fonction $y \mapsto \exp\left(2\int_0^y S_0(z)dz\right) c_{T,u}(y)$ soit majorée par une fonction intégrable sur $\mathbb{R}$ qui ne dépend pas de la variable T. On conclut par convergence dominée :

$$q_{S_{\nu,u}}(x_0) \xrightarrow[T\to\infty]{} q_{S_0}(x_0),$$

puis également

$$\sup_{\tilde{S}_T} |J_T(a,b) - K_T(a,b)| \xrightarrow[T\to\infty]{} 0. \tag{4.4}$$

A ce point de la démonstration, réécrivons

$$\rho_T^0(u) = \zeta_T \exp(-\sigma_\nu^2(u - \tilde{\Delta}_T)^2/2), \quad \zeta_T = \exp(\Delta_T^2/2\sigma_\nu^2), \quad \tilde{\Delta}_T = \Delta_T/\sigma_\nu^2$$

et posons $g_T = \varphi_T(\tilde{S}_T(x_0) - S_0(x_0))$, $\tilde{g}_T = g_T - \tilde{\Delta}_T$. Ainsi, on a les inégalités successives

$$\begin{aligned}
K_T(a,b) &= \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-b}^{b} v_a(u-g_T)\exp\left(-\sigma_\nu^2(u-\tilde{\Delta}_T)^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&= \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-b-\tilde{\Delta}_T}^{b-\tilde{\Delta}_T} v_a(u-\tilde{g}_T)\exp\left(-\sigma_\nu^2u^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&\geqslant \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} v_a(u-\tilde{g}_T)\exp\left(-\sigma_\nu^2u^2/2\right)\sqrt{2q_{S_0}(x_0)}du \\
&\geqslant \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} v_a(u)\exp\left(-\sigma_\nu^2u^2/2\right)\sqrt{2q_{S_0}(x_0)}du,
\end{aligned}$$

la deuxième provenant du Lemme 2.5.4.
D'où

$$K_T(a,b) \geqslant \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} |u|\exp\left(-\sigma_\nu^2u^2/2\right)\sqrt{2q_{S_0}(x_0)}du + \delta_T^1(a,b),$$

avec

$$\delta_T^1(a,b) = \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T\frac{1}{2b}\int_{-\sqrt{b}}^{\sqrt{b}} (v_a(u) - |u|)\exp\left(-\sigma_\nu^2u^2/2\right)\sqrt{2q_{S_0}(x_0)}du.$$

Un petit calcul fournissant

$$\lim_{T\to\infty} \mathbb{E}_{S_0}\mathbb{I}_{B_d}\zeta_T = 2\sigma_\nu(b-\sqrt{b})/\sqrt{2\pi},$$

on obtient pour tout $b>0$,

$$\lim_{a\to\infty}\liminf_{T\to\infty} \delta_T^1(a,b) = 0, \tag{4.5}$$

et

$$\liminf_{a\to\infty}\liminf_{T\to\infty}\inf_{\tilde{S}_T} K_T(a,b) \geqslant \frac{b-\sqrt{b}}{b}\frac{\sqrt{2q_{S_0}(x_0)}\sigma_\nu}{\sqrt{2\pi}}\int_{-\sqrt{b}}^{\sqrt{b}}|u|\exp(-u^2\sigma_\nu^2/2)du.$$

En se remémorant que $\sigma_\nu^2 = q_{S_0}(x_0)\int_{-1}^1 V_\nu^2(z)dz \xrightarrow[\nu\to 0]{} 2q_{S_0}(x_0)$, ensuite en passant à la limite $b\to\infty$ puis $\nu\to 0$ dans l'inégalité précédente, on en déduit que

$$\liminf_{\nu\to 0}\liminf_{b\to\infty}\liminf_{a\to\infty}\liminf_{T\to\infty}\inf_{\tilde{S}_T} K_T(a,b) \geqslant \mathbb{E}|\xi|, \quad \xi\sim\mathcal{N}(0,1).$$

Or

$$\inf_{\tilde{S}_T}\mathcal{R}_{\delta,\beta,T}(\tilde{S}_T) \geqslant -\sup_{\tilde{S}_T}|J_T(a,b)-K_T(a,b)| - \sup_{\tilde{S}_T}|\delta_T^0(a,b)| + \inf_{\tilde{S}_T}K_T(a,b).$$

Finalement, en appliquant (4.3), (4.4) et (4.5), nous avons montré que

$$\liminf_{T\to\infty}\inf_{\tilde{S}}\mathcal{R}_{\delta,\beta,T}(\tilde{S}_T,S_0) \geqslant \mathbb{E}|\xi|, \quad \xi\sim\mathcal{N}(0,1).$$

□

Lemme 4.3.2 *Pour tout $u\in\mathbb{R}$, on a*

$$\begin{aligned}\rho_T(u) &= \frac{d\mathbb{P}_{S_{\nu,u}}}{d\mathbb{P}_{S_0}} = \exp\left\{u\int_0^T D_\nu(X_t)dB_t - \frac{u^2}{2}\int_0^T D_\nu^2(X_t)dt\right\}\\ &= \exp\left\{u\Delta_T - \frac{u^2}{2}\sigma_\nu^2 + r_T(u)\right\},\end{aligned}$$

où $\Delta_T = \int_0^T D_\nu(X_t)dB_t$, $\sigma_\nu^2 = q_{S_0}(x_0)\int_{-1}^1 V_\nu^2(z)dz$ et $r_T(u) = \frac{u^2}{2}\sigma_\nu^2 - \frac{u^2}{2}\int_0^T D_\nu^2(X_t)dt$. De plus

$$\Delta_T \xrightarrow[T\to\infty]{\mathcal{L}_{\mathbb{P}_{S_0}}} \xi_\nu, \quad \xi_\nu\sim\mathcal{N}(0,\sigma_\nu^2) \quad et \quad r_T(u)\xrightarrow[T\to\infty]{\mathbb{P}_{S_0}} 0.$$

Démonstration: La première partie du lemme se trouve dans Liptser et Shiryaev (1978, p. 323). En ce qui concerne les différentes convergences, posons pour une fonction Φ intégrable contre q_{S_0} :

$$m(\Phi) = \int_{-\infty}^{\infty}\Phi(y)q_{S_0}(y)dy.$$

Soit $\varepsilon > 0$. On a

$$\begin{aligned}
\mathbb{P}_{S_0}\left(\left|\int_0^T D_\nu^2(X_t)dt - \sigma_\nu^2\right| > \varepsilon\right) &= \mathbb{P}_{S_0}\left(\left|\frac{1}{T}\int_0^T \frac{1}{h_T}V_\nu^2\left(\frac{X_t - x_0}{h_T}\right)dt - \sigma_\nu^2\right| > \varepsilon\right)\\
\leqslant\ \mathbb{P}_{S_0}\left(\left|\frac{1}{T}\int_0^T \frac{1}{h_T}V_\nu^2\left(\frac{X_t - x_0}{h_T}\right)dt - m(\Phi_T)\right| > \frac{\varepsilon}{2}\right) &+ \mathbb{P}_{S_0}\left(\left|m(\Phi_T) - \sigma_\nu^2\right| > \frac{\varepsilon}{2}\right)\\
\leqslant\ \mathbb{P}_{S_0}\left(\frac{1}{\sqrt{T}}\left|\Lambda_T(\Phi_T)\right| > \frac{\varepsilon}{2}\right) + \mathbb{P}_{S_0}\left(\left|m(\Phi_T) - \sigma_\nu^2\right| > \frac{\varepsilon}{2}\right),&
\end{aligned}$$

où $\Phi_T(y) = \dfrac{1}{h_T}V_\nu^2\left(\dfrac{y - x_0}{h_T}\right)$ et $\Lambda_T(\Phi_T) = \dfrac{1}{\sqrt{T}}\displaystyle\int_0^T \left(\Phi_T(X_t) - m(\Phi_T)\right)dt$.
Or d'une part, d'après Galtchouk et Pergamenshchikov (2001), pour toute fonction Φ_T vérifiant

$$\sup_{T\geqslant 1}\int_{-\infty}^{\infty} |\Phi_T(y)|dy < \infty,$$

il existe une constante $\kappa > 0$ telle que pour tout $\lambda > 0$,

$$\sup_{T\geqslant 1}\ \sup_{S_0\in\Sigma_{L,M}}\ \mathbb{P}_{S_0}\left(|\Lambda_T(\Phi_T)| > \lambda\right) \leqslant 2e^{-\kappa\lambda^2}.$$

Comme pour tout $T > 0$, on a

$$\int_{-\infty}^{\infty} \frac{1}{h_T}V_\nu^2\left(\frac{y - x_0}{h_T}\right)dy \leqslant \int_{-1}^{1} V_\nu^2(z)dz < \infty,$$

on obtient

$$\mathbb{P}_{S_0}\left(\frac{1}{\sqrt{T}}|\Lambda_T(\Phi_T)| > \frac{\varepsilon}{2}\right) \xrightarrow[T\to\infty]{} 0. \tag{4.6}$$

D'autre part,

$$\begin{aligned}
m(\Phi_T) - \sigma_\nu^2 &= \int_{-\infty}^{\infty} \frac{1}{h_T}V_\nu^2\left(\frac{y - x_0}{h_T}\right)q_{S_0}(y)dy - \sigma_\nu^2 = \int_{-\infty}^{\infty} V_\nu^2(z)q_{S_0}(x_0 + zh_T)dz - \sigma_\nu^2\\
&= \int_{-1}^{1} V_\nu^2(z)\left(q_{S_0}(x_0 + zh_T) - q_{S_0}(x_0)\right)dz = \int_{-1}^{1} V_\nu^2(z)\frac{z^2h_T^2}{2}q''_{S_0}(c_z)dz,
\end{aligned}$$

où c_z est compris entre x_0 et $x_0 + zh_T$.
Mais il existe $T_0 > 0$ tel que pour $T \geqslant T_0$, l'intervalle $[x_0 - h_T; x_0 + h_T]$ est compris dans $[0; 1]$. Il est alors clair qu'il existe une certaine constante $C = C(L, M)$ telle que

$$|m(\Phi_T) - \sigma_\nu^2| \leqslant h_T^2 C\int_{-1}^{1} V_\nu^2(z)dz.$$

On en déduit la convergence en $\mathbb{P}_{S_0}$–probabilité de $r_T(u)$ vers zéro quand $T \to \infty$.

Soulignons qu'en particulier, nous avons montré grâce à (4.6) que $(1/\sqrt{T})\Lambda_T(\Phi_T)$ converge en $\mathbb{P}_{S_0}$–probabilité vers zéro quand $T \to \infty$. En d'autres termes, on a

$$\mathbb{P}_{S_0} - \lim_{T\to\infty} \frac{1}{T}\int_0^T \frac{1}{h_T} V_\nu^2\left(\frac{X_t - x_0}{h_T}\right) dt = \mathbb{P}_{S_0} - \lim_{T\to\infty} \int_{-\infty}^{\infty} \frac{1}{h_T} V_\nu^2\left(\frac{y - x_0}{h_T}\right) q_{S_0}(y)dy.$$

On en déduit que

$$\begin{aligned}
\mathbb{P}_{S_0} - \lim_{T\to\infty} \int_0^T D_\nu^2(X_t)dt &= \mathbb{P}_{S_0} - \lim_{T\to\infty} \frac{1}{\varphi_T^2}\int_0^T V_\nu^2\left(\frac{X_t - x_0}{h_T}\right) dt \\
&= \mathbb{P}_{S_0} - \lim_{T\to\infty} \int_{-\infty}^{\infty} \frac{1}{h_T} V_\nu^2\left(\frac{y - x_0}{h_T}\right) q_{S_0}(y)dy \\
&= \lim_{T\to\infty} \int_{-1}^{1} V_\nu^2(z) q_{S_0}(x_0 + zh_T)dz \\
&= q_{S_0}(x_0)\int_{-1}^{1} V_\nu^2(z)dz = \sigma_\nu^2.
\end{aligned}$$

On peut alors appliquer le théorème central limite pour les intégrales stochastiques (cf. Kutoyants, 2004, p.43) pour obtenir

$$\mathcal{L}_{\mathbb{P}_{S_0}}(\Delta_T) \xrightarrow[T\to\infty]{} \mathcal{N}(0, \sigma_\nu^2).$$

$\square$

4.4 Estimateur asymptotiquement efficace

Afin de construire un estimateur de $S(x_0)$ asymptotiquement efficace, on commence par estimer la densité ergodique au point x_0 via les observations $\{X_t, t \leqslant t_0\}$, où $t_0 = T^{2\gamma}$, $\gamma_* < \gamma < 1/2$ et $\gamma_* = \frac{\alpha}{2\beta+1}$. Soit

$$\hat{q}_T(x_0) = \frac{1}{2t_0 l}\int_0^{t_0} Q\left(\frac{X_t - x_0}{l}\right) dt,$$

avec $Q = \mathbb{I}_{[-1;1]}$ et $l = l_T = o(1/\sqrt{T})$ quand $T \to \infty$.
Puis pour $H > 0$, on définit la procédure séquentielle $(\tau_H, S_T^*(x_0))$ comme

$$\begin{aligned}
\tau_H &= \inf\{t \geqslant t_0 : \int_{t_0}^{t} Q\left(\frac{X_t - x_0}{h}\right) dt \geqslant H\}, \\
S_T^*(x_0) &= \frac{1}{H}\int_{t_0}^{\tau_H} Q\left(\frac{X_t - x_0}{h}\right) dX_t \mathbb{I}_{\{\tau_H \leqslant T\}}.
\end{aligned} \tag{4.7}$$

On choisit la fenêtre $h = h_T = T^{-1/(2\beta+1)}$ et le niveau $H = H_T = (T - t_0)(2\tilde{q}_T(x_0) - \varepsilon_T)h_T$, où $\tilde{q}_T(x_0) = \max(\hat{q}_T(x_0), \nu_T^{-1/2})$, $\varepsilon_T = 1/(\nu_T T^{\gamma_*})$ et $\nu_T = \ln T$.

Les lemmes suivants provenant de Galtchouk et Pergamenshchikov (2006b) donnent quelques propriétés des variables aléatoires $\hat{q}_T(x_0)$, $\tilde{q}_T(x_0)$ et τ_H.

Lemme 4.4.1 *Il existe deux constantes $\kappa > 0$ et $T_* > 0$ telles que pour tout $T \geqslant T_*$ et tout $\lambda > 1/T$, on a*

$$\sup_{S \in \Sigma_{L+B,M-B}} \mathbb{P}_S\left(|\hat{q}_T(x_0) - q_S(x_0)| > \lambda\right) \leqslant 2e^{-\kappa\lambda^2 t_0}.$$

Lemme 4.4.2 *Il existe une constante $T_* > 0$ telle que*

$$\mu^* = \sup_{T \geqslant T_*} \sup_{S \in \Sigma_{L+B,M-B}} \sqrt{t_0}\mathbb{E}_S|\tilde{q}_T(x_0) - q_S(x_0)| < \infty.$$

Lemme 4.4.3 *On a*

$$\limsup_{T \to \infty} \sup_{S \in \Sigma_{L+B,M-B}} \mathbb{E}_S\left|\frac{1}{\tilde{q}_T(x_0)} - \frac{1}{q_S(x_0)}\right| = 0.$$

Lemme 4.4.4 *Pour tout $p > 0$,*

$$\lim_{T \to \infty} T^p \sup_{S \in \mathcal{U}_{\delta,\beta,T}(S_0)} \mathbb{P}_S(\tau_H > T) = 0.$$

Le centre du voisinage considéré doit vérifier une condition supplémentaire décrite par la définition suivante.

Définition 4.4.5 *Une fonction f satisfait à la condition "zéro-constante" de Hölder d'exposant $\iota > 0$ au point x_0 si*

$$\lim_{y \to x_0} \frac{f(y) - f(x_0)}{|y - x_0|^\iota} = 0.$$

Un exemple d'une telle fonction peut être trouvé dans Galtchouk et Pergamenshchikov (2006b).

Nous sommes maintenant en mesure de formuler le résultat sur la borne supérieure asymptotique du risque maximal de l'estimateur (4.7).

Théorème 4.4.6 *Soient $S_0 \in \Sigma_{L,M}$ et $\beta = 1 + \alpha$, $\alpha \in]0;1[$. Supposons que S_0' satisfait à la condition "zéro-constante" de Hölder d'exposant $\alpha > 0$ au point x_0. Alors nous avons*

$$\limsup_{\delta \to 0} \limsup_{T \to \infty} \mathcal{R}_{\delta,\beta,T}(S_T^*(x_0), S_0) \leqslant \mathbb{E}|\xi|, \qquad \xi \sim \mathcal{N}(0,1).$$

Démonstration: Soit $S \in \mathcal{U}_{\delta,\beta,T}(S_0)$. Décomposons l'erreur de l'estimation

$$S_T^*(x_0) - S(x_0) = \left(B_T - G_T + \frac{\xi_T}{\sqrt{H_T}}\right)\mathbb{I}_{\{\tau_H \leqslant T\}} - S(x_0)\mathbb{I}_{\{\tau_H > T\}},$$

où

$$\begin{aligned}
B_T &= \frac{1}{H_T}\int_{t_0}^{T} Q\left(\frac{X_t-x_0}{h_T}\right)(S(X_t)-S(x_0))dt,\\
G_T &= \frac{1}{H_T}\int_{\tau_H}^{T} Q\left(\frac{X_t-x_0}{h_T}\right)(S(X_t)-S(x_0))dt,\\
\xi_T &= \frac{1}{\sqrt{H_T}}\int_{t_0}^{T} Q\left(\frac{X_t-x_0}{h_T}\right)dB_t.
\end{aligned}$$

Montrons tout d'abord que

$$\lim_{\delta\to 0}\limsup_{T\to\infty}\sup_{S\in\mathcal{U}_{\delta,\beta,T}(S_0)}\mathbb{E}_S\sqrt{2q_S(x_0)}\varphi_T|B_T|\mathbb{I}_{\{\tau_H\leqslant T\}}=0. \tag{4.8}$$

Commençons par écrire

$$B_T=\frac{T}{(2\tilde{q}_T(x_0)-\varepsilon_T)(T-t_0)}\left(\frac{T-t_0}{T}m(f_h)+\frac{1}{\sqrt{T}}\Lambda_{t_0,T}(f_h)\right), \tag{4.9}$$

où $f_h(y)=\phi_h(y)(S(y)-S(x_0))$, $\phi_h(y)=\frac{1}{h_T}Q\left(\frac{y-x_0}{h_T}\right)$, $m(f)=\int_{\mathbb{R}}f(y)q_S(y)dy$ et $\Lambda_{t_0,T}(f)=\frac{1}{\sqrt{T}}\int_{t_0}^{T}(f(X_t)-m(f))dt$.

On peut transformer le terme $m(f_h)$ de la manière suivante

$$\begin{aligned}
m(f_h) &= \frac{1}{h_T}\int_{-\infty}^{\infty} Q\left(\frac{y-x_0}{h_T}\right)(S(y)-S(x_0))q_S(y)dy\\
&= \int_{-1}^{1}(S(x_0+h_Tz)-S(x_0))q_S(x_0+h_Tz)dz\\
&= \int_{-1}^{1}(S(x_0+h_Tz)-S(x_0))(q_S(x_0+h_Tz)-q_S(x_0))dz\\
&+ q_S(x_0)\int_{-1}^{1}(S(x_0+h_Tz)-S(x_0))dz\\
&=: m_1(h_T)+q_S(x_0)m_0(h_T).
\end{aligned}$$

Posons $r(y):=S_0(y)-S_0(x_0)-S_0'(x_0)(y-x_0)$. Puisqu'on a $\int_{-1}^{1}S_0'(x_0)zh_Tdz=0$ et $S\in\mathcal{U}_{\delta,\beta,T}(S_0)$, on obtient

$$\begin{aligned}
|m_0(h_T)| &= \left|\int_{-1}^{1}(S(x_0+zh_T)-S(x_0))dz\right|\\
&= \left|\int_{-1}^{1}r(x_0+zh_T)dz+\int_{-1}^{1}(D(x_0+zh_T)-D(x_0))dz\right|\\
&\leqslant 2\sup_{|u|\leqslant h}\frac{|r(x_0+u)|}{|u|^{\beta}}h_T^{\beta}+\delta h_T^{\beta}=:(2r^*(h_T)+\delta)h_T^{\beta}.
\end{aligned}$$

En utilisant la condition "zéro-constante" de Hölder, on peut prouver que $\lim_{h \to 0} r^*(h) = 0$.
En conséquence il vient $\varphi_T |m_0(h_T)| \leqslant (2r^*(h_T) + \delta)\varphi_T h_T^\beta = 2r^*(h_T) + \delta$ et

$$\lim_{\delta \to 0} \lim_{T \to \infty} \varphi_T |m_0(h_T)| = 0. \tag{4.10}$$

De plus, en notant $C = C(L, M, B)$ une constante dépendant de ces paramètres mais qui n'est pas nécessairement la même à chaque fois qu'elle apparaît, nous avons

$$\begin{aligned} \varphi_T |m_1(h_T)| &\leqslant \varphi_T \int_{-1}^{1} |S(x_0 + zh_T) - S(x_0)| \, |q_S(x_0 + zh_T) - q_S(x_0)| \, dz \\ &\leqslant C\varphi_T h_T^2 = C h_T^{2-\beta} \xrightarrow[T \to \infty]{} 0, \end{aligned} \tag{4.11}$$

car $|q_S(x_0+zh_T)-q_S(x_0)| = |zh_T q_S'(c_z)| \leqslant C|z|h_T$ et $|S(x_0+zh_T)-S(x_0)| \leqslant (L+B)|z|h_T$.
D'après l'inégalité (A.1) dans Galtchouk et Pergamenshchikov (2006b), il existe une constante $\kappa > 0$ telle que pour tout $\lambda > 0$:

$$\sup_{T \geqslant 1} \sup_{0 \leqslant t_0 \leqslant T} \sup_{S \in \Sigma_{L+B, M-B}} \mathbb{P}_S \left(|\Lambda_{t_0, T}(f_h)| > \lambda \right) \leqslant 2e^{-\kappa\lambda^2}. \tag{4.12}$$

Ainsi il est clair que $\varphi_T T^{-1/2} \mathbb{E}_S |\Lambda_{t_0,T}(f_h)|$ tend vers zéro quand $T \to \infty$ uniformément en $S \in \mathcal{U}_{\delta,\beta,T}(S_0)$.
Soit $q_*(x_0) := \inf_{S \in \Sigma_{L+B,M-B}} q_S(x_0)$. Alors $q_*(x_0) > 0$ et on peut écrire pour T suffisamment grand

$$\frac{1}{2\tilde{q}_T(x_0) - \varepsilon_T} \leqslant \frac{1}{\tilde{q}_T(x_0)} \leqslant \left| \frac{1}{\tilde{q}_T(x_0)} - \frac{1}{q_S(x_0)} \right| + \frac{1}{q_*(x_0)}.$$

Avec le Lemme 4.4.3, c'est un argument montrant que

$$\lim_{T \to \infty} \sup_{S \in \mathcal{U}_{\delta,\beta,T}(S_0)} \mathbb{E}_S \frac{\varphi_T}{\sqrt{T}(2\tilde{q}_T(x_0) - \varepsilon_T)} |\Lambda_{t_0,T}(f_h)| = 0. \tag{4.13}$$

De par ce même Lemme 4.4.3, (4.11) implique que

$$\lim_{T \to \infty} \sup_{S \in \mathcal{U}_{\delta,\beta}(S_0)} \mathbb{E}_S \frac{\varphi_T}{2\tilde{q}_T(x_0) - \varepsilon_T} |m_1(h_T)| = 0. \tag{4.14}$$

Finalement, puisque $\sup_{S \in \mathcal{U}_{\delta,\beta,T}(S_0)} q_S(x_0) < \infty$, (4.10) et le Lemme 4.4.3 nous amènent à

$$\lim_{\delta \to 0} \lim_{T \to \infty} \sup_{S \in \mathcal{U}_{\delta,\beta,T}(S_0)} \mathbb{E}_S \frac{\varphi_T}{2\tilde{q}_T(x_0) - \varepsilon_T} q_S(x_0) |m_0(h_T)| = 0. \tag{4.15}$$

L'assertion (4.8) provient alors de (4.13), (4.14) et (4.15).

On se propose désormais de montrer que pour tout $\delta \in (0, 1)$

$$\lim_{T \to \infty} \sup_{S \in \mathcal{U}_{\delta,\beta,T}(S_0)} \mathbb{E}_S \sqrt{2q_S(x_0)} \varphi_T |G_T| \mathbb{I}_{(\tau_H \leqslant T)} = 0. \tag{4.16}$$

Remarquons premièrement que

$$m(\phi_h) - 2q_S(x_0) = \int_{-1}^{1} (q_S(x_0 + uh_T) - q_S(x_0))\, du.$$

D'après la formule de Taylor, pour tout $u \in [-1,1]$ il existe $\theta \in (0,1)$ tel que

$$q_S(x_0 + uh_T) - q_S(x_0) = q'_S(x_0)uh_T + \frac{1}{2}u^2 h_T^2 q''_S(x_0 + \theta u h_T),$$

c'est pourquoi nous obtenons

$$\sup_{S\in\mathcal{U}_{\delta,\beta,T}(S_0)} |m(\phi_h) - 2q_S(x_0)| = \sup_{S\in\mathcal{U}_{\delta,\beta,T}(S_0)} \frac{1}{2}\left|\int_{-1}^{1} u^2 h_T^2 q''_S(x_0 + \theta u h_T) du\right| \leqslant C h_T^2, \quad (4.17)$$

pour une certaine constante $C = C(L, M, B)$.
Puis on a successivement

$$\begin{aligned}
|G_T| &\leqslant \frac{1}{H_T}\int_{\tau_H}^{T} h_T \phi_h(X_t)|S(X_t) - S(x_0)|dt \\
&\leqslant \frac{1}{H_T}\int_{\tau_H}^{T} h_T^2 (L+B)\phi_h(X_t)dt \leqslant \frac{(L+B)h_T^2}{H_T}\left(\int_{t_0}^{T}\phi_h(X_t)dt - \frac{H_T}{h_T}\right) \\
&\leqslant \frac{(L+B)h_T^2}{H_T}\left(\sqrt{T}\Lambda_{t_0,T}(\phi_h) + (T-t_0)m(\phi_h) - (T-t_0)(2\tilde{q}_T(x_0) - \varepsilon_T)\right) \\
&\leqslant \frac{(L+B)h_T^2}{H_T}\left(\sqrt{T}\Lambda_{t_0,T}(\phi_h) + (T-t_0)|2\tilde{q}_T(x_0) - m(\phi_h)| + \varepsilon_T(T-t_0)\right).
\end{aligned}$$

Alors (4.17) et la définition de H_T entrainent

$$|G_T| \leqslant \frac{(L+B)h_T}{\tilde{q}_T(x_0)}\left(\frac{\sqrt{T}\Lambda_{t_0,T}(\phi_h)}{T-t_0} + 2|\tilde{q}_T(x_0) - q_S(x_0)| + Ch_T^2 + \varepsilon_T\right).$$

En utilisant le Lemme 4.4.2 et la définition de $\tilde{q}_T(x_0)$, il vient

$$\mathbb{E}_S|G_T|\mathbb{I}_{(\tau_H\leqslant T)} \leqslant (L+B)\sqrt{\nu_T}\left(\frac{h_T\sqrt{T}}{T-t_0}\mathbb{E}_S|\Lambda_{t_0,T}(\phi_h)| + Ch_T^3 + h_T\varepsilon_T + 2\mu^*\frac{h_T}{\sqrt{t_0}}\right).$$

Puisque (4.12) implique que $\mathbb{E}_S|\Lambda_{t_0,T}(\phi_h)|$ est uniformément borné en $S \in \mathcal{U}_{\delta,\beta,T}(S_0)$ et $T > 0$, on prouve aisément (4.16).

Au final, comme ξ_T est une variable aléatoire gaussienne standard, on a

$$\begin{aligned}
&\left|\sqrt{2q_S(x_0)}\varphi_T\mathbb{E}_S\frac{|\xi_T|}{\sqrt{H_T}} - \mathbb{E}_S|\xi|\right| = \left|\frac{\sqrt{2q_S(x_0)}T^{\beta/(2\beta+1)}}{\sqrt{(T-t_0)(2\tilde{q}_T(x_0) - \varepsilon_T)h_T}} - 1\right|\mathbb{E}|\xi| \\
&= \left|\frac{\sqrt{2q_S(x_0)}(1-\frac{t_0}{T})^{-1/2}}{\sqrt{2\tilde{q}_T(x_0) - \varepsilon_T}} - \frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}} + \frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}} - 1\right|\mathbb{E}|\xi| \\
&\leqslant \left|\frac{\sqrt{2q_S(x_0)}(1-\frac{t_0}{T})^{-1/2}}{\sqrt{2\tilde{q}_T(x_0) - \varepsilon_T}} - \frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}}\right|\mathbb{E}|\xi| + \left|\frac{\sqrt{2q_S(x_0)}}{\sqrt{2\tilde{q}_T(x_0)}} - 1\right|\mathbb{E}|\xi|. \quad (4.18)
\end{aligned}$$

Or il est facile de voir que $2\tilde{q}_T(x_0) - \varepsilon_T \sim 2\tilde{q}_T(x_0)$ et $(1 - t_0/T)^{1/2} \sim 1$ quand $T \to \infty$. Par conséquent, la limite

$$\frac{(1-\frac{t_0}{T})^{-1/2}}{\sqrt{2\tilde{q}_T(x_0) - \varepsilon_T}} - \frac{1}{\sqrt{2\tilde{q}_T(x_0)}} = o\left(\frac{1}{\sqrt{\tilde{q}_T(x_0)}}\right) = o((\ln T)^{-1/2}) \xrightarrow[T\to\infty]{} 0$$

est uniforme sur $\mathcal{U}_{\delta,\beta,T}(S_0)$.
D'après le Lemme 4.4.3, la seconde partie de (4.18) tend vers zéro uniformément sur $\mathcal{U}_{\delta,\beta,T}(S_0)$ quand $T \to \infty$. En regroupant ceci avec (4.8), (4.16) et le Lemme 4.4.4, cela complète la démonstration du Théorème 4.4.6. $\square$

Chapitre 5

Simulations numériques

5.1 Résultats

On illustre dans ce chapitre les résultats obtenus au Chapitre 2. Les fonctions et les procédures utilisées ont été programmées sous Scilab et leurs codes sont donnés dans la prochaine section.

On se propose d'estimer en z_0 la fonction S définie sur $[0;1]$ par $S(x) = |x - z_0|^\beta/1000$. On vérifie que cette fonction appartient à $\mathcal{H}(\delta, \delta^{-1}, \beta)$ lorsque $\beta/1000 \leqslant \delta \leqslant 1000/\beta$. Les valeurs de z_0 et de β sont laissées au libre arbitre de l'utilisateur des programmes.

Pour l'écart-type des bruits g, on prend la fonction décrite en 2.9 :

$$\begin{aligned} g^2(x, S) &= G(x, S(x)) + \int_0^1 V(S(t))dt, \\ G(x, y) &= \alpha_0 + \alpha_1 x + \alpha_2 \sin^2 y, \\ V(y) &= \alpha_3 \sin^2 y. \end{aligned}$$

Les calculs présentés ici sont effectués avec un choix de constantes $\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 1$, mais l'utilisateur peut lui-même choisir ces valeurs.

On estime $S(x_0)$ par trois estimateurs construits à l'aide des observations $\{y_i, i = 1, \ldots, n\}$:

- notre estimateur à noyau $\hat{S}_n = (1/q_n) \sum_{k=1}^{n} Q\left(\frac{x_k - z_0}{h_n}\right) y_k$,
- un autre estimateur à noyau $\tilde{S}_n = (1/k_n) \sum_{k=1}^{n} K\left(\frac{x_k - z_0}{h_n}\right) y_k$,
- l'estimateur localement polynomial (cf. Tsybakov (2004)) $\hat{LP}_n = U^\top(0)\hat{\theta}_n$,

où

$$\begin{aligned}
x_i &= \frac{i}{n},\ h_n = n^{-1/(2\beta+1)}, \\
Q &= \mathbb{I}_{[-1;1]},\ q_n = \sum_{i=1}^{n} Q\left(\frac{x_i - x_0}{h_n}\right),\ k_n = \sum_{i=1}^{n} K\left(\frac{x_i - x_0}{h_n}\right), \\
U(u) &= (1,u)^\top,\ \hat{\theta}_n = B_n^{-1} a, \\
B_n &= \frac{1}{nh_n} \sum_{i=1}^{n} U\left(\frac{x_i - x_0}{h_n}\right) U^\top\left(\frac{x_i - x_0}{h_n}\right) K\left(\frac{x_i - x_0}{h_n}\right), \\
a &= \frac{1}{nh_n} \sum_{i=1}^{n} U\left(\frac{x_i - x_0}{h_n}\right) K\left(\frac{x_i - x_0}{h_n}\right) y_i,
\end{aligned}$$

avec K un noyau positif à support compact inclus dans $[-1;1]$ vérifiant pour tout $u \in \mathbb{R}$:

$$K_{min}\mathbb{I}_{\{|u| \leqslant \Delta\}} \leqslant K(u) \leqslant K_{max},$$

pour un certain $\Delta > 0$. On choisit pour noyau K le noyau d'Epanechnikov :

$$K(u) = \frac{3}{4}\max(0, 1 - u^2).$$

Les résultats numériques approximent les risques asymptotiques des trois estimateurs utilisés du fait du calcul d'une espérance (on effectue une moyenne pour $M = 200$ simulations) et du nombre fini d'observations n. On a choisi $x_0 = 0.8$ et $\beta = 1.36$. Précisons aussi que pour différentes valeurs de n, les simulations sont réinitialisées et donc ne courent pas sur les mêmes premiers aléas. Les valeurs sont à comparer à la constante asymptotique exacte $1/\sqrt{\pi} \simeq 0.564$.

Pour des variables aléatoires $\{\xi_i, i = 1, \ldots, n\}$ i.i.d gaussiennes standard, on a obtenu :

n	$\hat{S}_n$	$\tilde{S}_n$	$\hat{LP}_n$
10	0.667	0.721	0.801
100	0.533	0.550	0.549
1000	0.575	0.598	0.598
5000	0.528	0.557	0.557
10000	0.546	0.606	0.606

Pour des variables aléatoires $\{\xi_i, i = 1, \ldots, n\}$ i.i.d réduites à partir de variables aléatoires uniformes sur $[-1;1]$, on a obtenu :

n	$\hat{S}_n$	$\tilde{S}_n$	$\hat{LP}_n$
10	0.557	0.634	0.788
100	0.600	0.632	0.634
1000	0.550	0.626	0.626
5000	0.570	0.616	0.616
10000	0.560	0.600	0.600

Pour des variables aléatoires $\{\xi_i, i = 1, \ldots, n\}$ i.i.d centrées et réduites à partir de variables aléatoires exponentielles de paramètre 1, on a obtenu :

n	$\hat{S}_n$	$\tilde{S}_n$	$\hat{LP}_n$
10	0.644	0.718	0.815
100	0.570	0.601	0.618
1000	0.557	0.584	0.584
5000	0.592	0.636	0.636
10000	0.556	0.588	0.588

5.2 Programmes

Les fonctions et les procédures implémentées en Scilab sont données en italique :

– la fonction "signal" :

$function[y] = signal(x, z0, bet)$
$y = (1/1000) * (abs(x - z0))\,\hat{}\,bet$
$endfunction$

– les noyaux et quelques fonctions auxiliaires :

$function[y] = noyaurect(x)$
$if\ \ ((x > 1)|(x < -1))\ \ then\ \ y = 0$
$else\ \ y = 1$
end
$endfunction$

$function[W] = doubleverect(X, n, h)$
$W = zeros(1, n)$
$for\ \ i = 1 : n$
$W(i) = noyaurect((X(i) - z0)/h)$
end
$endfunction$

$function[y] = epanech(x)$
$y = 0.75 * max(0, 1 - x\,\hat{}\,2)$

$endfunction$

$function[W] = doubleveepanech(X, n, h)$
$W = zeros(1, n)$
$for \quad i = 1 : n$
$W(i) = epanech((X(i) - z0)/h)$
end
$endfunction$

$function[V] = U(x)$
$V = [1x]'$
$endfunction$

– la fonction représentant l'écart-type du bruit :

$function[y] = ecarttype(x, S, a0, a1, a2, a3, z0, bet)$
$y = sqrt(a0 + a1 * x + a2 * (sin(S(x, z0, bet)))\hat{}\,2$
$+ integrate('3 * (sin(S(t, z0, bet)))\hat{}\,2', 't', 0, 1));$
$endfunction$

– la procédure donnant les risques des trois estimateurs pour des bruits gaussiens :

$n = input("Rentrerlavaleurden : ");$
$z0 = input("Rentrer\ la\ valeur\ de\ z0 : ");$
$bet = input("Rentrer\ la\ valeur\ de\ beta : ");$
$a0 = input("Rentrer\ la\ valeur\ de\ a0 : ");$
$a1 = input("Rentrer\ la\ valeur\ de\ a1 : ");$
$a2 = input("Rentrer\ la\ valeur\ de\ a2 : ");$
$a3 = input("Rentrer\ la\ valeur\ de\ a3 : ");$
$M = input("Combien\ de\ simulations?");$
$X0 = linspace(0, 1, n + 1); X = X0(2 : n + 1); h = n\hat{}\,(-1/(2 * bet + 1));$
$WNW = doubleverect(X, n, h); qnNW = sum(WNW);$
$WEpa = doubleveepanech(X, n, h); qnEpa = sum(WEpa);$

$for \quad k = 1 : M$
$XI = grand(1, n, 'nor', 0, 1);$
$Y = signal(X, z0, bet) + ecarttype(X, signal1, a0, a1, a2, a3, z0, bet). * XI;$
$YYNW = (1/qnNW) * WNW. * Y;$
$NW(k) = sum(YYNW);$
$NWabs(k) = abs(NW(k));$
$YYEpa = (1/qnEpa) * WEpa. * Y;$
$Epa(k) = sum(YYEpa);$

$Epaabs(k) = abs(Epa(k));$

$i = 1; B = zeros(2,2); a = zeros(2,1);$
$while \quad i <= n$
$B = B + (epanech((X(i) - z0)/h)/(n * h)) * U((X(i) - z0)/h) * U((X(i) - z0)/h)';$
$a = a + (Y(i) * epanech((X(i) - z0)/h)/(n * h)) * U((X(i) - z0)/h);$
$i = i + 1;$
end
$teta = B \backslash a;$
$LPEpa(k) = U(0)' * teta;$
$LPEpaabs(k) = abs(LPEpa(k));$
end

$risqueNW = n\hat{}\,(bet/(2 * bet + 1)) * sum(NWabs)/$
$(M * ecarttype(z0, signal1, a0, a1, a2, a3, z0, bet))$
$risqueEpa = n\hat{}\,(bet/(2 * bet + 1)) * sum(Epaabs)/$
$(M * ecarttype(z0, signal1, a0, a1, a2, a3, z0, bet))$
$risqueLPEpa = n\hat{}\,(bet/(2 * bet + 1)) * sum(LPEpaabs)/$
$(M * ecarttype(z0, signal1, a0, a1, a2, a3, z0, bet))$

En ce qui concerne les autres lois des bruits, il faut remplacer la ligne

XI=grand(1,n,'nor',0,1) ;

par

XI=grand(1,n,'exp',1)-1 ;

pour des lois exponentielles recentrées, et par

*XI=sqrt(3)*grand(1,n,'unf',-1,1) ;*

pour des lois uniformes recentrées.

Bibliographie

Aït-Sahalia, Y. (2002) : Maximum likelihood estimation of discretely sampled diffusions : a closed-form approximation approach, *Econometrica* **70**(1), 223–262.

Banon, G. (1978) : Nonparametric identification for diffusion processes, *SIAM J. Control Optim.* **16**(3), 380–395.

Barron, A., Birgé, L. et Massart, P.(1999) : Risk bounds for model selection via penalization, *Probab. Theory Related Fields* **113**, 301–413.

Belomestny, D. et Reiß,M. (2006) : Spectral calibration of exponential Lévy models, *Finance Stoch.* **10**(4), 449–474.

Billingsley, P. (1999) : *Convergence of probability measures*, Wiley Series in Probability and Statistics : Probability and Statistics, John Wiley & Sons Inc., New York, A Wiley-Interscience Publication.

Bretagnolle, J. et Huber, C. (1979) : Estimation des densités : risque minimax, *Z. Wahrsch. Verw. Gebiete* **47**(2), 119–137.

Brown, L. D. et Low, M. G. (1996) : A constrained risk inequality with applications to nonparametric functional estimation, *Ann. Statist.* **24**(6), 2524–2535.

Brua, J.-Y. (2008a) : Asymptotically efficient estimators in nonparametric heteroscedastic regression models, *Statist. Methodol.* à paraître.

Brua, J.-Y. (2008b) : Adaptive estimators in nonparametric heteroscedastic regression models, en cours de révision, *J. Nonparametr. Statist.*

Brua J.-Y. (2008c) : Asymptotically efficient estimators of the drift function in ergodic diffusion processes, soumis.

Butucea, C. (2000a) : The adaptive rate of convergence in a problem of pointwise density estimation, *Statist. Probab. Lett.* **47**(1), 85–90.

Butucea, C. (2000b) : Two adaptive rates of convergence in pointwise density estimation, *Math. Methods Statist.* **9**(1), 39–64.

Cai, T. et Wang, L. (2008) : Adaptive variance function estimation in heteroscedastic nonparametric regression, *Ann. Statist.* **36**, 2025-2054.

Dalalyan, A. S. (2005) : Sharp adaptive estimation of the drift function for ergodic diffusions, *Ann. Statist.* **33**(6), 2507–2528.

Dalalyan, A. S. et Kutoyants, Yu. A. (2002) : Asymptotically efficient trend coefficient estimation for ergodic diffusion, *Math. Methods Statist.* **11**(4), 402–427.

Dalalyan, A. S. et Kutoyants, Yu. A. (2003) : Asymptotically efficient estimation of the derivative of the invariant density, *Stat. Inference Stoch. Process.* **6**(1), 89–107.

Donoho, D. L. (1994a) : Asymptotic minimax risk for sup-norm loss : solution via optimal recovery, *Probab. Theory Related Fields* **99**(2), 145–170.

Donoho, D.L. (1994b) : Statistical estimation and optimal recovery, *Ann. Statist.* **22**, 238–270.

Donoho, D. L., Johnstone, I. M., Kerkyacharian, G. et Picard, D. (1995) : Wavelet shrinkage : asymptopia ?, *J. Roy. Statist. Soc. Ser. B* **57**(2), 301–369.

Donoho, D.L. et Liu, R.C. (1991) : Geometrizing rates of convergence. II, III, *Ann. Statist.* **19**(2), 668–701.

Efroĭmovich, S. Yu. (1985) : Nonparametric estimation of a density of unknown smoothness, *Theory Probab. Appl.* **30**(3), 557–568.

Efroĭmovich, S. (1996) : On nonparametric regression for IID observations in a general setting, *Ann. Statist.* **24**(3), 1125–1144.

Efroĭmovich, S. : *Nonparametric Curve Estimation. Methods, Theory and Applications*, Springer, Berlin, New York, 1999.

Efroĭmovich, S. (2007) : Sequential design and estimation in heteroscedastic nonparametric regression, *Sequential Anal.* **26**(1), 3–25.

Efroĭmovich, S. Yu. et Pinsker, M. S. (1981) : Estimation of square-integrable density on the basis of a sequence of observations, *Problems Inform. Trans.* **17**(3), 182–196.

Efroĭmovich, S. Yu. et Pinsker, M. S. (1982) : Estimation of square-integrable probability density of a random variable, *Problems Inform. Trans.* **18**(3), 175–189.

Efroĭmovich, S. Yu. et Pinsker, M. S. (1984) : A self-training algorithm for nonparametric filtering, *Automat. Remote Control* **11**, 58–65.

Efroĭmovich, S. et Pinsker, M. (1996) : Sharp-optimal and adaptive estimation for heteroscedastic nonparametric regression, *Statist. Sinica* **6**(4), 925–942.

Fourdrinier, D. et Pergamenshchikov, S. (2007) : Improved model selection method for a regression function with dependent noise, *Ann. Inst. Statist. Math.* **59**(3), 435–464.

Freedman, D. (1971), *Brownian Motion and Diffusion*, Holden Day, San Francisco.

Galtchouk, L. et Pergamenshchikov, S. (2001) : Sequential nonparametric adaptive estimation of the drift coefficient in diffusion processes, *Math. Methods Statist.* **10**(3), 316–330.

Galtchouk, L. et Pergamenshchikov, S. (2004) : Nonparametric sequential estimation of the drift in diffusion processes via model selection, *Math. Methods Statist.* **13**(1), 25–49.

Galtchouk, L. et Pergamenshchikov, S. (2005a) : Efficient adaptive nonparametric estimation in heteroscedastic regression models, Prépublication de l'Université Louis Pasteur de Strasbourg, IRMA, disponible en ligne sur : http ://hal.archives-ouvertes.fr/hal-00129707/fr/

Galtchouk, L. et Pergamenshchikov, S. (2005b) : Nonparametric sequential minimax estimation of the drift coefficient in diffusion processes, *Sequential Anal.* **24**(3), 303–330.

Galtchouk, L. et Pergamenshchikov, S. (2006a) : Asymptotically efficient estimates for nonparametric regression models, *Statist. Probab. Lett.* **76**(8), 852–860.

Galtchouk, L. et Pergamenshchikov, S. (2006b) : Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes, *Stat. Inference Stoch. Process.* **9**(1), 1–16.

Gihman, I. I. et Skorohod, A. V. : *Stochastic Differential Equations, Naukova Dumka*, Kiev, 1968.

Goldenshluger, A. et Nemirovski, A. (1997) : On spatially adaptive estimation of nonparametric regression, *Math. Methods Statist.* **6**(2), 135–170.

Goldfeld, S. et Quandt, R. (1972) : *Nonlinear Methods in Econometrics*, North-Holland, Amsterdam, London.

Golubev, G. K. (1984) : Minimax estimation of regression, *Problems Inform. Trans.* **20**(1), 42–49.

Golubev, G. K. (1992) : Asymptotically minimax estimation of a regression function in an additive model, *Problems Inform. Trans.* **28**, 3–15.

Golubev, G. K. et Levit, B. Y. (1996) : Asymptotically efficient estimation for analytic distributions, *Math. Methods Statist.* **5**(3), 357–368.

Golubev, G. K., Levit, B. Y. et Tsybakov, A. B. (1996) : Asymptotically efficient estimation of analytic functions in Gaussian noise, *Bernoulli* **2**(2), 167–181.

Golubev, G. K. et Nussbaum, M. (1990a) : A risk bound in Sobolev class regression, *Ann. Statist.* **18**(2), 758–778.

Golubev, G. K. et Nussbaum, M. (1990b) : Nonparametric estimation of a regression function in L^2, *Problems Inform. Trans.* **26**(3), 213–225.

Golubev, G. K. et Nussbaum, M. (1992) : Adaptive spline estimates in a nonparametric regression model, *Theory Probab. Appl.* **37**(3), 521–529.

Guerre, E. et Tsybakov, A. B. (1998) : Exact asymptotic minimax constants for the estimation of analytical functions in L_p, *Probab. Theory Related Fields* **112**(1), 33–51.

Härdle, W. et Marron, J. S. (1985) : Optimal bandwidth selection in nonparametric regression function estimation, *Ann. Statist.* **13**(4), 1465–1481.

Ibragimov, I. A. et Has'minskiĭ, R. Z. (1980) : On the estimation of a signal, its derivatives and the maximum point for Gaussian observations, *Theory Probab. Appl.* **25**(4), 718–733.

Ibragimov, I. A. et Has'minskiĭ, R. Z. : *Statistical Estimation : Asymptotic Theory*, Springer, Berlin, New York, 1981.

Ibragimov, I. A. et Has'minskiĭ, R. Z. (1982a) : Bounds for the quality of nonparametric estimation of regression, *Theory Probab. Appl.* **27**(1), 81–94.

Ibragimov, I. A. et Has'minskiĭ, R. Z. (1982b) : An estimate of the density of a distribution belonging to a class of entire functions, *Theory Probab. Appl.* **27**(3),514–524.

Ibragimov, I. A. et Has'minskiĭ, R. Z. (1984) : Asymptotic bounds for the quality of nonparametric estimation of regression in $\mathcal{L}_p$, *J. Soviet Mathematics* **25**, 540–550 (Originally published in Russian in 1980).

Jiang, G. J. et Knight, J. L. (1997) : A nonparametric approach to the estimation of diffusion processes with an application to a short-term interest rate model, *Econometric Theory* **13**(5), 615–645.

Juditsky, A. (1997) : Wavelet estimators : adapting to unknown smoothness, *Math. Methods Statist.* **6**(1), 1–25.

Korostelev, A. P. (1993) : An asymptotically minimax regression estimator in the uniform norm up to a constant, *Theory Probab. Appl.* **38**(4), 737–743.

Korostelev, A. P. et Nussbaum, M. (1999) : The asymptotic minimax constant for supnorm loss in nonparametric density estimation, *Bernoulli* **5**(6), 1099–1118.

Kutoyants, Yu. A. (1998) : Efficient density estimation for ergodic diffusions, *Stat. Inference Stoch. Process.* **1**(2), 131–155.

Kutoyants, Yu. A. : *Statistical inference for ergodic diffusion processes*, Springer Series in Statistics, Springer-Verlag London Ltd., London, 2004.

Lepskiĭ, O. V. (1990) : A problem of adaptive estimation in Gaussian white noise, *Theory Probab. Appl.* **35**(3), 454–466.

Lepskiĭ, O. V. (1991) : Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates, *Theory Probab. Appl.* **36**(4), 682–697.

Lepskiĭ, O. V. (1992a) : Asymptotically minimax adaptive estimation. II. Schemes without optimal adaptation. Adaptive estimates, *Theory Probab. Appl.* **37**(3), 433–448.

Lepskiĭ, O. V. (1992b) : On problems of adaptive estimation in white Gaussian noise. In *Topics in Nonparametric Estimation.* **12**, 87–107. Amer. Math. Soc., Providence, RI.

Lepskiĭ, O. V., Mammen, E. et Spokoiny, V. G. (1997) : Optimal spatial adaptation to inhomogeneous smoothness : an approach based on kernel estimates with variable bandwidth selectors, *Ann. Statist.* **25**(3), 929–947.

Lepskiĭ, O. V. et Spokoiny, V. G. (1997) : Optimal pointwise adaptive methods in nonparametric estimation, *Ann. Statist.* **25**(6), 2512–2546.

Liptser, R. S. et Shiryaev, A. N. : *Statistics of random processes. I and II*, Springer-Verlag, New York, 1978.

Nemirovskiĭ, A. S. (1986) : Nonparametric estimation of smooth regression functions, *J. Comput. Systems Sci.* **23**(6), 1–11.

Nemirovskiĭ, A. S., Polyak, B. T. et Tsybakov, A. B. (1985) : The rate of convergence of nonparametric estimates of maximum likelihood type, *Problems Inform. Trans.* **21**(4), 17–33.

Nussbaum, M. : Spline smoothing in regression models and asymptotic efficiency in L_2, *Ann. Statist.* **13**(3), 984–997.

Pham D. T. (1981) : Nonparametric estimation of the drift coefficient in the diffusion equation, *Math. Operationsforsch. Statist. Ser. Statist.* **12**(1), 61–73.

Pinsker, M. S. (1980) : Optimal filtration of square-integrable signals in Gaussian noise, *Problems Inform. Trans.*, **16**(2), 120–133.

Sacks, J. et Strawderman, W. (1982) : Improvements on linear minimax estimates, *Statistical decision theory and related topics, III, Vol. 2*, 287–304.

Sacks, J. et Ylvisaker, D. (1978) : Linear estimation for approximately linear models, *Ann. Statist.* **6**(5), 1122–1137.

Sacks, J. et Ylvisaker, D. (1981) : Asymptotically optimum kernels for density estimation at a point, *Ann. Statist.* **9**(2), 334–346.

Speckman, P. (1985) : Spline smoothing and optimal rates of convergence in nonparametric regression models, *Ann. Statist.* **13**(3), 970–983.

Spokoiny, V. G. (2000) : Adaptive drift estimation for nonparametric diffusion model, *Ann. Statist.* **28**(3), 815–836.

Stone, Charles J. (1982) : Optimal global rates of convergence for nonparametric regression, *Ann. Statist.* **10**(4), 1040–1053.

Tsybakov, A. B. (1997) : Asymptotically efficient estimation of a signal in L_2 under general loss functions, *Problems Inform. Trans.* **33**(1), 78–88.

Tsybakov, A. B. (1998) : Pointwise and sup-norm sharp adaptive estimation of functions on the Sobolev classes, *Ann. Statist.* **26**(6), 2420–2469.

Tsybakov, A. B. : *Introduction à l'estimation non-paramétrique*, Springer-Verlag, Berlin, 2004.