



HAL
open science

Estimation non paramétrique pour les modèles autorégressifs

Ouerdia Arkoun

► **To cite this version:**

Ouerdia Arkoun. Estimation non paramétrique pour les modèles autorégressifs. Mathématiques [math]. Université de Rouen, 2009. Français. NNT : . tel-00464024

HAL Id: tel-00464024

<https://theses.hal.science/tel-00464024>

Submitted on 15 Mar 2010

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

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

THÈSE

en vue de l'obtention du titre de

Docteur de l'Université de Rouen

présentée par

Ouerdia ARKOUN

Discipline : Mathématiques

Spécialité : Statistique

**Estimation non paramétrique pour les modèles
autorégressifs**

Date de soutenance : 9 novembre 2009

Composition du Jury

Président	:	Dominique FOURDRINIER	Professeur, Université de Rouen
Rapporteurs	:	Delphine BLANKE	Professeur, Université d'Avignon
		Leonid GALTCHOUK	Professeur, Université de Strasbourg
Directeur de Thèse	:	Sergeï PERGAMENCHTCHIKOV	Professeur, Université de Rouen

**Thèse préparée à l'Université de Rouen
Laboratoire de Mathématiques Raphaël Salem, UMR-CNRS 6085**

Remerciements

Mes premiers remerciements s'adressent à Sergueï Pergamenchtchikov. Ce fut pour moi un réel plaisir de travailler sous sa direction tout au long de ces quatre années durant lesquelles j'ai eu à apprécier sa culture mathématique, ses grandes compétences, sa disponibilité et son efficacité. Je dois également saluer ses qualités pédagogiques, son optimisme (réaliste) qui, dans les périodes de doute, m'a permis de reprendre confiance et poussé à ne pas baisser les bras. Pour toutes ces raisons, je ne peux que me féliciter de l'avoir choisi pour m'encadrer.

Je tiens à remercier ensuite Leonid Galtchouk et Delphine Blanke d'avoir accepté la tâche de rapporter ma thèse. Les travaux de Galtchouk et Pergamenchtchikov ont initié chacune des deux parties de ma thèse, c'est donc pour moi un honneur qu'ils aient bien voulu en évaluer le contenu. Je suis par ailleurs d'autant plus honorée que Dominique Fourdrinier ait accepté de présider le jury.

J'en viens maintenant à ceux qui m'ont donné goût aux mathématiques et qui m'ont donné l'envie de poursuivre dans cette voie. Il est sans doute difficile de savoir à quel moment cela s'est produit, mais j'estime, sans nul doute, que je le dois en premier à Hamid Louni, mon enseignant de Statistiques durant ma licence, qui par ses talents de pédagogue, sa rigueur (ses blagues de matheux aussi), a suscité une vocation enfouie en moi et m'a soutenue dans ce que j'ai entrepris. Je ne le remercierai jamais assez.

Je pense ensuite à Fazia Bedouhene, mon enseignante de première année universitaire, qui s'investissait remarquablement dans l'enseignement, qui a toujours été présente et fait preuve de patience pour les nombreuses questions que j'ai pu lui poser. Je tiens à lui exprimer ma profonde gratitude.

Dans cet ordre d'idées, j'exprime mes remerciements aux membres du département de mathématiques de l'Université de Rouen, en particulier Paul Raynaud de Fitte, François Charlot, Thierry de La Rue, Élise Janvresse, Léo Glangetas et qui à un moment ou un autre ont tous plus ou moins contribué à mon épanouissement dans le domaine des mathématiques.

Merci également à Edwige Auvray et Marguerite Losada pour leur présence, leur efficacité et leur patience. Je veux également saluer le travail d'Isabelle Lamitte qui gère remarquablement la bibliothèque. Enfin mille mercis à Marc Jolly qui a réalisé l'impression et la reliure de cette thèse et sur qui j'ai toujours pu compter.

Je veux également remercier Claude Dellacherie, Thierry de La Rue et Élise Janvresse pour l'organisation de l'atelier des doctorants. J'y ai beaucoup appris autant sur le plan mathématique, sur la façon de communiquer que sur le plan culinaire.

Je remercie fortement Jean-Yves Brua pour ses conseils, son soutien moral et sa sympathie mais qui a surtout relu cette thèse avec beaucoup de minutie. Je lui dois de précieuses corrections.

Il m'est particulièrement agréable de dédier cette thèse de doctorat à mon oncle le penseur et historien Mohammed Arkoun dont le parcours intellectuel a, depuis mon jeune âge, constitué pour moi un exemple et une référence. Je ne le remercierai jamais assez pour avoir guidé mes premiers pas en France, pour les précieux conseils qu'il n'a cessé de me prodiguer, son indéfectible soutien moral et son assistance matérielle. Qu'il trouve ici le profond témoignage de ma reconnaissance et de mon affection.

Je dédie également ce travail de thèse à mes parents que je remercie pour l'enseignement qu'ils m'ont transmis et pour leur irremplaçable et inconditionnel soutien. Ils ont été présents pour écarter les doutes, soigner les blessures et partager les joies. Sans eux, je ne serais pas là où je suis aujourd'hui. Je remercie tout particulièrement ma mère Ghenima pour avoir été toujours à mes côtés tout au long de mes études et qui a toujours cru en ma volonté de réussir.

A mes beaux-parents, pour leur amour et la disponibilité dont ils ont constamment fait preuve à mon égard,

A ma grand-mère, décédée il y a peu et qui serait contente d'apprendre que sa petite fille a terminé le travail qu'elle avait commencé,

A Islam, mon époux, je dis "ce travail te doit beaucoup...qu'il soit pour toi le témoignage de mon infinie reconnaissance pour ces années de compréhension, de privations et d'efforts communs",

A mon très cher fils Rayan, conçu et né pendant la réalisation de cette thèse,

A mes trois frères Youcef, Brahim et Mohammed ainsi qu'à ma soeur Dahbia, son mari Nacer et leurs enfants Samih et Manel.

Toute mon amitié à Ali Righi avec qui j'ai partagé le bureau pendant ces années, avec qui j'ai eu beaucoup de plaisir à apprendre la programmation en Scilab et avec qui j'ai eu tant de discussions fructueuses.

Je ne saurais clore ces remerciements sans une pensée particulière à mon amie Assia Ghezali pour les trois années passées ensemble à Oued-Aissi dans le rire et la bonne humeur et sans saluer ici tous ceux qui ont partagé ou partagent encore mon bureau : Abdelatif et Nicolas. Je passe ensuite une dédicace spéciale à tous les doctorants que j'ai eu le plaisir de côtoyer durant ces quelques années à Rouen, avec qui j'ai passé de bons moments au RU, en salle de convivialité ou au soleil sur la terrasse, à savoir Aicha, Houda, Islam, Karima, les deux Olivier, Lahcen, Lamia, Manel, Vincent.

Enfin merci à toutes les personnes que je n'ai pas citées ici et qui j'espère se reconnaîtront dans ces quelques lignes.

Résumé

Cette thèse se consacre à l'estimation non paramétrique pour les modèles autorégressifs. Nous considérons le problème de l'estimation d'une fonction inconnue en un point fixe à l'aide de données régies par des modèles autorégressifs. Pour définir le risque associé à l'emploi d'un estimateur et ainsi mesurer la qualité de celui-ci, nous utilisons la fonction de perte liée à l'erreur absolue. Le travail de cette thèse suit l'approche minimax dont l'objectif est de trouver une borne inférieure asymptotique du risque minimax puis de construire un estimateur, dit asymptotiquement efficace, dont le risque maximal atteint asymptotiquement cette borne.

Pour un modèle autorégressif non paramétrique où la fonction autorégressive est supposée appartenir à une classe Höldérienne faible de régularité connue, nous montrons qu'un estimateur à noyau est asymptotiquement efficace. Lorsque la régularité de la fonction autorégressive est inconnue, nous obtenons la vitesse de convergence minimax adaptative des estimateurs sur une famille de classes Höldériennes.

Mots-clés : Efficacité asymptotique, Autorégression non paramétrique, Minimax, Estimateur à noyau, Estimation adaptative, Estimation séquentielle

Abstract

This thesis is devoted to nonparametric estimation for autoregressive models. We consider the problem of estimating an unknown function at a fixed point using data governed by autoregressive models. To define the risk associated with the use of an estimator and thus measure the quality of it, we use the loss function related to the absolute error. The work of this thesis follows the minimax approach for which the goal is to find a lower bound of the asymptotic minimax risk and then to construct an estimator, said asymptotically efficient, for which the maximum risk reaches asymptotically this bound.

For a nonparametric autoregressive model where the autoregressive function is supposed to belong to a weak Hölder class with known regularity, we show that a kernel estimator is asymptotically efficient. When the regularity of the autoregressive function is unknown, we get the minimax adaptive convergence rate of estimators on a family of Hölderian classes.

Key words : Asymptotical efficiency, Kernel estimates, Minimax, Nonparametric autoregression, Adaptive estimation, Sequential estimators.

Table des matières

1	Introduction	9
1.1	Problématique	9
1.1.1	Description générale	9
1.1.2	Approche minimax	10
1.1.3	Approche minimax adaptative	11
1.2	Description des résultats obtenus	13
1.2.1	Cas non adaptatif	13
1.2.2	Cas adaptatif	18
2	Modèles autorégressifs : cas non adaptatif	21
2.1	Introduction	21
2.2	Description du problème	24
2.3	Bornes asymptotiques pour des bruits de loi inconnue	26
2.3.1	Borne inférieure	26
2.3.2	Borne supérieure	31
2.4	Annexe A	34
3	Modèle autorégressif : cas adaptatif	39
3.1	Introduction	39
3.2	Description du problème	41
3.3	Borne inférieure	43
3.4	Estimation séquentielle adaptative (borne supérieure)	46
3.5	Annexe B	51
4	Simulations numériques	61
4.1	Cas non adaptatif	61
4.1.1	Résultats	61
4.1.2	Programmes	63
4.2	Cas adaptatif	65
4.2.1	Résultats	65
4.2.2	Programmes	66

Chapitre 1

Introduction

1.1 Problématique

1.1.1 Description générale

La tradition de considérer le problème de l'estimation statistique comme celui d'estimation d'un nombre fini de paramètres remonte à Fisher. Les modèles statistiques qui expliquent plus profondément les données sont d'habitude plus complexes : les inconnues de ces modèles sont, en général, des fonctions possédant certaines propriétés de régularité. Le problème de l'estimation non paramétrique consiste à estimer, à partir des observations, une fonction inconnue, élément d'une certaine classe fonctionnelle assez massive.

La théorie de l'estimation non paramétrique s'est développée considérablement ces deux dernières décennies, en se fixant pour objectif quelques thèmes principaux, en particulier, l'étude de l'optimalité des estimateurs et l'estimation adaptative. Ces deux thèmes occuperont la place centrale de cette thèse. En particulier nous nous intéresserons à l'optimalité des estimateurs lorsque la taille de l'échantillon tend vers l'infini. De tels estimateurs seront 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'minskiï [1981]). Nous nous sommes attachés dans cette thèse à montrer l'efficacité asymptotique de certains estimateurs à noyau pour les modèles autorégressifs non paramétriques.

Le modèle autorégressif non paramétrique considéré est le suivant. On dispose de n observations $(y_k)_{1 \leq k \leq n}$ régies par :

$$y_k = S(x_k)y_{k-1} + \xi_k, \quad 1 \leq k \leq n, \quad (1.1)$$

où $S(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ est la fonction inconnue à estimer à partir des observations, y_0 étant une constante. Les variables aléatoires $(\xi_k)_{1 \leq k \leq n}$ sont indépendantes, centrées, identiquement distribuées et de variance 1. Ce modèle est à pas fixe car nous supposerons que $x_k = k/n$ pour tout $k = 1, \dots, n$.

Pour ce modèle, on se propose d'estimer la fonction inconnue S en un point fixe z_0 en supposant qu'elle appartienne à une classe Hölderienne, puis on définit le risque associé à

cette classe. 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 le modèle considéré.

Définition 1.1.1 *Pour le modèle autorégressif (1.1), un estimateur de S au point z_0 est une variable aléatoire $\omega \mapsto \tilde{S}_n(z_0) = \tilde{S}_n(z_0, y_1, \dots, y_n)$ mesurable par rapport à la tribu engendrée par y_1, \dots, y_n .*

On définit le risque d'un estimateur \tilde{S}_n d'une fonction S appartenant à la classe fonctionnelle \mathcal{H} pour un z_0 fixé par $\mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|$, où \mathbf{E}_S désigne l'espérance quand l'aléa est déterminé par le modèle (1.1).

Le risque minimax sur une classe fonctionnelle \mathcal{H} est donné par

$$\mathcal{R}_n^* = \inf_{\tilde{S}_n} \sup_{S \in \mathcal{H}} \varphi_n \mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|,$$

l'infimum étant pris sur tous les estimateurs et la famille $(\varphi_n)_{n \in \mathbb{N}^*}$ est composée de réels strictement positifs ($\varphi_n \rightarrow +\infty$, quand $n \rightarrow +\infty$).

L'objectif premier de l'approche minimax est de trouver un estimateur \hat{S}_n dont le risque maximal est égal au risque minimax. Un tel estimateur est dit minimax. Un estimateur \hat{S}_n est dit asymptotiquement minimax si

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

lorsque la taille de l'échantillon n tend vers l'infini.

De nombreux travaux ont été fait pour les modèles de régression dans le cadre de l'estimation non paramétrique en un point fixe. Par exemple pour estimer 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. Dans notre cadre on considère le cas pour lequel le risque maximal d'un estimateur \tilde{S}_n est défini par

$$\mathcal{R}_n(\tilde{S}_n) := \sup_{S \in \mathcal{H}} \varphi_n \mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|.$$

Le but de l'approche est ainsi de trouver un estimateur S_n^* , des familles φ_n et des constantes $c > 0$ et $C < \infty$ telles que

$$\limsup_{n \rightarrow \infty} \mathcal{R}_n(S_n^*) \leq C \quad \text{et} \quad \liminf_{n \rightarrow \infty} \mathcal{R}_n^* \geq c. \quad (1.2)$$

Définition 1.1.2 La famille $(\varphi_n)_{n \in \mathbb{N}^*}$ est dite vitesse de convergence minimax des estimateurs sur \mathcal{H} si (1.2) est vérifiée.

Définition 1.1.3 Un estimateur S_n^* vérifiant $c \leq \liminf_{n \rightarrow \infty} \mathcal{R}_n^* \leq \limsup_{n \rightarrow \infty} \mathcal{R}_n(S_n^*) \leq C$, où $(\varphi_n)_{n \in \mathbb{N}^*}$ 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} .

Remarque 1.1.4 Pour montrer qu'un estimateur est asymptotiquement efficace, il suffit d'obtenir une borne inférieure et une borne supérieure égales ($C = c$ dans la Définition 1.1.3).

Définition 1.1.5 Un estimateur S_n^* est dit asymptotiquement efficace sur \mathcal{H} lorsque

$$\lim_{n \rightarrow \infty} \frac{\mathcal{R}_n(S_n^*)}{\mathcal{R}_n^*} = 1.$$

1.1.3 Approche minimax adaptative

L'approche minimax est dite adaptative lorsqu'un des paramètres définissant la classe fonctionnelle \mathcal{H} considérée est supposé inconnu, par exemple la régularité de la fonction autorégressive S dans le modèle (1.1). Notons alors $\mathcal{H}^{(\beta)}$ la classe fonctionnelle, où $\beta \in \mathcal{B}$, \mathcal{B} étant un intervalle quelconque et

$$\mathcal{R}_\beta(\tilde{S}_n, \phi_n(\beta)) = \sup_{S \in \mathcal{H}^{(\beta)}} \phi_n(\beta) \mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|,$$

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

La question que l'on se pose est l'existence d'un estimateur optimal adaptatif en vitesse de convergence, c'est-à-dire un estimateur indépendant de $\beta \in \mathcal{B}$ qui converge à cette vitesse 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 $\{\mathcal{H}^{(\beta)}\}_{\beta \in \mathcal{B}}$ s'il existe une constante $C > 0$ telle que :

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

où $\varphi_n(\beta)$ est la vitesse de convergence minimax sur la classe $\mathcal{H}^{(\beta)}$.

Puis, comme pour l'approche minimax non adaptative, 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 $\{\mathcal{H}^{(\beta)}\}_{\beta \in \mathcal{B}}$ s'il vérifie :*

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

Cependant, des estimateurs optimaux adaptatifs en vitesse de convergence n'existent pas toujours. En effet, Lepskiï [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 (1.3) et \mathcal{B} contenant au moins deux éléments. Néanmoins, il se peut qu'on ait une relation du type

$$\lim_{n \rightarrow \infty} \sup_{\beta \in \mathcal{B}} \sup \mathcal{R}_\beta(S_n^*, N_n(\beta)) \leq C,$$

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

Définition 1.1.8 *La famille $(N_n(\beta))_{n \in \mathbb{N}^*}$ 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 :

$$\lim_{n \rightarrow \infty} \sup_{\beta \in \mathcal{B}} \sup \mathcal{R}_\beta(S_n^*, N_n(\beta)) \leq C;$$

– il existe une constante $c > 0$ telle que :

$$\lim_{n \rightarrow \infty} \inf_{\tilde{S}_n} \inf_{\beta \in \mathcal{B}} \sup \mathcal{R}_\beta(\tilde{S}_n, N_n(\beta)) \geq c.$$

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

Définition 1.1.9 *Soient $L > 0$ et $\beta > 0$. La classe de Hölder $\Sigma(L, \beta)$ est définie par*

$$\Sigma(L, \beta) = \{S : \mathbb{R} \rightarrow \mathbb{R} : |S^{(m)}(x) - f^{(m)}(y)| \leq L|x - y|^{\beta - m}, \forall x, y \in \mathbb{R}\}, \quad (1.3)$$

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

Remarque 1.1.10 *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 Description des résultats obtenus

1.2.1 Cas non adaptatif

On considère le modèle autorégressif non paramétrique (1.1), la fonction autorégressive étant à estimer en un point fixe z_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 Arkoun et Pergamenshchikov [2008].

Le problème de l'estimation asymptotique efficace, pour lequel la fonction S appartient à la classe Höldérienne stable $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ de régularité $\beta = 1 + \alpha$, $\alpha \in]0; 1]$ définie par

$$\mathcal{H}^{(\beta)}(z_0, K, \varepsilon) = \left\{ S \in \Gamma_\varepsilon : \|\dot{S}\| \leq K \quad \text{et} \quad \sup_{x \in [0,1]} \frac{|\dot{S}(x) - \dot{S}(z_0)|}{|x - z_0|^\alpha} \leq K \right\}$$

avec l'ensemble de stabilité du modèle (1.1) par rapport à S

$$\Gamma_\varepsilon = \{S \in \mathbf{C}_1[0, 1] : \|S\| \leq 1 - \varepsilon\} \quad \text{où} \quad \|S\| = \sup_{0 \leq x \leq 1} |S(x)|,$$

reste ouvert. Dans notre cas, on étudie le risque minimax pris sur une classe plus large qu'une classe de Hölder, appelée classe Höldérienne faible au point z_0 et définie, pour $\delta \in]0; 1[$, par

$$\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon) = \left\{ S \in \Gamma_\varepsilon : \|\dot{S}\| \leq \delta^{-1} \quad \text{et} \quad \left| \int_{-1}^1 (S(z_0 + uh) - S(z_0)) du \right| \leq \delta h_n^\beta \right\}, \quad (1.4)$$

où le paramètre β est supposé connu et $h_n = n^{-1/(2\beta+1)}$, n étant le nombre d'observations dans notre modèle (1.1).

Remarque 1.2.1 *Remarquons que*

$$\int_{-1}^1 (S(z_0 + h_n u) - S(z_0)) du = \int_{-1}^1 \left(\int_{z_0}^{z_0 + u h_n} (\dot{S}(t) - \dot{S}(z_0)) dt \right) du, \quad (1.5)$$

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

On donne un exemple de fonctions appartenant à $\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$. Pour cela considérons une famille de fonctions $(S_\nu, 0 < \nu < 1/4)$ où,

$$S_\nu(x) = \varphi_n^{-1} V_\nu \left(\frac{x - z_0}{h_n} \right),$$

avec $\varphi_n = n^{\beta/(2\beta+1)}$ et $h_n = n^{-1/(2\beta+1)}$. On définit la fonction V_ν comme suit

$$V_\nu(x) = \nu^{-1} \int_{-\infty}^{\infty} \tilde{Q}_\nu(u) g\left(\frac{u-x}{\nu}\right) du,$$

$$\tilde{Q}_\nu(u) = \mathbb{I}_{\{|u| \leq 1-2\nu\}} + 2\mathbb{I}_{\{1-2\nu \leq |u| \leq 1-\nu\}}, \quad (1.6)$$

et g est une fonction paire positive, infiniment différentiable telle que $g(z) = 0$ pour $|z| \geq 1$ et $\int_{-1}^1 g(z) dz = 1$. Il est facile de voir que pour tout $0 < \nu < 1/4$

$$V_\nu(0) = 1 \quad \text{et} \quad \int_{-1}^1 V_\nu(x) dx = 2.$$

Alors

$$\int_{-1}^1 (S_\nu(z_0 + h_n u) - S_\nu(z_0)) du = 0.$$

Puisque

$$|\dot{S}_\nu(x)| = \varphi_n^{-1} h_n^{-1} \left| \dot{V}_\nu\left(\frac{x-z_0}{h_n}\right) \right|$$

$$\leq n^{-\alpha/(2\beta+1)} \nu^{-1} c^*, \quad c^* = 2 \int_{-1}^1 |\dot{g}(u)| du,$$

la fonction $S_\nu \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$, si on choisit $n \geq 1$ tel que

$$n^{-\alpha/(2\beta+1)} \nu^{-1} c^* \leq \delta^{-1}, \quad \text{i.e } n \geq (\delta c^* / \nu)^{(2\beta+1)/\alpha}$$

On peut remarquer que la constante de Hölder pour cette famille de fonctions $(S_\nu, 0 < \nu < 1/4)$ de régularité β , est donnée par

$$|\dot{S}_\nu(x) - \dot{S}_\nu(y)| = \varphi_n^{-1} h_n^{-1} \left| \dot{V}_\nu\left(\frac{x-z_0}{h_n}\right) - \dot{V}_\nu\left(\frac{y-z_0}{h_n}\right) \right|$$

$$= h_n^\alpha \left| \dot{V}_\nu\left(\frac{x-z_0}{h_n}\right) - \dot{V}_\nu\left(\frac{y-z_0}{h_n}\right) \right| \mathbf{1}_{|x-y| > h_n} + h_n^\alpha |\dot{V}_\nu(\theta)| \left| \frac{x-y}{h_n} \right| \mathbf{1}_{|x-y| \leq h_n}$$

$$\leq K_\nu^* |x-y|^\alpha,$$

où $(x-z_0)h_n^{-1} \leq \theta \leq (y-z_0)h_n^{-1}$ et $K_\nu^* = 2 \max_{|z| \leq 1} |\dot{V}_\nu(z)| + \max_{|z| \leq 1} |\ddot{V}_\nu(z)|$. Donc par (1.6), $K_\nu^* \approx \nu^{-2} \rightarrow \infty$ quand $\nu \rightarrow 0$. Ce qui veut dire qu'il n'existe pas de classe de Hölder $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$, contenant cette famille de fonctions $(S_\nu, 0 < \nu < 1/4)$.

Le risque d'un estimateur \tilde{S}_n de $S(z_0)$ est défini par

$$\mathcal{R}_n(\tilde{S}_n, S) = \sup_{p \in \mathcal{P}_{\sigma^*}} \mathbf{E}_{S,p} |\tilde{S}_n(z_0) - S(z_0)|,$$

où $\mathbf{E}_{S,p}$ est l'espérance calculée par rapport à la loi $\mathbf{P}_{S,p}$ correspondant à la fonction S dans le modèle (1.1) avec p la densité des variables aléatoires ξ_k prise dans l'ensemble \mathcal{P}_{σ^*} qu'on définira plus tard. Soulignons le fait que ce risque est robuste par rapport au bruit puisque on prend le supremum sur la famille de densité \mathcal{P}_{σ^*} .

Dans le cas d'un modèle de régression avec un bruit gaussien, c'est-à-dire quand $\xi_k \sim \mathcal{N}(0, \sigma^2)$ et pour $S \in C^1([0; 1])$, Galtchouk et Pergamenshchikov [2006a] ont obtenu la borne asymptotique inférieure exacte du risque minimax ainsi qu'un estimateur asymptotiquement efficace.

Théorème 1.2.2 *Galtchouk et Pergamenshchikov [2006a]*

Pour tout $\delta \in]0; 1[$, on a :

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \mathcal{R}_{z_0, \delta, n}(\tilde{S}_n) \geq \mathbb{E}|\eta|, \quad \eta \sim \mathcal{N}(0, \sigma^2/2).$$

où $\mathcal{R}_{z_0, \delta, n}(\tilde{S}_n)$ représente le risque maximal de l'estimateur \tilde{S}_n défini sur une classe fonctionnelle semblable à (1.4).

Théorème 1.2.3 *Galtchouk et Pergamenshchikov [2006a]*

Soit $Q = \mathbb{I}_{[-1; 1]}$. Alors l'estimateur à noyau Q défini par

$$\hat{S}_n = \left(\sum_{k=1}^n Q \left(\frac{x_k - z_0}{h_n} \right) \right)^{-1} \sum_{k=1}^n Q \left(\frac{x_k - z_0}{h_n} \right) y_k \quad (1.7)$$

est asymptotiquement efficace car il vérifie la relation

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \mathcal{R}_{z_0, \delta, n}(\hat{S}_n) \leq \mathbb{E}|\eta|, \quad \eta \sim \mathcal{N}(0, \sigma^2/2).$$

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

$$S_\nu(x) = \varphi_n^{-1} V_\nu \left(\frac{x - z_0}{h_n} \right),$$

où

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

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

La constante de Hölder des fonctions V_ν est de l'ordre de ν^{-2} quand le paramètre ν tend vers 0. Il n'existe donc pas de classe de Hölder $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ contenant toute la famille $\tilde{\Sigma}_n$.

Une démarche pour obtenir la borne inférieure du risque minimax nécessite dans Galtchouk et Pergamenshchikov [2006a] de faire tendre la constante de Hölder des fonctions considérées vers l'infini (quand $\nu \rightarrow 0$).

La constante δ majorant l'expression $\varphi_n \left| \int_{-1}^1 (S(x_0 + uh_n) - S(x_0)) du \right|$ dans la définition de $\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$, 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.

Par ailleurs la procédure décrite ci-dessus est robuste par rapport au bruit. En effet, notons $\mathcal{P}_{\varepsilon,L}$ l'ensemble des lois de probabilité de moyenne nulle et de variance 1 et telles que $\mathbb{E}|\xi|^{2+\varepsilon} \leq L$ 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 $\mathcal{P}_{\varepsilon,L}$. On définit alors le risque robuste d'un estimateur \tilde{S}_n par

$$\tilde{\mathcal{R}}_n(\tilde{S}_n) = \sup_{p \in \mathcal{P}_{\varepsilon,L}} \sup_{S \in \mathcal{U}_{z_0,\delta,n}} \varphi_n \mathbf{E}_{S,p} |\tilde{S}_n(z_0) - S(z_0)|, \quad \text{avec } \varphi_n = n^{\beta/2\beta+1}$$

Dans ce cas, la borne inférieure du risque minimax correspondant se déduit immédiatement du Théorème 1.2.2. Pour tout $\delta \in]0; 1[$, on a :

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \tilde{\mathcal{R}}_{z_0,\delta,n}(\tilde{S}_n) \geq \mathbf{E}|\eta|, \quad \eta \sim \mathcal{N}(0, \sigma^2/2).$$

Enfin, Galtchouk et Pergamenshchikov [2006a] montrent que la borne supérieure du risque maximal de l'estimateur (1.7) est identique :

Théorème 1.2.4 *Galtchouk et Pergamenshchikov [2006a]*

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \tilde{\mathcal{R}}_{z_0,\delta,n}(\hat{S}_n) \leq \mathbf{E}|\eta|, \quad \eta \sim \mathcal{N}(0, \sigma^2/2).$$

Nous allons maintenant décrire les résultats obtenus pour le modèle autorégressif (1.1) qui sont l'objet des théorèmes suivants démontrés au Chapitre 2.

En premier lieu on démontre que la suite φ_n définie ci-dessus est une vitesse de convergence optimale pour une classe Höldérienne $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$. Concernant la borne inférieure, on a :

Théorème 1.2.5 *Pour tout $K > 0$ et $0 < \varepsilon < 1$*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \varphi_n \mathcal{R}_n(\tilde{S}_n, S) > 0, \quad (1.8)$$

où l'infimum est pris sur tous les estimateurs \tilde{S}_n .

On obtient également une borne supérieure pour l'estimateur à noyau (2.2).

Théorème 1.2.6 *Pour tout $K > 0$ et $0 < \varepsilon < 1$ l'estimateur à noyau (2.2) avec les paramètres (2.4)–(2.6) satisfait l'inégalité suivante*

$$\limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \varphi_n \mathcal{R}_n(\hat{S}_n, S) < \infty. \quad (1.9)$$

Les Théorèmes 1.2.5 et 1.2.6 impliquent que la suite φ_n est une vitesse de convergence optimale (minimax) pour toute classe de Hölder stable de régularité β , i.e. l'estimateur (2.2) vérifiant les relations (2.4)–(2.6) est optimal en vitesse de convergence par rapport à la classe fonctionnelle (2.9).

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 $j = 1, 2, \dots, \lfloor \beta \rfloor$ et $\int Q(u) du < \infty$, où $\lfloor a \rfloor$ désigne le plus grand entier strictement plus petit que a .

Maintenant on va étudier les propriétés d'efficacité asymptotique pour l'estimateur optimal en vitesse de convergence. Pour cela comme dans Galtchouk et Pergamenschikov [2006a] nous utilisons la famille Hölderienne faible stable $\mathcal{U}_{\delta, n}^{(\beta)}(z_0, \varepsilon)$ au point z_0 .

De plus, on pose

$$\tau(S) = 1 - S^2(z_0). \quad (1.10)$$

Grâce à cette fonction nous décrivons la borne inférieure pour le risque minimax .

Théorème 1.2.7 *Pour tous $\delta > 0$ et $0 < \varepsilon < 1$*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \sup_{S \in \mathcal{U}_{\delta, n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \varphi_n \mathcal{R}_n(\tilde{S}_n, S) \geq \mathbf{E}|\eta|, \quad (1.11)$$

où η est une variable aléatoire gaussienne de paramètres $(0, 1/2)$.

Théorème 1.2.8 *L'estimateur à noyau $Q = 1_{[-1, 1]}$ (2.2) vérifiant les relations (2.4)–(2.5) satisfait l'inégalité suivante*

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{U}_{\delta, n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \varphi_n \mathcal{R}_n(\hat{S}_n, S) \leq \mathbf{E}|\eta|,$$

où η est une variable aléatoire gaussienne de paramètres $(0, 1/2)$.

Les théorèmes 1.2.7 et 1.2.8 impliquent que l'estimateur (2.2) est asymptotiquement efficace.

1.2.2 Cas adaptatif

Dans cette partie, on suppose la régularité de la fonction autorégressive S inconnue, et on se place sur la classe Höldérienne forte $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ correspondant à la vraie valeur du paramètre β qu'on situe dans un segment $[\beta_*; \beta^*]$ connu. Les résultats présentés ici seront démontrés au Chapitre 3 et sont relatés dans Arkoun [2009].

Le risque d'un estimateur \tilde{S}_n de $S(z_0)$ est défini par

$$\mathcal{R}_n(\tilde{S}_n) = \sup_{\beta \in [\beta_*; \beta^*]} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} N(\beta) \mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|,$$

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

Pour montrer que $N(\beta)$ est la vitesse de convergence minimax adaptative des estimateurs sur la famille de classe $(\mathcal{H}^{(\beta)}(z_0, K, \varepsilon))_{\beta \in [\beta_*; \beta^*]}$, on donne une borne inférieure du risque minimax adaptatif.

Théorème 1.2.9 *Le risque minimax adaptatif admet la borne inférieure suivante*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \mathcal{R}_n(\tilde{S}_n) \geq \frac{1}{4}.$$

où l'infimum est pris sur tous les estimateurs \tilde{S}_n .

Afin de construire un estimateur adaptatif en vitesse de convergence, nous ne pouvons plus considérer l'estimateur à noyau (2.2)

$$S_h^* = \frac{1}{A_n(h)} \sum_{k=1}^n Q\left(\frac{x_k - z_0}{h}\right) y_{k-1} y_k \mathbf{1}_{(A_n(h) \geq a^*)}, \quad A_n(h) = \sum_{k=1}^n Q\left(\frac{x_k - z_0}{h}\right) y_{k-1}^2$$

car sa fenêtre $h = h_n$ dépend de la régularité $\beta \in [\beta_*; \beta^*] \subset]0; 1]$ désormais inconnue. C'est pourquoi on procède suivant la méthode de Lepskiï [1990]. On partitionne l'intervalle $[\beta_*; \beta^*]$ de la manière suivante :

$$\beta_k = \beta_* + k \frac{\beta^* - \beta_*}{m}, \quad m = [\ln d_n] + 1, \quad k = 0, \dots, m \text{ et } d_n = \frac{n}{\ln n},$$

où $[a]$ désigne la partie entière du réel a .

A ces valeurs, on associe les fenêtres correspondantes

$$h_k = d_n^{-1/(2\beta_k+1)} \quad \text{et les vitesses} \quad N_k = d_n^{\beta_k/(2\beta_k+1)}.$$

Finalement, pour $\lambda > K + e\sqrt{4 + \frac{4}{2\beta_* + 1}}$, on pose

$$\omega(h_j) = \max_{0 \leq k \leq j} \left(|S_{h_j}^* - S_{h_k}^*| - \frac{\lambda}{N_{k+1}} \right)$$

pour définir l'indice optimal de la fenêtre comme

$$\widehat{k} = \inf \left\{ 0 \leq j \leq m : \omega(h_j) \geq \frac{\lambda}{N_j} \right\} - 1. \quad (1.12)$$

L'estimateur utilisé sera alors $\widehat{S}_n = S_{\widehat{h}}^*$ avec $\widehat{h} = h_{\widehat{k}}$.

Théorème 1.2.10 *Pour tout $0 < \varepsilon < 1$, on a*

$$\limsup_{n \rightarrow \infty} \mathcal{R}_n(\widehat{S}_n) < \infty,$$

ce qui fait que l'estimateur \widehat{S}_n est adaptatif en vitesse de convergence.

Chapitre 2

Modèles autorégressifs : cas non adaptatif

2.1 Introduction

On considère le problème de l'estimation de la fonction S en un point fixe $z_0 \in]0; 1[$, où l'on dispose des observations régies par le modèle autorégressif suivant

$$y_k = S(x_k)y_{k-1} + \xi_k, \quad 1 \leq k \leq n, \quad (2.1)$$

les régresseurs $x_k = k/n$ étant déterministes, y_0 étant une constante et les variables aléatoires ξ_k indépendantes, identiquement distribuées, avec $\mathbf{E}\xi_k = 0$ et $\mathbf{E}\xi_k^2 = 1$.

Le modèle (2.1) est une généralisation du processus autorégressif du premier ordre. De tels modèles sont utilisés dans les séries temporelles et leurs applications. L'intérêt des séries temporelles peut apparaître dans différents domaines : par exemple en finance pour décrire les prix d'actifs risqués et des indices, pour des processus de type GARCH, ARCH(1) et pour des modèles bilinéaires et ARMA (Embrechts, Kluppelberg et Mikosch [1997]). Un autre domaine d'application des modèles autorégressifs est l'économétrie avec les modèles à décalages temporels. En effet la théorie économique postule couramment non pas des effets synchrones mais des effets retardés et les modèles autorégressifs peuvent décrire des variables retardées, aussi bien des variables endogènes que des variables exogènes (Goldfeld et Quandt [1972]). On trouve aussi des applications des modèles autorégressifs en biologie. par exemple des études au sein du laboratoire de génétique moléculaire, évolutive et médicale (LGMEM) de L'INSERM de la faculté de médecine Necker à Paris ont permis à Guyon [2007] d'élaborer un modèle qui permet la détection du vieillissement cellulaire. En utilisant le Théorème central limite et la loi des grands nombres pour un processus stochastique il arrive à détecter le vieillissement cellulaire d'une bactérie "cherichia coli" grâce aux données expérimentales collectées dans le LGMEM et étudie la bifurcation des modèles autorégressifs.

Les processus autorégressifs ont été considérés aussi bien dans le cadre paramétrique que non paramétrique. Par exemple le problème de l'estimation de fonctions paramétriques

a été étudié dans Dahlhaus [1996b] où l'auteur s'est intéressé au comportement de l'estimateur du maximum de vraisemblance Gaussienne pour les séries temporelles qui ont un comportement localement stationnaire.

De plus, Dahlhaus [1996b] étudie les propriétés spectrales du processus stationnaire (2.1) avec la fonction non paramétrique S .

Ce chapitre traite l'estimation non paramétrique de la fonction autorégressive S en un point z_0 , où la régularité de S est supposée connue. Pour ce problème nous utilisons l'estimateur à noyau modifié

$$\hat{S}_n(z_0) = \frac{1}{A_n} \sum_{k=1}^n Q(u_k) y_{k-1} y_k \mathbf{1}_{(A_n \geq a^*)}, \quad (2.2)$$

où $Q(\cdot)$ est la fonction noyau,

$$A_n = \sum_{k=1}^n Q(u_k) y_{k-1}^2 \quad \text{avec} \quad u_k = \frac{x_k - z_0}{h_n};$$

a^* et h_n étant des paramètres positifs.

On suppose d'abord que la fonction inconnue S appartient à la classe Hölderienne locale stable au point z_0 avec une régularité connue $\beta \in]1, 2]$. Cette classe sera définie par la suite. On trouve une borne inférieure asymptotique positive (quand $n \rightarrow \infty$) pour le risque minimax avec le coefficient de normalisation

$$\varphi_n = n^{\frac{\beta}{2\beta+1}}. \quad (2.3)$$

Pour obtenir cette vitesse de convergence on utilise dans (2.2)

$$h_n = n^{-\frac{1}{2\beta+1}} \quad \text{et} \quad a^* = \kappa_n n h_n, \quad (2.4)$$

où la suite de nombres positifs $(\kappa_n)_{n \geq 1}$ vérifie

$$\lim_{n \rightarrow \infty} \kappa_n = 0 \quad \text{et} \quad \lim_{n \rightarrow \infty} \frac{h_n}{\kappa_n^2} = 0. \quad (2.5)$$

Pour la fonction noyau on suppose que

$$\int_{-1}^1 Q(z) dz > 0 \quad \text{et} \quad \int_{-1}^1 z Q(z) dz = 0. \quad (2.6)$$

Dans ce chapitre on prouve que l'estimateur (2.2) avec les relations (2.4)–(2.6) est asymptotiquement minimax, i.e. on démontre que la borne supérieure asymptotique du risque maximal de celui-ci par rapport à la classe Hölderienne est finie.

En deuxième lieu on étudie les propriétés asymptotiques pour l'estimateur minimax (2.2). Pour cela, comme dans Galtchouk et Pergamenshchikov [2006a] on introduit la

classe Hölderienne faible locale stable. Dans ce cas nous obtenons une borne asymptotique inférieure strictement positive pour le risque minimax avec la vitesse de convergence φ_n . Par ailleurs, on démontre que pour l'estimateur (2.2), avec les relations (2.4)–(2.5) et la fonction noyau $Q = \mathbf{1}_{[-1,1]}$, la borne supérieure asymptotique du risque maximal coïncide avec la borne inférieure, i.e. dans ce cas l'estimateur est asymptotiquement efficace.

Ce problème d'estimation a été étudié dans le cas d'une fonction de régression Höldérienne et cette dernière a été étudiée par de nombreux auteurs. Par exemple Sacks et Ylvisaker [1981] ont montré que l'estimateur linéaire minimax est un estimateur à noyau. Donoho et Liu [1991] ont ensuite 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.

Un exemple 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 puis 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 $n^{\beta/(2\beta+1)}$.

Un autre 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 , $p \in [1; \infty[$.

Belitser [2000b] considère le modèle précédent avec des conditions lipschitziennes. L'auteur propose un estimateur récursif pour le problème de l'estimation de la fonction autorégressive. Avec le risque quadratique, Belitser [2000b] établit une vitesse de convergence sans démontrer son optimalité.

Moulines et al. [2005] démontrent que la vitesse de convergence est optimale pour le risque quadratique en utilisant des méthodes récursives pour le modèle autorégressif non paramétrique d'ordre $d \geq 1$. Notons que dans ce chapitre nous établissons une vitesse de convergence optimale mais le risque considéré est différent de celui utilisé dans Moulines et al. [2005], et les hypothèses y sont plus faibles.

On traite ici de l'estimation non paramétrique d'une fonction autorégressive 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 $\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$ 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 (2.12)). Puis on définit le risque robuste $\mathcal{R}_n(\tilde{S}_n, S)$ d'un estimateur \tilde{S}_n de $S(z_0)$ et le risque minimax $\inf_{\tilde{S}_n} \mathcal{R}_n(\tilde{S}_n, S)$ (voir (2.7)).

La prochaine section présente le problème dans le cas de bruits de loi inconnue, 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. Enfin l'Annexe A contient des résultats techniques utiles dans les démonstrations.

2.2 Description du problème

En premier lieu on suppose que le bruit dans le modèle (2.1) est de loi inconnue, plus précisément les variables aléatoires $(\xi_k)_{1 \leq k \leq n}$ sont supposées indépendantes identiquement distribuées selon une densité p (par rapport à la mesure de Lebesgue) appartenant à la classe fonctionnelle \mathcal{P}_{σ^*} définie par

$$\mathcal{P}_{\sigma^*} := \left\{ p \geq 0 : \int_{-\infty}^{+\infty} p(x) dx = 1, \quad \int_{-\infty}^{+\infty} x p(x) dx = 0, \right. \\ \left. \int_{-\infty}^{+\infty} x^2 p(x) dx = 1 \quad \text{et} \quad \int_{-\infty}^{+\infty} |x|^4 p(x) dx \leq \sigma^* \right\}$$

avec $\sigma^* \geq 3$. Notons que la densité de la loi gaussienne standard appartient à \mathcal{P}_{σ^*} . Dans la suite on note cette densité par p_0 .

Le problème est d'estimer la fonction $S(\cdot)$ en un point fixé $z_0 \in]0, 1[$, i.e. la valeur $S(z_0)$. Pour ce problème nous utilisons le risque proposé dans Galtchouk et Pergamenshchikov [2006a]. En effet pour tout estimateur $\tilde{S}_n = \tilde{S}_n(z_0)$ (i.e. toute fonction mesurable par rapport aux observations $(y_k)_{1 \leq k \leq n}$), on pose

$$\mathcal{R}_n(\tilde{S}_n, S) = \sup_{p \in \mathcal{P}_{\sigma^*}} \mathbf{E}_{S,p} |\tilde{S}_n(z_0) - S(z_0)|, \quad (2.7)$$

où $\mathbf{E}_{S,p}$ est l'espérance prise par rapport à la distribution $\mathbf{P}_{S,p}$ du vecteur (y_1, \dots, y_n) dans (2.1) correspondant à la fonction S et la densité p de \mathcal{P}_{σ^*} .

Pour obtenir la stabilité (uniformément par rapport à la fonction S) du modèle (2.1) on suppose (voir Dahlhaus [1996a] et Dahlhaus [1996b]) que pour un certain $0 < \varepsilon < 1$ fixé la fonctions inconnue S appartient à l'ensemble stable

$$\Gamma_\varepsilon = \{S \in \mathbf{C}_1[0, 1] : \|S\| \leq 1 - \varepsilon\}, \quad (2.8)$$

où $\|S\| = \sup_{0 \leq x \leq 1} |S(x)|$. Ici $\mathbf{C}_1[0, 1]$ est l'espace de Banach des fonctions continûment différentiables $[0, 1] \rightarrow \mathbb{R}$. Pour une constante fixé $K > 0$ et $0 < \alpha \leq 1$ on définit la classe Höldérienne stable correspondante au point z_0 comme

$$\mathcal{H}^{(\beta)}(z_0, K, \varepsilon) = \left\{ S \in \Gamma_\varepsilon : \|\dot{S}\| \leq K \quad \text{et} \quad \Omega^*(z_0, S) \leq K \right\} \quad (2.9)$$

avec $\beta = 1 + \alpha$ et

$$\Omega^*(z_0, S) = \sup_{x \in [0, 1]} \frac{|\dot{S}(x) - \dot{S}(z_0)|}{|x - z_0|^\alpha}.$$

Tout d'abord on démontre que la suite φ_n est une vitesse de convergence optimale pour la classe fonctionnelle $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$.

Commençons par la borne inférieure.

Théorème 2.2.1 *Pour tous $K > 0$ et $0 < \varepsilon < 1$*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \varphi_n \mathcal{R}_n(\tilde{S}_n, S) > 0, \quad (2.10)$$

où l'infimum est pris sur tous les estimateurs.

Ensuite, nous obtenons une borne supérieure du risque pour l'estimateur à noyau \hat{S}_n .

Théorème 2.2.2 *Pour tous $K > 0$ et $0 < \varepsilon < 1$ l'estimateur à noyau \hat{S}_n vérifiant les relations (2.4)–(2.6) satisfait l'inégalité suivante*

$$\limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \varphi_n \mathcal{R}_n(\hat{S}_n, S) < \infty. \quad (2.11)$$

Les Théorèmes 2.2.1 et 2.2.2 impliquent que la suite φ_n est une vitesse de convergence optimale (minimax) pour la classe de Hölder stable de régularité β , i.e. l'estimateur \hat{S}_n est optimal en vitesse de convergence sur cette classe.

En deuxième lieu on étudie l'efficacité de l'estimateur \hat{S}_n .

Dans le cas où S appartient à la classe de Hölder $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ le problème de l'estimation asymptotiquement efficace reste ouvert. En conséquence on travaille avec un risque minimax pris sur une classe plus large, appelée classe Höldérienne faible. Pour l'estimateur \hat{S}_n , de la même manière que dans Galtchouk et Pergamenshchikov [2006a] nous utilisons la famille des classes Höldériennes faibles locales stables en un point z_0 . Ainsi pour tout $\delta > 0$ on définit une telle classe par

$$\mathcal{U}_{\delta, n}^{(\beta)}(z_0, \varepsilon) = \left\{ S \in \Gamma_\varepsilon : \|\dot{S}\| \leq \delta^{-1} \quad \text{et} \quad |\Omega_h(z_0, S)| \leq \delta h_n^\beta \right\}, \quad (2.12)$$

où

$$\Omega_{h_n}(z_0, S) = \int_{-1}^1 (S(z_0 + uh_n) - S(z_0)) du$$

et h_n est donné par (2.4). Précisons que la régularité β de la fonction S est supposée connue dans tout ce chapitre.

On définit

$$\tau(S) = 1 - S^2(z_0). \quad (2.13)$$

Grâce à cette fonction, nous obtenons une borne inférieure pour le risque minimax.

Théorème 2.2.3 *Pour tous $\delta > 0$ et $0 < \varepsilon < 1$*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \varphi_n \mathcal{R}_n(\tilde{S}_n, S) \geq \mathbf{E}|\eta|, \quad (2.14)$$

où η est une variable aléatoire gaussienne de paramètres $(0, 1/2)$.

Théorème 2.2.4 *L'estimateur \hat{S}_n à noyau $Q(z) = 1_{[-1,1]}$ vérifiant les relations (2.4)–(2.5) satisfait l'inégalité suivante*

$$\limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \varphi_n \mathcal{R}_n(\hat{S}_n, S) \leq \mathbf{E}|\eta|,$$

où η est une variable aléatoire gaussienne de paramètres $(0, 1/2)$.

Les Théorèmes 2.2.3 et 2.2.4 impliquent que l'estimateur $\hat{S}_n(z_0)$ est asymptotiquement efficace.

Remarque 2.2.5 *On peut démontrer que pour tous $0 < \delta < 1$ et $n \geq 1$*

$$\mathcal{H}^{(\beta)}(z_0, \delta, \varepsilon) \subset \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon).$$

Cela veut dire que le coefficient de normalisation "naturel" pour la classe fonctionnelle $\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$ n'est autre que la suite φ_n . On remarque qu'on conserve la même vitesse de convergence φ_n sur la classe $\mathcal{H}^{(\beta)}(z_0, \delta, \varepsilon)$ et sur $\mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$.

2.3 Bornes asymptotiques pour des bruits de loi inconnue

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

2.3.1 Borne inférieure

Preuve du Théorème 2.2.1

Pour démontrer (2.10) il suffit de prouver que

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \mathbf{E}_{S, p_0} \psi_n(\tilde{S}_n, S) > 0, \quad (2.15)$$

où

$$\psi_n(\tilde{S}_n, S) = \varphi_n |\tilde{S}_n(z_0) - S(z_0)|.$$

Nous utilisons une méthode similaire à celle proposée par Ibragimov et Has'minskiï [1981] pour obtenir une borne inférieure pour le problème de l'estimation d'une densité. D'abord nous choisissons la famille paramétrique correspondante dans $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$. Soit V une fonction continûment différentiable telle que $\int_{-1}^1 V(z) dz > 0$ et $V(z) = 0$ pour tout $|z| \geq 1$. On définit

$$S_u(x) = \frac{u}{\varphi_n} V\left(\frac{x - z_0}{h_n}\right), \quad (2.16)$$

où φ_n et h_n sont définies dans (2.3) et (2.4).

Il est clair que pour tout $z_0 - h_n \leq x \leq z_0 + h_n$,

$$\begin{aligned} |\dot{S}_u(x) - \dot{S}_u(z_0)| &= \frac{|u|}{h_n \varphi_n} \left| \dot{V}\left(\frac{x - z_0}{h_n}\right) - \dot{V}(0) \right| \\ &\leq \frac{|u|}{h_n \varphi_n} V''_* \left| \frac{x - z_0}{h_n} \right| \leq |u| V''_* |x - z_0|^\alpha, \end{aligned}$$

où $V''_* = \max_{|z| \leq 1} |\ddot{V}(z)|$. Donc, pour tout $0 < u \leq u^* = K/V''_*$ on obtient

$$\sup_{z_0 - h_n \leq x \leq z_0 + h_n} \frac{|\dot{S}_u(x) - \dot{S}_u(z_0)|}{|x - z_0|^\alpha} \leq K.$$

De plus, par la définition (2.16) pour tout $x > z_0 + h_n$

$$\dot{S}_u(x) = \dot{S}_u(z_0 + h_n) = 0$$

et pour tout $x < z_0 - h_n$

$$\dot{S}_u(x) = \dot{S}_u(z_0 - h_n) = 0.$$

Donc la dernière inégalité implique que

$$\sup_{|u| \leq u^*} \Omega^*(z_0, S_u) \leq K,$$

où la quantité $\Omega^*(z_0, S)$ est définie par (2.9).

Cela veut dire qu'il existe $n_{K, \varepsilon} > 0$ tel que $S_u \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ pour tous $|u| \leq u^*$ et $n \geq n_{K, \varepsilon}$. Donc pour tout $n \geq n_{K, \varepsilon}$ et pour tout estimateur \tilde{S}_n on peut écrire les minoration suivantes

$$\begin{aligned} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \mathbf{E}_{S, p_0} \psi_n(\tilde{S}_n, S) &\geq \sup_{|u| \leq u^*} \mathbf{E}_{S_u, p_0} \psi_n(\tilde{S}_n, S_u) \\ &\geq \frac{1}{2b} \int_{-b}^b \mathbf{E}_{S_u, p_0} \psi_n(\tilde{S}_n, S_u) du := I_n(b, \sigma) \end{aligned} \quad (2.17)$$

pour tout $0 < b \leq u^*$.

Notons que pour tout S la mesure \mathbf{P}_{S,p_0} est équivalente à la mesure \mathbf{P}_{0,p_0} , où \mathbf{P}_{0,p_0} est la distribution du vecteur (y_1, \dots, y_n) dans le modèle (2.1) correspondant à la fonction $S = 0$ et à la densité p_0 du bruit est gaussienne standard, i.e. les variables aléatoires (y_1, \dots, y_n) sont gaussiennes standard indépendantes identiquement distribuées par rapport à la mesure \mathbf{P}_{0,p_0} . Dans la suite on note \mathbf{P}_{0,p_0} par \mathbf{P} . Il très facile de voir que dans ce cas la dérivée de Radom-Nikodym peut s'écrire comme

$$\begin{aligned} \rho_n(u) &= \frac{d\mathbf{P}_{S_u,p_0}}{d\mathbf{P}}(y_1, \dots, y_n) \\ &= \exp \left\{ -\frac{1}{2} \sum_{k=1}^n \left((y_k - S_u(x_k)y_{k-1})^2 - y_k^2 \right) \right\} \\ &= \exp \left(u\varsigma_n\eta_n - \frac{u^2}{2}\varsigma_n^2 \right), \end{aligned}$$

où

$$\varsigma_n^2 = \frac{1}{\varphi_n^2} \sum_{k=1}^n V^2(u_k)y_{k-1}^2 \quad \text{et} \quad \eta_n = \frac{1}{\varsigma_n \varphi_n} \sum_{k=1}^n V(u_k) y_{k-1} y_k.$$

En utilisant la loi des grands nombres on obtient

$$\mathbf{P} - \lim_{n \rightarrow \infty} \varsigma_n^2 = \lim_{n \rightarrow \infty} \frac{1}{nh_n} \sum_{k=k_*}^{k^*} V^2(u_k)\xi_{k-1}^2 = \int_{-1}^1 V^2(u)du = \sigma^2,$$

où

$$k_* = [nz_0 - nh_n] + 1 \quad \text{et} \quad k^* = [nz_0 + nh_n]. \quad (2.18)$$

Ici $[a]$ désigne la partie entière de a .

Par le théorème central limite pour les martingales (voir Helland [1981] et Rebolledo [1980]), il est facile de voir que sous la mesure \mathbf{P}

$$\eta_n \implies \mathcal{N}(0, 1) \quad \text{quand} \quad n \rightarrow \infty.$$

Ainsi on réécrit la densité de Radom-Nikodym sous la forme asymptotique

$$\rho_n(u) = \exp \left(u\sigma\eta_n - \frac{u^2\sigma^2}{2} + r_n \right),$$

où r_n converge en \mathbf{P} -probabilité vers zéro.

En notant \mathbf{E} l'espérance correspondant à la mesure de probabilité \mathbf{P} , on a

$$I_n(b, \sigma) \geq \frac{1}{2b} \int_{-b}^b \mathbf{E} \mathbb{I}_{B_d} \psi_n(\tilde{S}_n, S_u) \varrho_n(u) du + \delta_n(b, \sigma) =: J_n(b, \sigma) + \delta_n(b, \sigma), \quad (2.19)$$

où

$$\begin{aligned} B_d &= \{|\eta_n| \leq d\} \text{ et } d = \sigma(b - \sqrt{b}), b > 1, \\ \varrho_n(u) &= \exp\left(u\sigma\eta_n - \frac{u^2\sigma^2}{2}\right), \\ \delta_n(b, \sigma) &= \frac{1}{2b} \int_{-b}^b \mathbf{E}\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\theta_n(u)du, \\ \theta_n(u) &= \rho_n(u) - \varrho_n(u). \end{aligned}$$

Remarquons que $\rho_n(u) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \rho_\infty(u) = \exp\left(u\sigma\eta - \frac{u^2\sigma^2}{2}\right)$. On montre aisément que $\mathbf{E}\rho_\infty(u) = 1$ et on a aussi $\mathbf{E}\rho_n(u) = 1$ car $\rho_n(u)$ est une densité. Donc, en utilisant le Lemme 2.4.4, $\{\rho_n(u), n \geq 1\}$ est uniformément intégrable. Comme $\varrho_n(u)$ est borné sur B_d , on obtient l'intégrabilité uniforme de la famille $\{\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\theta_n(u), n \geq 1\}$.

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

$$\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\theta_n(u) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0.$$

Puis il s'en suit : $\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\theta_n(u) \xrightarrow[n \rightarrow \infty]{\mathbb{L}^1} 0$ et $\mathbf{E}\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\theta_n(u) \xrightarrow[n \rightarrow \infty]{} 0$.

Finalement, par convergence dominée, il vient $\sup_{\tilde{S}_n} |\delta_n(b, \sigma)| \xrightarrow[n \rightarrow \infty]{} 0$ dans (2.19).

Intéressons-nous maintenant au terme $J_n(b, \sigma)$ dans (2.19). Réécrivons d'abord $\varrho_n(u) = \zeta_n e^{-\sigma^2(u - \tilde{\eta}_n)^2/2}$ avec $\zeta_n = e^{\eta_n^2/2}$ et $\tilde{\eta}_n = \frac{\eta_n}{\sigma}$. on obtient successivement

$$\begin{aligned} J_n(b, \sigma) &= \frac{1}{2b} \int_{-b}^b \mathbf{E}\mathbb{I}_{B_d}\psi_n(\tilde{S}_n, S_u)\varrho_n(u)du \\ &= \mathbf{E}\mathbb{I}_{B_d}\zeta_n \frac{1}{2b} \int_{-b}^b |u - c_n| \exp\left(-\frac{\sigma^2}{2}(u - \tilde{\eta}_n)^2\right) du \\ &= \mathbf{E}\mathbb{I}_{B_d}\zeta_n \frac{1}{2b} \int_{-b - \tilde{\eta}_n}^{b - \tilde{\eta}_n} |t - c_n + \tilde{\eta}_n| \exp\left(-\frac{\sigma^2}{2}t^2\right) dt \\ &\geq \mathbf{E}\mathbb{I}_{B_d}\zeta_n \frac{1}{2b} \int_{-\sqrt{b}}^{\sqrt{b}} |t - c_n + \tilde{\eta}_n| \exp\left(-\frac{\sigma^2}{2}t^2\right) dt, \end{aligned}$$

où $c_n = \varphi_n \tilde{S}_n$. En utilisant l'inégalité d'Anderson (voir le Lemme 2.4.5) on aura

$$J_n(b, \sigma) \geq \mathbf{E}\mathbb{I}_{B_d}\zeta \frac{1}{2b} \int_{-\sqrt{b}}^{\sqrt{b}} |t| \exp\left(-\frac{\sigma^2}{2}t^2\right) dt.$$

Ainsi, en utilisant le fait que $\mathbf{E}\mathbb{I}_{B_d}\zeta = 2\sigma(b - \sqrt{b})/\sqrt{2\pi}$, il vient

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} I_n(b, \sigma) \geq \frac{\sigma}{\sqrt{2\pi}} \frac{b - \sqrt{b}}{b} \int_{-\sqrt{b}}^{\sqrt{b}} |t| \exp\left(-\frac{\sigma^2}{2} t^2\right) dt =: A(b, \sigma), \quad (2.20)$$

cette dernière quantité étant strictement positive, on en déduit le théorème. \square

Preuve du Théorème 2.2.3

De façon similaire à la preuve du Théorème 2.2.1 on choisit la famille fonctionnelle paramétrique $S_{u,\nu}(\cdot)$ correspondant à la forme (2.16) avec la fonction $V = V_\nu$ définie comme suit

$$V_\nu(x) = \nu^{-1} \int_{-\infty}^{\infty} \tilde{Q}_\nu(u) g\left(\frac{u-x}{\nu}\right) du,$$

où $\tilde{Q}_\nu(u) = \mathbf{1}_{\{|u| \leq 1-2\nu\}} + 2\mathbf{1}_{\{1-2\nu \leq |u| \leq 1-\nu\}}$ avec $0 < \nu < 1/4$ et g est une fonction paire positive, infiniment différentiable telle que $g(z) = 0$ pour $|z| \geq 1$ et $\int_{-1}^1 g(z) dz = 1$.

Soient $b > 0$ et $\delta \in]0; 1[$. Notons, pour $x \in \mathbb{R}$ et $u \in [-b; b]$, $S_{u,\nu}(x) =: uS_\nu(x)$. D'après le Lemme 2.4.6, il existe un entier $n_* = n(b, \delta, \nu) > 0$ tel que $S_{u,\nu} \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$ pour tous $n \geq n_*$ et $u \in [-b; b]$. Par conséquent, si \tilde{S}_n est un estimateur de $S(z_0)$, on a pour $n \geq n_*$,

$$\begin{aligned} \varphi_n \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \mathcal{R}_n(\tilde{S}_n, S) &\geq \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \mathbf{E}_{S, p_0} \psi_n(\tilde{S}_n, S) \\ &\geq \tau_*(n, b) \frac{1}{2b} \int_{-b}^b \mathbf{E}_{S_{u,\nu}, p_0} \psi_n(\tilde{S}_n, S_{u,\nu}) du. \end{aligned}$$

où

$$\tau_*(n, b) = \inf_{|u| \leq b} \tau^{-1/2}(S_{u,\nu}).$$

Les définitions (2.13) et (2.16) impliquent que pour tout $b > 0$

$$\lim_{n \rightarrow \infty} \sup_{|u| \leq b} |\tau(S_{u,\nu}) - 1| = 0.$$

Donc, de la même manière que dans la démonstration du Théorème 2.2.1 on obtient que pour tous $b > 0$ et $0 < \nu < 1/4$

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \tau^{-1/2}(S) \varphi_n \mathcal{R}_n(\tilde{S}_n, S) \geq A(b, \sigma_\nu), \quad (2.21)$$

où la quantité $A(b, \sigma_\nu)$ est définie dans (2.20) avec $\sigma_\nu^2 = \int_{-1}^1 V_\nu^2(u) du$. Il est facile de voir que $\sigma_\nu^2 \rightarrow 2$ quand $\nu \rightarrow 0$. En passant à la limite $b \rightarrow \infty$ puis $\nu \rightarrow 0$ dans (2.21) on trouvera l'inégalité (2.14), ce qui conclut la démonstration du Théorème 2.2.3. \square

2.3.2 Borne supérieure

Preuve du Théorème 2.2.2

On rappelle que

$$A_n = \sum_{k=1}^n Q(u_k) y_{k-1}^2$$

et on note

$$\tilde{A}_n = \frac{A_n}{\varphi_n^2} \quad \text{et} \quad \hat{A}_n = \frac{1}{\tilde{A}_n} \mathbf{1}_{(\tilde{A}_n > \kappa_n)}. \quad (2.22)$$

où pour l'estimateur (2.2) on écrit l'erreur d'estimation comme

$$\hat{S}_n(z_0) - S(z_0) = -S(z_0) \mathbf{1}_{(\tilde{A}_n \leq \kappa_n)} + \frac{1}{\varphi_n} \hat{A}_n \zeta_n + \frac{1}{\varphi_n} \hat{A}_n B_n, \quad (2.23)$$

avec

$$\zeta_n = \frac{1}{\varphi_n} \sum_{k=1}^n Q(u_k) y_{k-1} \xi_k \quad \text{et} \quad B_n = \frac{1}{\varphi_n} \sum_{k=1}^n Q(u_k) (S(x_k) - S(z_0)) y_{k-1}^2.$$

Notons que le premier terme dans la quantité de droite de (2.23) est étudié dans le lemme 2.4.3.

Pour estimer le second terme nous utilisons le lemme 2.4.2 qui implique directement

$$\limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} \zeta_n^2 < \infty$$

et, donc par (2.35) du Lemme 2.4.3 nous obtenons

$$\limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} |\hat{A}_n| |\zeta_n| < \infty.$$

Estimons maintenant le dernier terme de droite de (2.23). Pour ce dernier on a besoin de démontrer que

$$\lim_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} B_n^2 < \infty. \quad (2.24)$$

Pour commencer, posons $r_k = S(x_k) - S(z_0) - \dot{S}(z_0)(x_k - z_0)$. Par la formule de Taylor on réécrit B_n comme

$$B_n = \frac{h_n}{\varphi_n} \dot{S}(z_0) \tilde{B}_n + \frac{1}{\varphi_n} \hat{B}_n,$$

où $\tilde{B}_n = \sum_{k=1}^n Q(u_k) u_k y_{k-1}^2$ et $\hat{B}_n = \sum_{k=1}^n Q(u_k) r_k y_{k-1}^2$. On rappelle que par la condition

(2.6), $\int_{-1}^1 u Q(u) du = 0$. Donc par le Lemme 2.4.2 on obtient

$$\lim_{n \rightarrow \infty} \frac{h_n^2}{\varphi_n^2} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} \tilde{B}_n^2 = 0.$$

De plus, pour toute fonction $S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ et pour $k_* \leq k \leq k^*$ (k_* et k^* sont donnés dans (2.18))

$$|r_k| = \left| \int_{z_0}^{x_k} \left(\dot{S}(u) - \dot{S}(z_0) \right) du \right| \leq K |x_k - z_0|^\beta \leq Kh^\beta = K\varphi_n^{-1},$$

par conséquent $\widehat{B}_n \leq \varphi_n \widetilde{A}_n$. Donc, par le lemme 2.4.2, il vient

$$\lim_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \frac{1}{\varphi_n^2} \mathbf{E}_{S,p} \widehat{B}_n^2 < \infty,$$

ce qui implique (2.24). D'où le Théorème 2.2.2. \square

Preuve du Théorème 2.2.4

Suivant la démonstration du Lemme A.2 de Galtchouk et Pergamenschikov [2006a] et en utilisant les deux lemmes 2.4.1 et 2.4.2 on démontre que

$$\sqrt{\frac{\tau(S)}{2}} \zeta_n \implies \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty$$

uniformément en $S \in \Gamma_\varepsilon$ et $p \in \mathcal{P}$. Donc, par le lemme 2.4.2 nous obtenons le résultat de convergence uniforme en $S \in \Gamma_\varepsilon$ et $p \in \mathcal{P}$

$$\tau^{-1/2}(S) \widehat{A}_n \zeta_n \implies \mathcal{N}(0, 1/2) \quad \text{quand } n \rightarrow \infty.$$

Par ailleurs, en utilisant l'inégalité de Burkholder et le lemme 2.4.2 pour la martingale ζ_n on déduit que

$$\lim_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} \zeta_n^4 < \infty.$$

Donc, l'inégalité (2.35) implique que la suite $(\widehat{A}_n \zeta_n)_{n \geq 1}$ est uniformément intégrable. Cela veut dire que

$$\lim_{n \rightarrow \infty} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \sup_{p \in \mathcal{P}} \left| \tau^{-1/2}(S) \mathbf{E}_{S,p} |\widehat{A}_n \zeta_n| - \mathbf{E}|\eta| \right| = 0,$$

où η est une variable aléatoire gaussienne de paramètres $(0, 1/2)$. Pour finir cette preuve il nous reste à démontrer que

$$\lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{S \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} B_n^2 = 0. \quad (2.25)$$

En effet, en posant $f_S(u) = S(z_0 + h_n u) - S(z_0)$ nous réécrivons B_n comme suit

$$B_n = \frac{1}{\varphi_n} \sum_{k=k_*}^{k^*} f_S(u_k) y_{k-1}^2 = \varphi_n \varrho_n(f_S, S) + \frac{\varphi_n}{\tau(S)} \Omega_h(z_0, S), \quad (2.26)$$

où

$$\varrho_n(f, S) = \frac{\sum_{k=1}^n f(u_k) y_{k-1}^2}{\varphi_n^2} - \frac{1}{\tau(S)} \int_{-1}^1 f(u) du$$

et $\Omega_{h_n}(z_0, S)$ est définie dans (2.12). La définition (2.13) implique que pour tout $S \in \Gamma_\varepsilon$

$$\varepsilon^2 \leq \tau(S) \leq 1. \quad (2.27)$$

Ainsi par la définition (2.12) nous obtenons

$$|B_n| \leq \varphi_n |\varrho_n(f_S, S)| + \frac{\delta}{\varepsilon^2}.$$

De plus, pour tout $S \in \mathcal{U}_{\delta, n}^{(\beta)}(z_0, \varepsilon)$ la fonction f_S satisfait l'inégalité suivante

$$\|f_S\| + \|\dot{f}_S\| \leq \delta^{-1} h_n.$$

Notons aussi que $\varphi_n h_n^2 \rightarrow 0$ quand $n \rightarrow \infty$. Donc, en utilisant le Lemme 2.4.2 avec $R = h_n/\delta$ nous obtenons (2.25) et ainsi le Théorème 2.2.4. \square

2.4 Annexe A

Dans ce paragraphe nous étudions les propriétés des distributions des processus stationnaires (2.1).

Lemme 2.4.1 *Pour tout $0 < \varepsilon < 1$ les variables aléatoires dans le modèle (2.1) satisfont l'inégalité suivante*

$$m^* = \sup_{n \geq 1} \sup_{0 \leq k \leq n} \sup_{S \in \Gamma_\varepsilon} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} y_k^4 < \infty. \quad (2.28)$$

DÉMONSTRATION: On peut réécrire les y_k du modèle (2.1) comme suit

$$y_k = y_0 \prod_{j=1}^k S(x_j) + \sum_{i=1}^k \prod_{l=i+1}^k S(x_l) \xi_i.$$

Nous pouvons en déduire avec $S \in \Gamma_\varepsilon$ que pour tout $1 \leq k \leq n$

$$y_k^4 \leq \left((1-\varepsilon)^k |y_0| + \sum_{j=1}^k (1-\varepsilon)^{k-j} |\xi_j| \right)^4 \leq 8y_0^4 + 8 \left(\sum_{j=1}^k (1-\varepsilon)^{k-j} |\xi_j| \right)^4.$$

Par l'inégalité de Hölder avec $q = 4/3$ et $p = 4$

$$y_k^4 \leq 8|y_0|^4 + \frac{8}{\varepsilon^3} \sum_{j=1}^k (1-\varepsilon)^{k-j} \xi_j^4.$$

Donc, pour tout $p \in \mathcal{P}$

$$\mathbf{E}_{S,p} y_k^4 \leq 8|y_0|^4 + \frac{8}{\varepsilon^4} \sigma_*.$$

D'où le lemme 2.4.1. □

Maintenant pour tous $K > 0$ et $0 < \varepsilon < 1$ notons

$$\Theta_{K,\varepsilon} = \{S \in \Gamma_\varepsilon : \|\dot{S}\| \leq K\}. \quad (2.29)$$

Lemme 2.4.2 *Soit f une fonction deux fois continûment dérivable dans $[-1, 1]$, tel que $f(u) = 0$ pour $|u| > 1$. Alors*

$$\limsup_{n \rightarrow \infty} \sup_{R > 0} \frac{1}{(Rh_n)^2} \sup_{\|f\|_1 \leq R} \sup_{S \in \Theta_{K,\varepsilon}} \sup_{p \in \mathcal{P}} \mathbf{E}_{S,p} \varrho_n^2(f, S) < \infty, \quad (2.30)$$

où $\|f\|_1 = \|f\| + \|\dot{f}\|$ et $\varrho_n(f, S)$ est défini dans (2.26).

DÉMONSTRATION: Premièrement, notons

$$\sum_{k=1}^n f(u_k) y_{k-1}^2 = T_n + a_n, \quad (2.31)$$

où

$$T_n = \sum_{k=k_*}^{k^*} f(u_k) y_k^2 \quad \text{et} \quad a_n = \sum_{k=k_*}^{k^*} (f(u_k) - f(u_{k-1})) y_{k-1}^2 - f(u_{k^*}) y_{k^*}^2$$

avec k^* et k_* définis dans (2.18). De plus, le modèle (2.1) donne

$$T_n = I_n(f) + \sum_{k=k_*}^{k^*} f(u_k) S^2(x_k) y_{k-1}^2 + M_n,$$

où

$$I_n(f) = \sum_{k=k_*}^{k^*} f(u_k) \quad \text{et} \quad M_n = \sum_{k=k_*}^{k^*} f(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k)$$

avec $\eta_k = \xi_k^2 - 1$. On note aussi

$$C_n = \sum_{k=k_*}^{k^*} (S^2(x_k) - S^2(z_0)) f(u_k) y_{k-1}^2 \quad \text{et} \quad D_n = \sum_{k=k_*}^{k^*} f(u_k) (y_{k-1}^2 - y_k^2).$$

Tenant compte également du fait que $\varphi_n^2 = nh_n$, on obtient

$$\frac{1}{nh_n} T_n = \frac{1}{\tau(S)} \frac{I_n(f)}{nh_n} + \frac{1}{\tau(S)} \frac{\Delta_n}{nh_n} \quad (2.32)$$

avec $\Delta_n = M_n + C_n + S^2(z_0) D_n$. On obtient

$$\begin{aligned} \frac{I_n(f)}{\varphi_n^2} &= \int_{-1}^1 f(t) dt + \sum_{k=k_*}^{k^*} \int_{u_{k-1}}^{u_k} f(u_k) dt - \int_{-1}^1 f(t) dt \\ &= \sum_{k=k_*}^{k^*} \int_{u_{k-1}}^{u_k} (f(u_k) - f(t)) dt + \int_{u_{k^*-1}}^{u_{k^*}} f(t) dt - \int_{-1}^1 f(t) dt. \end{aligned}$$

Rappelons que $\|f\| + \|\dot{f}\| \leq R$. Donc

$$\left| \frac{1}{nh_n} \sum_{k=k_*}^{k^*} f(u_k) - \int_{-1}^1 f(t) dt \right| \leq \frac{R}{nh_n}.$$

En considérant ce résultat dans (2.32) et la borne inférieure pour $\tau(S)$ donnée par (2.27), on obtient

$$\left| \frac{T_n}{\varphi_n^2} - \frac{1}{\tau(S)} \int_{-1}^1 f(t) dt \right| \leq \frac{1}{\varepsilon^2} \left(\frac{R}{nh_n} + \frac{M_n}{nh_n} + \frac{C_n}{nh_n} + \frac{D_n}{nh_n} \right). \quad (2.33)$$

On note que M_n est le dernier terme d'une martingale $(G_j)_{k_* \leq j \leq k^*}$ de carré intégrable, où

$$G_j = \sum_{k=k_*}^j f(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k).$$

Donc

$$\begin{aligned} \mathbf{E}_{S,p} \left(\frac{1}{nh_n} M_n \right)^2 &= \frac{1}{(nh_n)^2} \mathbf{E}_{S,p} \sum_{k=k_*}^{k^*} f^2(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k)^2 \\ &\leq \frac{4R^2 (4\sqrt{m^*} + \sigma^*)}{nh_n}, \end{aligned}$$

où m^* est donné par (2.28). Or, pour tout $S \in \Theta_{K,\varepsilon}$ on a $|S(x_k) - S(z_0)| \leq K|x_k - z_0|$ et comme $k^* - k_* \leq 2nh_n$ il vient

$$\begin{aligned} \frac{1}{(nh_n)^2} \mathbf{E}_{S,p} C_n^2 &\leq \frac{2}{nh_n} \sum_{k=k_*}^{k^*} |(S^2(x_k) - S^2(z_0))|^2 f^2(u_k) \mathbf{E}_{S,p} y_{k-1}^4 \\ &\leq 16 R^2 K^2 m^* h_n^2. \end{aligned}$$

Considérons maintenant le dernier terme de la quantité de droite dans l'inégalité (2.33). Pour cela nous réécrivons D_n comme

$$D_n = \sum_{k=k_*}^{k^*} ((f(u_k) - f(u_{k-1})) y_{k-1}^2 + f(u_{k_*-1}) y_{k_*-1}^2 - f(u_{k^*}) y_{k^*}^2).$$

Donc, sachant que $\|f\| + \|\dot{f}\| \leq R$, on obtient

$$\mathbf{E}_{S,p} D_n^2 \leq 3R^2 \mathbf{E}_{S,p} \left(\frac{2}{nh_n} \sum_{k=k_*}^{k^*} y_{k-1}^4 + y_{k^*}^4 + y_{k_*-1}^4 \right) \leq 18 R^2 m^*.$$

De la même manière nous estimons le second terme de la quantité de droite de (2.31), ce qui prouve le Lemme 2.4.2. □

Lemme 2.4.3 *Les suites $(\tilde{A}_n)_{n \geq 1}$ et $(\hat{A}_n)_{n \geq 1}$ définies dans (2.22) satisfont les propriétés suivantes*

$$\limsup_{n \rightarrow \infty} \frac{1}{h_n^2} \sup_{S \in \Theta_{K,\varepsilon}} \sup_{p \in \mathcal{P}} \mathbf{P}_{S,p} (\tilde{A}_n \leq \kappa_n) < \infty \quad (2.34)$$

et

$$\limsup_{n \rightarrow \infty} \sup_{S \in \Theta_{K,\varepsilon}} \sup_{p \in \mathcal{P}_{\sigma^*}} \mathbf{E}_{S,p} \hat{A}_n^4 < \infty. \quad (2.35)$$

DÉMONSTRATION: Il est clair que l'assertion (2.34) découle du Lemme 2.4.2. Nous vérifions maintenant l'assertion (2.35). On note $\gamma_* = \varepsilon^{-2} \int_{-1}^1 Q(u) du$ et on obtient

$$\begin{aligned} \mathbf{E}_{S,p} \widehat{A}_n^4 &= 4 \int_0^\infty t^3 \mathbf{P}_{S,p} \left(\widetilde{A}_n \leq t^{-1}, \widetilde{A}_n > \kappa_n \right) dt \\ &\leq 4 \int_0^{\kappa_n^{-1}} t^3 \mathbf{P}_{S,p} \left(\varrho_n(Q, S) + \gamma_* \leq t^{-1} \right) dt \\ &\leq \left(\frac{2}{\gamma_*} \right)^4 + \frac{1}{\kappa_n^4} \mathbf{P}_{S,p} \left(|\varrho_n(Q, S)| \geq \gamma_*/2 \right). \end{aligned}$$

En utilisant le Lemme 2.4.2 avec la condition (2.5) on obtient l'inégalité (2.35). \square

Lemme 2.4.4 [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

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \quad \text{et} \quad \mathbf{E}(X_n) \xrightarrow[n \rightarrow \infty]{} \mathbf{E}(X),$$

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

Lemme 2.4.5 [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

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

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

Alors pour tout $y \in \mathbb{R}^d$, on a

$$\mathbf{E}l(X+y) \geq \mathbf{E}l(X).$$

Lemme 2.4.6 Soient $\delta \in]0; 1[$ et $\nu \in]0; 1/4[$. Alors il existe un entier $n_{\delta,\nu} > 0$ tel que si $n \geq n_{\delta,\nu}$, on a $S_\nu \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$.

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

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

Par ailleurs, on a de suite

$$|S'_\nu(x)| = \varphi_n^{-1} h_n^{-1} \left| V'_\nu \left(\frac{x - z_0}{h_n} \right) \right| \leq n^{-\alpha/(2\beta+1)} 2\nu^{-1} \int_{-1}^1 |g'(z)| dz.$$

Donc $S_\nu \in \mathcal{U}_{\delta,n}^{(\beta)}(z_0, \varepsilon)$ pour

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

D'où le lemme. □

Lemme 2.4.7 [Freedman, 1971, pp. 90-91]

Soient $\delta \in]0; 1[$ et $r > 0$. Supposons que $(u_k)_{k \geq 0}$ est une "martingale difference" par rapport à la filtration $(\mathcal{F}_k)_{k \geq 0}$ telle que $|u_k| \leq \delta$ pour tout k et $\sum_{k=1}^{\infty} \mathbf{E}(u_k^2 | \mathcal{F}_{k-1}) \geq r$.

Soit $\tau = \inf \left\{ n : \sum_{k=1}^n \mathbf{E}(u_k^2 | \mathcal{F}_{k-1}) \geq r \right\}$.

Alors il existe une fonction $\rho :]0; +\infty[\rightarrow [0; 2]$, qui ne dépend pas de la distribution de la "martingale difference", telle que $\lim_{x \rightarrow 0} \rho(x) = 0$ et

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

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

Chapitre 3

Modèle autorégressif : cas adaptatif

3.1 Introduction

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

$$y_k = S(x_k)y_{k-1} + \xi_k, \quad 1 \leq k \leq n, \quad (3.1)$$

où $x_k = k/n$, $(\xi_k)_{k \in \{1, \dots, n\}}$ sont des variables aléatoires indépendantes et identiquement distribuées selon la loi gaussienne standard. On s'intéresse à l'estimation de la fonction autorégressive S en un point fixe $z_0 \in]0; 1[$.

On suppose que la fonction S appartient à une classe Höldérienne forte mais sa régularité β est inconnue. L'objectif est de trouver une vitesse de convergence adaptative et pour celle-ci de construire un estimateur séquentiel adaptatif. Parce que β est inconnu, cette vitesse diffèrera ici de la vitesse de convergence obtenue dans le cas contraire.

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 Efroïmovich et Pinsker [1984] pour un modèle de bruit blanc gaussien, Härdle et Marron [1985] pour un modèle de régression et Efroïmovich [1985] pour l'estimation d'une densité.

Belitser [2000a] considère le modèle (3.1) avec des conditions lipschitziennes, propose un estimateur récursif et étudie le problème d'estimation non adaptative. En utilisant le risque quadratique, l'auteur établit la vitesse de convergence. Dans Galtchouk et Pergamenschikov [2005b] les auteurs décrivent une méthode séquentielle pour le problème d'estimation non paramétrique du processus de la dérive du coefficient de diffusion. Dans Lepskiï [1990] l'auteur considère le problème adaptatif, dans un modèle de bruit blanc gaussien, de l'estimation d'un signal appartenant à une classe Höldérienne donnée $\Sigma(m + \alpha, L)$, où $m + \alpha$ et L sont des constantes connues. Fourdrinier, Konev et Pergamenchikov [2009] proposent une procédure séquentielle tronquée qui permet de

considérer le problème de l'estimation du paramètre du processus autorégressif du premier ordre avec bruit dépendant.

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.

Dans ce chapitre on considère le cas adaptatif avec β inconnu. Notre construction est basée sur celle que l'on peut trouver dans Lepskiï [1990] et Galtchouk et Pergamenshchikov [2001] pour l'estimation séquentielle 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 d'un modèle de régression homoscédastique et non adaptatif. L'approche séquentielle c'est celle qu'on trouve dans Borisov et Konev [1977] mais dans le cas paramétrique. La procédure de Lepskiï s'applique à des estimateurs pour lesquels la queue de la distribution a le même comportement asymptotique qu'une variable aléatoire gaussienne, usuellement cette procédure s'utilise dans le cas i.i.d gaussien. Pour notre problème l'estimateur à noyau non séquentiel ne possède pas cette propriété, par contre l'approche séquentielle aboutit dans le cadre du modèle (3.1) en réalisant la procédure adaptative de Lepskiï.

On s'intéresse à l'estimation de la fonction autorégressive S en un point fixé $z_0 \in]0; 1[$. Nous supposons que la fonction S appartient à une classe Höldérienne mais la régularité β reste inconnue. Dans le chapitre précédent on s'est intéressé au cas non adaptatif, pour ce problème on a utilisé l'estimateur à noyau modifié. On réduit le dénominateur de l'estimateur défini dans (2.2) par la constante $H > 0$, tel que pour $0 \leq \alpha_H \leq 1$

$$\sum_{j=1}^{\tau_H-1} Q(u_j) y_{j-1}^2 + \alpha_H Q(u_{\tau_H}) y_{\tau_H-1}^2 = H,$$

où τ_H est le temps d'arrêt défini comme suit

$$\tau_H = \inf\{1 \leq k \leq n : \sum_{j=1}^k Q(u_j) y_{j-1}^2 \geq H\}. \quad (3.2)$$

Notons

$$A_k = \sum_{j=1}^k Q(u_j) y_{j-1}^2 \quad \text{avec} \quad u_j = \frac{x_j - z_0}{h_n}.$$

Ainsi l'estimateur à noyau s'écrit sous la forme suivante

$$S_{H,h_n}^*(z_0) = \frac{1}{H} \left(\sum_{j=1}^{\tau_H-1} Q(u_j) y_{j-1} y_j + \alpha_H Q(u_{\tau_H}) y_{\tau_H-1} y_{\tau_H} \right) \mathbf{1}_{(A_n \geq H)}, \quad (3.3)$$

où le noyau $Q(\cdot)$ est la fonction indicatrice de l'intervalle $[-1; 1]$. Un tel estimateur s'avère très commode pour calculer la quantité $\mathbf{E} |S_{H,h_n}^*(z_0) - S(z_0)|$.

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 de l'estimateur à noyau (3.3)

3.2 Description du problème

Le problème est d'estimer la fonction S en un point fixé $z_0 \in]0, 1[$, i.e. la valeur $S(z_0)$. Pour ce problème nous utilisons le risque proposé dans le chapitre 2. En effet, pour tout estimateur $\tilde{S}_n = \tilde{S}_n(z_0)$ (i.e. toute fonction mesurable par rapport aux observations $(y_k)_{1 \leq k \leq n}$), le risque est défini sur le voisinage $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ par

$$\mathcal{R}_n(\tilde{S}_n) = \sup_{\beta \in [\beta_*; \beta^*]} \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} N(\beta) \mathbf{E}_S |\tilde{S}_n(z_0) - S(z_0)|, \quad (3.4)$$

où $N(\beta) = \left(\frac{n}{\ln n}\right)^{\beta/(2\beta+1)}$ correspond à la vitesse de convergence adaptative des estimateurs sur la classe $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$. Ici \mathbf{E}_S dénote toujours l'espérance prise par rapport à la distribution \mathbf{P}_S du vecteur (y_1, \dots, y_n) dans (3.1) correspondant à la fonction S .

On considère le modèle (3.1) où $S \in \mathbf{C}_1([0, 1], \mathbb{R})$ est la fonction inconnue. On veut estimer la fonction autorégressive S au point fixé z_0 . On obtient la stabilité (uniformément par rapport à la fonction S) du modèle (3.1) en supposant que pour un certain $0 < \varepsilon < 1$ fixé la fonction inconnue S appartient à l'ensemble stable

$$\Gamma_\varepsilon = \{S \in \mathbf{C}_1([0, 1], \mathbb{R}) : \|S\| \leq 1 - \varepsilon\}, \quad (3.5)$$

où $\|S\| = \sup_{0 \leq x \leq 1} |S(x)|$. Ici $\mathbf{C}_1([0, 1], \mathbb{R})$ est l'espace de Banach des fonctions continûment différentiables de $[0, 1]$ dans \mathbb{R} . Pour des constantes fixées $K > 0$ et $\beta \in [\beta_*; \beta^*] \subset]0, 1]$ on définit la classe Höldérienne stable correspondante au point z_0 par

$$\mathcal{H}^{(\beta)}(z_0, K, \varepsilon) = \{S \in \Gamma_\varepsilon : \Omega^*(z_0, S) \leq K\} \quad (3.6)$$

avec

$$\Omega^*(z_0, S) = \sup_{x \in [0, 1]} \frac{|S(x) - S(z_0)|}{|x - z_0|^\beta}.$$

La régularité β est supposée inconnue mais l'intervalle $[\beta_*; \beta^*]$ est considéré comme connu.

En premier lieu on donne la borne inférieure pour le risque minimax. On montre qu'avec la vitesse de convergence $N(\beta)$, la borne inférieure pour le risque minimax est strictement positive.

Théorème 3.2.1 *Le risque adaptatif admet la borne inférieure suivante*

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{S}_n} \mathcal{R}_n(\tilde{S}_n) \geq \frac{1}{4}.$$

où l'infimum est pris sur tous les estimateurs \tilde{S}_n .

Maintenant on donne la borne supérieure pour le risque maximal de l'estimateur séquentiel adaptatif défini dans (3.3). Tenant compte du fait que β est inconnu, on ne peut pas utiliser cet estimateur tel quel car la fenêtre h_n dépend de β . C'est la raison pour laquelle on partitionne l'intervalle $[\beta_*; \beta^*]$ pour suivre une procédure de Lepskiï. Pour cela posons

$$d_n = n/\ln n \quad \text{et} \quad h(\beta) = \left(\frac{1}{d_n}\right)^{\frac{1}{2\beta+1}}. \quad (3.7)$$

On définit la grille sur l'intervalle $[\beta_*; \beta^*]$ par les points :

$$\beta_k = \beta_* + \frac{k}{m}(\beta^* - \beta_*), \quad k = 0, \dots, m \quad \text{avec} \quad m = \lfloor \ln d_n \rfloor + 1. \quad (3.8)$$

On note

$$N_k = N(\beta_k) \quad \text{et} \quad h_k = h(\beta_k),$$

et aussi

$$\omega(h_j) = \max_{0 \leq k \leq j} \left(|S_{h_j}^* - S_{h_k}^*| - \frac{\lambda}{N_{k+1}} \right)$$

pour définir l'indice optimal de la fenêtre comme

$$\hat{k} = \inf \left\{ 0 \leq j \leq m : \omega(h_j) \geq \frac{\lambda}{N_j} \right\} - 1. \quad (3.9)$$

Si l'ensemble ci-dessus est vide, on pose $\hat{k} = m - 1$.

On remarque que $\omega(h_0) = -\lambda/N_1$, donc $\hat{k} \geq 0$. Le paramètre positif λ est choisi tel que

$$\lambda > K + e\sqrt{4 + \frac{4}{2\beta_* + 1}}.$$

L'estimateur adaptatif est maintenant défini comme

$$\hat{S}_n = S_{H, \hat{h}}^* \quad \text{avec} \quad \hat{h} = h_{\hat{k}}. \quad (3.10)$$

Le résultat suivant donne la borne supérieure pour le risque maximal de l'estimateur séquentiel adaptatif ci-dessus.

Théorème 3.2.2 *Pour tout $0 < \varepsilon < 1$, on a*

$$\limsup_{n \rightarrow \infty} \mathcal{R}_n(\hat{S}_n) < \infty, \quad (3.11)$$

donc \hat{S}_n est un estimateur adaptatif en vitesse de convergence.

3.3 Borne inférieure

On montre qu'avec la vitesse $N(\beta)$, la borne inférieure du risque minimax est strictement positive.

Preuve du Théorème 3.2.1

Afin de simplifier les notations on note $N(\beta_*) = N_*$, $N(\beta^*) = N^*$ et $h(\beta_*) = h_*$.

On note

$$S(y) = \frac{1}{N_*} V \left(\frac{y - z_0}{h_*} \right),$$

où V est une fonction de classe C^∞ à support compact $[-1, 1]$ telle que

$$\int_{-1}^1 V^2(u) du = \frac{\bar{\beta}}{2} \quad \text{avec} \quad \bar{\beta} = \frac{\beta^* - \beta_*}{(2\beta^* + 1)(2\beta_* + 1)},$$

$V(0) = 1$ et $V(u) = 0$ pour $|u| \geq 1$.

On montre aisément que pour un réel K suffisamment grand, $S \in \mathcal{H}^{(\beta_*)}(z_0, K, \varepsilon)$.

Notons que pour tout S la mesure \mathbf{P}_S est équivalente à la mesure \mathbf{P}_0 , où \mathbf{P}_0 est la distribution du vecteur (y_1, \dots, y_n) dans (3.1) correspondant à la fonction $S_0 = 0$. Il est facile de voir que dans ce cas la densité de Radom-Nikodym peut s'écrire

$$\begin{aligned} \rho_n &:= \frac{d\mathbf{P}_0}{d\mathbf{P}_S}(y_1, \dots, y_n) \\ &= \exp \left\{ -\frac{1}{2} \sum_{k=1}^n (y_k^2 - (y_k - S(x_k)y_{k-1})^2) \right\} \\ &= \exp \left(-\varsigma_n \eta_n - \frac{1}{2} \varsigma_n^2 \right), \end{aligned}$$

avec

$$\varsigma_n^2 = \frac{1}{d_n h_*} \sum_{k=1}^n V^2 \left(\frac{x_k - z_0}{h_*} \right) y_{k-1}^2 \quad \text{et} \quad \eta_n = \frac{1}{\sqrt{d_n h_*} \varsigma_n} \sum_{k=1}^n V \left(\frac{x_k - z_0}{h_*} \right) y_{k-1} \xi_k.$$

D'après le Lemme 3.5.2 il vient

$$\begin{aligned} \mathbf{P}_S - \lim_{n \rightarrow \infty} \frac{d_n}{n} \varsigma_n^2 &= \mathbf{P}_S - \lim_{n \rightarrow \infty} \left(\frac{1}{n h_*} \sum_{k=1}^n V^2 \left(\frac{x_k - z_0}{h_*} \right) y_{k-1}^2 \right) \\ &= \mathbf{P}_S - \lim_{n \rightarrow \infty} \frac{1}{\tau(S)} \int_0^1 V^2 \left(\frac{x - z_0}{h_*} \right) dx \\ &= \int_{-1}^1 V^2(u) du = \frac{\bar{\beta}}{2} = \varsigma_*^2, \end{aligned}$$

car $\tau(S) = 1 - S^2(z_0) = 1 - \frac{1}{N_*^2}$.

De plus, en utilisant un théorème central limite pour les martingales (cf. le Lemme 3.5.6), il est facile de voir que sous la mesure \mathbf{P}_S

$$\eta_n \implies \mathcal{N}(0, 1) \quad \text{quand } n \rightarrow \infty.$$

En effet, on peut réécrire η_n sous la forme suivante :

$$\eta_n = \sqrt{\frac{n}{d_n}} \frac{\varsigma_*}{\varsigma_n} \sum_{k=1}^n u_{k,n},$$

avec

$$u_{k,n} = \frac{1}{\varsigma_* \sqrt{nh_*}} V \left(\frac{x_k - z_0}{h_*} \right) y_{k-1} \xi_k.$$

Intéressons-nous à la première condition du Lemme 3.5.6. Pour la vérifier il suffit de montrer

$$\mathbf{E}_S \sum_{k=1}^n \mathbf{E}_S(u_{k,n}^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)} | \mathcal{F}_{k-1,n}) \xrightarrow{n \rightarrow \infty} 0.$$

On a

$$\begin{aligned} \mathbf{E}_S \sum_{k=1}^n \mathbf{E}_S(u_{k,n}^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)} | \mathcal{F}_{k-1,n}) &= \sum_{k=1}^n \mathbf{E}_S(u_{k,n}^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)}) \\ &= \frac{1}{\varsigma_*^2 nh_*} \sum_{k=k_*}^{k^*} V^2 \left(\frac{x_k - z_0}{h_*} \right) \mathbf{E}_S(y_{k-1}^2 \xi_k^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)}), \end{aligned} \quad (3.12)$$

avec

$$\begin{aligned} \mathbf{E}_S(y_{k-1}^2 \xi_k^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)}) &\leq \sqrt{\mathbf{E}_S y_{k-1}^4 \mathbf{E}_S \xi_k^4} \sqrt{\mathbf{P}_S(|u_{k,n}| > \varepsilon)} \\ &\leq \sqrt{\mathbf{E}_S y_{k-1}^4 \mathbf{E}_S \xi_k^4} \sqrt{\frac{1}{\varepsilon^2} \mathbf{E}_S u_{k,n}^2} \\ &\leq C_1 \sqrt{\frac{\mathbf{E}_S y_{k-1}^2 \xi_k^2}{nh_*}} \leq \frac{C_2}{\sqrt{nh_*}}, \end{aligned}$$

où C_1 et C_2 sont des constantes indépendantes de n . Donc le terme (3.12) est majoré par

$$\mathbf{E} \sum_{k=1}^n \mathbf{E}(u_{k,n}^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)} | \mathcal{F}_{k-1,n}) \leq \frac{C_3}{nh_*} \sum_{k=k_*}^{k^*} \frac{1}{\sqrt{nh_*}},$$

où C_3 est une nouvelle constante, ce dernier terme tendant alors vers zéro quand $n \rightarrow \infty$.

Il est très facile de voir pour la deuxième condition que

$$\begin{aligned}
\sum_{k=1}^n \mathbf{E}_S (u_{k,n}^2 | \mathcal{F}_{k-1,n}) &= \frac{1}{\zeta_*^2 n h_*} \sum_{k=1}^n V^2 \left(\frac{x_k - z_0}{h_*} \right) \mathbf{E}(y_{k-1}^2 \xi_k^2 | \mathcal{F}_{k-1,n}) \\
&= \frac{1}{\zeta_*^2 n h_*} \sum_{k=1}^n V^2 \left(\frac{x_k - z_0}{h_*} \right) y_{k-1}^2 \\
&= \frac{d_n}{n} \frac{\zeta_n^2}{\zeta_*^2} \xrightarrow[n \rightarrow \infty]{\mathbf{P}_S} 1.
\end{aligned}$$

En notant $\theta_n = N_* |\tilde{S}_n|$, on a

$$\begin{aligned}
\mathcal{R}_n(\tilde{S}_n) &\geq \max \left(\mathbf{E}_{S_0} N^* |\tilde{S}_n|, \mathbf{E}_S N_* |\tilde{S}_n - S(z_0)| \right) \\
&= \max \left(\mathbf{E}_{S_0} \frac{N^*}{N_*} |\theta_n|, \mathbf{E}_S |1 - \theta_n| \right) \\
&\geq \frac{1}{2} \mathbf{E}_S \left(\frac{N^*}{N_*} |\theta_n| \frac{d\mathbf{P}_0}{d\mathbf{P}_S}(y) + |1 - \theta_n| \right)
\end{aligned} \tag{3.13}$$

En posant $\gamma_n = \frac{N^*}{N_*}$, on peut minorer (3.13) de la manière suivante :

$$\mathcal{R}_n(\tilde{S}_n) \geq \frac{1}{2} \mathbf{E}_S (\gamma_n \rho_n |\theta_n| + |1 - \theta_n|).$$

Soient $B_n = \{\eta_n \leq 0\}$ et $C_n = \{\frac{d_n}{n} \zeta_n^2 < \bar{\beta}\}$. Il est clair que lorsque $B_n \cap C_n$ est réalisé, on a

$$\gamma_n \rho_n \geq \exp\left\{\bar{\beta} \ln d_n - \frac{\bar{\beta}}{2} \frac{n}{d_n}\right\},$$

cette dernière expression tendant vers ∞ quand n tend vers ∞ . Ce qui veut dire que pour n suffisamment grand

$$\begin{aligned}
\mathcal{R}_n(\tilde{S}_n) &\geq \frac{1}{2} \mathbf{E}_S \mathbf{1}_{B_n \cap C_n} (\gamma_n \rho_n |\theta_n| + |1 - \theta_n|) \\
&\geq \frac{1}{2} \mathbf{E}_S \mathbf{1}_{B_n \cap C_n} (|\theta_n| + 1 - |\theta_n|) \\
&= \frac{1}{2} \mathbf{P}_S(B_n \cap C_n).
\end{aligned} \tag{3.14}$$

Or,

$$\mathbf{P}_S(B_n \cap C_n) = \mathbf{P}_S(B_n) - \mathbf{P}_S(B_n \cap C_n^c)$$

et

$$\mathbf{P}_S(B_n \cap C_n^c) \leq \mathbf{P}_S(C_n^c) = \mathbf{P}_S\left(\frac{d_n}{n} \zeta_n \geq \bar{\beta}\right).$$

Comme

$$\frac{d_n}{n} \zeta_n \xrightarrow[n \rightarrow \infty]{\mathbf{P}_S} \frac{\bar{\beta}}{2},$$

il vient

$$\mathbf{P}_S(C_n^c) \xrightarrow[n \rightarrow \infty]{} 0.$$

Puisque $\mathbf{P}_S(B_n) = 1/2$, on en déduit que $\mathbf{P}_S(B_n \cap C_n) \xrightarrow[n \rightarrow \infty]{} 1/2$.

En passant à la limite quand $n \rightarrow \infty$ dans (3.14), on obtient le Théorème. \square

3.4 Estimation séquentielle adaptative (borne supérieure)

Preuve du Théorème 3.2.2

On procède suivant une méthode basée sur l'analyse séquentielle. En premier lieu, nous réécrivons l'erreur de l'estimation sous la forme suivante :

$$S_{H,h}^*(z_0) - S(z_0) = -S(z_0) \mathbf{1}_{(A_n < H)} + B_H(h) \mathbf{1}_{(A_n \geq H)} + \frac{1}{\sqrt{H}} \zeta_H(h) \mathbf{1}_{(A_n \geq H)}, \quad (3.15)$$

où

$$B_H(h) = \frac{1}{H} \left(\sum_{j=1}^{\tau_H-1} Q(u_j) (S(x_j) - S(z_0)) y_{j-1}^2 + \alpha_H Q(u_{\tau_H}) (S(x_{\tau_H}) - S(z_0)) y_{\tau_H-1}^2 \right)$$

et

$$\zeta_H(h) = \frac{1}{\sqrt{H}} \left(\sum_{j=1}^{\tau_H-1} Q(u_j) y_{j-1} \xi_j + \alpha_H Q(u_{\tau_H}) y_{\tau_H-1} \xi_{\tau_H} \right).$$

Notons que le premier terme dans la quantité de droite de (3.15) est étudié dans le Lemme 3.5.3. On peut montrer directement que pour tout $S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$

$$|B_H(h)| \leq Kh^\beta \quad (3.16)$$

et aussi, en utilisant le lemme 3.5.5 on a

$$\sup_{n \geq 1} \sup_{h_* \leq h \leq h^*} \mathbf{E}_S |\zeta_H(h)| < \infty, \quad (3.17)$$

où $h_* = h(\beta_*)$ et $h^* = h(\beta^*)$. Maintenant on choisit $H = nh$ et

$$\iota = \inf\{k \geq 0 : \beta_k \geq \beta\} - 1.$$

Ce qui veut dire

$$\beta_\iota < \beta \leq \beta_{\iota+1} \quad \text{et} \quad h_\iota < h(\beta) \leq h_{\iota+1}.$$

Dorénavant, on notera $S_h^*(z_0) = S_{H,h}^*(z_0)$. On a maintenant

$$|S_{h_\iota}^*(z_0) - S(z_0)| \leq \mathbf{1}_{(A_n(h_\iota) < nh_\iota)} + K(h(\beta_\iota))^\beta + \frac{1}{\sqrt{nh_\iota}} |\zeta_H(h_\iota)|$$

et

$$|S_{h_{\iota-1}}^*(z_0) - S(z_0)| \leq \mathbf{1}_{(A_n(h_{\iota-1}) < nh_{\iota-1})} + K(h(\beta_{\iota-1}))^\beta + \frac{1}{\sqrt{nh_{\iota-1}}} |\zeta_H(h_{\iota-1})|.$$

L'inégalité (3.17) implique

$$\limsup_{n \rightarrow \infty} \sup_{\beta_* \leq \beta \leq \beta^*} N(\beta) \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \mathbf{E}_S \varpi(\iota, z_0) < \infty, \quad (3.18)$$

où

$$\varpi(\iota, z_0) = |S_{h_{\iota-1}}^*(z_0) - S(z_0)| + |S_{h_\iota}^*(z_0) - S(z_0)|.$$

Maintenant en considérant l'estimateur \widehat{S}_n , on peut écrire

$$|\widehat{S}_n(z_0) - S(z_0)| \leq I_1 + I_2 + \varpi(\iota, z_0), \quad (3.19)$$

où

$$I_1 = |\widehat{S}_n(z_0) - S(z_0)| \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} \quad \text{et} \quad I_2 = |\widehat{S}_n(z_0) - S(z_0)| \mathbf{1}_{\{\widehat{k} \leq \iota-2\}}.$$

Occupons-nous d'abord du premier terme dans cette inégalité. On a

$$|\widehat{S}_n(z_0) - S(z_0)| \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} \leq |S_{\widehat{h}}^*(z_0) - S_{h_\iota}^*(z_0)| \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} + |S_{h_\iota}^*(z_0) - S(z_0)| \mathbf{1}_{\{\widehat{k} \geq \iota+1\}}.$$

De plus,

$$\begin{aligned} |S_{\widehat{h}}^*(z_0) - S_{h_\iota}^*(z_0)| \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} &\leq \omega(h_{\widehat{k}}) \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} + \frac{\lambda}{N_{\iota+1}} \\ &\leq \frac{\lambda}{N_{\widehat{k}}} \mathbf{1}_{\{\widehat{k} \geq \iota+1\}} + \frac{\lambda}{N_{\iota+1}} \leq \frac{2\lambda}{N_{\iota+1}} \leq \frac{2\lambda}{N(\beta)}. \end{aligned}$$

Ce qui implique directement

$$\limsup_{n \rightarrow \infty} \sup_{\beta_* \leq \beta \leq \beta^*} N(\beta) \sup_{S \in \mathcal{H}^{(\beta)}(z_0, K, \varepsilon)} \mathbf{E}_S I_1 < \infty. \quad (3.20)$$

Ensuite majorons le second terme dans (3.19) :

$$I_2 \leq \left(\mathbf{1}_{(A_n(h_{\widehat{k}}) < nh_{\widehat{k}})} + K(h(\beta_{\widehat{k}}))^\beta + \frac{1}{\sqrt{nh_{\widehat{k}}}} \zeta^* \right) \mathbf{1}_{\{\widehat{k} \leq \iota-2\}},$$

où

$$\zeta^* = \max_{1 \leq j \leq m} |\zeta_{H_j}(h_j)|. \quad (3.21)$$

Notons que

$$\{\widehat{k} \leq \iota - 2\} = \bigcup_{j=1}^{\iota-1} \{\omega(h_j) \geq \lambda/N_j\}.$$

De plus,

$$\begin{aligned} \{\omega(h_j) \geq \lambda/N_j\} &= \bigcup_{l=0}^{j-1} \left\{ |S_{h_j}^*(z_0) - S_{h_l}^*(z_0)| \geq \lambda/N_j + \lambda/N_{l+1} \right\} \\ &\subseteq \bigcup_{l=0}^{j-1} \left(\{|S_{h_j}^*(z_0) - S(z_0)| \geq \lambda/N_j\} \cup \{|S_{h_l}^*(z_0) - S(z_0)| \geq \lambda/N_{l+1}\} \right). \end{aligned} \quad (3.22)$$

On note que pour $j \leq \iota - 1$

$$N_j (h_j)^\beta \leq \exp\left\{-\frac{\ln d_n}{(2\beta^* + 1)m}\right\} \leq 1.$$

Pour $l \leq \iota - 1$

$$N_{l+1} (h_l)^\beta \leq \exp\left\{-\frac{\ln d_n}{(2\beta^* + 1)m}\right\} \leq 1$$

et

$$\frac{N_l}{N_{l+1}} \geq \exp\left\{-\frac{\ln d_n}{m}\right\} = e^{-1}.$$

Dans le premier terme de (3.22), par le Lemme 3.5.2 on prouve que pour n suffisamment grand et pour $\lambda > K + e\sqrt{4 + \frac{4}{2\beta_* + 1}}$

$$\begin{aligned} \{|S_{h_j}^*(z_0) - S(z_0)| \geq \lambda/N_j\} &\subseteq \left\{ K(h_j)^\beta + \frac{1}{\sqrt{nh_j}} |\zeta_n(h_j)| \geq \lambda/N_j \right\} \\ &\subseteq \left\{ |\zeta_n(h_j)| \geq \sqrt{nh_j} \left(\frac{\lambda}{N_j} - K(h_j)^\beta \right) \right\}. \end{aligned}$$

Il suffit juste de remarquer que $(1/d_n)^{\beta/(2\beta+1)} \sqrt{nh} = \sqrt{n/d_n}$ et cette dernière inclusion devient

$$\{|S_{h_j}^*(z_0) - S(z_0)| \geq \lambda/N_j\} \subseteq \left\{ |\zeta_n(h_j)| \geq (\lambda - K) \sqrt{\frac{n}{d_n}} \right\}.$$

De la même manière pour le second terme dans (3.22) on aura

$$\{|S_{h_l}^*(z_0) - S(z_0)| \geq \lambda/N_{l+1}\} \subseteq \left\{ |\zeta_n(h_l)| \geq (\lambda - K)/e \sqrt{\frac{n}{d_n}} \right\}.$$

Finalement,

$$\{\hat{k} \leq \iota - 2\} \subseteq \{\zeta^* \geq \lambda_1 \sqrt{n/d_n}\},$$

avec $\lambda_1 = (\lambda - K)/e$. Ainsi

$$I_2 \leq \mathbf{1}_{(A_n(h_{\hat{k}}) < nh_{\hat{k}})} + \frac{K}{N(\beta)} + \frac{1}{\sqrt{nh_*}} \zeta^* \mathbf{1}_{\{\zeta^* \geq \lambda_1 \sqrt{n/d_n}\}}. \quad (3.23)$$

En utilisant le Lemme 3.5.2 pour $t \geq 2$, on peut facilement estimer le premier terme de cette inégalité par

$$\begin{aligned} \mathbf{P}_S(A_n(h_{\hat{k}}) < nh_{\hat{k}}) &= \sum_{l=1}^m \mathbf{P}_S(A_n(h_l) < nh_l, \hat{k} = l) \\ &\leq \sum_{l=1}^m \mathbf{P}_S(A_n(h_l) < nh_l) \\ &= \sum_{l=1}^m \mathbf{P}_S\left(\frac{1}{\tau(S)} \int_{-1}^1 Q(u) du + \Delta_n(Q, h_l) < 1\right) \\ &= \sum_{l=1}^m \mathbf{P}_S\left(\Delta_n(Q, h_l) < 1 - \frac{2}{\tau(S)}\right) \\ &\leq \sum_{l=1}^m \mathbf{P}_S(|\Delta_n(Q, h_l)| > 1) \\ &\leq \sum_{l=1}^m \mathbf{E}_S \Delta_n^{2t}(Q, h_l) \leq ([\ln d_n] + 1) C_1 R^{2t} (h^*)^{2t\beta}. \end{aligned}$$

Intéressons-nous désormais au dernier terme du membre de droite de l'inégalité (3.23) :

$$\begin{aligned} \mathbf{E}_S \zeta^* \mathbf{1}_{\{\zeta^* \geq \lambda_1 \sqrt{\ln n}\}} &= \int_0^{+\infty} \mathbf{P}_S(\zeta^* \mathbf{1}_{\{\zeta^* \geq \lambda_1 \sqrt{\ln n}\}} \geq z) dz \\ &= \int_0^{+\infty} \mathbf{P}_S(\zeta^* \geq z, \zeta^* \geq \lambda_1 \sqrt{\ln n}) dz \\ &= \lambda_1 \sqrt{\ln n} \mathbf{P}_S(\zeta^* \geq \lambda_1 \sqrt{\ln n}) + \int_{\lambda_1 \sqrt{\ln n}}^{+\infty} \mathbf{P}_S(\zeta^* \geq z) dz. \end{aligned}$$

En utilisant (3.21) et le Lemme 3.5.5, on a

$$\begin{aligned} \mathbf{P}_S(\zeta^* \geq z) &= \mathbf{P}_S(\max_{1 \leq j \leq m} |\zeta_n(h_j)| \geq z) \\ &= \sum_{j=1}^m \mathbf{P}_S(|\zeta_n(h_j)| \geq z) \\ &\leq 2m e^{-z^2/8}. \end{aligned}$$

Donc

$$\begin{aligned}
\mathbf{E}_S \zeta^* \mathbf{1}_{\{\zeta^* \geq \lambda_1 \sqrt{\ln n}\}} &\leq 2m \lambda_1 \sqrt{\ln n} e^{-\frac{1}{8} \lambda_1^2 \ln n} + 2m \int_{\lambda_1 \sqrt{\ln n}}^{+\infty} e^{-z^2/8} dz \\
&\leq 2m \lambda_1 \sqrt{\ln n} e^{-\frac{1}{8} \lambda_1^2 \ln n} + 2m \int_{\lambda_1 \sqrt{\ln n}}^{+\infty} z e^{-z^2/8} dz \\
&\leq \left(\lambda_1 \sqrt{\ln n} + 4 \right) 2m n^{-\lambda_1^2/8}.
\end{aligned}$$

Ce qui implique l'inégalité (3.11) puis le Théorème 3.2.2. □

3.5 Annexe B

Dans cette section on étudie les propriétés du processus stationnaire dans le modèle (3.1).

Lemme 3.5.1 *Pour tout $t \in \mathbb{N}^*$ et $0 < \varepsilon < 1$, les variables aléatoires dans (3.1) satisfont l'inégalité suivante*

$$r^* = \sup_{n \geq 1} \sup_{0 \leq k \leq n} \sup_{S \in \Gamma_\varepsilon} \mathbf{E}_S y_k^{2t} < \infty. \quad (3.24)$$

DÉMONSTRATION: On suppose pour ce Lemme que $y_0 = 0$ et les y_k du modèle (3.1) deviennent

$$y_k = \sum_{i=1}^k \prod_{l=i+1}^k S(x_l) \xi_i,$$

on déduit avec $S \in \Gamma_\varepsilon$ et pour tout $1 \leq k \leq n$

$$y_k^{2t} \leq \left(\sum_{j=1}^k (1 - \varepsilon)^{k-j} |\xi_j| \right)^{2t}.$$

De plus, par l'inégalité de Hölder avec $p = 2t$,

$$\begin{aligned} y_k^{2t} &\leq \left(\sum_{j=1}^k (1 - \varepsilon)^{k-j} \right)^{2t-1} \left(\sum_{j=1}^k (1 - \varepsilon)^{k-j} \xi_j^{2t} \right) \\ &\leq \left(\frac{1}{\varepsilon} \right)^{2t-1} \left(\sum_{j=1}^k (1 - \varepsilon)^{k-j} \xi_j^{2t} \right). \end{aligned}$$

Il s'en suit que

$$\mathbf{E}_S y_k^{2t} \leq \frac{(2t)!}{2^t t!} \left(\frac{1}{\varepsilon} \right)^{2t}.$$

D'où le Lemme 3.5.1. □

Introduisons la notation suivante

$$\Delta_n(f, h) = \frac{1}{nh} \sum_{k=1}^n f(u_k) y_{k-1}^2 - \frac{1}{\tau(S)} \int_{-1}^1 f(u) du,$$

où

$$\tau(S) = 1 - S^2(z_0). \quad (3.25)$$

Lemme 3.5.2 *Soit f une fonction deux fois continûment différentiable dans $[-1, 1]$, telle que $f(u) = 0$ pour $|u| > 1$. Alors pour tout $t \in \mathbb{N}^*$*

$$\limsup_{n \rightarrow \infty} \sup_{h_* \leq h \leq h^*} \sup_{R > 0} \frac{1}{R^{2t} h^{2t\beta}} \sup_{\|f\|_1 \leq R} \sup_{S \in \mathcal{H}^\beta(z_0, K, \varepsilon)} \mathbf{E}_S \Delta_n^{2t}(f, h) \leq C_1, \quad (3.26)$$

où $\|f\|_1 = \|f\| + \|\dot{f}\|$ et $C_1 = 2^{4t} K^{2t} (r^*)^2$.

DÉMONSTRATION: Premièrement réécrivons

$$\sum_{k=1}^n f(u_k) y_{k-1}^2 = T_n + a_n, \quad (3.27)$$

où

$$T_n = \sum_{k=k_*}^{k^*} f(u_k) y_k^2 \quad \text{et} \quad a_n = \sum_{k=k_*}^{k^*} (f(u_k) - f(u_{k-1})) y_{k-1}^2 - f(u_{k^*}) y_{k^*}^2,$$

les entiers k^* et k_* étant définis dans (2.18). L'écriture du modèle (3.1) nous donne

$$T_n = I_n(f) + \sum_{k=k_*}^{k^*} f(u_k) S^2(x_k) y_{k-1}^2 + M_n,$$

où

$$I_n(f) = \sum_{k=k_*}^{k^*} f(u_k) \quad \text{et} \quad M_n = \sum_{k=k_*}^{k^*} f(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k)$$

avec $\eta_k = \xi_k^2 - 1$. En notant

$$C_n = \sum_{k=k_*}^{k^*} (S^2(x_k) - S^2(z_0)) f(u_k) y_{k-1}^2 \quad \text{et} \quad D_n = \sum_{k=k_*}^{k^*} f(u_k) (y_{k-1}^2 - y_k^2),$$

on obtient

$$\frac{1}{nh} T_n = \frac{1}{\tau(S)} \frac{I_n(f)}{nh} + \frac{1}{\tau(S)} \frac{H_n}{nh} \quad (3.28)$$

avec $H_n = M_n + C_n + S^2(z_0) D_n$. De plus, il est facile de voir que

$$\begin{aligned} \frac{I_n(f)}{nh} &= \int_{-1}^1 f(t) dt + \sum_{k=k_*}^{k^*} \int_{u_{k-1}}^{u_k} f(u_k) dt - \int_{-1}^1 f(t) dt \\ &= \sum_{k=k_*}^{k^*} \int_{u_{k-1}}^{u_k} (f(u_k) - f(t)) dt + \int_{u_{k^*-1}}^{u_{k^*}} f(t) dt - \int_{-1}^1 f(t) dt. \end{aligned}$$

On rappelle que $\|f\| + \|\dot{f}\| \leq R$. Donc

$$\left| \frac{1}{nh} \sum_{k=k_*}^{k^*} f(u_k) - \int_{-1}^1 f(t) dt \right| \leq \frac{R}{nh}.$$

En tenant compte de (3.27) et de la borne inférieure pour $\tau(S)$ donnée dans (2.27), on prouve que

$$\left| \frac{T_n}{nh} - \frac{1}{\tau(S)} \int_{-1}^1 f(t) dt \right| \leq \frac{1}{\varepsilon^2} \left(\frac{R}{nh} + \frac{M_n}{nh} + \frac{C_n}{nh} + \frac{D_n}{nh} \right). \quad (3.29)$$

On note que M_n est le dernier terme d'une martingale $(G_j)_{k_* \leq j \leq k^*}$ de carré intégrable, où

$$G_j = \sum_{k=k_*}^j f(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k).$$

Donc en appliquant l'inégalité de Burkholder, il vient

$$\begin{aligned} \mathbf{E}_S \left(\frac{1}{nh} M_n \right)^{2t} &\leq \frac{A_{2t}^{2t}}{(nh)^{2t}} \mathbf{E}_S \left(\sum_{k=k_*}^{k^*} f^2(u_k) (2S(x_k) y_{k-1} \xi_k + \eta_k)^2 \right)^t \\ &\leq A_{2t}^{2t} \frac{R^t}{(nh)^{t+1}} \mathbf{E}_S \sum_{k=k_*}^{k^*} (2S(x_k) y_{k-1} \xi_k + \eta_k)^{2t} \\ &\leq \frac{R^t}{(nh)^t} 2^{4t-2} A_{2t}^{2t} \left(\frac{(2t)!}{2^{2t}} \left(2r^* + \frac{(2t)!}{2^{2t}} \right) + 1 \right) \end{aligned}$$

où $A_{2t} = 18(2t)^{3/2}/(2t-1)^{1/2}$ et r^* est donné dans (3.24). Or, $|S(x_k) - S(z_0)| \leq K|x_k - z_0|^\beta$ pour tout $S \in \mathcal{H}^\beta(z_0, K, \varepsilon)$. En appliquant alors l'inégalité de Hölder pour $p = 2t$ et $q = 2t/(2t-1)$, on obtient

$$\begin{aligned} \frac{1}{(nh)^{2t}} \mathbf{E}_S C_n^{2t} &\leq \frac{1}{(nh)^{2t}} \left(\sum_{k=k_*}^{k^*} |(S^2(x_k) - S^2(z_0))|^q \mathbf{1}_{|u_k| \leq 1} \right)^{2t/q} \sum_{k=k_*}^{k^*} f^{2t}(u_k) \mathbf{E}_S y_{k-1}^{4t} \\ &\leq 2^{4t} R^{2t} K^{2t} (r^*)^2 h^{2t\beta}. \end{aligned}$$

Considérons maintenant le dernier terme du membre de droite de l'inégalité (3.28). Pour cela on représente D_n comme

$$D_n = \sum_{k=k_*}^{k^*} ((f(u_k) - f(u_{k-1})) y_{k-1}^2 + f(u_{k_*-1}) y_{k_*-1}^2 - f(u_{k^*}) y_{k^*}^2).$$

Donc, puisque $\|f\| + \|\dot{f}\| \leq R$ on a

$$\mathbf{E}_S D_n^{2t} \leq 2^{4t-2} R^{2t} \mathbf{E}_S \left(\frac{1}{nh} \sum_{k=k_*}^{k^*} y_{k-1}^{4t} + y_{k^*}^{4t} + y_{k_*-1}^{4t} \right) \leq 2^{4t} R^{2t} (r^*)^2.$$

De la même façon on majore le second terme de droite dans l'expression (3.26). D'où le Lemme 3.5.2. \square

Lemme 3.5.3 *Le temps d'arrêt τ_H défini dans (3.2) satisfait la propriété suivante, pour $H = nh$*

$$\mathbf{P}_S(\tau_H > n) \leq C_1 (Rh)^{2t\beta}$$

où C_1 est défini dans (3.26).

DÉMONSTRATION: En tenant compte du fait que $\tau(S) \leq 1$ on obtient

$$\begin{aligned} \mathbf{P}_S(\tau_H > n) &= \mathbf{P}_S\left(\frac{1}{nh} \sum_{k=1}^n Q(u_k) y_{k-1}^2 < \frac{H}{nh}\right) \\ &= \mathbf{P}_S\left(\frac{1}{\tau(S)} \int_{-1}^1 Q(u) du + \Delta_n(Q, h) < 1\right) \\ &= \mathbf{P}_S\left(\Delta_n(Q, h) < 1 - \frac{2}{\tau(S)}\right) \\ &\leq \mathbf{P}_S(|\Delta_n(Q, h)| > 1) \leq \mathbf{E}_S \Delta_n^{2t}(Q, h) \leq C_1 R^{2t} h^{2t\beta}, \end{aligned}$$

cette dernière inégalité provient du Lemme 3.5.2 \square

Afin de prouver le Lemme 3.5.5, nous avons besoin du Lemme suivant montré dans Liptser et Shiryaev [1978] p.234-235.

Lemme 3.5.4 *Soit le processus de Wiener $W = (W_t, \mathcal{F}_t), t \geq 0$, donné sur un espace probabilisé et soit un processus aléatoire $f = (f_t, \mathcal{F}_t), t \geq 0$, tel que :*

$$(1) \quad P\left(\int_0^T f_t^2 dt < \infty\right) = 1, \quad 0 < T < \infty,$$

$$(2) \quad P\left(\int_0^\infty f_t^2 dt = \infty\right) = 1.$$

Alors le processus aléatoire $z = (z_s, \Gamma_s), s \geq 0$, avec $z_s = \int_0^{\tau_s} f_t dW_t, \Gamma_s = \mathcal{F}_{\tau_s}$, où $\tau_s = \inf(t : \int_0^t f_u^2 du > s)$, est un processus de Wiener.

Lemme 3.5.5 *Pour tout $z \geq 2$,*

$$\mathbf{P}_S(|\zeta_H(h)| > z) \leq 2e^{-z^2/8}. \quad (3.30)$$

DÉMONSTRATION: Le mouvement brownien $(W_t)_{t \geq 0}$ est un processus stochastique dont les accroissements disjoints sont indépendants et tel que $W_{t+s} - W_t$ suit une loi normale de moyenne nulle et de variance s . Donc dans notre cas on peut écrire

$$\xi_k = W_k - W_{k-1} \sim \mathcal{N}(0, 1).$$

On rappelle que

$$\zeta_H = \frac{1}{\sqrt{H}} \left(\sum_{j=1}^{\tau_H-1} Q(u_j) y_{j-1} \xi_j + \alpha_H Q(u_{\tau_H}) y_{\tau_H-1} \xi_{\tau_H} \right) \mathbf{1}_{(A_n \geq H)}.$$

Ainsi

$$\mathbf{P}_S(|\zeta_H| > z \mathbf{1}_{(A_n \geq H)}) = \mathbf{P}_S(|\zeta_H| > z, \mathbf{1}_{(A_n \geq H)}) = \mathbf{P}_S(|\tilde{\zeta}_H| > z, \mathbf{1}_{(A_n \geq H)}),$$

où

$$\tilde{\zeta}_H(h) = \frac{1}{\sqrt{H}} \left(\sum_{k=1}^{\tilde{\tau}_H-1} \delta_k \xi_k + \tilde{\alpha}_{\tilde{\tau}_H} \delta_{\tilde{\tau}_H} \xi_{\tilde{\tau}_H} \right),$$

et

$$\sum_{k=1}^{\tilde{\tau}_H-1} \delta_k^2 + \tilde{\alpha}_{\tilde{\tau}_H}^2 \delta_{\tilde{\tau}_H}^2 = H,$$

avec $\delta_k = Q(u_k) y_{k-1} \mathbf{1}_{(k \leq k^*)} + \mathbf{1}_{(k > k^*)}$ et

$$\tilde{\tau}_H = \inf \left\{ k \geq 1 : \sum_{j=1}^k \delta_j^2 \geq H \right\}.$$

On peut voir que

$$\mathbf{P}_S(|\tilde{\zeta}_H(h)| > z) = \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \left| \int_0^{\tilde{\tau}_H} f_t dW_t \right| > z \right),$$

où

$$f_t = \sum_{j=1}^{\infty} \delta'_j \mathbf{1}_{[j-1, j]}(t)$$

avec

$$\delta'_j = \begin{cases} \delta_j & j < \tilde{\tau}_H \\ \tilde{\alpha}_{\tilde{\tau}_H} \delta_{\tilde{\tau}_H} & j = \tilde{\tau}_H \\ 0 & j > \tilde{\tau}_H. \end{cases} \quad (3.31)$$

En effet,

$$\begin{aligned} \int_0^{\tilde{\tau}_H} f_t dw_t &= \sum_{j=1}^{\tilde{\tau}_H} \int_{j-1}^j f_t dw_t = \sum_{j=1}^{\tilde{\tau}_H} \delta'_j [w_j - w_{j-1}] \\ &= \sum_{j=1}^{\tilde{\tau}_H} \delta'_j \xi_j = \sum_{j=1}^{\tilde{\tau}_H-1} \delta_j \xi_j + \tilde{\alpha}_{\tilde{\tau}_H} \delta_{\tilde{\tau}_H} \xi_{\tilde{\tau}_H}. \end{aligned}$$

En posant

$$g_t = \sum_{j=1}^{\infty} \delta_j'' \mathbf{1}_{[j-1, j]}(t)$$

avec

$$\delta_j'' = \begin{cases} \delta_j & j < \tilde{\tau}_H \\ \sqrt{\tilde{\alpha}_{\tilde{\tau}_H}} \delta_{\tilde{\tau}_H} & j = \tilde{\tau}_H \\ 0 & j > \tilde{\tau}_H, \end{cases} \quad (3.32)$$

il vient,

$$\begin{aligned} \int_0^{\tilde{\tau}_H} g_t^2 dt &= \sum_{j=1}^{\tilde{\tau}_H} \int_{j-1}^j g_t^2 dt \\ &= \sum_{j=1}^{\tilde{\tau}_H-1} \delta_j^2 + \tilde{\alpha}_{\tilde{\tau}_H} \delta_{\tilde{\tau}_H}^2 = H. \end{aligned}$$

Par le lemme 3.5.4, on obtient

$$\eta = \frac{1}{\sqrt{H}} \int_0^{\tilde{\tau}_H} g_t dW_t \sim \mathcal{N}(0, 1).$$

Or,

$$\begin{aligned} & \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \left| \int_0^{\tilde{\tau}_H} f_t dW_t \right| > z \right) \\ & \leq \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \left| \int_0^{\tilde{\tau}_H} g_t dW_t \right| > \frac{z}{2} \right) + \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \left| \int_0^{\tilde{\tau}_H} (f_t - g_t) dW_t \right| > \frac{z}{2} \right) \\ & \leq \mathbf{P}_S \left(|\eta| > \frac{z}{2} \right) + \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \left| \sqrt{\alpha_{\tilde{\tau}_H}} - \alpha_{\tilde{\tau}_H} \right| |\delta_{\tilde{\tau}_H} \xi_{\tilde{\tau}_H}| > \frac{z}{2} \right) \\ & \leq \mathbf{P}_S \left(|\eta| > \frac{z}{2} \right) + \mathbf{P}_S \left(\frac{1}{\sqrt{H}} \sqrt{\alpha_{\tilde{\tau}_H}} |\delta_{\tilde{\tau}_H}| |\xi_{\tilde{\tau}_H}| > \frac{z}{2} \right) \\ & = \mathbf{P}_S \left(|\eta| > \frac{z}{2} \right) + \mathbf{P}_S \left(\frac{1}{H} \alpha_{\tilde{\tau}_H} \delta_{\tilde{\tau}_H}^2 \xi_{\tilde{\tau}_H}^2 > \frac{z^2}{4} \right) \\ & \leq \mathbf{P}_S \left(|\eta| > \frac{z}{2} \right) + \mathbf{P}_S \left(\xi_{\tilde{\tau}_H}^2 > \frac{z^2}{4} \right). \end{aligned} \quad (3.33)$$

Comme η est une variable aléatoire gaussienne standard on peut écrire pour tout $z \geq 2$

$$\begin{aligned} \mathbf{P}_S \left(|\eta| > \frac{z}{2} \right) &= \sqrt{\frac{2}{\pi}} \int_{z/2}^{+\infty} e^{-t^2/2} dt \\ &\leq \sqrt{\frac{2}{\pi}} \int_{z/2}^{+\infty} t e^{-t^2/2} dt = \sqrt{\frac{2}{\pi}} e^{-z^2/8}. \end{aligned}$$

On peut écrire le second terme de (3.33) comme

$$\begin{aligned} \mathbf{P}_S \left(\xi_{\tilde{\tau}_H}^2 > \frac{z^2}{4} \right) &= \sum_{l=1}^{+\infty} \mathbf{P}_S \left(\xi_l^2 > \frac{z^2}{4}, \tilde{\tau}_H = l \right) \\ &= \sum_{l=1}^{+\infty} \mathbf{P}_S \left(\xi_l^2 > \frac{z^2}{4}, \sum_{j=1}^{l-1} \delta_j^2 < H, \sum_{j=1}^l \delta_j^2 \geq H \right) \\ &= \sum_{l=1}^{+\infty} \mathbf{P}_S \left(|\xi_l| > \frac{z}{2} \right) \mathbf{P}_S(\tilde{\tau}_H = l) \\ &\leq \sqrt{\frac{2}{\pi}} e^{-z^2/8} \sum_{l=1}^{+\infty} \mathbf{P}_S(\tilde{\tau}_H = l) = \sqrt{\frac{2}{\pi}} e^{-z^2/8}. \end{aligned}$$

Donc pour tout $z \geq 2$, (3.33) implique

$$\mathbf{P}_S \left(\frac{1}{\sqrt{nh}} \left| \int_0^{\tilde{\tau}_H} f_t dw_t \right| > z \right) \leq 2 \sqrt{\frac{2}{\pi}} e^{-z^2/8}.$$

□

Lemme 3.5.6 [Helland , 1981, pp. 80-82]

Soient $(u_{k,n})_{1 \leq k \leq n}$ une "martingale difference" définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbf{P})$ et des filtrations $\{\mathcal{F}_{k,n}, k \in \mathbb{N}\}$ de \mathcal{F} , $n \in \mathbb{N}^*$ telles que $u_{k,n}$ est $\mathcal{F}_{k,n}$ -mesurable. On suppose que les deux conditions suivantes sont satisfaites :

$$\sum_{k=1}^n \mathbf{E}(u_{k,n}^2 \mathbf{1}_{(|u_{k,n}| > \varepsilon)} | \mathcal{F}_{k-1,n}) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0, \quad \text{pour tout } \varepsilon > 0,$$

$$\sum_{k=1}^n \mathbf{E}(u_{k,n}^2 | \mathcal{F}_{k-1,n}) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 1.$$

Alors

$$\sum_{k=1}^n u_{k,n} \implies \mathcal{N}(0, 1).$$

Conclusion

- Dans le cas non adaptatif, on a développé l'approche robuste pour les modèles autorégressifs. On a aussi construit une procédure minimax pour trouver une vitesse de convergence minimax pour les classes Höldériennes fortes. Et on a montré que l'estimateur à noyau rectangulaire est asymptotiquement efficace.
- Puis on a suivi une procédure adaptative sur la base des estimateurs séquentiels qui nous fournit la vitesse de convergence adaptative.
- Enfin tous les résultats obtenus sont illustrés par les simulations numériques effectuées.

Chapitre 4

Simulations numériques

4.1 Cas non adaptatif

4.1.1 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 $\beta/1000 \leq \delta \leq 1000/\beta$.

On estime $S(z_0)$ par le même estimateur considéré au chapitre 2 construit à l'aide des observations $\{y_i, i = 1, \dots, n\}$ avec différents noyaux :

- notre estimateur à noyau

$$\hat{S}_n = (1/A_n) \sum_{k=1}^n Q \left(\frac{x_k - z_0}{h_n} \right) y_{k-1} y_k \mathbf{1}_{(A_n \geq a^*)},$$

- le même estimateur avec le noyau d'Epanechnikov

$$\tilde{S}_n = (1/B_n) \sum_{k=1}^n K \left(\frac{x_k - z_0}{h_n} \right) y_{k-1} y_k \mathbf{1}_{(B_n \geq a^*)},$$

- le troisième avec le noyau gaussien

$$S_n^* = (1/C_n) \sum_{k=1}^n T \left(\frac{x_k - z_0}{h_n} \right) y_{k-1} y_k \mathbf{1}_{(C_n \geq a^*)},$$

où

$$x_k = \frac{k}{n}, \quad h_n = n^{-1/(2\beta+1)}, \quad a^* = \kappa_n n h_n \quad \text{et} \quad \kappa_n = 1/\ln n,$$

$$Q = \mathbf{1}_{[-1;1]}, \quad K(u) = \frac{3}{4} \max(0, 1 - u^2) \mathbf{1}_{[-1,1]}(u) \quad \text{et} \quad T(u) = 1/\sqrt{2\pi} \exp(-u^2/2),$$

$$A_n = \sum_{k=1}^n Q \left(\frac{x_k - z_0}{h_n} \right) y_{k-1}^2,$$

$$B_n = \sum_{k=1}^n K \left(\frac{x_k - z_0}{h_n} \right) y_{k-1}^2,$$

et

$$C_n = \sum_{k=1}^n T \left(\frac{x_k - z_0}{h_n} \right) y_{k-1}^2.$$

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 . Ici on calcule pour chaque estimateur la quantité $\frac{\varphi_n}{M} \sum_{k=1}^M |\hat{S}_n(z_0) - S(z_0)|$. On a choisi $z_0 = 0.8$ et $\beta = 1.30$. Précisons aussi que pour différentes valeurs de n , les simulations sont réinitialisées et donc ne courent pas sur les mêmes premiers aléas.

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

n	\hat{S}_n	\tilde{S}_n	S_n^*
10	0.533	0.454	0.326
100	0.581	0.605	0.512
1000	0.525	0.607	0.428
5000	0.536	0.551	0.410
10000	0.584	0.639	0.452

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

n	\hat{S}_n	\tilde{S}_n	S_n^*
10	0.608	0.604	0.400
100	0.642	0.672	0.521
1000	0.588	0.639	0.475
5000	0.567	0.612	0.432
10000	0.589	0.564	0.395

Pour des variables aléatoires $\{\xi_i, i = 1, \dots, n\}$ 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	S_n^*
10	0.559	0.389	0.215
100	0.597	0.636	0.489
1000	0.557	0.599	0.458
5000	0.539	0.597	0.395
10000	0.573	0.572	0.423

4.1.2 Programmes

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

- la fonction "S" :

```
function[y] = S(x, z0, beta1)
y = (1/1000) * (abs(x - z0)) ^ beta1
endfunction
```

- les fonctions noyaux considérées :

```
function[y] = noyau(x)
if((x > 1)|(x < -1)) then y = 0
else y = 1
end
endfunction
```

```
function[y] = noyau1(x)
y = 0.75 * max(0, 1 - x ^ 2);
endfunction
```

```
function[y] = noyau2(x)
y = 1/sqrt(2 * %pi) * exp(-x ^ 2/2);
endfunction
```

- la procédure donnant pour chaque estimateur le calcul de la quantité

$$\frac{\varphi_n}{M} \sum_{k=1}^M |\hat{S}_n(z_0) - S(z_0)|, \text{ pour des bruits gaussiens :}$$

```
n = input("Rentrer la valeur de n :");
z0 = 0.8; M = 200; beta1 = 1.30 : ";
y0 = 0; X0 = linspace(0, 1, n + 1); X = X0(2 : n + 1);
phi = n ^ (beta1/(2 * beta1 + 1)); h = n ^ (-1/(2 * beta1 + 1));
kappa = 1/log(n); a ^ * = kappa * n * h;
```



```

for simulation = 1 : M
XI = grand(1, n, 'nor', 0, 1);
Y(1) = S(X(1), z0, beta1) * y0 + XI(1);
for k = 2 : n
Y(k) = S(X(k), z0, beta1) * Y(k - 1) + XI(k);
end;

S0(simulation) = 0;
S1(simulation) = 0;
S2(simulation) = 0;
An(simulation) = 0;
Bn(simulation) = 0;
Cn(simulation) = 0;

S0(simulation) = S0(simulation) + noyau((X(1) - z0)/h) * y0 * Y(1);
S1(simulation) = S1(simulation) + noyau1((X(1) - z0)/h) * y0 * Y(1);
S2(simulation) = S2(simulation) + noyau2((X(1) - z0)/h) * y0 * Y(1);

An(simulation) = An(simulation) + noyau((X(k) - z0)/h) * y0 ^ 2;
Bn(simulation) = Bn(simulation) + noyau1((X(k) - z0)/h) * y0 ^ 2;
Cn(simulation) = Cn(simulation) + noyau2((X(k) - z0)/h) * y0 ^ 2;

for k = 2 : n
S0(simulation) = S0(simulation) + noyau((X(k) - z0)/h) * Y(k - 1) * Y(k);
An(simulation) = An(simulation) + noyau((X(k) - z0)/h) * Y(k - 1) ^ 2;
S1(simulation) = S1(simulation) + noyau1((X(k) - z0)/h) * Y(k - 1) * Y(k);
Bn(simulation) = Bn(simulation) + noyau1((X(k) - z0)/h) * Y(k - 1) ^ 2;
S2(simulation) = S2(simulation) + noyau2((X(k) - z0)/h) * Y(k - 1) * Y(k);
Cn(simulation) = An(simulation) + noyau2((X(k) - z0)/h) * Y(k - 1) ^ 2;
end;(fin de la boucle k=2 :n)

if(An(simulation) >= a ^ *) then S0(simulation) = S0(simulation)/An(simulation);
else S0(simulation) = 0;
end;

if(Bn(simulation) >= a ^ *) then S1(simulation) = S1(simulation)/Bn(simulation);
else S1(simulation) = 0;
end;

if(Cn(simulation) >= a ^ *) then S2(simulation) = S2(simulation)/Cn(simulation);
else S2(simulation) = 0;

```

end;

end; (fin de la boucle k=1 :M)

*Risque1 = phi * mean(abs(S0))*

*Risque2 = phi * mean(abs(S1))*

*Risque3 = phi * mean(abs(S2))*

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 réduites.

4.2 Cas adaptatif

4.2.1 Résultats

On illustre les résultats obtenus au Chapitre 3. 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 $S(x) = |x - z_0|^\beta$. On vérifie que cette fonction appartient à $\mathcal{H}^{(\beta)}(z_0, K, \varepsilon)$ lorsque $K \geq 1$. On prend toujours $z_0 = 0.8$, puis $\beta_* = 0.6$ comme valeur inférieure de la régularité et $\beta^* = 0.8$ comme valeur supérieure.

On a simulé n données à partir de la fonction $S(x) = |x - z_0|^\beta$ pour $\beta = 0.7$. On a obtenu une estimation en construisant l'estimateur \hat{S}_n défini en (3.10) par la procédure de Lepskiï qui nous fournit la fenêtre optimale par l'indice \hat{k} défini en (3.9).

En faisant varier le nombre d'observations n , on obtient différentes estimations répertoriées dans le tableau suivant :

n	1000	5000	10000	30000	50000	100000
\hat{S}_n	0.180	0.086	0.103	0.090	0.039	0.05

En effectuant ces simulations on a remarqué que l'indice \hat{k} donné par la procédure de Lepskiï valait toujours $m - 1$ qui est le nombre de points sur la grille $[\beta_*, \beta^*]$ hormis les extrémités. Cette particularité était attendue puisqu'il faut a priori un très grand nombre d'observations pour que cet indice soit strictement inférieur à $m - 1$ d'après la définition (3.9).

4.2.2 Programmes

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

– la fonction "S" :

```
function[y] = S(x)
y = (abs(x - z0)) ^ beta1
endfunction
```

– les noyaux et quelques fonctions auxiliaires :

```
function[y] = noyau(x)
if((x > 1)|(x < -1)) then y = 0
else y = 1
end
endfunction
```

```
function[y] = D(n)
y = n/log(n);
endfunction
```

```
function[y] = H(beta1);
y = (1/D(n)) ^ (1/(2 * beta1 + 1));
endfunction
```

```
function[y] = N(beta1);
y = (D(n)) ^ (beta1/(2 * beta1 + 1));
endfunction
```

```
function[y] = betaf(k);
y = betainf + (k/m) * (betasup - betainf);
endfunction
```

– la procédure donnant l'indice optimale \hat{k} et l'estimateur séquentiel \hat{S}_n :

```
n = input("entrer la valeur de n :");
M = 1; z0 = 0.8; epsilon = 0.2; beta1 = 0.7;(le beta choisi dans la fonction S)
betainf = 0.6; betasup = 0.8; y0 = 0;
A = zeros(1, n); m = floor(log(D(n))) + 1;
X0 = linspace(0, 1, n + 1); X = X0(2 : n + 1);
lambda0 = (1 + epsilon) + exp(1) * sqrt(4 + 4/(2 * betainf + 1));

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

```
Y(1) = S(X(1)) * y0 + XI(1);
```

```
for k = 2 : n
Y(k) = S(X(k)) * Y(k - 1) + XI(k);
end
```

– le calcul de \hat{k} :

```
SS = zeros(1, m + 1);
lambda = lambda0;
```

```
for kk = 0 : m
h0 = H(betaf(kk));
H0 = n * h0;
```

– calcul de τ_H :

```
tauH = n;
A(1) = noyau((X(1) - z0)/h0) * y0 ^ 2;
if(A(1) >= H0) then tauH = 1;
alphaH = H0/A(1);
else
```

```
for j = 2 : n
A(j) = A(j - 1) + noyau((X(j) - z0)/h0) * Y(j - 1) ^ 2;
if(A(j) >= H0) then tauH = j;
alphaH = (H0 - A(j - 1))/(A(j) - A(j - 1));
break; end;
end;
end;
```

```
if (tauH <= 1) then S0 = 0
else S0 = 0;
```

```
for k = 2 : tauH - 1
S0 = S0 + noyau((X(k) - z0)/h0) * Y(k - 1) * Y(k);
end;
```

```
S0 = S0 + alphaH * noyau((X(tauH) - z0)/h0) * Y(tauH - 1) * Y(tauH);
end (condition sur  $\tau_H$ )
```

```
S0 = S0/H0;
SS(kk + 1) = S0;
```

```
for jj = 0 : m
vals = zeros(1, jj + 1);

for k = 0 : jj
vals(k + 1) = abs(SS(jj + 1) - SS(k + 1)) - lambda/N(betaf(k + 1));
end; fin de calcul

W(jj + 1) = max(vals);
end; (fin de boucle pour les W)
end (fin de la boucle kk=1 :m)

kchapeau = m - 1;

for j = 0 : m
if(W(j + 1) >= lambda/N(betaf(j))) then kchapeau = j - 1;
break;
end; end;

Estimation = SS(kchapeau + 1)
```


Bibliographie

- Arkoun, O. and Pergamenchtchikov, S. : Sequential estimators of nonparametric estimation for an autoregressive model. (En préparation)
- Arkoun, O. (2009) : Sequential adaptive estimators in nonparametric autoregressive models (Soumis à la publication)
- Arkoun, O. and Pergamenchtchikov, S. (2008) : Nonparametric Estimation for an Autoregressive Model. Vestnik of Tomsk State University, Ser. *Mathematics et Mechanics* **2** (3), 20 - 30.
- Belitser, E. (2000a) : Local minimax pointwise estimation of a multivariate density, *Statisti. Nederletica* **54** (3), 351-365.
- Belitser, E. (2000b) : Recursive estimation of a drifted autoregressive parameter, *The annals of Statistics* **26** (3), 860-870.
- Billingsley, P. (1999) : *Convergence of probability measures*, Wiley Series in Probability and Statistics : Probability et Statistics, John Wiley & Sons Inc., New York, A Wiley-Interscience Publication.
- Borisov, V.Z. and Konev, V.V. (1977) : Sequential Estimation of Parameters of Discrete Processes, *Automat. and Remote control* **10**, 58-64.
- Brua, J.-Y. (2008) : Estimation non paramétrique pour des modèles de diffusion et de régression. Université Louis Pasteur-Strasbourg I [tel-00338286 - version 1].
- Brua, J.-Y. (2009a) : Asymptotically efficient estimators in nonparametric heteroscedastic regression models, *Statist. Methodol.* **6** (1), 47-60.
- Brua, J.-Y. (2009b) : Adaptive estimators in nonparametric heteroscedastic regression models, *J. Nonparametr. Statist.* **21** (8), 991-1002.
- Dahlhaus, R. (1996a) : On the Kullback-Leibler information divergence of locally stationary processes, *Stochastic Process. Appl.* **62** (1), 139–168.
- Dahlhaus, R. (1996b) : Maximum likelihood estimation and model selection for locally stationary processes, *J. Nonparametr. Statist.* **6** (2-3), 171–191.

- Donoho, D.L. and Liu, R.C. (1991) : Geometrizing rates of convergence. II, III, *Ann. Statist.* **19** (2), 668-701.
- 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. and Picard, D. (1995) : Wavelet shrinkage : asymptopia?, *J. Roy. Statist. Soc. Ser. B* **57** (2), 301–369.
- Efroïmovich, S. Yu. (1985) : Nonparametric estimation of a density of unknown smoothness, *Theory Probab. Appl.* **30** (3), 557–568.
- Efroïmovich, S. (1996) : On nonparametric regression for IID observations in a general setting, *Ann. Statist.* **24** (3), 1125–1144.
- Efroïmovich, S. : *Nonparametric Curve Estimation. Methods, Theory and Applications*, Springer, Berlin, New York, 1999.
- Efroïmovich, S. (2007) : Sequential design and estimation in heteroscedastic nonparametric regression, *Sequential Anal.* **26** (1), 3–25.
- Efroïmovich, S. and Pinsker, M. (1996) : Sharp-optimal et adaptive estimation for heteroscedastic nonparametric regression, *Statist. Sinica* **6** (4), 925–942.
- Efroïmovich, S. Yu. and Pinsker, M. S. (1984) : A self-training algorithm for nonparametric filtering, *Automat. Remote Control* **11**, 58–65.
- Embrechts, P., Kluppelberg, C. and Mikosch, T : Modelling extremal events. *Springer* 1997.
- Emery, M., Nemirovskiï, A. S. and Voiculescu, D. : Lectures on probability theory and statistic. Lectures from the 28th Summer School on Probability Theory held in Saint-Flour, August 17-September 3, 1998. *Lecture Notes in Mathematics, 1738*. Springer-Verlag, Berlin, 2000.
- Fourdrinier, D., Konev, V.V. and Pergamenchtchikov, S. (2009) : Truncated Sequential Estimation of the Parameter of a First Order Autoregressive Process with Dependent Noises, *Mathematical Methods of Statistics* **18** (1), 43-58.
- Fourdrinier, D. and 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. and Pergamenshchikov, S. (2001) : Sequential nonparametric adaptive estimation of the drift coefficient in diffusion processes, *Math. Methods Statist.* **10** (3), 316–330.
- Galtchouk, L. and 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. and Pergamenshchikov, S. (2005b) : Nonparametric sequential minimax estimation of the drift coefficient in diffusion processes, *Sequential Anal.* **24** (3), 303–330.
- Galtchouk, L. and Pergamenshchikov, S. (2006a) : Asymptotically efficient estimates for nonparametric regression models, *Statist. Probab. Lett.* **76** (8), 852–860.
- Galtchouk, L. and Pergamenshchikov, S. (2006b) : Asymptotically efficient sequential kernel estimates of the drift coefficient in ergodic diffusion processes, *Stat. Inference Stoch. Process.* **9** (1), 1–16.
- Goldenshluger, A. and Nemirovski, A. (1997) : On spatially adaptive estimation of nonparametric regression, *Math. Methods Statist.* **6** (2), 135–170.
- Goldfeld, S. and Quandt, R. (1972) : *Nonlinear Methods in Econometrics*, North-Hollet, Amsterdam, London.
- Golubev, G. K. (1992) : Asymptotically minimax estimation of a regression function in an additive model, *Problems Inform. Trans.* **28**, 3–15.
- Golubev, G. K. and Levit, B. Y. (1996) : Asymptotically efficient estimation for analytic distributions, *Math. Methods Statist.* **5** (3), 357–368.
- Golubev, G. K., Levit, B. Y. and Tsybakov, A. B. (1996) : Asymptotically efficient estimation of analytic functions in Gaussian noise, *Bernoulli* **2** (2), 167–181.
- Golubev, G. K. and Nussbaum, M. (1990b) : Nonparametric estimation of a regression function in L^2 , *Problems Inform. Trans.* **26** (3), 213–225.
- Golubev, G. K. and Nussbaum, M. (1992) : Adaptive spline estimates in a nonparametric regression model, *Theory Probab. Appl.* **37** (3), 521–529.
- Guerre, E. and 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.
- Guyon, J. (2007) : Limit Theorems for bifurcating Markov chains. Application to the detection of cellular aging, *The Annals of Applied Probability.* **17** (5/6), 1538–1569.

- Härdle, W. et Marron, J. S. (1985) : Optimal bandwidth selection in nonparametric regression function estimation, *Ann. Statist.* **13** (4), 1465–1481.
- Helland, I. S. (1981) : Central limit theorems for martingales with discrete or continuous time. *Scet. J. Statist.* **9** (2), 79–94.
- Ibragimov, I. A. and 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. and Has'minskiĭ, R. Z. : *Statistical Estimation : Asymptotic Theory*, Springer, Berlin, New York, 1981.
- Ibragimov, I. A. and 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. and 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).
- 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. and Nussbaum, M. (1999) : The asymptotic minimax constant for sup-norm loss in nonparametric density estimation, *Bernoulli* **5** (6), 1099–1118.
- 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. and Spokoiny, V. G. (1997) : Optimal pointwise adaptive methods in nonparametric estimation, *Ann. Statist.* **25** (6), 2512–2546.
- Liptser, R. S. and Shiryaev, A. N. : *Statistics of random processes. I and II*, Springer-Verlag, New York, 1978.
- Moulines et al. (2005) : On recursive estimation for time varying autoregressive processes, *The annals of Statistics* **33** (6), 2610–2654.

- Nemirovskii, A. S. (1986) : Nonparametric estimation of smooth regression functions, *J. Comput. Systems Sci.* **23** (6), 1–11.
- Rebolledo, R. (1980) : Central limit theorems for local martingales, *Z. Wahrsch. Verw. Gebiete* **51** (3), 269–286.
- Sacks, J. and Strawderman, W. (1982) : Improvements on linear minimax estimates, *Statistical decision theory and related topics, III, Vol. 2*, 287–304.
- Sacks, J. and Ylvisaker, D. (1978) : Linear estimation for approximately linear models, *Ann. Statist.* **6** (5), 1122–1137.
- Sacks, J. and Ylvisaker, D. (1981) : Asymptotically optimum kernels for density estimation at a point, *Ann. Statist.* **9** (2), 334–346.
- Shiryaev, A.N. : Probability, Second Edition, *Springer*, 1996.
- Spokoiny, V. G. (2000) : Adaptive drift estimation for nonparametric diffusion model, *Ann. Statist.* **28** (3), 815–836.
- 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.